

Probabilistic Verification via Category Theory:

Categorical Generalization of

Fair Simulation and Ranking Function by

Kleisli Coalgebras, and Its Concretization

圏論による確率的検証：

クライスリ圏の余代数による

公平模倣とランキング関数の圏論的一般化と具体化

by

Natsuki Urabe

ト部夏木

A Doctor Thesis

博士論文

Submitted to

the Graduate School of the University of Tokyo

on

in Partial Fulfillment of the Requirements

for the Degree of Doctor of Information Science and

Technology

in Computer Science

Thesis Supervisor: Naoki Kobayashi 小林直樹

Professor of Computer Science

## ABSTRACT

*Formal verification* is a technique for quality assurance that uses mathematical methods to prove that a system satisfies a given property. In this thesis, we will focus on *quantitative verification*. As computers become ubiquitous in the world, computer systems tend to involve quantitative behaviors like probabilities or energy consumptions. For such quantitative systems, we can consider quantitative properties like “the probability where the program does not terminate is at most 10%,” or “the amount of gas consumed by the system is at most 1.0L.” In quantitative verifications, we prove quantitative properties for quantitative systems. Compared to that about qualitative verifications for systems like nondeterministic systems, the number of studies about quantitative verification is relatively small. This thesis aims to introduce notions that we can use for quantitative verification. We will mainly focus on probabilistic systems in this thesis.

We introduce notions for *quantitative* verifications by modifying existing notions used for *qualitative* verifications. In this procedure, the *category theory* and the notion of *coalgebra* play essential roles. Category theory is a general and abstract mathematical theory which is understood as a *theory of structures*. Coalgebra is one of the fundamental notions in category theory that can give general characterizations for dynamics of transition systems. Using category theory and the theory of coalgebras, we first *generalize* the notion for qualitative verification by extracting its categorical essence. We then prove the correctness of the generalized notion in terms of category theory. The categorical generalization is *concretized* for quantitative systems and a notion for quantitative verifications is obtained. The correctness of the obtained notion is automatically inherited from the categorical level.

We have applied the above “generalize-concretize” scheme to two existing notions used for verifications of nondeterministic systems: *fair simulation* and *ranking function*.

*Simulation* is a notion commonly used for proving behavioral inclusion between transition systems. More concretely, a simulation from one system to another system implies the behavior of the former system is included in that of the latter system. Therefore, by constructing a simulation from some system to another system whose safety (i.e. that the system does not exhibit a bad behavior) is already known, we can conclude safety of the former system, because all of its behaviors are included in those of the latter, which never include bad events. Similarly, we can use simulation for proving liveness (i.e. that the system does exhibit some good behavior) by constructing a simulation from a system whose liveness is known to the system in question. Fair simulation is a simulation notion for *Büchi automata*, transition systems that accept infinite words according to the *Büchi condition*. We say that an infinite path on an automaton satisfies the Büchi condition if it visits accepting states infinitely many times.

Via category theory, we induce a fair simulation notion for *probabilistic Büchi automata*, a quantitative variant of Büchi automata. To this end, we first have to categorically characterize behaviors of Büchi automata and probabilistic Büchi automata in a unified manner. We will introduce two characterizations. One of them is used for categorically generalizing the notion of fair simulation. We use the other characterization in the correctness proof of the first characterization. We expect that the latter characterization also serves as a basis for categorically generalizing other simulation notions for Büchi automata than fair simulation in the future.

Ranking function, the second notion our framework is applied to, is commonly used to prove termination of nondeterministic systems. Ranking functions are known to be especially effective for verifications of infinite-state systems like while programs. The key notion in categorically generalizing it was a categorical notion called *corecursive algebra*. We then concretize it for probabilistic systems. In fact, a ranking function-like notion was already known for probabilistic systems under the name of *ranking supermartingale*. It is known to be useful for proving *almost-sure termination*, i.e. that the system terminates in probability 1. However, it turned out that our categorical generalization of ranking function does not instantiate to ranking supermartingale. Instead, we found that it induces two new ranking function-like notions for probabilistic systems. We named them *distribution-valued ranking function* and  $\gamma$ -*scaled submartingale*. Unlike ranking supermartingale, we can use the new ranking function-like notions for *quantitative reasoning*: they give lower bounds for termination probabilities.

For  $\gamma$ -scaled submartingales, we also provide algorithms for finding  $\gamma$ -scaled submartingales for probabilistic programs and give implementations. We found that existing template-based synthesis algorithms for ranking supermartingales can be adapted for  $\gamma$ -scaled submartingales with little modification. We first implemented a linear template-based algorithm. It fixes a linear template for a  $\gamma$ -scaled submartingale, reduces the axioms of  $\gamma$ -scaled submartingales to a linear programming (LP) problem, and solves it using an LP solver. We will compare our algorithm with an existing algorithm proposed by Chatterjee et al. in 2017 for the same purpose, i.e. underapproximating termination probabilities. We have also implemented a polynomial template-based algorithm that reduces the problem to a semi-definite programming (SDP) problem and solves it with an SDP solver. Although this implementation did not work well because of numerical errors caused by the SDP solver, we will give a concrete description of the algorithm and experimental results for the record.

## 論文要旨

形式検証は数学的手法を用いてシステムが期待される性質を満たすことを証明する品質保証の手法である。本論文では、定量的検証に着目する。コンピュータが世界中のあらゆる場所で使われるようになるにつれ、コンピュータシステムが確率やエネルギー消費などの定量的なふるまいを含むようになってきている。このような定量的なシステムに対しては、「プログラムが停止する確率は10%以下」、或いは「システムが消費するガソリンは1.0L以下」といったような定量的な性質を考えることができる。定量的システムが定量的性質を満たすことを証明するのが定量的検証である。非決定的システムのようなシステムに対する定性的検証に比べ、定量的検証に関する研究の数は比較的少ない。本論文では主に確率的システムに焦点を置き、定量的検証に利用できる概念の導入を目指す。

本研究では、定性的検証に用いられる既存の概念をもとに定量的検証のための概念を導出する。この手法においては、圏論、及び余代数の概念が主要な役割を果たす。圏論は一般的かつ抽象的な数学の理論であり、直感的には構造の理論と捉えられる。一方余代数は圏論の主要な概念の一つであり、遷移系のダイナミクスに一般的な特徴づけを与えることができる。圏論と余代数の理論を用い、本研究ではまず定性的検証に用いられる概念をその圏論的な本質を抜き出すことで一般化する。次にその一般化された概念の正しさに圏論的な証明を与える。最後にそれを定量的システムに対して具体化することで定量的検証に有用な概念を導出する。導出された概念の正しさは圏論的なレベルから受け継がれる。

本研究では上述の「一般化・具体化」スキームを公平模倣とランキング関数とよばれる、非決定的システムの検証に使われる2種類の既存の概念に対して適用した。

模倣は、通常遷移系のふるまいの間の包含関係を示すために用いられる概念である。具体的には、あるシステムから別のシステムへ模倣が存在する場合、前者のシステムのふるまいは後者のシステムのふるまいに含まれることが導かれる。したがって、あるシステムから別の安全性（即ち何らかの危険なふるまいをしないこと）がすでに保証されているシステムへの模倣を構成することにより、前者のシステムもその安全性を満たすことを示すことができる。なぜなら前者のシステムのすべてのふるまいが、後者のシステムの、決して危険なものを含まないふるまいに含まれるからである。同様に、ある活性（即ち何らかの望ましいふるまいをすること）がすでに保証されているシステムから別のシステムへの模倣を構成することにより後者のシステムもその活性を満たすことを示すこともできる。公平模倣は、Büchiオートマトンとよばれる、Büchi条件という受理条件に従って無限長文字列を受理する遷移系に対する模倣の概念である。オートマトン状の路が受理状態を無限回訪れる時、その路はBüchi条件を満たす、という。

本研究では圏論を用いて、確率的BüchiオートマトンというBüchiオートマトンの定量的変種に公平模倣の概念を導出する。このためにはまずBüchiオートマトンと確率的Büchiオートマトンのふるまいを統一されたやり方で圏論的に特徴付けする必要がある。ここでは2種類の特徴付けを行う。そのうち1つは公平模倣の概念を圏論的に一般化するために使われる。もう1つの特徴付けは前者の特徴付けの正しさの証明に用いられる。また、将来的には、後者の特徴付けも公平模倣以外のBüchiオートマトンに対する模倣の概念を圏論的に一般化するために用いることができることを期待している。

本研究の枠組みが適用される2つ目の概念であるランキング関数は、通常非決定的なシステムの停止性を検証するために用いられる。ランキング関数はwhileプログラムのよう

な無限個の状態をもつシステムの検証に特に有用であることが知られている。これを圏論的に一般化する上では余再帰的代数とよばれる圏論の概念が重要な役割を果たす。我々は得られたランキング関数の圏論的一般化を確率的システムに対して具体化した。実は確率的システムに対してはランキング優マルチンゲールとよばれるランキング関数のような概念がすでに知られている。ランキング優マルチンゲールを用いると確率的システムが殆ど確実に停止すること、即ち確率1で停止することを検証できる。しかし研究の結果、ランキング優マルチンゲールは、ランキング関数の圏論的一般化を具体化することで得られるものではないことがわかった。一方で代わりに、ランキング関数の圏論的一般化を具体化することで確率的システムに対しランキング関数に似た2つの新しい概念を得ることができた。本研究ではこれらの得られた概念を分布値ランキング優マルチンゲール及び $\gamma$ 縮尺劣マルチンゲールと名付けた。ランキング優マルチンゲールとは異なり、これらの新しいランキング関数的概念は停止確率の下界を与えるため、これらを用いると定量的な議論を行うことができる。

また $\gamma$ 縮尺劣マルチンゲールについては、確率的プログラムに対してこれを見つけるアルゴリズムも提案し、その実装を与える。本研究では、既存の、テンプレートを用いたランキング優マルチンゲールの合成アルゴリズムが少しの改造で $\gamma$ 縮尺劣マルチンゲールの合成にも応用できることがわかった。本研究ではまず線形テンプレートを用いたアルゴリズムを実装した。このアルゴリズムは $\gamma$ 縮尺劣マルチンゲールに対し線形のテンプレートを定め、 $\gamma$ 縮尺劣マルチンゲールの公理を線形計画(LP)の問題に帰着し、それをLPソルバを用いて解く。このアルゴリズムについては、Chatterjeeらにより2017年に提案された同じ目的、即ち停止確率に下界を与えるための既存のアルゴリズムとの比較も行う。また我々は多項式テンプレートを用いた、問題を半正定値計画(SDP)の問題に帰着し、それをSDPソルバを用いて解くアルゴリズムも実装した。この実装はSDPソルバの数値誤差のために正しく動作しなかったが、それでも記録のため、アルゴリズムの説明と実験結果を述べる。

## Acknowledgements

My supervisors Ichiro Hasuo and Naoki Kobayashi gave me helpful comments and suggestions. The chief examiner Masami Hagiya and vice examiners Noboru Kunihiro, Yusuke Miyao, Koki Nishizawa and Akihiko Takano also helped me to improve this thesis. Papers that this thesis is based on are coauthored with Masaki Hara, Ichiro Hasuo, Yuichiro Oyabu, Shunsuke Shimizu, and Toru Takasaka. Takamasa Okudono mentioned me papers about using SDP solvers for verifications.

I was supported by JSPS KAKENHI Grant Number 16J08157, JSPS's Research Fellowship for Young Scientists, and JST ERATO HASUO Metamathematics for Systems Design Project (No. JPMJER1603) during the course of my PhD. Google Travel Grants for PhD Students in Japan and South Korea supported me to attend a conference for presenting the paper that Sections 3.2–3.3 are based on.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Qualitative and Quantitative Verification . . . . .	1
1.2	Categorical Generalization and Concretization . . . . .	2
1.3	Backgrounds . . . . .	4
1.3.1	Coalgebra . . . . .	4
1.3.2	Fixed Point Logic . . . . .	5
1.4	Contributions . . . . .	6
1.5	Organization . . . . .	7
<b>2</b>	<b>Preliminaries</b>	<b>9</b>
2.1	Notations . . . . .	9
2.2	Transition Systems . . . . .	10
2.2.1	Nondeterministic Parity Tree Automaton . . . . .	10
2.2.2	Probabilistic Parity Tree Automaton . . . . .	12
2.2.3	Two-player Game . . . . .	14
2.2.4	Probabilistic Transition System . . . . .	16
2.3	Fixed Point Logic . . . . .	16
2.3.1	Fixed Point Theorems . . . . .	16
2.3.2	Hierarchical Equation System . . . . .	18
2.4	Categorical Preliminaries . . . . .	19
2.4.1	Preliminaries on Basic Category Theories . . . . .	19
2.4.2	Algebra and Coalgebra . . . . .	23
<b>3</b>	<b>Categorical Trace Semantics for Büchi and Parity Automata</b>	<b>26</b>
3.1	Kleisli Approach for Finite and Infinitary Trace Semantics . . . . .	26
3.1.1	Monad and Kleisli Category . . . . .	27
3.1.2	Kleisli Approach . . . . .	28
3.2	Categorical Representation of Büchi Automata . . . . .	31
3.3	Characterization via Logical Fixed Point . . . . .	31
3.4	Characterization via Categorical Fixed Points . . . . .	32
3.4.1	Alternating Fixed Point of Functor . . . . .	32
3.4.2	Lifting $F^+$ and $F^\oplus$ over $\mathcal{Kl}(T)$ . . . . .	34
3.4.3	Decorated Trace Semantics . . . . .	36
3.5	Logical Fixed Point vs. Categorical Fixed Point . . . . .	37
3.6	Extension to Parity Automata . . . . .	40
3.6.1	Categorical Representation of parity Automata . . . . .	40
3.6.2	Characterization via Logical Fixed Point: Parity Case . . . . .	41
3.6.3	Characterization via Categorical Fixed Point: Parity Case . . . . .	41
3.6.4	Logical Fixed Point vs. Categorical Fixed Point: Parity Case . . . . .	43
3.7	Extension to Nondeterministic Parity Tree Automata . . . . .	57
3.7.1	Trace Semantics of NPTA via Logical Fixed Point . . . . .	57
3.7.2	Trace Semantics of NPTA via Categorical Fixed Point . . . . .	57

3.8	Extension to Probabilistic Automata . . . . .	62
3.8.1	Trace Semantics of PPTA via Logical Fixed Point . . . . .	62
3.8.2	Trace Semantics of PPTA via Categorical Fixed Point . . . . .	62
3.9	Conclusion and Related Work . . . . .	65
<b>4</b>	<b>Categorical Fair Simulation</b>	<b>67</b>
4.1	Simulation . . . . .	67
4.2	Kleisli Simulation . . . . .	69
4.3	Kleisli Fair Simulation . . . . .	69
4.3.1	Partially Additive Monad . . . . .	69
4.3.2	Lattice-Theoretic Progress Measures . . . . .	71
4.3.3	Kleisli Fair Simulation with Dividing . . . . .	74
4.3.4	Kleisli Fair Simulation without Dividing . . . . .	79
4.4	Kleisli Fair Simulation for NBTAs . . . . .	82
4.5	Kleisli Fair Simulation for PBTAs . . . . .	86
4.5.1	Kleisli Fair Simulation with Dividing for PBTAs . . . . .	86
4.5.2	Kleisli Fair Simulation without Dividing for PBTAs . . . . .	87
4.6	Conclusion and Related Work . . . . .	91
<b>5</b>	<b>Categorical Ranking Function</b>	<b>93</b>
5.1	Ranking Function . . . . .	93
5.2	Modalities and Fixed-Point Properties . . . . .	94
5.3	Categorical Generalization of Ranking Function . . . . .	95
5.4	Concretization to Reachability Games . . . . .	98
5.5	Concretization to PTSs . . . . .	102
5.5.1	Distribution-valued Ranking Supermartingale . . . . .	103
5.5.2	$\gamma$ -Scaled Submartingale . . . . .	108
5.6	Conclusion and Related Work . . . . .	113
<b>6</b>	<b><math>\gamma</math>-Scaled Submartingale for Probabilistic Programs and its Synthesis</b>	<b>115</b>
6.1	Linear Template-Based Algorithm . . . . .	116
6.1.1	Syntax of Probabilistic Programs . . . . .	116
6.1.2	Problem . . . . .	117
6.1.3	Algorithm . . . . .	119
6.1.4	Implementation . . . . .	123
6.1.5	Experiments I: Probabilistic Programs in the Literature . . . . .	124
6.1.6	Experiments II: Comparison with Existing Work . . . . .	125
6.2	Polynomial Template-Based Algorithm . . . . .	129
6.2.1	Syntax of Probabilistic Programs . . . . .	130
6.2.2	Problem . . . . .	130
6.2.3	Algorithm . . . . .	131
6.2.4	Improvement of Polynomial Template-based Algorithm . . . . .	133
6.2.5	Implementation . . . . .	134
6.2.6	Experiments . . . . .	134
6.2.7	(Failed) Attempt to Remedy the Situation . . . . .	135
6.3	Conclusion and Related Work . . . . .	135
<b>7</b>	<b>Related Work</b>	<b>137</b>
<b>8</b>	<b>Conclusion and Future Work</b>	<b>139</b>
	<b>References</b>	<b>142</b>



# Chapter 1

## Introduction

Because computer systems are omnipresent, ensuring that they work as expected is a very important problem. *Formal verification* is a technique of using mathematical methods for quality assurance of systems. In the most basic setting of formal verification, we focus on *qualitative specifications* of *qualitative systems*. However, in the verification of systems in the real worlds, we are sometimes interested in *quantitative* systems and *quantitative* specifications. For example, some security protocols exhibit probabilistic behaviors (see e.g. [88, 70]). Another example is *cyber-physical systems* like cars, for which we can consider many quantitative specifications like “fuel consumption is no greater than 20.0km/L.”

The main goal of this thesis is to introduce techniques for formal verification of probabilistic systems.

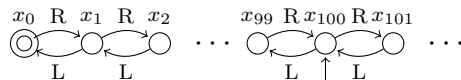
### 1.1 Qualitative and Quantitative Verification

We shall show examples of qualitative and quantitative formal verification.

**Qualitative Model Checking** *Model checking* is one of the most popular approaches for formal verification. A standard framework of model checking is as follows: we first mathematically model the system whose quality we wish to assure. There exist several choices for mathematical models. One possible model is *nondeterministic automata*, a kind of labeled transition system.

For example, suppose that we are given program code shown on the right. It shows a program with a variable  $x$ . It starts with  $x = 100$ , and in each loop, it decreases or increases  $x$  by 1 depending on the input. The system terminates if  $x$  gets 0. We can model the program as the following transition system.

```
x := 100;
while 0 < x do
  y = input();
  if y = 'L' then
    x := x - 1
  else // (if y = 'R')
    x := x + 1
  fi
od
```



(1.1) Figure 1.1: nondeterministic walk

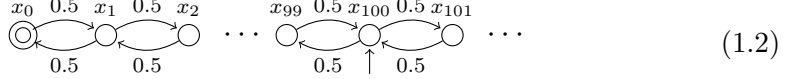
We next mathematically characterize the specification that we wish to check whether the system satisfies or not. For the code in Figure 1.1, we can consider the following specification for example: “if we feed inputs to the program appropriately, we can make the program terminate.” The specification corresponds to a property that “there exists a run from  $x_{100}$  to  $x_0$  in the system (1.1).”

We then check whether the model of the system satisfies the property induced from the specification using mathematical methods.

**Probabilistic Model Checking** An example of a quantitative system is given on the right. It is similar to the one in Figure 1.1, except that whether  $x$  is incremented or decremented is determined in a randomized manner: each of them is chosen with a probability 0.5. The code induces the following *probabilistic transition system*.

```
x := 100;
while 0 < x do
  if prob(0.5) then
    x := x - 1
  else
    x := x + 1
  fi
od
```

Figure 1.2: random walk



For probabilistic systems, we can consider both qualitative and quantitative specifications. An example of a qualitative specification for the code in Figure 1.2 is: “the program terminates with probability 1.” (This property is called *almost-sure termination*.) In contrast, an example of a quantitative specification is: “the program terminates in probability greater than 0.9.” It is not hard to represent these specifications as properties of the probabilistic transition system (1.2).

## 1.2 Categorical Generalization and Concretization

In this thesis, we introduce two novel techniques for probabilistic verification. The techniques are derived with the help of *category theory*.

**Category Theory** Briefly speaking, category theory is a general mathematical theory of *structures*. A *category* consists of a collection of *objects* and a collection of *arrows* between objects. An example of a category is the *category Sets*, whose objects are all the sets and arrows are functions between them. Another example is the *category Meas* of *measurable spaces*, whose objects are all the measurable spaces and arrows are measurable functions between them.

A good point of category theory is that it allows us to capture multiple analogous notions in a unified manner. For example, for two sets  $X$  and  $Y$ , their *product* is a set  $X \times Y := \{(x, y) \mid x \in X, y \in Y\}$ . An analogous notion exists for measurable spaces: for measurable spaces  $(X, \mathfrak{F}_X)$  and  $(Y, \mathfrak{F}_Y)$ , their *product* is commonly defined as a measurable space  $(X \times Y, \mathfrak{F}_{X \times Y})$  where the first component is a product of sets and  $\mathfrak{F}_{X \times Y}$  is the smallest  $\sigma$ -algebra containing  $\{A \times B \mid A \in \mathfrak{F}_X, B \in \mathfrak{F}_Y\}$ . Category theory can describe these two notions of “product” in one definition. This means that once we develop some theory at the categorical level, it freely applies to both sets and measurable spaces.

This good point provides a method to introduce new quantitative verification techniques. It consists of the following two steps. Firstly, making use of the generality of category theory, we *generalize* an existing verification technique for qualitative systems with the help of category theory. We then concretize the generalized verification technique for quantitative systems. This results in a verification technique for the quantitative systems, which is sometimes novel.

An advantage of this “generalize-and-concretize” strategy is that the generality of category theory sometimes help us to understand why the existing qualitative verification technique is sound in a way that allows us to transfer it for probabilistic systems. Another advantage is that a categorically generalized verification technique can induce more than one new verification techniques. Indeed, the categorical generalization of an existing verification technique that we develop in Chapter 6 instantiates to two probabilistic techniques.

Prior to this thesis, this “generalize-and-concretize” strategy has achieved success and induced a novel verification technique called *matrix simulation* [43,

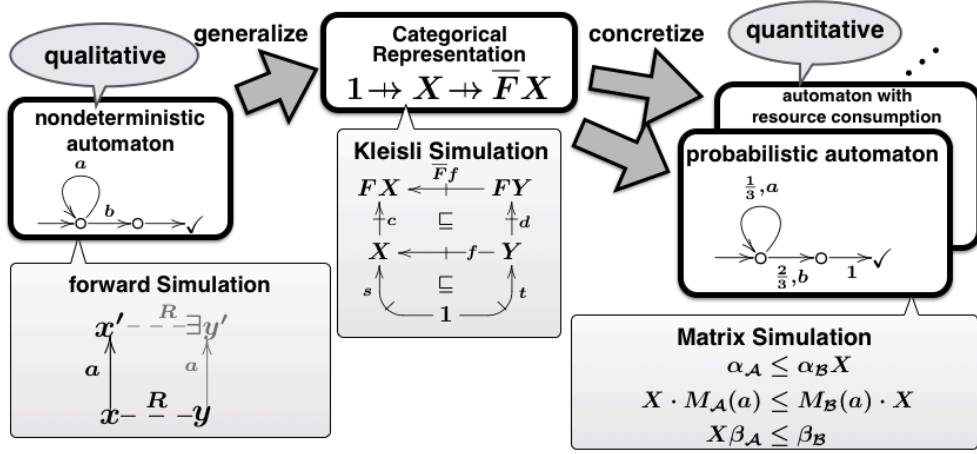


Figure 1.3: generalization and concretization in [43, 111]

111]. We hereby review it to explain the strategy (see also Figure 1.3).

**An Example: Matrix Simulation** *Simulation* is a well-known technique for proving behavioral inclusion between transition systems. In a simulation-based verification, we prove behavioral inclusion by proving that one system can “simulate” the system. We can use simulation for proving a safety property of a system. To do so, it suffices to pick another system that is known to satisfy the property, and show that the former system is simulated by the latter. We can similarly use simulation for proving a liveness property.

Various notions of simulations have been introduced [74, 75, 66, 55, 50, 32]. Among them are *forward and backward simulations* [74] that we can use for proving finite language inclusion between nondeterministic automata. Their definitions are well-explained in terms of two-player games: for example, a forward simulation from a nondeterministic automaton  $\mathcal{A}$  to  $\mathcal{B}$  exists if every transition of  $\mathcal{A}$  can be “simulated” by a transition on  $\mathcal{B}$ .

Using category theory, Hasuo generalized the notions of forward and backward simulations and named them *Kleisli simulations* [43]. In [111], it was concretized for two types of quantitative systems—automata weighted with a semiring  $([0, \infty], +, 0, \times, 1)$ , which model probabilistic systems, and those weighted with a semiring  $([-\infty, \infty], \max, -\infty, +, 0)$ , which model systems with rewards or resource consumption. The concretization resulted in new simulation notions for those weighted automata, which were named *matrix simulation*. The resulting simulation notions are defined as a matrix  $X$  satisfying linear inequalities over the corresponding semirings. They allow us to check a quantitative language inclusion between weighted automata, i.e., that one automaton assigns a smaller weight for each finite word than the other one. The linear inequalities defining a matrix simulation are solvable with numerical methods (for example, it reduces to a linear programming problem for  $([0, \infty], +, 0, \times, 1)$ -weighted automata). Programs searching for a matrix simulation were implemented and evaluated in [111].

In this thesis, we apply the same “generalize-and-concretize” strategy to two existing verification techniques: *fair simulation* and *ranking function*. As a result, we obtain their probabilistic variants that are novel.

### 1.3 Backgrounds

We have explained the overall framework of this thesis in the previous section. In this section, we shall explain the notions that we use throughout this thesis.

#### 1.3.1 Coalgebra

*Coalgebra* is a basic notion in category theory. It is a categorical model of dynamics of state-based systems. Formally, a coalgebra is an arrow of a form  $c: X \rightarrow FX$  where  $F$  is a *functor*, an operation that maps objects to objects and arrows to arrows. Various transition systems are representable as coalgebras. For example, a labeled transition system is modeled as  $c: X \rightarrow \mathbf{A} \times X$ , a coalgebra of a functor  $\mathbf{A} \times (\_)$ . Indeed, we can regard this  $c$  as a transition function of a labeled transition system. Other examples are a deterministic automaton as  $c: X \rightarrow 2 \times X^{\mathbf{A}}$ , a Mealy machine as  $c: X \rightarrow (\mathcal{O} \times X)^{\mathcal{I}}$ , and so on.

A *final* coalgebra is a coalgebra  $\zeta$  that admits a unique homomorphism (an arrow  $u$  that makes the diagram on the right commute) from an arbitrary coalgebra  $c$ . Final coalgebras play an important role in the theory of coalgebra because the unique homomorphism towards a final coalgebra often captures *behaviors* of systems represented as coalgebras. For example, for a  $(\mathbf{A} \times (\_))$ -coalgebra  $c: X \rightarrow \mathbf{A} \times X$  (recall that it models a labeled transition system), the carrier of the final coalgebra is given by the set  $\mathbf{A}^\omega$  of infinite words over  $\mathbf{A}$ . The unique homomorphism  $X \rightarrow \mathbf{A}^\omega$  from  $c$  to the final coalgebra is a function mapping each  $x \in X$  to the unique infinite word  $a_0a_1 \dots \in \mathbf{A}^\omega$  such that there exists a sequence  $x_0, x_1, \dots \in X$  satisfying  $x_0 = x$  and  $c(x_i) = (a_i, x_{i+1})$ . It is natural to call this function a “behavior” of  $c$ .

**Kleisli Approach** For systems with side-effects like nondeterminism, the framework above can fail due to the lack of a final coalgebra. For example, a nondeterministic automaton is representable as a coalgebra  $X \rightarrow \mathcal{P}(\mathbf{A} \times X)$  where  $\mathcal{P}$  is the *powerset functor*. However, a final coalgebra does not exist for the functor.

One known solution is the so-called *Kleisli approach* [85, 58, 46, 47], which was used to categorically generalize forward and backward simulations (see Section 1.2). There, we separate the functor into two parts, a functor  $T$  representing the *branching type* and a functor  $F$  representing the *transition type*. For  $c: X \rightarrow \mathcal{P}(\mathbf{A} \times X)$ ,  $T$  is  $\mathcal{P}$  and  $F$  is  $\mathbf{A} \times (\_)$ . The key is that the branching type-part  $T$  often constitutes a *monad*, a functor with special structures. Indeed,  $\mathcal{P}$  constitutes the *powerset monad*. The monad structure allows us to consider the *Kleisli category*, a category whose arrows are arrows of a form  $X \rightarrow \mathcal{P}Y$ . We represent a system as an  $F$ -coalgebra in the Kleisli category. Then a final coalgebra, or sometimes a weakly final coalgebra (a coalgebra admitting a not necessarily unique homomorphism) often exist and a homomorphism to it captures some behavior of the system, although in the latter case we have to introduce some mechanism like an order to choose a homomorphism. The captured behavior of transition systems varies depending on the choice of a (weakly) final coalgebra. For  $c: X \rightarrow \mathcal{P}(\mathbf{A} \times X)$ , one possible choice is a weakly final coalgebra having  $\mathbf{A}^\omega$  as its carrier. Then the behavior captured by the weakly final coalgebra is:

$$x \mapsto \{a_0a_1 \dots \in \mathbf{A}^\omega \mid \exists x_0, x_1, \dots \in X. x_0 = x \text{ and } \forall i. (a_i, x_{i+1}) \in c(x_i)\}.$$

### 1.3.2 Fixed Point Logic

The notions of least and greatest fixed point are important in the theoretical computer science. It is because *reachability* and *unreachability*—basic specifications in model checking—are characterized as the least and the greatest fixed point of a certain function.

For example, suppose that we are given a nondeterministic transition system  $\mathcal{A} = (X, \tau: X \rightarrow \mathcal{P}X, \text{Acc} \subseteq X)$  with accepting states. We define a function  $\diamond: \mathcal{P}X \rightarrow \mathcal{P}X$  by  $\diamond(A) := \{x \in X \mid \tau(x) \cap A \neq \emptyset\} \cup \text{Acc}$ . This function is a monotone function with respect to the inclusion order, and has the least and the greatest fixed points. The least fixed point  $\mu\diamond \in \mathcal{P}X$  collects states from which  $\text{Acc}$  is reachable. In contrast, the greatest fixed point  $\nu\diamond \subseteq X$  collects states from which an infinite run can be constructed (i.e. no dead-end is reachable).

Well-known theorems in fixed point-theory provide us with means for calculating a lower or upper bound of the least or the greatest fixed point. For underapproximating the least fixed point, we can refer to the Kleene fixed point theorem (see e.g. [101]). It claims that if a monotone function  $f: (L, \leq) \rightarrow (L, \leq)$  is  $\omega$ -continuous (i.e. it preserves supremums of increasing chains), then the least fixed point of  $f$  is given by the supremum of an increasing chain  $\perp \leq f(\perp) \leq f^2(\perp) \leq \dots$ . The theorem implies that for  $i \in \mathbb{N}$ ,  $f^i(\perp)$  underapproximates  $\mu f$ . When  $f$  is not  $\omega$ -continuous, we can use the result in [27].

We can overapproximate the least fixed point using the Knaster-Tarski theorem (see e.g. [105]): if  $a \in L$  is a pre-fixed point (i.e.  $a \leq f(a)$ ) then  $\mu f \leq a$ .

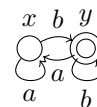
For over-/underapproximating the greatest fixed point, we can use the dual of the above two statements. Throughout this paper, we shall extensively use these two principles and their duals.

**Alternating Fixed Point** In the examples so far, we focused on rather simple properties like termination. Of course, we can consider more complex specifications. For example, for the program in Figure 1.1, we can consider the following specification: “if we feed inputs to the program appropriately, we can make  $x$  become 1 infinitely often without making the program terminate.”

We can use the *Büchi condition* [15] to represent such a property. An infinite run is said to satisfy the Büchi condition if it visits accepting states infinitely many times. A nondeterministic automaton with the Büchi acceptance condition is called a *Büchi automaton*. For example, in the Büchi automaton on the right above, an infinite run satisfies the Büchi condition if it visits an accepting state  $y$  infinitely many times. We define the *language* of the Büchi automaton as a function assigning each state the set of words having an accepting run, i.e. a function  $\{x, y\} \rightarrow \{a, b\}^\omega$  that assigns the set of words containing infinitely many  $b$ 's to both  $x$  and  $y$ .

In order to fixed point-theoretically deal with the Büchi condition, least or greatest fixed points are not enough. Instead, we use *alternating fixed points*.

Suppose that we are given a nondeterministic transition system  $\mathcal{A} = (X, \tau)$  equipped with a partition  $X = X_1 + X_2$  of the state space. Then we can naturally divide the function  $\diamond: \mathcal{P}X \rightarrow \mathcal{P}X$  defined above into two parts:  $\diamond_1: \mathcal{P}X \rightarrow \mathcal{P}X_1$  and  $\diamond_2: \mathcal{P}X \rightarrow \mathcal{P}X_2$ . To capture the Büchi condition, we take the least fixed point for  $\diamond_1$  and the greatest fixed point for  $\diamond_2$ . More concretely, we first calculate the least fixed point of a function  $u_1 \mapsto \diamond_1(u_1 + u_2) \in \mathcal{P}X_1$  regarding  $u_2 \in \mathcal{P}X_2$  a parameter. We then calculate the greatest fixed point of  $u_2 \mapsto \diamond_2(u_1 + u_2) \in \mathcal{P}X_2$  using the fixed point. This results in an “alternating” fixed point  $A \in \mathcal{P}X$  of  $\diamond$  that is not the least or the greatest. It indeed captures the Büchi condition: it is



given by the set of states from which  $X_2$  can be visited infinitely many times.

Hence the Büchi condition is characterized by alternating fixed points.

The *parity condition* (see e.g. [33]) is a generalization of the Büchi condition. It can represent more complex specifications than the Büchi condition. By using alternating fixed points, we can also deal with the parity condition.

**Categorical Fixed Point** In this thesis, we use fixed point theory together with category theory. There mainly exist two ways to introduce a notion of fixed point into category theory: introduce it as a *logical fixed point* or introduce as a *categorical fixed point*. Their difference is well-illustrated in the contrast between “the category of pre-ordered sets” and “a pre-ordered set as a category.”

Pre-ordered sets and monotone functions between them constitute a category **PreOrd**. We can naturally make the category an *order-enriched* one by introducing an order  $\leq$  to each homset  $\mathbf{PreOrd}((X, \leq), (Y, \leq))$  by extending the preorder over  $(Y, \leq)$  in a pointwise manner. Hence for an endofunction over  $\mathbf{PreOrd}((X, \leq), (Y, \leq))$ , we can consider its least, greatest and alternating fixed points. We call such a fixed point a “logical fixed point.”

In contrast, a “categorical fixed point” is defined by regarding a category as a generalization of a pre-ordered set. A preordered set  $(X, \leq)$  induces a category  $\mathbb{X}$  whose objects are given by the elements of  $X$  and arrows are given by  $\mathbb{X}(x, y) := \{*\}$  if  $x \leq y$  and  $\emptyset$  otherwise. Moreover, many categorical notions can be explained as generalizations of notions in preordered sets: functors generalize monotone functions, coalgebras generalize post-fixed points, and so on.

The notion of final coalgebra also has its counterpart in pre-ordered sets. Suppose that a monotone endofunction  $f : (X, \leq) \rightarrow (X, \leq)$  has the greatest fixed point  $\nu f$ . By the (dual of the) Knaster-Tarski theorem,  $f$  is the supremum of the set  $\{x \in X \mid f(x) \geq x\}$  of post-fixed points. When we see a pre-ordered set as a category, it means that the greatest fixed point of  $f$  is a final  $f$ -algebra. Hence a final coalgebra is a categorical generalization of the greatest fixed point. Similarly, an initial algebra, a dual notion of final coalgebra, is a generalization of the least fixed point. We call a categorical greatest fixed point for a final coalgebra and a categorical least fixed point for an initial algebra. In this thesis, we will also see that we can define a notion of “categorical alternating fixed point.”

## 1.4 Contributions

Contributions of this thesis are the following two: i) categorical generalization of fair simulation and concretization to probabilistic systems; and ii) those of ranking function. More concretely, our contributions are summarized as follows:

**Categorical Characterization of Parity Languages** We extend the Kleisli approach (see Section 1.3.1) for Büchi and parity automata. We will introduce two categorical characterizations for their languages. Both of them make use of the relationship between parity automata and alternating fixed points, but they differ in how to categorically reflect the notion of fixed point (see Section 1.3.2).

One characterization considers logical fixed points. Here we calculate an alternating fixed point of a certain function in a homset of the Kleisli category which is assumed to carry an order. The characterization is a direct translation of the fixed point-theoretic characterization of parity languages. The other characterization considers categorical fixed points. For capturing behaviors of parity automata, we will define datatypes as an alternating fixed point of a functor.

The latter characterization is more complicated than the former one and characterizes nonstandard “languages” of parity automata. We will categorically relate these two characterizations and show that we can regard the latter as a characterization of ordinary languages of parity automata as well.

**Generalization and Concretization I: Fair Simulation** Using the categorical characterization of Büchi languages by logical fixed points, we will categorically generalize *fair simulation* [50, 32].

Fair simulation is a simulation notion for Büchi automata. Our categorical generalization is inspired by those of forward and backward simulations reviewed in Section 1.2, but is much more complicated. We concretize it for *probabilistic Büchi automata* and obtain a new simulation notion for them [110].

**Generalization and Concretization II: Ranking Function** We also generalize *ranking function*. Unlike the case of fair simulation, we have used an existing standard framework for categorically capturing behaviors of systems.

Ranking functions are commonly used for checking termination of nondeterministic systems [35]. We categorically generalized ranking function and concretize it for probabilistic transition systems. In the generalization, a categorical notion called *corecursive algebra* played an important role. As a result of the concretization, we have obtained two probabilistic ranking function-like notions. A ranking function-like notion called *ranking supermartingale* [19] were known for probabilistic systems, but the induced notions were new and different from it.

We call the induced notions *distribution-valued ranking function* and  $\gamma$ -scaled *submartingale*. They have different characteristics from ranking supermartingale. That is, we can use them for *quantitative reasoning* in the sense that they give lower bounds for termination probabilities of probabilistic systems.

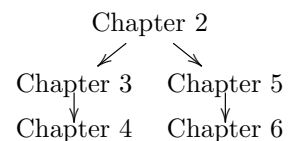
**Implementation of Probabilistic Ranking Function** We have implemented a program that underapproximates a termination probability of a probabilistic program by synthesizing a  $\gamma$ -scaled submartingale. We implemented two algorithms: a linear template-based one and a polynomial template-based one. The former uses a linear programming (LP) solver while the latter uses a semidefinite programming (SDP) solver. The algorithms are adapted from existing template-based algorithms for synthesizing ranking supermartingales [19, 23, 21].

We conducted some experiments with the implementations. We first tested the linear template-based implementation for several probabilistic programs, some of which are taken from literature. We have also compared it with existing algorithm in [23]. We found that there exist probabilistic programs where our algorithm can give better lower bound for the termination probability.

We also conducted experiments on the polynomial-template based implementation. However, we found that it does not work well because of numerical errors. We nevertheless show the algorithm and the experimental results for records.

## 1.5 Organization

The rest of this thesis is organized as follows. In Chapter 2, we give preliminaries. Chapter 3 is devoted to developing categorical characterizations of behaviors of Büchi and parity automata. The developed characterizations are used in Chapter 4 to categorically generalize the notion of fair simula-



tion and induce its probabilistic variant. In Chapter 5 we categorically generalize ranking functions, concretize it for probabilistic systems, and induce probabilistic verification methods. In Chapter 6 we give an algorithm and an implementation based on one of the notions induced in the previous chapter, and present experimental results. We discuss related work in Chapter 7, and conclude in Chapter 8.

**First Appearance** Most results of this thesis were previously published. Chapter 3 is based on [114, 112], Chapter 4 is based on [110], and most part of Chapter 5 is based on [109]. Some results in Chapter 6 constitute a part of [102].



# Chapter 2

## Preliminaries

This chapter is devoted to preliminaries. In Section 2.1 we introduce notations. In Section 2.2 we define various transition systems. In this thesis, the fixed point logic plays an important role. Preliminaries on it are in Section 2.3. Categorical preliminaries are in Section 2.4.

We assume that readers are familiar with the basic measure theory, and omit preliminaries on it. See e.g. [30, 8, 103] for the detail.

### 2.1 Notations

We first introduce notations that are used throughout this thesis.

For  $a, b \in \mathbb{R}$ ,  $[a, b]$  denotes  $\{r \in \mathbb{R} \mid a \leq r \leq b\}$ ,  $(a, b)$  denotes  $\{r \in \mathbb{R} \mid a < r < b\}$ ,  $[a, b)$  denotes  $\{r \in \mathbb{R} \mid a \leq r < b\}$  and  $(a, b]$  denotes  $\{r \in \mathbb{R} \mid a < r \leq b\}$ .

Let  $f: X \rightarrow Y$ . For  $A \subseteq X$ , we define  $f|_A: A \rightarrow Y$  by  $f|_A(x) := f(x)$ . For  $B \subseteq Y$ , we write  $f^{-1}(B)$  for  $\{x \in X \mid f(x) \in B\}$ . For  $a \in X$  and  $b \in Y$ , we define  $f[a \mapsto b]: X \rightarrow Y$  by  $f[a \mapsto b](x) := b$  if  $x = a$  and  $f(x)$  otherwise.

For a partial function  $g: X \rightarrow Y$ , we write  $g(x) = \perp$  if  $g(x)$  is undefined.

We write  $X^*$ ,  $X^+$ ,  $X^\omega$  and  $X^\infty$  for the sets of finite, non-empty finite, infinite and possibly infinite words over  $X$  respectively. For  $w \in X^*$  and  $w' \in X^\infty$ , we write  $ww'$  for their concatenation. We write  $\langle \rangle$  for the empty word. For  $w, w' \in X^\infty$ , we write  $w \preceq w'$  when  $w$  is a prefix of  $w'$ .

For a set  $I$ , we write  $\prod_{i \in I} X_i$  for a product  $\{(x_i)_{i \in I} \mid \forall i. x_i \in X_i\}$  and  $\coprod_{i \in I} X_i$  for a disjoint sum  $\{(i, x) \mid i \in I, x \in X_i\}$ . If  $X_i \cap X_j = \emptyset$  for  $i, j \in I$ , we write  $x$  for  $(i, x) \in \coprod_{i \in I} X_i$ . For  $i \in I$ ,  $\pi_i: \prod_{i \in I} X_i \rightarrow X_i$  denotes the canonical projection and  $\kappa_i: X_i \rightarrow \prod_{i \in I} X_i$  denotes the canonical injection.

For a poset  $(X, \leq)$  and an endofunction  $f: (X, \leq) \rightarrow (X, \leq)$ , we write  $\mu f$  (resp.  $\nu f$ ) for the least (resp. greatest) fixed point of  $f$ , if it exists.

For a set  $X$ ,  $\mathcal{P}X$  denotes the set of subsets of  $X$  (i.e.  $\mathcal{P}X := \{A \subseteq X\}$ ),  $\mathcal{D}X$  denotes the set of distributions over  $X$  (i.e.  $\mathcal{D}X := \{d: X \rightarrow [0, 1] \mid |\{x \in X \mid d(x) > 0\}| \text{ is countable and } \sum_{x \in X} d(x) = 1\}$ ) and  $\mathcal{D}_s X$  denotes the set of subdistributions over  $X$ .

Otherwise specified, we assume that  $\mathbb{R}$  and a subset  $A$  of  $\mathbb{R}$  are equipped with the standard  $\sigma$ -algebras. For a measurable space  $(X, \mathfrak{F}_X)$ , we write  $\mathcal{G}(X, \mathfrak{F}_X)$  for a measurable space  $(\mathcal{G}X, \mathfrak{F}_{\mathcal{G}X})$ , where  $\mathcal{G}X$  is the set of probability measures over  $(X, \mathfrak{F}_X)$ , and  $\mathfrak{F}_{\mathcal{G}X}$  is the smallest  $\sigma$ -algebra that makes a function  $\text{ev}_A: \mathcal{G}X \rightarrow [0, 1]$  defined by  $\text{ev}_A(\alpha) := \alpha(A)$  measurable for each  $A \in \mathfrak{F}_X$  (see also [39] and Definition 2.4.6). Similarly, we write  $\mathcal{G}_s(X, \mathfrak{F}_X)$  for a measurable space  $(\mathcal{G}_s X, \mathfrak{F}_{\mathcal{G}_s X})$  where  $\mathcal{G}_s X$  is the set of sub-probability measures over  $(X, \mathfrak{F}_X)$  and  $\mathfrak{F}_{\mathcal{G}_s X}$  is a  $\sigma$ -algebra defined as above. For  $x \in X$ ,  $\delta_x$  denotes the *Dirac measure* at  $x$ , i.e. a probability measure such that  $\delta_x(A) = 1$  if  $x \in A$  and 0 otherwise.

Note that when  $X$  is equipped with the discrete  $\sigma$ -algebra  $\mathcal{P}X$ ,  $\mathcal{G}X \cong \mathcal{D}X$  and  $\mathcal{G}_s X \cong \mathcal{D}_s X$ . We sometimes implicitly use these isomorphisms. For example, we might write  $[0 \mapsto \frac{1}{2}, 1 \mapsto \frac{1}{2}]$  for a probability measure over  $(\mathbb{N}, \mathcal{P}\mathbb{N})$ . If  $\mathfrak{F}_X$  is clear from the context, we sometimes write  $X$  for a measurable space  $(X, \mathfrak{F}_X)$ . Hence we may write  $\mathcal{G}X$  for both a measurable space  $\mathcal{G}(X, \mathfrak{F}_X)$  and its underlying set.

For measurable spaces  $(X_1, \mathfrak{F}_{X_1}), \dots, (X_n, \mathfrak{F}_{X_n})$ , we write  $(X_1, \mathfrak{F}_{X_1}) \times \dots \times (X_n, \mathfrak{F}_{X_n})$  for a measurable space  $(X_1 \times \dots \times X_n, \mathfrak{F}_{X_1 \times \dots \times X_n})$  where  $\mathfrak{F}_{X_1 \times \dots \times X_n}$  is the smallest  $\sigma$ -algebra containing  $A_1 \times \dots \times A_n$  for each  $A_1 \in \mathfrak{F}_{X_1}, \dots, A_n \in \mathfrak{F}_{X_n}$ . For a family  $((X_i, \mathfrak{F}_{X_i}))_{i \in I}$  of measurable spaces,  $\coprod_{i \in I} (X_i, \mathfrak{F}_{X_i})$  denotes  $(\coprod_{i \in I} X_i, \mathfrak{F}_{\coprod_{i \in I} X_i})$  where  $\mathfrak{F}_{\coprod_{i \in I} X_i} := \{\bigcup_{i \in I} \{(i, x) \mid x \in A_i\} \mid \forall i \in I. A_i \in \mathfrak{F}_{X_i}\}$ .

## 2.2 Transition Systems

In this section, we define four types of transition systems and related notions. The first two notions (in Sections 2.2.1–2.2.2) are used in Chapters 3–4 while the other two (in Sections 2.2.3–2.2.4) are mainly used in Chapters 5–6.

### 2.2.1 Nondeterministic Parity Tree Automaton

*Nondeterministic tree automata* (NTA) and *nondeterministic parity tree automata* (NPTA) are systems that accept possibly infinite-depth trees. The former accept trees with respect to the *trivial (acceptance) condition* while the latter accept trees with respect to the *parity condition*. We also define *nondeterministic Büchi tree automata* (NBTA), which is a special case of NPTA.

We first review several notions regarding *trees*.

**Definition 2.2.1** (ranked alphabet). A *ranked alphabet* is a pair  $\Sigma = (\Sigma, |\_|)$  of a set  $\Sigma$  and a function  $|\_| : \Sigma \rightarrow \mathbb{N}$ . For  $a \in \Sigma$ ,  $|a|$  is called the *arity* of  $a$ . For  $n \in \mathbb{N}$ , we write  $\Sigma_n$  for  $\{a \in \Sigma \mid |a| = n\}$ .

Let  $\Sigma$  be a ranked alphabet and  $X$  be a set. Otherwise specified, we regard  $\Sigma \times X$  as a ranked alphabet by letting  $|(\sigma, x)| = |\sigma|$  for  $\sigma \in \Sigma$  and  $x \in X$ , and regard  $\Sigma + X$  and  $X + \Sigma$  as a ranked alphabet by letting  $|x| = |0|$  for  $x \in X$ .

There are several definitions for tree. We follow a standard definition which is used in [17] for example.

**Definition 2.2.2** (tree and branch). Let  $\Sigma$  be a ranked alphabet. A (*possibly infinite*)  $\Sigma$ -*labeled tree* is a pair  $t = (D_t, l_t)$  of a non-empty set  $D_t \subseteq \mathbb{N}^*$  and a function  $l_t : D \rightarrow \Sigma$  that satisfies the following conditions.

1.  $D_t$  is *prefix-closed*, that is, for each  $w, w' \in \mathbb{N}^*$ ,  $ww' \in D_t$  implies  $w \in D_t$ .
2.  $l_t$  respects the arities, that is, for each  $w \in D_t$  and  $i \in \mathbb{N}$ ,  $wi \in D_t$  if and only if  $i \leq |l_t(w)| - 1$ .

A tree  $t = (D_t, l_t)$  is said to be *finite* (resp. *infinite*) if  $D_t$  is a finite (resp. infinite) set. We write  $\text{Tree}_{\Sigma}^{\infty}$  for the set of possibly infinite  $\Sigma$ -labeled trees, and  $\text{Tree}_{\Sigma}^*$  for the set of finite  $\Sigma$ -labeled trees. A *branch* over a tree  $t = (D_t, l_t)$  is an element  $w \in D_t$ . A branch is said to be *finite* (resp. *infinite*) if it is a finite (resp. infinite) sequence. For a finite branch  $w = w_0 \dots w_n \in D_t$  over  $t = (D_t, l_t)$ , we define the  $w$ -*th subtree*  $t_w = (D_{t,w}, l_{t,w})$  of  $t$  by  $D_{t,w} := \{w' \in \mathbb{N}^* \mid ww' \in D_t\}$  and  $l_{t,w}(w') := l_t(ww')$ .

**Remark 2.2.3.** For the sake of notational simplicity, we identify a  $\Sigma$ -labeled tree with a  $\Sigma$ -term in a natural manner. For example, an  $\{a, b\}$ -term  $(a, (b, b))$  denotes an  $\{a, b\}$ -labeled finite tree  $t = (\{\langle \rangle, 0, 1\}, [\langle \rangle \mapsto a, 0 \mapsto b, 1 \mapsto b])$ . Moreover, for  $\{a, b, c\}$ -labeled trees  $t_0 = (D_0, l_0)$  and  $t_1 = (D_1, l_1)$ , we write  $(c, t_0, t_1)$  for a tree  $t = (\{\langle \rangle \cup \{0w \mid w \in D_0\} \cup \{1w \mid w \in D_1\}, [\langle \rangle \mapsto c, 0w \mapsto l_0(w), 1w \mapsto l_1(w)])$ .

We are now ready to define the notions of NTA, NPTA and NBTA.

**Definition 2.2.4** (NTA, NPTA and NBTA). Let  $\Sigma$  be a ranked alphabet. A ( $\Sigma$ -labeled) *nondeterministic tree automaton* (NTA) is a pair  $\mathcal{A} = (X, \tau)$  consisting of a *state space*  $X$  and a *transition function*  $\tau: X \rightarrow \mathcal{P}(\prod_{i \in \omega} \Sigma_i \times X^i)$ . A ( $\Sigma$ -labeled) *nondeterministic parity tree automaton* (NPTA) is a triple  $\mathcal{A} = (X, \tau, \mathbf{p})$  such that the first two components constitute an NTA and  $\mathbf{p}: X \rightarrow \{1, \dots, 2n\}$  is a *priority function* where  $n$  is a positive integer. A ( $\Sigma$ -labeled) *nondeterministic Büchi tree automaton* (NBTA) is a triple  $\mathcal{A} = (X, \tau, \text{Acc})$  where the first two components are the same, and  $\text{Acc} \subseteq X$  is the set of *accepting states*.

We next define *languages* of NTAs, NPTAs and NBTA. We define them as functions that map each state to the set of trees accepted from the state. To this end, we first define functions that map each state to the set of accepted *run trees*.

**Definition 2.2.5** (run tree). Let  $\mathcal{A} = (X, \tau)$  be a  $\Sigma$ -labeled NTA. We regard  $X \times \Sigma$  as a ranked alphabet as in Definition 2.2.1. For  $x \in X$ , a (*possibly infinite*) *run tree* over  $\mathcal{A}$  from  $x$  is a  $(X \times \Sigma)$ -labeled tree  $\rho = (D_\rho, l_\rho)$  such that: i)  $\pi_1(l_\rho(\langle \rangle)) = x$ ; and ii) for each  $w \in D_r$ , if  $l_\rho(w) = (x, a)$  and  $l_\rho(wi) = (x_i, a_i)$  for each  $i \in \{0, \dots, |a| - 1\}$  then  $(a, x_0, \dots, x_{|a|-1}) \in \tau(x)$ . A run tree is called *finite* if it is a finite tree. We write  $\text{Run}_\mathcal{A}^\infty(x)$  (resp.  $\text{Run}_\mathcal{A}^*(x)$ ) for the set of possibly infinite (resp. finite) run trees from  $x$ , and let  $\text{Run}_\mathcal{A}^\infty(A) := \cup_{x \in A} \text{Run}_\mathcal{A}^\infty(x)$  and  $\text{Run}_\mathcal{A}^*(A) := \cup_{x \in A} \text{Run}_\mathcal{A}^*(x)$  for  $A \subseteq X$ . A run tree over NPTA or NBTA is similarly defined.

A run tree and a tree is connected by the following function.

**Definition 2.2.6** ( $\text{DelSt}(r)$ ). Let  $\mathcal{A} = (X, \tau)$  be a  $\Sigma$ -labeled NTA. We define a function  $\text{DelSt}: \text{Run}_\mathcal{A}^\infty(X) \rightarrow \text{Tree}_\Sigma^\infty$  by  $\text{DelSt}((D_\rho, l_\rho)) := (D_\rho, l'_\rho)$  where  $l'_\rho(w) := \pi_2(l_\rho(w))$ . A function  $\text{DelSt}$  for an NPTA or an NBTA is similarly defined.

All the run trees are accepted by NTA. In contrast, run trees on NPTAs (resp. NBTA) are accepted if they satisfy the *parity* (resp. *Büchi*) condition.

**Definition 2.2.7** (accepting run). Let  $\mathcal{A} = (X, \tau, \mathbf{p})$  be a  $\Sigma$ -labeled NPTA. A run tree  $\rho = (D_\rho, l_\rho)$  over  $\mathcal{A}$  *satisfies the parity (acceptance) condition* (or simply is *accepting*) if for each infinite branch  $w_0 w_1 w_2 \dots$  over  $\rho$ ,  $\limsup_{i \rightarrow \infty} \mathbf{p}(l_\rho(w_0 \dots w_i))$  is even. We write  $\text{Run}_\mathcal{A}^{\text{Acc}}(x)$  for the set of accepting runs in  $\text{Run}_\mathcal{A}(x)$ .

A run tree  $\rho = (D_\rho, l_\rho)$  over a  $\Sigma$ -labeled NBTA  $\mathcal{A} = (X, \tau, \text{Acc})$  *satisfies the Büchi (acceptance) condition* (or is *accepting*) if for each infinite branch  $w_0 w_1 w_2 \dots$ ,  $l_\rho(w_0 \dots w_i) \in \text{Acc}$  for infinitely many  $i$ 's. We define  $\text{Run}_\mathcal{A}^{\text{Acc}}$  similarly.

**Remark 2.2.8.** An NBTA  $\mathcal{A} = (X, \tau, \text{Acc})$  induces an NPTA  $(X, \tau, \mathbf{p})$  with  $n = 1$  where  $\mathbf{p}(x) = 1$  if  $x \notin \text{Acc}$  and 2 otherwise. It is easy to see that a run tree on an NBTA is accepting if and only if it is accepting on the induced NPTA. Hence NBTA is a special case of NPTA.

We can now define languages of NTAs, NPTAs and NBTA.

**Definition 2.2.9** ( $L_{\mathcal{A}}^*$ ,  $L_{\mathcal{A}}^\infty$ ,  $L_{\mathcal{A}}^p$ ,  $L_{\mathcal{A}}^B$ ).

- Let  $\mathcal{A} = (X, \tau)$  be a  $\Sigma$ -labeled NTA. The *finite language* of  $\mathcal{A}$  is a function  $L_{\mathcal{A}}^* : X \rightarrow \mathcal{P}(\text{Tree}_\Sigma^*)$  defined by  $L_{\mathcal{A}}^*(x) := \{\text{DelSt}(\rho) \mid \rho \in \text{Run}_{\mathcal{A}}^*(x)\}$ . The finite language of an NPTA or NBTA is defined in the same way.
- Let  $\mathcal{A} = (X, \tau)$  be a  $\Sigma$ -labeled NTA. The *infinitary*<sup>1</sup> *language* of  $\mathcal{A}$  is a function  $L_{\mathcal{A}}^\infty : X \rightarrow \mathcal{P}(\text{Tree}_\Sigma^\infty)$  defined by  $L_{\mathcal{A}}^\infty(x) := \{\text{DelSt}(\rho) \mid \rho \in \text{Run}_\tau^\infty(x)\}$ . The infinitary language of an NPTA or NBTA is defined in the same way.
- Let  $\mathcal{A} = (X, \tau, \mathfrak{p})$  be a  $\Sigma$ -labeled NPTA. The *(parity) language* of  $\mathcal{A}$  is a function  $L_{\mathcal{A}}^p : X \rightarrow \mathcal{P}(\text{Tree}_\Sigma^\infty)$  defined by  $L_{\mathcal{A}}^p(x) := \{\text{DelSt}(\rho) \mid \rho \in \text{Run}_{\mathcal{A}}^{\text{Acc}}(x)\}$ . The *(Büchi) language*  $L_{\mathcal{A}}^B$  of an NBTA  $\mathcal{A}$  is similarly defined.

**Remark 2.2.10.** For an NPTA  $\mathcal{A} = (X, \tau, \mathfrak{p})$ , we have the following inclusions for each  $x \in X$ :  $\text{Run}_{\mathcal{A}}^*(x) \subseteq \text{Run}_{\mathcal{A}}^{\text{Acc}}(x) \subseteq \text{Run}_{\mathcal{A}}^\infty(x)$  and  $L_{\mathcal{A}}^*(x) \subseteq L_{\mathcal{A}}^p(x) \subseteq L_{\mathcal{A}}^\infty(x)$ . From a fixed point-logical perspective,  $\text{Run}_{\mathcal{A}}^*(x)$  and  $L_{\mathcal{A}}^*(x)$  are the *least fixed-points* of certain functions,  $\text{Run}_{\mathcal{A}}^\infty(x)$  and  $L_{\mathcal{A}}^\infty(x)$  are the *greatest fixed-points* of the same functions, and  $\text{Run}_{\mathcal{A}}^{\text{Acc}}(x)$  and  $L_{\mathcal{A}}^p(x)$  are *alternating fixed-points* between them. In Chapter 3, we will characterize  $L_{\mathcal{A}}^p$  by categorically translating these fixed point-logical characterizations of  $\text{Run}_{\mathcal{A}}^{\text{Acc}}$  and  $L_{\mathcal{A}}^p$ .

**Remark 2.2.11.** If we regard sets  $A$  and  $\{\checkmark\} + A$  as ranked alphabets by letting  $|\checkmark| = 0$  and  $|a| = 1$  for each  $a \in A$ , then there exist canonical isomorphisms  $\text{Tree}_A^\infty \cong A^\omega$ ,  $\text{Tree}_{\{\checkmark\}+A}^\infty \cong A^\infty$ ,  $\text{Tree}_A^* \cong \emptyset$  and  $\text{Tree}_{\{\checkmark\}+A}^* \cong A^*$ .

When  $\Sigma = A$ , we can identify a  $\Sigma$ -labeled NTA with an  $A$ -labeled nondeterministic *word automaton* (NWA). For an NWA  $\mathcal{A} = (X, \tau)$  we have  $\text{Run}_{\mathcal{A}}^\infty \subseteq (X \times A)^\omega$  and  $\text{Run}_{\mathcal{A}}^* \cong \emptyset$ , and therefore  $L_{\mathcal{A}}^\infty$  and  $L_{\mathcal{A}}^*$  are considered to take values in  $A^\omega$  and  $\emptyset$  respectively. Similarly, we can identify an  $A$ -labeled NPTA (resp. NBTA) with an  $A$ -labeled nondeterministic parity (resp. Büchi) word automaton (NPWA, resp. NBWA).

When  $\Sigma = \{\checkmark\} + A$ , we identify a  $\Sigma$ -labeled NTA with an  $A$ -labeled nondeterministic *word automaton with accepting states* by regarding a state  $x \in X$  as *accepting* if  $\checkmark \in \tau(x)$ . We have  $\text{Run}_{\mathcal{A}}^\infty \subseteq (X \times A)^\infty$  and  $\text{Run}_{\mathcal{A}}^* \subseteq (X \times A)^*X$ , and therefore  $L_{\mathcal{A}}^\infty$  and  $L_{\mathcal{A}}^*$  take values in  $A^\infty$  and  $A^*$  respectively. Note that the notion of “accepting state” in this remark is different from the one on NBTA (Definition 2.2.5).

## 2.2.2 Probabilistic Parity Tree Automaton

We next define *probabilistic tree automaton* and *probabilistic parity tree automaton*. They are quantitative variants of NTA and NPTA respectively.

**Definition 2.2.12** (PTA, PPTA, PBTA). Let  $\Sigma$  be a ranked alphabet. A *(generative) probabilistic tree automaton* (PTA) is a pair  $\mathcal{A} = ((X, \mathfrak{F}_X), \xi)$  consisting of a standard Borel space  $(X, \mathfrak{F}_X)$  called a *state space* and a measurable function

$$\xi : (X, \mathfrak{F}_X) \rightarrow \mathcal{G}_s \coprod_{n \in \omega} \left( (\Sigma_n, \mathcal{P}\Sigma_n) \times \underbrace{(X, \mathfrak{F}_X) \times \cdots \times (X, \mathfrak{F}_X)}_n \right)$$

called a *transition function*. A *(generative) probabilistic parity tree automaton* (PPTA) is a triple  $\mathcal{A} = ((X, \mathfrak{F}_X), \xi, \mathfrak{p})$  such that  $((X, \mathfrak{F}_X), \xi)$  is a PTA

<sup>1</sup>We use a term “infinitary” to mean “possibly infinite.”

and  $p: X \rightarrow \{1, \dots, 2n\}$  is a measurable function called a *priority function*, where  $n$  is a positive integer. A *probabilistic Büchi tree automaton* (PBTA)  $\mathcal{A} = ((X, \mathfrak{F}_X), \xi, \text{Acc})$  is defined in a similar manner (cf. Definition 2.2.4).

A  $\Sigma$ -labeled generative PTA generates a possibly infinite  $\Sigma$ -labeled tree in a probabilistic manner. To define a language of a PTA, we have to introduce a  $\sigma$ -algebra into the set of  $\Sigma$ -labeled trees. As usual, it is defined using *cylinders*.

**Definition 2.2.13** ( $\text{cyl}(t)$  and  $\mathfrak{F}_{\text{Tree}_\Sigma^\infty}$ ). Let  $\Sigma$  be a ranked alphabet. We define a ranked alphabet  $\{\perp\} + \Sigma$  by adding a letter  $\perp$  such that  $|\perp| = 0$ . For  $t = (D_t, l_t) \in \text{Tree}_{\{\perp\} + \Sigma}^*$  and  $t' = (D_{t'}, l_{t'}) \in \text{Tree}_\Sigma^\infty$ ,  $t$  is a *prefix* of  $t'$  if  $D_t \subseteq D_{t'}$  and  $l_t(w) \neq \perp$  implies  $l_t(w) = l_{t'}(w)$  for each  $w \in D_t$ . We write  $t \preceq t'$  in this case.

For  $t \in \text{Tree}_{\{\perp\} + \Sigma}^*$ , a set  $\{t' \in \text{Tree}_\Sigma^\infty \mid t \preceq t'\}$  of possibly infinite trees is called the *cylinder set* generated by  $t$ , and denoted by  $\text{cyl}(t)$ . We write  $\mathfrak{F}_{\text{Tree}_\Sigma^\infty}$  for the smallest  $\sigma$ -algebra containing  $\{\text{cyl}(t) \subseteq \text{Tree}_\Sigma^\infty \mid t \in \text{Tree}_{\{\perp\} + \Sigma}^*\}$ , and call it the  *$\sigma$ -algebra generated by the cylinder sets*. We define a  $\sigma$ -algebra  $\mathfrak{F}_{\text{Tree}_\Sigma^*}$  over  $\text{Tree}_\Sigma^*$  by  $\mathfrak{F}_{\text{Tree}_\Sigma^*} := \{A \cap \text{Tree}_\Sigma^* \mid A \in \mathfrak{F}_{\text{Tree}_\Sigma^\infty}\}$ .

As we have done in Section 2.2.1, we first define the notion of run trees to define languages of PTAs, PPTAs and PBTAs.

**Definition 2.2.14** ( $\text{Run}_\mathcal{A}^\infty$ ,  $\text{Run}_\mathcal{A}^*$ ). Let  $\mathcal{A} = ((X, \mathfrak{F}_X), \xi)$  be a  $\Sigma$ -labeled PTA. We regard  $X \times \Sigma$  and  $\{\perp\} + X \times \Sigma$  as ranked alphabets as in Definition 2.2.1. A *run tree* over  $\mathcal{A}$  is a  $(X \times \Sigma)$ -labeled tree. For  $x \in X$ , we write  $\text{Run}_\mathcal{A}^\infty(x)$  for the set of possibly infinite run trees over  $\mathcal{A}$  such that the first component of its root node is labeled by  $x$ . We define a  $\sigma$ -algebra  $\mathfrak{F}_{\text{Run}_\mathcal{A}^\infty(x)}$  over  $\text{Run}_\mathcal{A}^\infty(x)$  by  $\{A \cap \text{Run}_\mathcal{A}^\infty(x) \mid A \in \mathfrak{F}_{\text{Tree}_{X \times \Sigma}^*}\}$ . We write  $\text{Run}_\mathcal{A}^\infty(A)$  for  $\bigcup_{x \in A} \text{Run}_\mathcal{A}^\infty(x)$ , and  $\mathfrak{F}_{\text{Run}_\mathcal{A}^\infty(A)}$  for  $\{\bigcup_{x \in A} B_i \mid B_i \in \mathfrak{F}_{\text{Tree}_{X \times \Sigma}^*}(x)\}$ . We similarly define sets  $\text{Run}_\mathcal{A}^*(x)$  and  $\text{Run}_\mathcal{A}^*$  of finite run trees and  $\sigma$ -algebras  $\mathfrak{F}_{\text{Run}_\mathcal{A}^\infty(x)}$  and  $\mathfrak{F}_{\text{Run}_\mathcal{A}^*}$  on them. The notions are defined for PPTAs and PBTAs similarly.

A PTA  $\mathcal{A}$  induces a probability measure  $L_\mathcal{A}^{\text{Run}}$  over the set of run trees. To define it, we first have to calculate the probability where  $\mathcal{A}$  does not diverge. This probability is obtained by: i) calculating a probability  $\text{NoDiv}_k(x)$  where  $\mathcal{A}$  does not diverge for  $k$  steps starting from  $x$ ; and ii) take the limit of  $k \rightarrow \infty$ .

**Definition 2.2.15** ( $L_\mathcal{A}^{\text{Run}}$ ). Let  $\mathcal{A} = ((X, \mathfrak{F}_X), \xi)$  be a  $\Sigma$ -labeled PTA. We inductively define a probability measure  $L_\mathcal{A}^{\text{Run}}(x)$  over  $(\text{Run}_\mathcal{A}^\infty(x), \mathfrak{F}_{\text{Run}_\mathcal{A}^\infty(x)})$  as follows:

- For each  $k \in \omega$ , we inductively define a function  $\text{NoDiv}_k: X \rightarrow [0, 1]$  as:

$$\begin{aligned} & - \text{NoDiv}_0(x) := 1; \text{ and} \\ & - \text{NoDiv}_{k+1}(x) := \int_{(a, x_1, \dots, x_n) \in \prod_{n \in \omega} \Sigma_n \times X^n} \prod_{i=1}^n \text{NoDiv}_k(x_i) d\xi(x). \end{aligned}$$

It is easy to see that  $k \leq k'$  implies  $\text{NoDiv}_k(x) \geq \text{NoDiv}_{k'}(x)$ . For a run tree  $\rho' = (D', l') \in \text{Tree}_{\{\perp\} + X \times \Sigma}^*(x)$  such that  $D' = \{\langle \rangle\}$  and  $l'(\langle \rangle) := \perp$ , we let  $L_\mathcal{A}^{\text{Run}}(x)(\text{cyl}(\rho')) := \lim_{k \rightarrow \infty} \text{NoDiv}_k(x)$ .

- For  $\rho' = (D', l') \in \text{Tree}_{\{\perp\} + X \times \Sigma}^*(x)$  such that  $l'(\langle \rangle) = a \in \Sigma_n$ , we let

$$L_\mathcal{A}^{\text{Run}}(x)(\text{cyl}(\rho')) := \int_{(a, x_0, \dots, x_{n-1}) \in \{a\} \times X^n} \prod_{i=1}^n L_\mathcal{A}^{\text{Run}}(x_i)(\text{cyl}(\rho'_i)) d\xi(x).$$

The probability measure  $L_\mathcal{A}^{\text{Run}}$  is well-defined: by the Kolmogorov extension theorem (see e.g. [103]), a probability measure satisfying the above conditions is unique. We define  $L_\mathcal{A}^{\text{Run}}$  for PPTAs and PBTAs in the same way.

We can now define languages of PTAs, PPTAs and PBTA.

**Definition 2.2.16** ( $L_{\mathcal{A}}^*$ ,  $L_{\mathcal{A}}^\infty$ ,  $L_{\mathcal{A}}^P$ ).

- Let  $\mathcal{A} = ((X, \mathfrak{F}_X), \xi)$  be a  $\Sigma$ -labeled PTA. The *finite language* of  $\mathcal{A}$  is a function  $L_{\mathcal{A}}^*: X \rightarrow \mathcal{G}(\text{Tree}_{\Sigma}^*, \mathfrak{F}_{\text{Tree}_{\Sigma}^*})$  defined by  $L_{\mathcal{A}}^*(x)(A) := L_{\mathcal{A}}^{\text{Run}}(x) \left( \{ \rho \in \text{Tree}_{X \times \Sigma}^\infty(x) \mid \text{DelSt}(\rho) \in A \} \right)$ . Define  $L_{\mathcal{A}}^*$  for PPTAs and PBTA similarly.
- Let  $\mathcal{A} = ((X, \mathfrak{F}_X), \xi)$  be a  $\Sigma$ -labeled PTA. The *infinitary language* of  $\mathcal{A}$  is  $L_{\mathcal{A}}^\infty: X \rightarrow \mathcal{G}(\text{Tree}_{\Sigma}^\infty, \mathfrak{F}_{\text{Tree}_{\Sigma}^\infty})$  defined by  $L_{\mathcal{A}}^\infty(x)(A) := L_{\mathcal{A}}^{\text{Run}}(x) \left( \{ \rho \in \text{Tree}_{X \times \Sigma}^\infty(x) \mid \text{DelSt}(\rho) \in A \} \right)$ . Define  $L_{\mathcal{A}}^\infty$  for PPTAs and PBTA similarly.
- Let  $\mathcal{A} = ((X, \mathfrak{F}_X), \xi, \rho)$  be a  $\Sigma$ -labeled PPTA. The *parity language* of  $\mathcal{A}$  is a function  $L_{\mathcal{A}}^P: X \rightarrow \mathcal{G}(\text{Tree}_{\Sigma}^\infty, \mathfrak{F}_{\text{Tree}_{\Sigma}^\infty})$  defined by  $L_{\mathcal{A}}^P(x)(A) := L_{\mathcal{A}}^{\text{Run}}(x) \left( \{ \rho \in \text{Tree}_{X \times \Sigma}^\infty(x) \mid \text{DelSt}(\rho) \in A \} \cap \text{Run}_\xi^{\text{Acc}}(x) \right)$ . The *(Büchi) language*  $L_{\mathcal{A}}^B$  of a  $\Sigma$ -labeled PBTA is similarly defined.

### 2.2.3 Two-player Game

We define two types of two-player games: a *reachability game* and a *parity game*. The notion of *ranking function* [35] that we will categorically generalize was defined for the former. The latter was used in the formalization of fair simulation [50, 32, 117].

We first define notions that are common to reachability games and parity games. We start with the notion of *game structure*, on which a game is played.

**Definition 2.2.17** (game structure). An (alternating) game structure is a quadruple  $G = (X^{\text{Max}}, X^{\text{Min}}, E^{\text{Max}}, E^{\text{Min}})$  consisting of a set  $X^{\text{Max}}$  of *Player Max's states*, a set  $X^{\text{Min}}$  of *Player Min's states* and transition relations  $E^{\text{Max}} \subseteq X^{\text{Max}} \times X^{\text{Min}}$  and  $E^{\text{Min}} \subseteq X^{\text{Min}} \times X^{\text{Max}}$ .

A *run* is a sequence of states respecting the transition relations.

**Definition 2.2.18** ( $\text{Run}_G^\infty$ ). Let  $G = (X^{\text{Max}}, X^{\text{Min}}, E^{\text{Max}}, E^{\text{Min}})$  be a game structure. An *infinite run* over  $G$  is an infinite sequence  $x_0 x'_0 x_1 x'_1 \dots \in (X^{\text{Max}} \times X^{\text{Min}})^\omega$  such that  $(x_i, x'_i) \in E^{\text{Max}}$  and  $(x'_i, x_{i+1}) \in E^{\text{Min}}$  for each  $i \in \omega$ . A *partial run* over  $G$  is a possibly infinite sequence

$$\begin{aligned} x_0 x'_0 x_1 x'_1 \dots \in & (X^{\text{Max}} \times X^{\text{Min}})^* \times X^{\text{Max}} \times \{\perp_{\text{Max}}\} \\ & \cup (X^{\text{Max}} \times X^{\text{Min}})^+ \times \{\perp_{\text{Min}}\} \cup (X^{\text{Max}} \times X^{\text{Min}})^\omega \end{aligned}$$

that is a prefix of an infinite run over  $G$  if we ignore  $\perp_{\text{Max}}$  or  $\perp_{\text{Min}}$ . For  $x \in X^{\text{Max}}$ , we write  $\text{Run}_G^\infty(x)$  for the set of runs whose first component is  $x$  and  $\text{Run}_G^\infty(A)$  for  $\bigcup_{x \in X} \text{Run}_G^\infty(x)$ .

We can now define the notion of two-player game.

**Definition 2.2.19** (two-player game). A *two-player game* is a quadruple  $\mathfrak{G} = (X^{\text{Max}}, X^{\text{Min}}, E^{\text{Max}}, E^{\text{Min}}, W)$  where the first four components constitute a game structure and  $W \subseteq \text{Run}_{(X^{\text{Max}}, X^{\text{Min}}, E^{\text{Max}}, E^{\text{Min}})}^\infty(X)$  is the set of *winning runs*.

We next define the “winner” of a game. Briefly speaking, the winner is a player who has a strategy such that the player can win the game regardless of the opponent’s strategy as long as the player follows the strategy.

**Definition 2.2.20** (strategies). Let  $\mathfrak{G} = (X^{\text{Max}}, X^{\text{Min}}, E^{\text{Max}}, E^{\text{Min}}, W)$  be a two-player game. *Player Max's strategy* for  $\mathfrak{G}$  is a partial function  $\mathfrak{s}^{\text{Max}} : (X^{\text{Max}} \times X^{\text{Min}})^* \times X^{\text{Max}} \rightarrow X^{\text{Min}}$  such that if  $\mathfrak{s}^{\text{Max}}(x_0y_0 \dots x_{i-1}y_{i-1}x_i) = y \neq \perp$  then  $(x_i, y) \in E^{\text{Max}}$ . *Player Min's strategy*  $\mathfrak{s}^{\text{Min}} : (X^{\text{Max}} \times X^{\text{Min}})^* \times X^{\text{Max}} \times X^{\text{Min}} \rightarrow X^{\text{Max}}$  for  $\mathfrak{G}$  is similarly defined. We write  $\mathfrak{S}_{\mathfrak{G}}^{\text{Max}}$  and  $\mathfrak{S}_{\mathfrak{G}}^{\text{Min}}$  for the sets of Player Max's strategies and Player Min's strategies respectively. A strategy  $\mathfrak{s}^{\text{Max}} \in \mathfrak{S}_{\mathfrak{G}}^{\text{Max}}$  is called *positional* if there exists a function  $s : X^{\text{Max}} \rightarrow X^{\text{Min}}$  such that  $\mathfrak{s}^{\text{Max}}(x_0y_0 \dots x_{i-1}y_{i-1}x_i) = s(x_i)$  for each  $x_0y_0 \dots x_{i-1}y_{i-1}x_i \in (X^{\text{Max}} \times X^{\text{Min}})^* \times X^{\text{Max}}$ . The notion of Player Min's positional strategy is similar.

An initial state and strategies of Max and Min uniquely determine a run.

**Definition 2.2.21** ( $\rho_{x, \mathfrak{s}^{\text{Max}}, \mathfrak{s}^{\text{Min}}}$ ). Let  $\mathfrak{G} = (X^{\text{Max}}, X^{\text{Min}}, E^{\text{Max}}, E^{\text{Min}}, W)$  be a two-player game. For  $x \in X^{\text{Max}}$ ,  $\mathfrak{s}^{\text{Max}} \in \mathfrak{S}_{\mathfrak{G}}^{\text{Max}}$  and  $\mathfrak{s}^{\text{Min}} \in \mathfrak{S}_{\mathfrak{G}}^{\text{Min}}$ , we inductively define a run  $\rho_{x, \mathfrak{s}^{\text{Max}}, \mathfrak{s}^{\text{Min}}} = x_0y_0x_1y_1 \dots \in \text{Run}_{\mathfrak{G}}^{\infty}(x)$  as follows.

- $x_0 := x$ ;
- If  $\mathfrak{s}^{\text{Max}}(x_i) = y \neq \perp$  then  $y_i := y$ . Otherwise, we let  $y_i := \perp_{\text{Max}}$  and the run ends here.
- If  $\mathfrak{s}^{\text{Min}}(y_i) = x' \neq \perp$  then  $x_{i+1} := x'$ . Otherwise, we let  $x_{i+1} := \perp_{\text{Min}}$  and the run ends here.

We are now ready to define the “winner” of a game. Player Max wins if a winning run  $\rho \in W$  is constructed or Player Min gets stuck. Player Min wins otherwise.

**Definition 2.2.22** (winning strategy/player/region). Let  $\mathfrak{G} = (X^{\text{Max}}, X^{\text{Min}}, E^{\text{Max}}, E^{\text{Min}}, W)$  be a two-player game. A run  $\rho = x_0y_0x_1y_1 \dots \in \text{Run}_{\mathfrak{G}}^{\infty}$  is *winning* if  $\rho \in W$  or it is a finite sequence whose last component is  $\perp_{\text{Min}}$ . For  $x \in X^{\text{Max}}$ , Player Max's strategy  $\mathfrak{s}^{\text{Max}} \in \mathfrak{S}_{\mathfrak{G}}^{\text{Max}}$  is *winning* from  $x$  if for each  $\mathfrak{s}^{\text{Min}} \in \mathfrak{S}_{\mathfrak{G}}^{\text{Min}}$  the run  $\rho_{x, \mathfrak{s}^{\text{Max}}, \mathfrak{s}^{\text{Min}}}$  is winning. Similarly, Player Min's strategy  $\mathfrak{s}^{\text{Min}} \in \mathfrak{S}_{\mathfrak{G}}^{\text{Min}}$  is *winning* from  $x$  if for each  $\mathfrak{s}^{\text{Max}} \in \mathfrak{S}_{\mathfrak{G}}^{\text{Max}}$  the run  $\rho_{x, \mathfrak{s}^{\text{Max}}, \mathfrak{s}^{\text{Min}}}$  is not winning. Player Max (resp. Player Min) is *winning* from  $x$  if there exists a winning strategy for Player Max (resp. Player Min) from  $x$ . The *winning region* (for Player Max) is the set  $\text{Win}_G \subseteq X^{\text{Max}}$  of states where Player Max is winning.

Note that it is *not* that if Player Max is not winning from  $x$  then Player Max is winning from  $x$ . A two-player game that satisfies this condition and its opposite is especially said to be *determinate*. In fact, all the two-player games that we consider in this thesis satisfy the following stronger determinacy.

**Definition 2.2.23** (positional determinacy). A two-player game  $\mathfrak{G} = (X^{\text{Max}}, X^{\text{Min}}, E^{\text{Max}}, E^{\text{Min}}, W)$  is *positionally determinate* if the following conditions are satisfied for each  $x \in X^{\text{Max}}$ .

- If Player Max is winning from  $x$ , then there exists a positional strategy  $\mathfrak{s}^{\text{Max}} \in \mathfrak{S}_{\mathfrak{G}}^{\text{Max}}$  that is winning from  $x$ .
- If Player Min is winning from  $x$ , then there exists a positional strategy  $\mathfrak{s}^{\text{Min}} \in \mathfrak{S}_{\mathfrak{G}}^{\text{Min}}$  that is winning from  $x$ .

A reachability game and a parity game are both defined as game structures equipped with additional data. Both of them induce a two-player game. They are known to be positionally determinate [41].

**Definition 2.2.24** (reachability game). A *reachability game* is a tuple  $\mathcal{T} = (X^{\text{Max}}, X^{\text{Min}}, E^{\text{Max}}, E^{\text{Min}}, \text{Acc})$  where the first four components constitute a game structure and  $\text{Acc} \subseteq X^{\text{Max}}$  is the set of *accepting states*. A reachability game induces a two-player game  $\mathfrak{G}_{\mathcal{T}} = (X^{\text{Max}}, X^{\text{Min}}, E^{\text{Max}}, E^{\text{Min}}, W)$  where  $W$  is the set of runs  $x_0y_0x_1y_1\dots$  such that  $x_i \in \text{Acc}$  for some  $i$ .

**Definition 2.2.25** (parity game [31]). Let  $n$  be a positive integer. A *parity game* is a tuple  $\mathcal{S} = (X^{\text{Max}}, X^{\text{Min}}, E^{\text{Max}}, E^{\text{Min}}, \mathfrak{p})$  where the first four components constitute a game structure and  $\mathfrak{p} : X^{\text{Max}} \rightarrow \{1, \dots, 2n\}$  is a *priority function*. A parity game  $G = (X^{\text{Max}}, X^{\text{Min}}, E^{\text{Max}}, E^{\text{Min}}, \mathfrak{p})$  induces a two-player game  $\mathfrak{G}_{\mathcal{S}} = (X^{\text{Max}}, X^{\text{Min}}, E^{\text{Max}}, E^{\text{Min}}, W)$  where  $W$  is the set of infinite runs  $x_0y_0x_1y_1\dots$  such that  $\limsup_{i \rightarrow \infty} \mathfrak{p}(x_i)$  is even.

We shall identify  $\mathcal{T}$  and  $\mathfrak{G}_{\mathcal{T}}$ , and  $\mathcal{S}$  and  $\mathfrak{G}_{\mathcal{S}}$  respectively.

## 2.2.4 Probabilistic Transition System

We define *probabilistic transition systems*, for which we will later introduce probabilistic ranking function notions. They are unlabeled systems.

**Definition 2.2.26** (probabilistic transition system). A *probabilistic transition system* (PTS) is a triple  $\mathcal{T} = ((X, \mathfrak{F}_X), \xi, \text{Acc})$  of a standard Borel space  $(X, \mathfrak{F}_X)$  called a *state space*, a measurable function  $\xi : (X, \mathfrak{F}_X) \rightarrow \mathcal{G}(X, \mathfrak{F}_X)$  called a *transition function*, and a measurable set  $\text{Acc} \subseteq X$  of *accepting states*.

We next define the notion of *reachability probability*. It is defined in a dual manner to a function  $\lim_{k \rightarrow \infty} \text{NoDiv}_k(x)$  in Definition 2.2.15.

**Definition 2.2.27** ( $\text{Reach}_{\mathcal{T}}$ ). Let  $\mathcal{T} = ((X, \mathfrak{F}_X), \xi, \text{Acc})$  be a PTS and  $A \subseteq X$ . For each  $k \in \omega$ , we inductively define a function  $\text{Reach}_{\mathcal{T}, A}^k : X \rightarrow [0, 1]$  as follows:

- $\text{Reach}_{\mathcal{T}, A}^0(x) := 0$ ; and
- $\text{Reach}_{\mathcal{T}, A}^{k+1}(x) := 1$  if  $x \in A$  and  $\int_{x' \in X} \text{Reach}_{\mathcal{T}, A}^k(x') d\xi(x)$  otherwise.

It is easy to see that  $k \leq k'$  implies  $\text{Reach}_{\mathcal{T}, A}^k(x) \leq \text{Reach}_{\mathcal{T}, A}^{k'}(x)$ . We define a *reachability probability function*  $\text{Reach}_{\mathcal{T}, A} : X \rightarrow [0, 1]$  as follows:

$$\text{Reach}_{\mathcal{T}, A}(x) := \lim_{k \rightarrow \infty} \text{Reach}_{\mathcal{T}, A}^k(x).$$

We write  $\text{Reach}_{\mathcal{T}}(x)$  for  $\text{Reach}_{\mathcal{T}, \text{Acc}}(x)$ .

By an easy induction on  $k$ , we can prove that  $\text{Reach}_{\mathcal{T}}$  is a measurable function.

## 2.3 Fixed Point Logic

In this thesis, the notion of *fixed point* plays a very important role.

### 2.3.1 Fixed Point Theorems

We shall review some theorems regarding fixed points. Some of them use the following transfinite-inductive construction of fixed points.

**Definition 2.3.1.** Let  $(L, \sqsubseteq)$  be a poset and  $f : L \rightarrow L$  be monotone function. For a post-fixed point  $l \in L$  of  $f$  (i.e.  $l \leq f(l)$ ) and an ordinal  $\mathfrak{a}$ , we define  $f^{\mathfrak{a}}(l) \in L$  by the transfinite induction as follows:



- $f^0(x) := l$ .
- If  $\alpha$  is a successor ordinal then  $f^\alpha(l) := f(f^{\alpha-1}(l))$ .
- If  $\alpha$  is a limit ordinal then  $f^\alpha(l) := \bigsqcup_{\alpha' < \alpha} f^{\alpha'}(l)$ , where  $\bigsqcup$  denotes the supremum. If such a supremum does not exist, then it is undefined.

Note that by the monotonicity of  $f$ ,  $\alpha < \alpha'$  implies  $f^\alpha(l) \sqsubseteq f^{\alpha'}(l)$ . We define  $f^\alpha(l)$  similarly when  $l \in L$  is a pre-fixed point, i.e.  $f(l) \leq l$ .

The following three constructions of least fixed points are well-known.

- Theorem 2.3.2.**
1. (Knaster-Tarski, see e.g. [105]) Assume that  $(L, \sqsubseteq)$  is a complete lattice. Then the set  $\{l \in L \mid f(l) \sqsubseteq l\}$  of pre-fixed points forms a complete lattice. Moreover, its least element is the least fixed point of  $f$ .
  2. (Kleene, see e.g. [101]) Assume that  $(L, \sqsubseteq)$  is  $\omega$ -complete (i.e. each increasing chain has the supremum) and  $f$  is  $\omega$ -continuous (i.e. it preserves  $\omega$ -supremums). If  $x$  is a post-fixed point of  $f$  then  $f^\omega(x)$  is a fixed point of  $f$ . Especially, if  $l$  is the least element then it is the least fixed point.
  3. (Cousot-Cousot, [27]) Assume  $(L, \sqsubseteq)$  be (upward) directed complete (i.e. each directed subset has the supremum). Let  $x$  be a post-fixed point of  $f$ . Then for each ordinal  $\alpha$ ,  $f^\alpha(x)$  is defined. Moreover, if  $|L| < |\alpha|$  then  $f^\alpha(x)$  is a fixed point of  $f$ . If  $l$  is the least element then it is the least fixed point.  $\square$

The above theorems and their dual statements provide us with the following well-known statements for over/under-approximating the least fixed point.

**Corollary 2.3.3.** Let  $l \in L$ . Assume that  $f$  has the least fixed point  $\mu f$ .

1. If  $(L, \sqsubseteq)$  is a complete lattice;  $f$  is  $\omega^{\text{op}}$ -continuous; or  $(L, \sqsubseteq)$  is downward directed complete, then  $f(l) \sqsubseteq l$  implies  $\mu f \sqsubseteq l$ .
2. Assume  $L$  has the least element  $\perp$ . Then for each ordinal  $\alpha$  such that  $f^\alpha(\perp)$  is defined, we have  $f^\alpha(\perp) \sqsubseteq \mu f$ .  $\square$

Corollary 2.3.3.1 is also called the Knaster-Tarski theorem. The dual theorems provide methods for over/under-approximating the greatest fixed point.

We end this section with the following theorem about the preservation of least fixed points. The result is new to us (hence we give a proof), but hardly original.

**Theorem 2.3.4.** Let  $(M, \sqsubseteq_M)$  and  $(N, \sqsubseteq_N)$  be posets,  $g : M \rightarrow M$  and  $h : N \rightarrow N$  be monotone endofunctions, and  $k : M \rightarrow N$  be a monotone function. Assume that  $g$  and  $h$  have the least fixed points  $\mu g$  and  $\mu h$  respectively. If the following conditions are satisfied, then we have  $k(\mu g) \sqsubseteq_N \mu h$ .

1.  $(M, \sqsubseteq_M)$  has the least element  $\perp_M$ . Moreover,  $k$  is strict, i.e.  $k(\perp_M)$  is the least element in  $(N, \sqsubseteq_N)$ .
2.  $k(g(l)) \sqsubseteq_N h(k(l))$  for each  $l \in M$ .
3. Either of the following conditions are satisfied:
  - $(M, \sqsubseteq_M)$  has the least element and is  $\omega$ -complete, and  $g$  is  $\omega$ -continuous. Moreover,  $k$  is  $\omega$ -continuous, i.e. for each sequence  $l_0 \sqsubseteq_M l_1 \sqsubseteq_M l_2 \sqsubseteq_M \dots$  in  $M$ , we have  $g(\bigsqcup_{i \in \omega} l_i) = \bigsqcup_{i \in \omega} g(l_i)$ .

- $(M, \sqsubseteq_M)$  has the least element and is upward directed complete. Moreover,  $k$  is (upward) directed-continuous, i.e. for each directed subset  $A \subseteq M$ , we have  $g(\bigsqcup_{l \in A} l) = \bigsqcup_{l \in A} g(l)$ .

Especially, if  $k(g(l)) = h(k(l))$  for each  $l \in M$  then  $k(\mu g) = \mu h$ .

**Proof.** We prove the statement when  $(M, \sqsubseteq_M)$  has the least element and is upward directed complete. The proof for the other case is similar.

Let  $\mathfrak{m}$  be an ordinal such that  $|M| < |\mathfrak{m}|$ . By Theorem 2.3.2.3,  $g^\alpha(\perp_M) \in M$  is defined for each an ordinal  $\alpha$ , and  $g^{\mathfrak{m}}(\perp_M) = \mu g$ . We prove  $k(g^\alpha(\perp_M)) \sqsubseteq_N \mu h$  by the transfinite induction on  $\alpha$ .

If  $\alpha = 0$ , by the assumptions, we have:

$$k(g^\alpha(\perp_M)) = k(\perp_M) = \perp_N \sqsubseteq_N \mu h.$$

Let  $\alpha$  be a successor ordinal and assume  $k(g^{\alpha-1}(\perp_M)) \sqsubseteq_N \mu h$ . Then we have,

$$k(g^\alpha(\perp_M)) = k(g(g^{\alpha-1}(\perp_M))) \sqsubseteq_N h(k(g^{\alpha-1}(\perp_M))) \sqsubseteq_N h(\mu h) = \mu h.$$

Let  $\alpha$  be a limit ordinal and assume  $k(g^{\alpha'}(\perp_M)) \sqsubseteq_N \mu h$  for each  $\alpha' < \alpha$ . Then,

$$k(g^\alpha(\perp_M)) = k\left(\bigsqcup_{\alpha' < \alpha} g^{\alpha'}(\perp_M)\right) \sqsubseteq_N k\left(\bigsqcup_{\alpha' < \alpha} \mu h\right) = h(\mu h) = \mu h.$$

Hence we have  $k(\mu g) = k(g^{\mathfrak{m}}(\perp_M)) \sqsubseteq_N \mu h$ .

Assume that  $k(g(l)) = h(k(l))$  for each  $l \in M$ . Then

$$h(k(\mu g)) = k(g(\mu g)) = k(\mu g).$$

Hence  $k(\mu g)$  is a fixed point of  $h$ , and together with  $k(\mu g) \sqsubseteq_N \mu h$ , we have that  $k(\mu g) = \mu h$ .  $\square$

By reversing the order, we can prove its dual statement about the preservation of greatest fixed points.

### 2.3.2 Hierarchical Equation System

A *hierarchical equation systems* (HES for short) is a representation of an alternating fixed point used in [26, 6].

**Definition 2.3.5** (HES). Let  $m \in \mathbb{N}$ . A *hierarchical equation system* (HES) is a finite family of equations of the following form.

$$E = \begin{cases} u_1 =_{\eta_1} f_1(u_1, \dots, u_m) \in (L_1, \sqsubseteq_1) \\ u_2 =_{\eta_2} f_2(u_1, \dots, u_m) \in (L_2, \sqsubseteq_2) \\ \vdots \\ u_m =_{\eta_m} f_m(u_1, \dots, u_m) \in (L_m, \sqsubseteq_m) \end{cases}$$

Here for each  $i \in \{1, \dots, m\}$ ,  $(L_i, \sqsubseteq_i)$  is a poset,  $u_i$  is a variable that ranges over  $L_i$ ,  $\eta_i \in \{\mu, \nu\}$  and  $f_i : L_1 \times \dots \times L_m \rightarrow L_i$  is a monotone function.

We write  $\sqsubseteq$  for  $\sqsubseteq_i$  if no confusion is likely.

An HES is commonly defined over complete lattices (see e.g. [26, 6, 49]). However, in this thesis, we do not assume  $L_i$  to be a complete lattice. See Remark 3.8.2 for an example of HES over posets that are not complete lattices.

An HES determines a fixed point of a function  $f_1 \times \dots \times f_m : L_1 \times \dots \times L_m \rightarrow L_1 \times \dots \times L_m$ .

**Definition 2.3.6** (solution of HES). Let  $E$  be an HES as in Def. 2.3.5. For each  $i \in \{1, \dots, m\}$  and  $j \in \{1, \dots, i\}$  we inductively define functions  $f_i^\ddagger : L_i \times \dots \times L_m \rightarrow L_i$  and  $l_j^{(i)} : L_{i+1} \times \dots \times L_m \rightarrow L_j$  as follows (we do not have to distinguish the base case from the step case):

- $f_i^\ddagger(u_i, \dots, u_m) := f_i(l_1^{(i-1)}(u_i, \dots, u_m), \dots, l_{i-1}^{(i-1)}(u_i, \dots, u_m), u_i, \dots, u_m)$ .
- $l_i^{(i)}(u_{i+1}, \dots, u_m)$  is the least (resp. greatest) fixed point of  $f_i^\ddagger(\_, u_{i+1}, \dots, u_m) : L_i \rightarrow L_i$  if  $\eta_i = \mu$  (resp.  $\nu$ ). If such a least or greatest fixed point does not exist, then  $l_i^{(i)}(u_{i+1}, \dots, u_m)$  is undefined. For  $j < i$ , we let  $l_j^{(i)}(u_{i+1}, \dots, u_m) := l_j^{(i-1)}(l_i^{(i)}(u_{i+1}, \dots, u_m), u_{i+1}, \dots, u_m)$ .

We call  $(l_1^{(i)}, \dots, l_i^{(i)})$  the  $i$ -th *intermediate solution* (if it exists). Note that  $l_i^{(m)}$  has a type  $1 \rightarrow L_i$  where 1 denotes a singleton. The *solution* of the HES  $E$  is a tuple  $(u_1^{\text{sol}}, \dots, u_m^{\text{sol}}) \in L_1 \times \dots \times L_m$  defined by  $u_i^{\text{sol}} := l_i^{(m)}(*)$  for each  $i$ , where  $*$  denotes the unique element in 1.

## 2.4 Categorical Preliminaries

Category theory plays the central role in this thesis. In this chapter we review categorical notions.

### 2.4.1 Preliminaries on Basic Category Theories

We first review the notions of *category*, *functor*, *natural transformation*, *limit*, *colimit*, *product*, *coproduct*, *final object* and *initial object*. They are all basic. See also [76, 10, 60].

**Definition 2.4.1** (category). A (locally small) *category* is a triple  $\mathbb{C} = (|\mathbb{C}|, \text{arr}(\mathbb{C}), \circ)$  consisting of the following components.

- A class<sup>2</sup>  $|\mathbb{C}|$  of *objects*. We shall write  $X \in \mathbb{C}$  to mean  $X \in |\mathbb{C}|$ .
- A family  $\text{arr}(\mathbb{C}) = (\mathbb{C}(X, Y))_{X, Y \in |\mathbb{C}|}$  of (small) sets. An element in  $\mathbb{C}(X, Y)$  is called an *arrow*. When  $f \in \mathbb{C}(X, Y)$ , we write  $f : X \rightarrow Y$  and call  $X$  and  $Y$  the *domain* and the *codomain* of  $f$  respectively.
- A family  $\circ = (\_ \circ_{X, Y, Z} \_ : \text{arr}(\mathbb{C})(Y, Z) \times \text{arr}(\mathbb{C})(X, Y) \rightarrow \text{arr}(\mathbb{C})(X, Z))_{X, Y, Z \in |\mathbb{C}|}$  of *composite functions*. We often omit subscripts and just write  $\circ$  for  $\circ_{X, Y, Z}$ .

We require that the following conditions are satisfied.

- For each  $X \in |\mathbb{C}|$ , there exists an *identity arrow*  $\text{id}_X : X \rightarrow X$  such that:  $\text{id}_X \circ f = f$  holds for each  $Y \in |\mathbb{C}|$  and  $f : Y \rightarrow X$ , and  $g \circ \text{id}_X = g$  holds for each  $Z \in |\mathbb{C}|$  and  $g : Z \rightarrow X$ . We sometimes write  $\text{id}$  for  $\text{id}_X$ .
- For each  $X, Y, Z, W \in |\mathbb{C}|$ ,  $f : X \rightarrow Y$ ,  $g : Y \rightarrow Z$  and  $h : Z \rightarrow W$ ,  $(h \circ g) \circ f = h \circ (g \circ f)$ .

It is easy to see that an identity arrow  $\text{id}_X$  is unique for each  $X$ .

**Example 2.4.2.** Here are some examples of categories.

<sup>2</sup>Briefly speaking, a *class* is a collection of mathematical objects specified by some property. See e.g. [65] for more details.

- The category **Sets** of sets. Its objects are given by all the sets, its arrows are functions between them, and its compositions are given by the usual compositions of functions.
- The category **Meas** of measurable spaces. Its objects are given by all the measurable sets its arrows are measurable functions between them, and its compositions are given by the usual compositions of functions.
- The category **SB** of standard Borel spaces. Its objects are all the standard Borel spaces, and its arrows and their compositions are defined as in **Meas**.

The objects of **SB** constitute a subclass of **Meas**, and for each  $X, Y \in \mathbf{SB}$ ,  $\mathbf{SB}(X, Y) = \mathbf{Meas}(X, Y)$ . We say that **SB** is a *full subcategory* of **Meas** for this situation. Throughout this thesis, we shall use **SB** instead of **Meas** to characterize probabilistic systems.

Two sets  $X$  and  $Y$  are *isomorphic* if there exists a bijection  $b: X \xrightarrow{\cong} Y$ . The notion of isomorphism can be also defined for measurable spaces. The following definition generalizes them.

**Definition 2.4.3** (isomorphism). Let  $\mathbb{C}$  be a category and  $X, Y \in \mathbb{C}$ . We say  $X$  is *isomorphic* to  $Y$  and write  $X \cong Y$  if there exists  $f: X \rightarrow Y$  and  $g: Y \rightarrow X$  such that  $g \circ f = \text{id}_X$  and  $f \circ g = \text{id}_Y$ .

A *functor* is an operation that maps objects to objects and arrows to arrows.

**Definition 2.4.4** (functor). Let  $\mathbb{C}$  and  $\mathbb{D}$  be categories. A *functor* from  $\mathbb{C}$  to  $\mathbb{D}$  is a pair  $F = (F_{\text{obj}}, F_{\text{arr}})$  of functions  $F_{\text{obj}}: |\mathbb{C}| \rightarrow |\mathbb{D}|$  and  $F_{\text{arr}}: \text{arr}(\mathbb{C}) \rightarrow \text{arr}(\mathbb{D})$  that satisfy the following conditions:

- if  $f: X \rightarrow Y$  then  $F_{\text{arr}}f: F_{\text{obj}}X \rightarrow F_{\text{obj}}Y$ ;
- for  $X \in \mathbb{C}$ ,  $F_{\text{arr}}\text{id}_X = \text{id}_{F_{\text{obj}}X}$ ; and
- for  $X, Y, Z \in \mathbb{C}$ ,  $f: X \rightarrow Y$  and  $g: Y \rightarrow Z$ ,  $(F_{\text{arr}}g) \circ (F_{\text{arr}}f) = F_{\text{arr}}(g \circ f)$ .

We write  $F: \mathbb{C} \rightarrow \mathbb{D}$  when  $F$  is a functor from  $\mathbb{C} \rightarrow \mathbb{D}$ . We often omit subscripts and just write  $F$  for both  $F_{\text{obj}}$  and  $F_{\text{arr}}$ . A functor from a category to the same category is called an *endofunctor*.

Here are some spacial functors.

**Definition 2.4.5** ( $\text{id}_{\mathbb{C}}$ ,  $\Delta_X$ ,  $FG$ ,  $F^n$ ). For a category  $\mathbb{C}$ , we write  $\text{id}_{\mathbb{C}}$  for the *identity functor* that maps objects and arrows to themselves, and omit the subscript if no confusion is likely. For categories  $\mathbb{C}, \mathbb{D}$  and  $X \in \mathbb{C}$ ,  $\Delta_X: \mathbb{D} \rightarrow \mathbb{C}$  denotes the *constant functor* that maps each object to  $X$  and each arrow to  $\text{id}_X$ . For functors  $F: \mathbb{C} \rightarrow \mathbb{D}$  and  $G: \mathbb{D} \rightarrow \mathbb{E}$ , we define the *composite functor*  $GF: \mathbb{C} \rightarrow \mathbb{E}$  by  $GF X := G(FX)$  for each an object  $X \in \mathbb{C}$  and  $GF f := G(Ff)$  for each an arrow  $f$  in  $\mathbb{C}$ . For an endofunctor  $F: \mathbb{C} \rightarrow \mathbb{C}$ , we write  $F^n$  for  $\underbrace{FF \dots F}_n: \mathbb{C} \rightarrow \mathbb{C}$

Here are examples of functors.

**Definition 2.4.6.** Let  $\Sigma$  be a ranked alphabet.

- The *powerset functor*  $\mathcal{P}: \mathbf{Sets} \rightarrow \mathbf{Sets}$  is defined by  $\mathcal{P}X := \{A \subseteq X\}$  for  $X \in \mathbf{Sets}$  and  $\mathcal{P}f(A) := \{f(x) \in Y \mid x \in A\}$  for  $f: X \rightarrow Y$ .

- The *lift functor*  $\mathcal{L} : \mathbf{Sets} \rightarrow \mathbf{Sets}$  is defined by  $\mathcal{L}X := \{\perp\} + X$  for  $X \in \mathbf{Sets}$  and  $\mathcal{L}f(x) := f(x)$  if  $x \in X$  and  $\perp$  if  $x = \perp$  for  $f: X \rightarrow Y$ .
- The *distribution functor*  $\mathcal{D} : \mathbf{Sets} \rightarrow \mathbf{Sets}$  is defined as follows. For  $X \in \mathbf{Sets}$ ,  $\mathcal{D}X$  is defined as in Section 2.1. Moreover,  $\mathcal{D}f(\xi)(y) := \sum_{x \in f^{-1}(y)} \xi(x)$  for  $f: X \rightarrow Y$ ,  $\xi \in \mathcal{D}X$  and  $y \in Y$ .
- The *subdistribution functor*  $\mathcal{D}_s : \mathbf{Sets} \rightarrow \mathbf{Sets}$  is similarly defined.
- The *Giry functor*  $\mathcal{G} : \mathbf{SB} \rightarrow \mathbf{SB}$  [39, 28] is defined as follows. For an object  $(X, \mathfrak{F}_X) \in \mathbf{SB}$ ,  $\mathcal{G}(X, \mathfrak{F}_X)$  is defined as in Section 2.1. For an arrow  $f : (X, \mathfrak{F}_X) \rightarrow (Y, \mathfrak{F}_Y)$ ,  $\mathcal{G}f(\alpha)(B) := \alpha(f^{-1}(B))$  for  $\alpha \in \mathcal{G}X$  and  $B \in \mathfrak{F}_Y$ .
- The *sub-Giry functor*  $\mathcal{G}_s : \mathbf{SB} \rightarrow \mathbf{SB}$  is similarly defined.
- We define  $F_\Sigma : \mathbf{Sets} \rightarrow \mathbf{Sets}$  by  $F_\Sigma X := \coprod_{i \in \omega} \Sigma_i \times X^i$  for a set  $X$  and  $F_\Sigma f(a, x_0, \dots, x_{n-1}) := (a, f(x_0), \dots, f(x_{n-1}))$  for a function  $f: X \rightarrow Y$ .
- We define  $F_\Sigma : \mathbf{SB} \rightarrow \mathbf{SB}$  as follows: for objects,  $F_\Sigma(X, \mathfrak{F}_X) := (F_\Sigma X, \mathfrak{F}_{F_\Sigma X})$  where  $\mathfrak{F}_{F_\Sigma X}$  is the smallest  $\sigma$ -algebra including  $\{\{a\} \times A_1 \times \dots \times A_n \mid n \in \omega, a \in \Sigma_n, A_1, \dots, A_n \in \mathfrak{F}_X\}$ . For arrows, it is defined as above. It is easy to prove measurability of the resulting functions.
- We define  $F_g : \mathbf{Sets} \rightarrow \mathbf{Sets}$  by  $F_g X := \mathcal{P}^2 X \times \{0, 1\}$  for a set  $X$  and  $F_g f(\Gamma, t) := (\{\{f(x) \mid x \in A\} \mid A \in \Gamma\}, t)$  for a function  $f: X \rightarrow Y$ . We use this functor to model a reachability game later.
- We define  $F_p : \mathbf{SB} \rightarrow \mathbf{SB}$  by  $F_p X := \mathcal{G}(X, \mathfrak{F}_X) \times (\{0, 1\}, \mathcal{P}\{0, 1\})$  for a measurable space  $(X, \mathfrak{F}_X)$  and  $F_p f(\delta, t) := (\mathcal{G}f(\delta), t)$  for a measurable function  $f : (X, \mathfrak{F}_X) \rightarrow (Y, \mathfrak{F}_Y)$ , where  $\mathcal{G}$  is defined as above. We use this functor to model a PTS later.

A *natural transformation* is a morphism from a functor to a functor.

**Definition 2.4.7** (natural transformation). Let  $\mathbb{C}$  and  $\mathbb{D}$  be categories and  $F, G : \mathbb{C} \rightarrow \mathbb{D}$  be functors. A *natural transformation* from  $F$  to  $G$  is a family  $\alpha = (\alpha_X : FX \rightarrow GX)_{X \in |\mathbb{C}|}$  of arrows in  $\mathbb{D}$  that satisfies  $\alpha_Y \circ Ff = Gf \circ \alpha_X$  for each  $X, Y \in \mathbb{C}$  and  $f: X \rightarrow Y$ . The last equality is called *naturality*, and pictorially,

$$\begin{array}{ccc} FY & \xrightarrow{\alpha_Y} & GY \\ \uparrow Ff & & \uparrow Gf \\ FX & \xrightarrow{\alpha_X} & GX \end{array} \quad \text{in } \mathbb{D}.$$

We write  $\alpha : F \Rightarrow G$  when  $\alpha$  is a natural transformation from  $F$  to  $G$ .

We review the notions of *(co)cone* and *(co)limit*. A cone over a diagram is a pair of an object  $L$  and arrows from  $L$  to the diagram that is “compatible” with the diagram in a certain sense. A limit is a special cone that is most “universal.”

**Definition 2.4.8** ((co)cone and (co)limit). Let  $G : \mathbb{D} \rightarrow \mathbb{C}$  be a functor. Recall that  $\Delta_X$  denotes the constant functor.

- A *cone* over  $G$  is a pair  $(X \in \mathbb{C}, \alpha : \Delta_X \Rightarrow G)$ . A *limit* over  $G$  is a cone  $(L, \gamma)$  over  $G$  such that for an arbitrary cone  $(X, \alpha)$  over  $G$  there exists a unique arrow  $m: X \rightarrow L$  such that  $\gamma_D \circ m = \alpha_D$  for each  $D \in \mathbb{D}$ .

- A *cocone* over  $G$  is a pair  $(X \in \mathbb{C}, \beta : G := \Delta_X)$ . A *colimit* over  $G$  is a cocone  $(L, \xi)$  over  $G$  such that for an arbitrary cocone  $(X, \beta)$  over  $G$  there exists a unique arrow  $m : L \rightarrow X$  such that  $m \circ \xi_D = \beta_D$  for each  $D \in \mathbb{D}$ .

In both cases, the unique arrow  $m$  is called the *mediating arrow*.

It is known that a (co)limit is unique up to isomorphism if it exists.

Suppose that we are given a diagram on a category  $\mathbb{C}$ , i.e. a subset of objects and some arrows between them. We can define the smallest category that contains the objects and the arrows, and there exists the canonical inclusion functor  $I$  from the category to  $\mathbb{C}$ . We can thus regard a cone over  $I$  as a cone over the diagram, and when the functor has a limit, we can call it a *limit of the diagram*. A *colimit of a diagram* is similarly defined.

A functor  $F : \mathbb{C} \rightarrow \mathbb{E}$  maps a cone over a functor  $G : \mathbb{D} \rightarrow \mathbb{C}$  to a cone over a functor  $FG : \mathbb{D} \rightarrow \mathbb{E}$ . We are sometimes interested in a situation where if the former is a limit then also the latter is.

**Definition 2.4.9.** Let  $(L, \gamma)$  be a limit of a functor  $G : \mathbb{D} \rightarrow \mathbb{C}$ . We say that an endofunctor  $F : \mathbb{C} \rightarrow \mathbb{E}$  *preserves the limit* if a cone  $(FL, (G\gamma_D)_{D \in \mathbb{D}})$  over  $FG : \mathbb{D} \rightarrow \mathbb{C}$  is a limit. Preservation of a colimit is similarly defined.

We conclude this section by reviewing *(co)products* and *final* and *initial* objects. They are special (co)limits.

**Definition 2.4.10** ((co)product). Let  $\mathbb{C}$  be a category and  $(X_i)_{i \in I}$  be a family of objects.

- A *product* of  $(X_i)_{i \in I}$  is a limit over the discrete diagram consisting of  $(X_i)_{i \in I}$ , and it is denoted by  $(\prod_{i \in I} X_i, (\pi_i)_{i \in I})$ . When  $I$  is a finite set  $\{1, \dots, n\}$ ,  $\prod_{i \in I} X_i$  is also denoted by  $X_1 \times \dots \times X_n$ . Moreover, for a family  $(f_i : Y \rightarrow X_i)_{i \in I}$  of arrows, we write  $\langle f_i \rangle_{i \in I}$  for the mediating arrow from a cone  $(Y, (f_i)_{i \in \omega})$  to  $(\prod_{i \in I} X_i, (\pi_i)_{i \in I})$ .
- A *coproduct* of  $(X_i)_{i \in I}$  is a colimit over the discrete diagram consisting of  $(X_i)_{i \in I}$ , and it is denoted by  $(\coprod_{i \in I} X_i, (\kappa_i)_{i \in I})$ . When  $I$  is a finite set  $\{1, \dots, n\}$ ,  $\coprod_{i \in I} X_i$  is also denoted by  $X_1 + \dots + X_n$ . Moreover, for a family  $(f_i : X_i \rightarrow Y)_{i \in I}$  of arrows, we write  $[f_i]_{i \in I}$  for the mediating arrow from  $(\coprod_{i \in I} X_i, (\pi_i)_{i \in I})$  to a cocone  $(Y, (f_i)_{i \in \omega})$ .

**Definition 2.4.11** (final/initial object). Let  $\mathbb{C}$  be a category. A *final object* is an object  $1 \in \mathbb{C}$  such that for each  $Y \in \mathbb{C}$  there exists a unique arrow  $!_Y : Y \rightarrow 1$ . An *initial object* is an object  $0 \in \mathbb{C}$  such that for each  $Y \in \mathbb{C}$  there exists a unique arrow  $i_Y : 0 \rightarrow Y$ . We sometimes omit subscripts and just write  $!$  and  $i$ .

**Example 2.4.12.** In the category **Sets**, a product  $X \times Y$  is given by a set-theoretic product, a coproduct  $X + Y$  is given by a disjoint sum, a final object is a singleton and an initial object is the empty set.

In **SB**, a product of  $(X, \mathfrak{F}_X)$  and  $(Y, \mathfrak{F}_Y)$  is  $(X \times Y, \mathfrak{F}_{X \times Y})$  where  $\mathfrak{F}_{X \times Y}$  is the smallest  $\sigma$ -algebra containing  $\{A \times B \mid A \in \mathfrak{F}_X, B \in \mathfrak{F}_Y\}$ . A coproduct of  $(X, \mathfrak{F}_X)$  and  $(Y, \mathfrak{F}_Y)$  is  $(X + Y, \mathfrak{F}_{X+Y})$  where  $\mathfrak{F}_{X+Y}$  is the smallest  $\sigma$ -algebra containing  $\mathfrak{F}_X$  and  $\mathfrak{F}_Y$ . A final object is a singleton equipped with the trivial  $\sigma$ -algebra, and an initial object is the empty set with the trivial  $\sigma$ -algebra.

## 2.4.2 Algebra and Coalgebra

The following notions, especially that of *coalgebra*, are central in this thesis.

**Definition 2.4.13** ((co)algebra and homomorphism). Let  $F : \mathbb{C} \rightarrow \mathbb{C}$ .

- An  $F$ -*algebra* is an arrow of a form  $a : FX \rightarrow X$ . We call  $X$  the *carrier* of  $a$ . For  $F$ -algebras  $a : FX \rightarrow X$  and  $b : FY \rightarrow Y$ , a homomorphism from  $a$  to  $b$  is an arrow  $f : X \rightarrow Y$  such that  $f \circ a = b \circ Ff$ .
- An  $F$ -*coalgebra* is an arrow of a form  $c : X \rightarrow FX$ . We call  $X$  the *carrier* of  $c$ . For  $F$ -coalgebras  $c : X \rightarrow FX$  and  $d : D \rightarrow FD$ , a homomorphism from  $c$  to  $d$  is an arrow  $g : X \rightarrow Y$  such that  $Ff \circ c = d \circ f$ .

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ \downarrow a & & \downarrow b \\ FX & \xrightarrow{Ff} & FY \end{array} \qquad \begin{array}{ccc} FX & \xrightarrow{Fg} & FY \\ \uparrow c & & \uparrow d \\ X & \xrightarrow{g} & Y \end{array}$$

Various transition systems are representable as coalgebras. For example, using the functors in Definition 2.4.6, we can model the following transition systems.

**Example 2.4.14.** Let  $\Sigma$  be a ranked alphabet.

- A  $\Sigma$ -labeled NTA  $\mathcal{A} = (X, \tau)$  (Definition 2.2.4) can be modeled as a  $\mathcal{P}F_{\Sigma}$ -coalgebra  $c : X \rightarrow \mathcal{P} \coprod_{i=0}^{\infty} \Sigma_i \times X^i$  defined by  $c := \tau$ .
- A  $\Sigma$ -labeled PTA  $\mathcal{A} = ((X, \mathfrak{F}_X), \xi)$  (Definition 2.2.12) can be modeled as a  $\mathcal{G}F_{\Sigma}$ -coalgebra  $c : (X, \mathfrak{F}_X) \rightarrow \mathcal{G} \coprod_{i=0}^{\infty} \Sigma_i \times (X, \mathfrak{F}_X)^i$  defined by  $c := \xi$ .
- A reachability game  $\mathcal{T} = (X^{\text{Max}}, X^{\text{Min}}, E^{\text{Max}}, E^{\text{Min}}, \text{Acc})$  (Definition 2.2.24) is modeled as an  $F_{\mathfrak{g}}$ -coalgebra  $c_{\mathcal{T}} : X^{\text{Max}} \rightarrow \mathcal{P}^2 X^{\text{Max}} \times \{0, 1\}$  defined by  $c_{\mathcal{T}}(x) := (\Gamma, t)$  where

$$\Gamma = \left\{ \{x' \mid (y, x') \in E^{\text{Min}}\} \mid (x, y) \in E^{\text{Max}} \right\} \quad \text{and} \quad t = \begin{cases} 1 & (x \in \text{Acc}) \\ 0 & (x \notin \text{Acc}). \end{cases}$$

- A PTS  $\mathcal{T} = ((X, \mathfrak{F}_X), \xi)$  (Definition 2.2.26) is modeled as an  $F_{\mathfrak{p}}$ -coalgebra  $c_{\mathcal{T}} : (X, \mathfrak{F}_X) \rightarrow \mathcal{G}(X, \mathfrak{F}_X) \times \{0, 1\}$  defined by  $c_{\mathcal{T}}(x) := (\xi(x), t)$  where  $t$  is defined as above.

Algebras and homomorphisms constitute a category. Its initial object is called an *initial algebra*. We are also interested in its dual notion, *final coalgebra*.

**Definition 2.4.15** (initial algebra & final coalgebra). Let  $F : \mathbb{C} \rightarrow \mathbb{C}$ .

- An algebra  $a : FA \rightarrow A$  is *initial* if for an arbitrary algebra  $b : FB \rightarrow B$  there exists a unique homomorphism from  $a$  to  $b$ .
- A coalgebra  $c : X \rightarrow FX$  is *final* if for an arbitrary coalgebra  $d : Y \rightarrow FY$  there exists a unique homomorphism from  $d$  to  $c$ .

When we regard categories as generalizations of preordered sets (see Section 1.3.2), an  $F$ -algebra is understood as a *pre-fixed point* of  $F$ , and an algebra's being initial means that it is the *least* pre-fixed point. Similarly, an  $F$ -coalgebra is a *post-fixed point* of  $F$ , and a final coalgebra is the *greatest* post-fixed point.

It is not hard to prove that if  $x$  and  $y$  are both least pre-fixed (or greatest post-fixed) points, then  $x \leq y$  and  $y \leq x$ . Moreover, by the Knaster-Tarski theorem (Theorem 2.3.2.1), the least pre-fixed (resp. greatest post-fixed) point is the least (resp. greatest) fixed point. The following proposition categorically generalizes them.

**Proposition 2.4.16** (see e.g. [60]). *Let  $F : \mathbb{C} \rightarrow \mathbb{C}$ .*

1. *Initial algebras are unique up-to isomorphism. That is, if  $\iota : FX \rightarrow X$  and  $\iota' : FX' \rightarrow X'$  are initial  $F$ -algebras then there exists an isomorphism  $f : X \xrightarrow{\cong} X'$  such that  $f \circ \iota = \iota' \circ Ff$ .*
2. *An initial algebra is an isomorphism.*
3. *Final coalgebras are unique up-to isomorphism. That is, if  $\zeta : X \rightarrow FX$  and  $\zeta' : X' \rightarrow FX'$  are final  $F$ -coalgebras then there exists an isomorphism  $f : X \xrightarrow{\cong} X'$  such that  $Ff \circ \zeta = \zeta' \circ f$ .*
4. *A final coalgebra is an isomorphism.* □

The following theorem generalizes the Kleene fixed-point theorem (Theorem 2.3.2.2), and shows a way to construct initial algebras and final coalgebras.

**Theorem 2.4.17** ([3]). *Let  $F : \mathbb{C} \rightarrow \mathbb{C}$ .*

- *Assume that  $\mathbb{C}$  has an initial object  $0$ ; an initial sequence  $0 \xrightarrow{i_{F0}} F0 \xrightarrow{F!_{F0}} F^2 1 \xrightarrow{F^2!_{F0}} \dots$  has a colimit  $(A, (\xi_i : A \rightarrow F^i 1)_{i \in \omega})$ ; and  $F$  preserves the colimit. Then the unique mediating arrow  $\iota^F : FA \rightarrow A$  from the colimit  $(FA, (F\xi_i : F^{i+1}0 \rightarrow A))$  to a cocone  $(Z, (\xi_{i+1} : F^{i+1}0 \rightarrow A)_{i \in \omega})$  is an initial  $F$ -algebra.*
- *Assume that  $\mathbb{C}$  has a final object  $1$ ; a final sequence  $1 \xleftarrow{!_{F1}} F1 \xleftarrow{F!_{F1}} F^2 1 \xleftarrow{F^2!_{F1}} \dots$  has a limit  $(Z, (\gamma_i : Z \rightarrow F^i 1)_{i \in \omega})$ ; and  $F$  preserves the limit. Then the unique mediating arrow  $\zeta^F : Z \rightarrow FZ$  from a cone  $(Z, (\gamma_{i+1} : Z \rightarrow F^{i+1}1)_{i \in \omega})$  to the limit  $(FZ, (F\gamma_i : FZ \rightarrow F^{i+1}1))$  is a final  $F$ -coalgebra.*

$$\begin{array}{ccc}
 & & Z \\
 & \nearrow^{\xi_0} & \uparrow^{\pi_2} \\
 0 & \xrightarrow{i} F0 & \xrightarrow{F!} F^2 0 \xrightarrow{\dots} \cong \iota \\
 & \searrow_{F\xi_0} & \downarrow_{F\xi_1} \\
 & & FZ
 \end{array}
 \qquad
 \begin{array}{ccc}
 & & Z \\
 & \nwarrow_{\gamma_0} & \downarrow^{\pi_2} \\
 1 & \xleftarrow{!} F1 & \xleftarrow{F!} F^2 1 \xleftarrow{\dots} \cong \zeta \\
 & \nwarrow_{F\gamma_0} & \downarrow_{F\gamma_1} \\
 & & FZ
 \end{array}
 \quad \square$$

**Example 2.4.18** (see e.g. [113]). Define  $F_\Sigma : \mathbf{Sets} \rightarrow \mathbf{Sets}$  as in Definition 2.4.6. The carriers of an initial  $F_\Sigma$ -algebra and a final  $F_\Sigma$ -coalgebra are isomorphic to  $\text{Tree}_\Sigma^*$  and  $\text{Tree}_\Sigma^\infty$  respectively (recall that an initial algebra and a final coalgebra are unique up to isomorphisms). When  $F_\Sigma : \mathbf{SB} \rightarrow \mathbf{SB}$ , the carriers are isomorphic to measurable sets  $(\text{Tree}_\Sigma^*, \mathfrak{F}_{\text{Tree}_\Sigma^*})$  and  $(\text{Tree}_\Sigma^\infty, \mathfrak{F}_{\text{Tree}_\Sigma^\infty})$  where the  $\sigma$ -algebras are given as in Section 2.2.4.

Final coalgebras are very important in the theory of coalgebra, because the unique homomorphism to it often captures “behaviors” of a coalgebra.

**Example 2.4.19.** We continue Examples 2.4.14. An  $F_\Sigma$ -coalgebra has a type  $c : X \rightarrow \coprod_{i=0}^\infty \Sigma_i \times X^i$ , and can be understood as a  $\Sigma$ -labeled deterministic tree automaton. Especially, if  $F = A \times (\_)$ , then an  $F$ -coalgebra is understood as an  $A$ -labeled deterministic word automaton (cf. Remark 2.2.11).

For an  $F_\Sigma$ -coalgebra  $c : X \rightarrow F_\Sigma X$ , the unique homomorphism  $X \rightarrow \text{Tree}_\Sigma^\infty$  from  $c$  to  $\zeta$  assigns each  $x$  the unique tree  $t = (D, l)$  that satisfies the following condition: there exists a  $X \times \Sigma$ -labeled tree  $\rho = (D', l')$  such that i)  $\pi_1(l_\rho(\langle \rangle)) = x$ ; and ii) for each  $w \in D_r$ , if  $l_\rho(w) = (x, a)$  and  $l_\rho(wi) = (x_i, a_i)$  for each  $i \in \{0, \dots, |a| - 1\}$  then  $(a, x_0, \dots, x_{|a|-1}) = c(x)$  (cf. Definition 2.2.5). This function can be regarded as a “language” of  $c$ .



However, this framework to characterize behaviors via final coalgebras does work for some systems. One of the reasons is that a final coalgebra does not exist for some functors. A counterexample is as follows.

**Example 2.4.20.** Assume that there exists a final  $\mathcal{P}$ -coalgebra  $\zeta^{\mathcal{P}}: X \rightarrow \mathcal{P}X$ . By Proposition 2.4.16.4, we have  $X \cong \mathcal{P}X$ , but it is a well-known fact that such a set  $X$  does not exist.

Similarly if  $\Sigma_i \neq \emptyset$  for some  $i > 0$  then functors  $\mathcal{P}F_\Sigma$ ,  $\mathcal{G}_s F_\Sigma$ ,  $F_g$  and  $F_p$  (see Example 2.4.14) do not have final coalgebras. We will later review two types of extension of the “final coalgebra framework” that work for those functors (Section 3.1 and Section 5.2).

We conclude this section with a notion of *coalgebra-algebra homomorphism*. As its name suggests, it is a homomorphism from a coalgebra to algebra.

**Definition 2.4.21.** Let  $F: \mathbb{C} \rightarrow \mathbb{C}$ . A *coalgebra-algebra homomorphism* from an  $F$ -coalgebra  $c: X \rightarrow FX$  and an  $F$ -algebra  $\sigma: F\Omega \rightarrow \Omega$ , is an arrow  $f: X \rightarrow \Omega$  such that  $f = \sigma \circ Ff \circ c$ .

$$\begin{array}{ccc} FX & \xrightarrow{Ff} & F\Omega \\ c \uparrow & \bar{f} & \downarrow \sigma \\ X & \xrightarrow{f} & \Omega \end{array}$$

The above notion is useful when we are combining category theory with fixed point logic. The main reason is that a coalgebra-algebra homomorphism is a fixed point of the following function.

**Definition 2.4.22.** Let  $F: \mathbb{C} \rightarrow \mathbb{C}$ . For an  $F$ -coalgebra  $c: X \rightarrow FX$  and an  $F$ -algebra  $\sigma: F\Omega \rightarrow \Omega$ , we define a function  $\Phi_{c,\sigma}: \mathbb{C}(X, \Omega) \rightarrow \mathbb{C}(X, \Omega)$  by  $\Phi_{c,\sigma}(f) := \sigma \circ Ff \circ c$ , i.e.

$$\Phi_{c,\sigma} : \left( X \xrightarrow{f} \Omega \right) \mapsto \left( \begin{array}{ccc} FX & \xrightarrow{Ff} & F\Omega \\ c \uparrow & & \downarrow \sigma \\ X & & \Omega \end{array} \right).$$

**Lemma 2.4.23.** *An arrow  $f: X \rightarrow \Omega$  is a coalgebra-algebra homomorphism from  $c$  to  $\sigma$  if and only if it is a fixed point of  $\Phi_{c,\sigma}$ .  $\square$*

## Chapter 3

# Categorical Trace Semantics for Büchi and Parity Automata

One of the main goals of this thesis is to categorically generalize fair simulation, a simulation notion for NBTAs. To state correctness of our categorical generalization, we first have to categorically characterize languages  $L_{\mathcal{A}}^{\text{B}}$  of NBTAs (Definition 2.2.9). In this chapter, we achieve this goal by extending an existing framework called *Kleisli approach*. We will also categorically characterize languages of NPTAs.

We will introduce two categorical characterizations for languages of Büchi and parity automata. We shall call the characterizations a *logical fixed point-based* characterization and a *categorical fixed point-based* characterization. Both of them make use of the well-known relationship between Büchi and parity automata and alternating fixed points, but they differ in how the notion of alternating fixed point is categorically reflected.

The Büchi acceptance condition is a special case of the parity acceptance condition, and all the categorical results in this chapter are fully applicable for the parity acceptance condition. However, as the discussions for the parity condition is very complicated and difficult, we will first present our results for the Büchi condition for explaining the intuition. We then discuss the parity condition. As the latter encompasses the former, all the proofs are omitted in the former part.

This chapter is organized as follows. We first review the so-called Kleisli approach (Section 3.1). In Section 3.2 we give our categorical characterization of Büchi automata. We give a logical fixed point-based characterization for languages of Büchi automata in Section 3.3, and give categorical fixed point-based one in Section 3.4. In Section 3.5, we investigate the relationship between the two characterizations. In Section 3.6, we extend the framework in Sections 3.2–3.5 to parity automata. In Sections 3.7–3.8, we instantiate the categorical frameworks to NPTAs (Definition 2.2.4) and PPTAs (Definition 2.2.12).

The result on the logical fixed point-based characterization and that on the categorical fixed point-based one first that has appeared in [114] and [112], respectively.

### 3.1 Kleisli Approach for Finite and Infinitary Trace Semantics

In this section, we review the Kleisli approach following [85, 58, 46, 47]. In the next section we will extend it for Büchi and parity automata.

At the last of Section 2.4.2, we have seen that the “final coalgebra framework” does not work for  $\mathcal{P}F_{\Sigma}$ - and  $\mathcal{G}_s F_{\Sigma}$ -coalgebras. The key of Kleisli approach is to focus on that  $\mathcal{P}$  and  $\mathcal{G}_s$  constitute *monads* and consider coalgebras in the Kleisli categories (Definition 3.1.3).

### 3.1.1 Monad and Kleisli Category

We review notions of *monad* and *Kleisli category*. See also [60].

A *monad* is a special functor equipped with structures called a *unit* and a *multiplication*. Monads are often used to characterize *side-effects* like *nondeterminism* or *probabilistic branchings*.

**Definition 3.1.1** (monad). Let  $\mathbb{C}$  be a category. A *monad* over  $\mathbb{C}$  is a triple  $T = (T, \eta^T, \mu^T)$  consisting of an endofunctor  $T : \mathbb{C} \rightarrow \mathbb{C}$  and natural transformations  $\eta^T : \text{id} \Rightarrow T$  and  $\mu^T : T^2 \Rightarrow T$  called the *unit* and the *multiplication* respectively.<sup>1</sup> We require that the following equalities are satisfied for each  $X \in \mathbb{C}$ .

$$\mu_X \circ \eta_{TX} = \text{id}_X, \quad \mu_X \circ T\eta_X = \text{id}_X \quad \text{and} \quad \mu_X \circ \mu_{TX} = \mu_X \circ T\mu_X.$$

Pictorially, they mean that the following diagrams commute.

$$\begin{array}{ccc} TX & \xrightarrow{\eta_{TX}^T} & T^2X & \xleftarrow{T\eta_X^T} & TX \\ & \searrow \text{id}_X & \mu_X^T \downarrow & \swarrow \text{id}_X & \\ & & TX & & \end{array} \quad \begin{array}{ccc} T^2X & \xrightarrow{\mu_X^T} & TX \\ \uparrow \mu_{TX}^T & T\mu_X^T & \uparrow \mu_X^T \\ T^3X & \xrightarrow{\quad} & T^2X \end{array}$$

In this thesis, we mainly use the following monads.

**Definition 3.1.2.** The functors  $\mathcal{P}$ ,  $\mathcal{L}$ ,  $\mathcal{D}_s$  and  $\mathcal{G}_s$  in Definition 2.4.6 are extended to monads as follows.

- The *powerset monad*  $\mathcal{P} = (\mathcal{P}, \eta^{\mathcal{P}}, \mu^{\mathcal{P}})$  is defined as follows: for  $X \in \mathbf{Sets}$ ,  $\eta_X^{\mathcal{P}} : X \rightarrow \mathcal{P}X$  is given by  $\eta_X^{\mathcal{P}}(x) := \{x\}$  and  $\mu_X^{\mathcal{P}} : \mathcal{P}^2X \rightarrow \mathcal{P}X$  is given by  $\mu_X^{\mathcal{P}}(\Gamma) := \bigcup_{A \in \Gamma} A$ .
- The *lift monad*  $\mathcal{L} = (\mathcal{L}, \eta^{\mathcal{L}}, \mu^{\mathcal{L}})$  is defined as follows: for  $X \in \mathbf{Sets}$ ,  $\eta_X^{\mathcal{L}} : X \rightarrow \mathcal{L}X$  is given by  $\eta_X^{\mathcal{L}}(x) := x$  and  $\mu_X^{\mathcal{L}} : \mathcal{L}^2X \rightarrow \mathcal{L}X$  is given by  $\mu_X^{\mathcal{L}}(x) := x$  if  $x \in X$  and  $\perp$  if  $x = \perp$ .
- The *subdistribution monad*  $\mathcal{D}_s = (\mathcal{D}_s, \eta^{\mathcal{D}_s}, \mu^{\mathcal{D}_s})$  is defined as follows: for  $X \in \mathbf{Sets}$ ,  $\eta_X^{\mathcal{D}_s} : X \rightarrow \mathcal{D}_sX$  is given by  $\eta_X^{\mathcal{D}_s}(x)(x') := 1$  if  $x = x'$  and 0 otherwise, and  $\mu_X^{\mathcal{D}_s} : \mathcal{D}_s^2X \rightarrow \mathcal{D}_sX$  is given by  $\mu_X^{\mathcal{D}_s}(\Delta)(x) := \sum_{\xi \in \Delta} \xi(x)$ .
- The *sub-Giry monad*  $\mathcal{G}_s = (\mathcal{G}_s, \eta^{\mathcal{G}_s}, \mu^{\mathcal{G}_s})$  [39] is defined by: for  $(X, \mathfrak{F}_X) \in \mathbf{SB}$ ,  $\eta_{(X, \mathfrak{F}_X)}^{\mathcal{G}_s} : X \rightarrow \mathcal{G}_sX$  is given by  $\eta_{(X, \mathfrak{F}_X)}^{\mathcal{G}_s}(x) := \delta_x$  (i.e. the Dirac measure at  $x$ ), and  $\mu_{(X, \mathfrak{F}_X)}^{\mathcal{G}_s} : \mathcal{G}_s^2X \rightarrow \mathcal{G}_sX$  is given by  $\mu_{(X, \mathfrak{F}_X)}^{\mathcal{G}_s}(\Psi)(A) := \int_{\mathcal{G}_sX} \text{ev}_A d\Psi$ .

A monad induces a category called *Kleisli category*. It is possible that an endofunctor can be *lifted* to the Kleisli category.

**Definition 3.1.3** ( $\mathcal{Kl}(T)$ ,  $J$ ,  $U$  and  $\overline{F}$ ). Let  $T = (T, \eta, \mu)$  be a monad on  $\mathbb{C}$ . The *Kleisli category*  $\mathcal{Kl}(T)$  is defined by:  $|\mathcal{Kl}(T)| = |\mathbb{C}|$  and  $\mathcal{Kl}(T)(X, Y) = \mathbb{C}(X, TY)$  for  $X, Y \in |\mathcal{Kl}(T)|$ . An arrow  $f \in \mathcal{Kl}(T)(X, Y)$  is called a *Kleisli arrow*, and we write  $f : X \dashrightarrow Y$  for distinction. A composition of arrows  $f : X \dashrightarrow Y$  and  $g : Y \dashrightarrow Z$  is defined by  $\mu_Z \circ Tg \circ f$ , and denoted by  $g \circ f$ . The *lifting functor*  $J : \mathbb{C} \rightarrow \mathcal{Kl}(T)$  is defined by:  $JX := X$  for  $X \in \mathbb{C}$  and  $J(f) := \eta_Y \circ f$  for  $f : X \rightarrow Y$ . The *forgetful functor*  $U : \mathcal{Kl}(T) \rightarrow \mathbb{C}$  is defined by:  $UX := TX$  for  $X \in \mathcal{Kl}(T)$  and  $U(g) := \mu_Y \circ Tg$  for  $g : X \dashrightarrow Y$ . A functor  $\overline{F} : \mathcal{Kl}(T) \rightarrow \mathcal{Kl}(T)$  is called a *lifting* of  $F : \mathbb{C} \rightarrow \mathbb{C}$  if  $\overline{F}J = JF$ .

<sup>1</sup>We use the same symbol for a monad and its first component. A confusion is unlikely.

It is well-known that there is a bijective correspondence between a lifting of a functor and a natural transformation called *distributive law*.

**Definition 3.1.4.** A *distributive law* from  $T$  to  $F$  is a natural transformation  $\lambda : FT \Rightarrow TF$  that makes the following diagrams commute for each  $X$ .

$$\begin{array}{ccc} FTX & \xrightarrow{\lambda_X} & TFX \\ \uparrow F\eta_X & \nearrow \eta_{FX} & \\ FX & & \end{array} \quad (3.1) \qquad \begin{array}{ccc} FTX & \xrightarrow{\lambda_X} & TFX \\ \uparrow F\mu_X & & \uparrow \mu_{FX} \\ FT^2X & \xrightarrow{\lambda_{TX}} & TFTX \xrightarrow{T\lambda_X} T^2FX \end{array} \quad (3.2)$$

**Proposition 3.1.5** (see e.g. [60]). *Let  $T$  be a monad and  $F$  be an endofunctor on a category  $\mathbb{C}$ . If  $\overline{F} : \mathcal{Kl}(T) \rightarrow \mathcal{Kl}(T)$  is a lifting of  $F$ , then  $(FTX \xrightarrow{\overline{F}\text{id}_{TX}} TFX)_{X \in \mathbb{C}} : FT \Rightarrow TF$  is a distributive law from  $T$  to  $F$ . Conversely if  $\lambda : FT \Rightarrow TF$  is a distributive law from  $T$  to  $F$  then a functor  $\overline{F} : \mathcal{Kl}(T) \rightarrow \mathcal{Kl}(T)$  defined by  $\overline{F}X := X$  for  $X \in \mathbb{C}$  and  $\overline{F}f := (FX \xrightarrow{Ff} FTY \xrightarrow{\lambda_Y} TFY)$  for  $f : X \rightarrow Y$  is a lifting of  $F$ . Moreover, these constructions constitute a bijection.  $\square$*

It is known for  $(T, F) \in \{(\mathcal{P}, F_\Sigma), (\mathcal{L}, F_\Sigma), (\mathcal{D}_s, F_\Sigma), (\mathcal{G}_s, F_\Sigma)\}$  (see Definition 2.4.6 and 3.1.1), a distributive law exists (see e.g. [47, 113]), and hence we can lift the functors to the Kleisli categories.

**Example 3.1.6** (see e.g. [47, 113]). When  $T = \mathcal{P}$  and  $F = F_\Sigma$ , a lifting  $\overline{F}_\Sigma : \mathcal{Kl}(\mathcal{P}) \rightarrow \mathcal{Kl}(\mathcal{P})$  is given by  $\overline{F}_\Sigma X = F_\Sigma X$  for  $X \in \mathcal{Kl}(\mathcal{P})$  and  $\overline{F}_\Sigma f(a, x_1, \dots, x_n) = \{(a, y_1, \dots, y_n) \mid y_i \in f(x_i)\}$  for  $f : X \rightarrow Y$ .

When  $T = \mathcal{G}_s$  and  $F = F_\Sigma$ ,  $\overline{F}_\Sigma : \mathcal{Kl}(\mathcal{G}_s) \rightarrow \mathcal{Kl}(\mathcal{G}_s)$  is given by  $\overline{F}_\Sigma X := F_\Sigma X$  for  $X \in \mathcal{Kl}(\mathcal{G}_s)$  and

$$\overline{F}_\Sigma f(a, x_1, \dots, x_n)(\{a'\} \times A_1 \times \dots \times A_n) := \begin{cases} f(x_1)(A_1) \cdot \dots \cdot f(x_n)(A_n) & (a = a') \\ 0 & (a \neq a') \end{cases}$$

for  $f : X \rightarrow Y$  (by the Kolmogorov extension theorem, this  $\overline{F}_\Sigma f$  is well-defined).

### 3.1.2 Kleisli Approach

Throughout this section, let  $\mathbb{C}$  be a category,  $T$  be a monad and  $F$  be an endofunctor over  $\mathbb{C}$ . In the Kleisli approach, we model a system as follows.

**Definition 3.1.7** ( $(T, F)$ -system). A  $(T, F)$ -system is a pair  $(X, c)$  of an object  $X \in \mathbb{C}$  and a  $TF$ -coalgebra  $c : X \rightarrow TFX$ .

Intuitively,  $T$  represents the *branching type* and  $F$  represents the *transition type* of the system. As we have seen in Example 2.4.14, various transition systems can be represented as  $(T, F)$ -systems by choosing suitable  $T$  and  $F$ .

Assume that the functor  $F$  is equipped with a lifting  $\overline{F} : \mathcal{Kl}(T) \rightarrow \mathcal{Kl}(T)$ . Then  $c : X \rightarrow TFX$  can be regarded as an  $\overline{F}$ -coalgebra in  $\mathcal{Kl}(T)$ . It is possible that a final  $\overline{F}$ -coalgebra exists even if a final  $TF$ -coalgebra does not exist.

**Definition 3.1.8** ( $\text{tr}(c)$ , [47]). We say that  $T$  and  $F$  constitute a *finite trace situation* if the following conditions are satisfied.

- $F$  has a lifting  $\overline{F} : \mathcal{Kl}(T) \rightarrow \mathcal{Kl}(T)$ .
- An initial  $F$ -algebra  $\iota : F(\mu F) \rightarrow \mu F$  exists.
- An  $\overline{F}$ -coalgebra  $J\iota^{-1} : \mu F \rightarrow \overline{F}(\mu F)$  is final (cf. Proposition 2.4.16.2).

$$\begin{array}{ccc} \overline{F}X & \xrightarrow{\overline{F}(\text{tr}(c))} & \overline{F}\mu F \\ c \uparrow & = & \cong \uparrow J\iota^{-1} \\ X & \xrightarrow{\text{tr}(c)} & \mu F \end{array} \text{ final}$$

For  $c: X \rightarrow \overline{F}X$ , the unique homomorphism from  $c$  to  $J(\iota^F)^{-1}$  is called the *(coalgebraic) finite trace semantics* of  $c$  and denoted by  $\text{tr}(c): X \rightarrow \mu F$ .

**Example 3.1.9.** We continue Example 2.4.14. It is known that  $\mathcal{P}$  and  $F_\Sigma$  constitute a finite trace situation [47]. Recall that a carrier of an initial  $F_\Sigma$ -algebra  $\iota^{F_\Sigma}$  is given by  $\text{Tree}_\Sigma^*$  (Example 2.4.18). Hence  $\text{tr}(c_A)$  is a Kleisli arrow  $X \rightarrow \text{Tree}_\Sigma^*$ , i.e. a function of a type  $X \rightarrow \mathcal{P}\text{Tree}_\Sigma^*$ . It is given by  $L_A^*$  (Definition 2.2.9).

Similarly,  $\mathcal{G}_s$  and  $F_\Sigma$  also constitute a finite trace situation (see a discussion later in this section). The unique homomorphism  $\text{tr}(c_{\mathcal{A}})$  has a type  $X \rightarrow \mathcal{G}_s\text{Tree}_\Sigma^*$ , and is given by  $L_{\mathcal{A}}^*$  (Definition 2.2.16).

Hence  $\text{tr}(c)$  categorically characterizes of *finite languages*. We next give a categorical characterization of *infinitary languages* such as  $L_A^\infty$ . In the finite case, we have lifted an initial  $F$ -algebra to  $\mathcal{Kl}(T)$  to obtain a final  $\overline{F}$ -coalgebra. In the infinitary case, we lift a *final  $F$ -coalgebra* to  $\mathcal{Kl}(T)$ . For example, when  $F = F_\Sigma$ , the carrier of a final  $F$ -coalgebra  $\zeta^F$  is isomorphic to  $\text{Tree}_\Sigma^\infty$  (Example 2.4.18), and therefore an  $\overline{F}$ -coalgebra provides a good datatype for characterizing an infinitary language  $L_A^\infty: X \rightarrow \mathcal{P}\text{Tree}_\Sigma^\infty$ . A problem is that  $J\zeta^F$  is *not* a final coalgebra in general, but a *weakly* final coalgebra that can admit multiple homomorphisms. In [58], the problem is solved by introducing *partial orders* to homsets and choosing the *greatest* homomorphism.

**Definition 3.1.10** ( $\text{tr}^\infty(c)$ , [58]). Assume that each homset of  $\mathcal{Kl}(T)$  carries a partial order  $\sqsubseteq$ . We say that  $F$  and  $T$  constitute an *infinitary trace situation* if the following conditions are satisfied:

- $F$  has a lifting  $\overline{F}: \mathcal{Kl}(T) \rightarrow \mathcal{Kl}(T)$ .
- A final  $F$ -coalgebra  $\zeta^F: \nu F \rightarrow F(\nu F)$  exists.
- $J\zeta^F: \nu F \rightarrow \overline{F}\nu F$  is a weakly final  $\overline{F}$ -coalgebra that admits the greatest homomorphism. That is, for an arbitrary  $\overline{F}$ -coalgebra  $c: X \rightarrow \overline{F}X$ , there exists the greatest homomorphism from  $c$  to  $J\zeta^F$  with respect to  $\sqsubseteq$ .

$$\begin{array}{ccc} \overline{F}X & \xrightarrow{\overline{F}(\text{tr}^\infty(c))} & \overline{F}\nu F \\ c \uparrow & \dashv_{\nu} \cong \uparrow & J\zeta^F \\ X & \xrightarrow{\text{tr}^\infty(c)} & \nu F \end{array} \begin{array}{l} \text{weakly} \\ \text{final} \end{array}$$

The greatest homomorphism from  $c$  to  $J\zeta^F$  is called the *(coalgebraic) infinitary trace semantics* of  $c$  and denoted by  $\text{tr}^\infty(c): X \rightarrow \nu F$ .

**Example 3.1.11.** We continue Example 3.1.9. It is known that  $\mathcal{P}$  and  $F_\Sigma$  constitute an infinitary trace situation [58, 113]. The order  $\sqsubseteq$  on each homset  $\mathcal{Kl}(\mathcal{P})(X, Y)$  of the Kleisli category is given as follows: for  $X, Y \in \mathbf{Sets}$ ,  $f \sqsubseteq g \stackrel{\text{def.}}{\iff} \forall x \in X. f(x) \subseteq g(x)$ . The greatest homomorphism  $\text{tr}^\infty(c_A)$  has a type  $X \rightarrow \mathcal{P}\text{Tree}_\Sigma^\infty$ , and is given by  $L_A^\infty$  (Definition 2.2.9).

Similarly,  $\mathcal{G}_s$  and  $F_\Sigma$  constitute an infinitary trace situation [113]. For  $(X, \mathfrak{F}_X)$ ,  $(Y, \mathfrak{F}_Y) \in \mathbf{SB}$ , an order  $\sqsubseteq$  on  $\mathcal{Kl}(\mathcal{G}_s)((X, \mathfrak{F}_X), (Y, \mathfrak{F}_Y))$  is defined by  $f \sqsubseteq g \stackrel{\text{def.}}{\iff} \forall x \in X. \forall A \in \mathfrak{F}_Y. f(x)(A) \leq g(x)(A)$ . The greatest homomorphism  $\text{tr}^\infty(c_S): X \rightarrow \mathcal{G}_s\text{Tree}_\Sigma^\infty$  is given by  $L_S^\infty$  (Definition 2.2.16).

A sufficient condition for a monad  $T$  and functor  $F$  to constitute a finite trace situation is known. The condition uses the notion of **Cppo**-enriched category and **Cppo**-enriched functor, which are instances of categorical notions  $\mathbb{V}$ -enriched category and  $\mathbb{V}$ -enriched functor (see e.g. [14]).

**Definition 3.1.12** (**Cppo**-enriched category). A category  $\mathbb{C}$  is **Cppo**-enriched if it satisfies the following conditions:

1. Each homset  $\mathbb{C}(X, Y)$  carries a partial order  $\sqsubseteq_{X, Y}$ . Moreover each homset  $\mathbb{C}(X, Y)$  is a *pointed  $\omega$ -cpo* with respect to the order, i.e. it has the least element  $\perp_{X, Y}$  and each increasing sequence  $f_0 \sqsubseteq_{X, Y} f_1 \sqsubseteq_{X, Y} \dots \in \mathbb{C}(X, Y)$  has the supremum  $\bigsqcup_{i < \omega} f_i : X \rightarrow Y$ .
2. For each  $X, Y, Z \in \mathbb{C}$ , the composition  $(\_ \circ \_) : \mathbb{C}(Y, Z) \times \mathbb{C}(X, Y) \rightarrow \mathbb{C}(X, Z)$  is monotone with respect to the product order.
3. The composition  $\circ$  is  $\omega$ -continuous, i.e. for  $g : Z \rightarrow X$ ,  $h : Y \rightarrow W$  and an increasing sequence  $f_0 \sqsubseteq_{X, Y} f_1 \sqsubseteq_{X, Y} \dots : X \rightarrow Y$  of arrows,

$$\left(\bigsqcup_{i < \omega} f_i\right) \circ g = \bigsqcup_{i < \omega} (f_i \circ g) \quad \text{and} \quad h \circ \left(\bigsqcup_{i < \omega} f_i\right) = \bigsqcup_{i < \omega} (h \circ f_i). \quad (3.3)$$

Let  $\mathbb{C}$  be a **Cppo**-enriched category. A functor  $F : \mathbb{C} \rightarrow \mathbb{C}$  is called a **Cppo**-enriched functor if it satisfies the following conditions.

- (a) It is *locally monotone*, i.e. for each  $X, Y \in \mathbb{C}$  and  $f, g : X \rightarrow Y$ ,  $f \sqsubseteq_{X, Y} g$  implies  $Ff \sqsubseteq_{FX, FY} Fg$ .
- (b) It is *locally  $\omega$ -continuous*, i.e. for each  $X, Y \in \mathbb{C}$  and an increasing sequence  $f_0 \sqsubseteq_{X, Y} f_1 \sqsubseteq_{X, Y} \dots \in \mathbb{C}(X, Y)$ , we have  $F(\bigsqcup_{i < \omega} f_i) = \bigsqcup_{i < \omega} (Ff_i)$ .

If no confusion is likely, we write  $\sqsubseteq$  for  $\sqsubseteq_{X, Y}$ .

**Theorem 3.1.13** ([47]). *If the following conditions are satisfied,  $F$  and  $T$  constitute a finite trace situation.*

- The functor  $F$  preserves  $\omega$ -colimits in  $\mathbb{C}$ .
- $Kl(T)$  is a **Cppo**-enriched category.
- $\bar{F}$  is a **Cppo**-enriched functor.

It is known that  $(T, F) \in \{(\mathcal{P}, F_\Sigma), (\mathcal{D}_s, F_\Sigma), (\mathcal{L}, F_\Sigma)\}$  (see Definition 2.4.6 and 3.1.2) satisfies the conditions of Theorem 3.1.13 [47]. By the result for  $(T, F) = (\mathcal{D}_s, F_\Sigma)$ , we can easily see that  $(T, F) = (\mathcal{G}_s, F_\Sigma)$  also satisfies the conditions. Hence each of them constitutes a finite trace situation.

It is also known that  $(T, F) \in \{(\mathcal{P}, F_\Sigma), (\mathcal{D}_s, F_\Sigma), (\mathcal{L}, F_\Sigma), (\mathcal{G}_s, F_\Sigma)\}$  constitutes an infinitary trace situation. However, sufficient conditions for constituting an infinitary trace situation are not unified. In [113], two sufficient conditions are given. One is applicable for  $(T, F) = (\mathcal{P}, F_\Sigma)$ , and the other is applicable for  $(T, F) \in \{(\mathcal{D}_s, F_\Sigma), (\mathcal{L}, F_\Sigma)\}$ . No condition is known for  $(T, F) = (\mathcal{D}_s, F_\Sigma)$ .

It is known that under the assumptions of Theorem 3.1.13, if we take the *least* homomorphism instead of the greatest one in Definition 3.1.10 then we obtain a Kleisli arrow that characterizes the same data as  $\text{tr}(c)$ , in the following sense.

**Proposition 3.1.14** ([47]). *Assume that  $T$  and  $F$  satisfy the conditions in Theorem 3.1.13, and hence constitute a finite trace situation. Assume also that they constitute an infinitary trace situation. Let  $p : \mu F \rightarrow \nu F$  be the unique homomorphism from  $(\iota^F)^{-1}$  to  $\zeta^F$ . For  $c : X \rightarrow \bar{F}X$ , we define  $\text{tr}^*(c) : X \rightarrow \nu F$  by  $\text{tr}^*(c) := Jp \circ \text{tr}(c)$ . Then  $\text{tr}^*(c)$  is the least homomorphism from  $c$  to  $J\zeta^F$ .*

$$\begin{array}{ccccc}
& & & \bar{F} \text{tr}^*(c) & \\
& & & \uparrow & \\
\bar{F}X & \xrightarrow{\quad \bar{F} \text{tr}(c) \quad} & \bar{F}\mu F & \xrightarrow{\quad \bar{F}Jp \quad} & \bar{F}\nu F \\
\uparrow c & & \uparrow J(\iota^F)^{-1} & & \uparrow J\zeta^F \\
X & \xrightarrow{\quad \text{tr}(c) \quad} & \mu F & \xrightarrow{\quad Jp \quad} & \nu F \\
& & & \uparrow & \\
& & & \text{tr}^*(c) & 
\end{array}$$

For example, when  $T = \mathcal{P}$  and  $F = F_\Sigma$ , we have  $\mu F \cong \text{Tree}_\Sigma^*$  and  $\nu F \cong \text{Tree}_\Sigma^\infty$ , and  $p : \mu F \rightarrow \nu F$  is given by the natural injection.

### 3.2 Categorical Representation of Büchi Automata

We have reviewed that systems are modeled as  $(T, F)$ -systems (Definition 3.1.7) in the Kleisli approach. We extend the modeling so that we can deal with the Büchi condition.

**Definition 3.2.1** (Büchi  $(T, F)$ -system). A *Büchi  $(T, F)$ -system* is a triple  $\mathcal{X} = (X, c, (X_1, X_2))$  of an object  $X \in \mathbb{C}$ , a  $TF$ -coalgebra  $c : X \rightarrow TF X$  and a pair of objects  $X_1, X_2 \in \mathbb{C}$  such that  $X = X_1 + X_2$ . For  $i \in \{1, 2\}$ , we write  $c_i$  for  $c \circ \kappa_i : X_i \rightarrow \overline{F} X$ .

Intuitively,  $X_1$  collects nonaccepting states and  $X_2$  collects accepting states.

**Example 3.2.2.** Let  $F_A := A \times (-)$ . An  $A$ -labeled nondeterministic Büchi word automaton (see Remark 2.2.11)  $\mathcal{A} = (X, \tau, \text{Acc})$  induces a Büchi  $(\mathcal{P}, F_A)$ -system  $\mathcal{X}_\mathcal{A} = (X, c_\mathcal{A}, (X_1, X_2))$  defined by  $c_\mathcal{A} = \tau$ ,  $X_1 = X \setminus \text{Acc}$  and  $X_2 = \text{Acc}$ .

### 3.3 Characterization via Logical Fixed Point

We give one of the extensions of the Kleisli approach, which uses *logical fixed point*. It considers an alternating fixed point in a homset of a Kleisli category whose homsets carry partial orders. The extension is inspired by the well-known relationship between Büchi automata and fixed point logic (see e.g. [115]). We characterize languages as solutions of HESs over homsets of Kleisli categories.

**Definition 3.3.1** ( $\text{tr}_i^B(c)$ ). Let  $F$  be an endofunctor and  $T$  be a monad on a category  $\mathbb{C}$ . Assume that each homset of  $\mathcal{Kl}(T)$  carries a partial order  $\sqsubseteq$ . We say that  $F$  and  $T$  constitute a *Büchi trace situation* with respect to  $\sqsubseteq$  if:

- $F$  has a lifting  $\overline{F} : \mathcal{Kl}(T) \rightarrow \mathcal{Kl}(T)$ ;
- A final  $F$ -coalgebra  $\zeta^F : \nu F \rightarrow F(\nu F)$  exists; and
- For each Büchi  $(T, F)$ -system  $(X, c, (X_1, X_2))$ , the following HES has a solution.

$$\begin{cases} u_1 & =_\mu J(\zeta^F)^{-1} \odot \overline{F}[u_1, u_2] \odot c_1 & \in (\mathcal{Kl}(T)(X_1, \nu F), \sqsubseteq) \\ u_2 & =_\nu J(\zeta^F)^{-1} \odot \overline{F}[u_1, u_2] \odot c_2 & \in (\mathcal{Kl}(T)(X_2, \nu F), \sqsubseteq) \end{cases} \quad (3.4)$$

Recall from Definition 3.2.1 that  $c_i$  denotes  $c \circ \kappa_i : X_i \rightarrow \overline{F} X$ . The solution  $(u_1^{\text{sol}} : X_1 \rightarrow \nu F, u_2^{\text{sol}} : X_2 \rightarrow \nu F)$  of the HES above is called the *(coalgebraic) Büchi trace semantics* of  $c$ . We write  $\text{tr}_i^B(c)$  for  $u_i^{\text{sol}}$  ( $i \in \{1, 2\}$ ), and  $\text{tr}^B(c)$  for  $[\text{tr}_1^B(c), \text{tr}_2^B(c)] : X \rightarrow \nu F$ . Pictorially,

$$\begin{array}{ccc} \overline{F}[\text{tr}_1^B(c), \text{tr}_2^B(c)] & & \overline{F}[\text{tr}_1^B(c), \text{tr}_2^B(c)] \\ F(X_1 + X_2) \longrightarrow & \longrightarrow & F(X_1 + X_2) \longrightarrow F(\nu F) \\ \begin{array}{ccc} c_1 \uparrow & \xrightarrow{=} & J\zeta^F \uparrow \\ X_1 \longrightarrow & \text{tr}_1^B(c) & \longrightarrow \nu F \end{array} & \cong & \begin{array}{ccc} c_2 \uparrow & \xrightarrow{=} & J\zeta^F \uparrow \\ X_2 \longrightarrow & \text{tr}_2^B(c) & \longrightarrow \nu F \end{array} \end{array} \quad (3.5)$$

**Example 3.3.2.** We continue Example 3.6.2. We define an order on each homset of  $\mathcal{Kl}(\mathcal{P})$  as in Example 3.1.11, and a lifting  $\overline{F}_A : \mathcal{Kl}(\mathcal{P}) \rightarrow \mathcal{Kl}(\mathcal{P})$  of  $F_A$  as in Example 3.1.6. Then  $\mathcal{P}$  and  $F_A$  constitute a Büchi trace situation with respect to them. The carrier set of a final  $F_A$ -coalgebra is isomorphic to  $A^\omega$ . Hence  $\text{tr}^B(c_\mathcal{A})$  has a type  $X \rightarrow A^\omega$  (i.e.  $X \rightarrow \mathcal{P}(A^\omega)$ ), and is given by  $\text{tr}^B(c_\mathcal{A}) = L_\mathcal{A}^B$ .

### 3.4 Characterization via Categorical Fixed Points

We describe the other categorical characterization of behaviors of Büchi automata. In this characterization, we introduce a notion of *(categorical) decorated trace semantics*, which is a variant of categorical Büchi trace semantics (Definition 3.3.1). For nondeterministic Büchi word automata (Example 3.2.2), decorated trace semantics is given as follows.

**Definition 3.4.1** ( $\text{DecL}_{\mathcal{A}}^{\text{B}}$ ). Let  $\mathcal{A} = (X, \tau, \text{Acc})$  be a Büchi automaton. We define a set  $\text{AccRun}(\mathcal{A}) \subseteq (\mathbf{A} \times \{\circ, \odot\})^\omega$  as follows:

$$\text{AccRun}(\mathcal{A}) := \{(a_0, \bullet_0)(a_1, \bullet_1) \dots \in (\mathbf{A} \times \{\circ, \odot\})^\omega \mid \bullet_i = \odot \text{ for infinitely many } i\}.$$

We define a function  $\mathbf{p}: X \rightarrow \{\circ, \odot\}$  by  $\mathbf{p}(x) := \circ$  if  $x \notin \text{Acc}$  and  $\odot$  if  $x \in \text{Acc}$ . For each  $x \in X$ , we define  $\mathbf{p}: \text{Run}_{\mathcal{A}}^\infty(x) \rightarrow (\mathbf{A} \times \{\circ, \odot\})^\omega$  by  $\mathbf{p}((a_0, x_0)(a_1, x_1) \dots) := (a_0, \mathbf{p}(x_0))(a_1, \mathbf{p}(x_1)) \dots$ . The *decorated trace semantics* of  $\mathcal{A}$  is a function  $\text{DecL}_{\mathcal{A}}^{\text{B}}: X \rightarrow \mathcal{P}(\text{AccRun}(\mathcal{A}))$  that is defined as follows:

$$\text{DecL}_{\mathcal{A}}^{\text{B}} : x \mapsto \{\mathbf{p}(\rho) \mid \rho \in \text{Run}_{\mathcal{A}}^\infty(x), \rho \text{ satisfies the Büchi acceptance condition}\}.$$

In this section, we define a *decorated trace semantics*  $\text{dtr}_1(c)$  and  $\text{dtr}_2(c)$  as a categorical generalization of  $\text{DecL}_{\mathcal{A}}^{\text{B}}$ . It makes use of *categorical fixed points*, that is, fixed points of functors in the sense of Section 1.3.2.

Recall that an initial algebra is understood as the least fixed point of a functor while a final coalgebra is the greatest fixed point of a functor. From this perspective,  $\text{tr}_1^{\text{B}}(c)$  and  $\text{tr}_2^{\text{B}}(c)$  in the previous section are described as follows: i)  $\text{tr}_1^{\text{B}}(c): X_1 \rightarrow \nu F$  and  $\text{tr}_2^{\text{B}}(c): X_2 \rightarrow \nu F$  constitute an alternating fixed point; and ii) their codomain  $\nu F$  is the greatest fixed point of  $F$ . In contrast,  $\text{dtr}_1(c)$  and  $\text{dtr}_2(c)$  are described as follows: i)  $\text{dtr}_1(c): X_1 \rightarrow U_1^{\text{sol}}$  and  $\text{dtr}_2(c): X_2 \rightarrow U_2^{\text{sol}}$  constitute the greatest fixed points; and ii) their codomains  $U_1^{\text{sol}}$  and  $U_2^{\text{sol}}$  are an alternating fixed point.

#### 3.4.1 Alternating Fixed Point of Functor

We first define datatypes  $U_1^{\text{sol}}$  and  $U_2^{\text{sol}}$  categorically. Intuitively speaking, they are the solution of the following ‘‘HES.’’

$$\begin{cases} U_1 =_\mu F(U_1 + U_2) & \in \mathbb{C} \\ U_2 =_\nu F(U_1 + U_2) & \in \mathbb{C} \end{cases} \quad (3.6)$$

When we try to solve the above ‘‘HES’’ in the same manner as ordinary HESs (Definition 2.3.6), we first calculate the intermediate solution for the first equation. It is the ‘‘least fixed point’’ of  $U_1 \mapsto F(U_1 + U_2)$  regarding an object  $U_2$  a parameter. Such a parameterized least (or greatest) fixed point of a functor is formally defined as follows.

**Definition 3.4.2** ( $F^+, F^\oplus$ ). Let  $F: \mathbb{C} \rightarrow \mathbb{C}$ . Note that for each  $X \in \mathbb{C}$ , we can define a functor  $F(\_ + X): \mathbb{C} \rightarrow \mathbb{C}$ .

- Assume that an initial  $F(\_ + X)$ -algebra exists for each  $X$ . We define a functor  $F^+: \mathbb{C} \rightarrow \mathbb{C}$  as follows:
  - for  $X \in \mathbb{C}$ ,  $F^+X$  is given by the carrier of (a choice of) an initial  $F(\_ + X)$ -algebra  $\iota_X^F$ ; and
  - for  $f: X \rightarrow Y$ ,  $F^+f: F^+X \rightarrow F^+Y$  is given by the unique homomorphism from  $\iota_X^F$  to  $\iota_Y^F \circ F(\text{id}_{F^+Y} + f)$  (see the left diagram below).



- Assume that a final  $F(\_ + X)$ -coalgebra exists for each  $X$ . We define a functor  $F^\oplus : \mathbb{C} \rightarrow \mathbb{C}$  as follows:

- for  $X \in \mathbb{C}$ ,  $F^\oplus X$  is given by the carrier of (a choice of) a final  $F(\_ + X)$ -algebra  $\zeta_X^F$ ; and
- for  $f : X \rightarrow Y$ ,  $F^\oplus f : F^\oplus X \rightarrow F^\oplus Y$  is the unique homomorphism from  $F(\text{id}_{F^\oplus X} + f) \circ \zeta_X^F$  to  $\zeta_Y^F$  (see the right diagram below).

$$\begin{array}{ccc}
F(F^+X + X) & \xrightarrow{F(F^+f+\text{id})} & F(F^+Y + X) & & F(F^\oplus X + Y) & \xrightarrow{F(F^\oplus f+\text{id})} & F(F^\oplus Y + Y) \\
\downarrow \iota_X^F & & \downarrow F(\text{id}+f) & & F(\text{id}+f) \uparrow & & \uparrow \zeta_Y^F \\
F^+X & \xrightarrow{F^+f} & F^+Y & & F(F^\oplus X + X) & = & \zeta_X^F \uparrow \cong \\
& & \downarrow \iota_Y^F & & \downarrow \zeta_X^F & & \downarrow \cong \\
& & F^+Y & & F^\oplus X & \xrightarrow{F^\oplus f} & F^\oplus Y
\end{array}$$

We can think of  $(\_)^+$  and  $(\_)^\oplus$  as operations that transform a functor to a functor. This means that it is possible to apply  $(\_)^+$  or  $(\_)^\oplus$  to a functor multiple times and consider functors like  $(F^+)^\oplus$ . In the rest of this thesis, we shall omit parentheses and just write  $F^{+\oplus}$  for  $(F^+)^\oplus$  for simplicity.

Using the defined data, we can solve the ‘‘HES’’ (3.6) as follows. By Definition 3.4.2, the intermediate solution for the first equation of (3.6) is given by  $F^+U_2$ . We next substitute  $U_1$  in the second equation with  $F^+U_2$ , and get an equation  $U_2 =_\nu F(F^+U_2 + U_2)$ . Note that the right-hand side is isomorphic to  $F^+U_2$ . Hence the greatest fixed point of this equation is given by the carrier of the final  $F^+$ -coalgebra, i.e.  $F^{+\oplus}0$ . We finally substitute  $U_2$  in the intermediate solution  $F^+U_2$  with  $F^{+\oplus}0$ , and obtain  $F^+(F^{+\oplus}0)$ .

The datatype  $F^{+\oplus}0$  and  $\text{AccRun}(\mathbf{A})$  in Definition 3.4.1 are related as follows.

**Example 3.4.3.** For  $F_{\mathbf{A}} = \mathbf{A} \times (\_)$ , we have  $F_{\mathbf{A}}^+ X \cong \mathbf{A}^+ X$ ,  $F_{\mathbf{A}}^\oplus X \cong \mathbf{A}^+ X + \mathbf{A}^\omega$ ,  $F_{\mathbf{A}}^{+\oplus} 0 \cong (\mathbf{A}^+)^\omega$  and  $F_{\mathbf{A}}^+(F_{\mathbf{A}}^{+\oplus} 0) \cong \mathbf{A}^+(\mathbf{A}^+)^\omega$ . An element  $(a_{00}a_{01} \dots a_{0n_0})(a_{10}a_{11} \dots a_{1n_1}) \dots \in F^{+\oplus}0 \cong (\mathbf{A}^+)^\omega$  is identified with the following sequence.

$$\begin{aligned}
& (a_{00}, \odot)(a_{01}, \circ) \dots (a_{0n_0}, \circ)(a_{10}, \odot)(a_{11}, \circ) \dots (a_{1n_1}, \circ) \dots \\
& \in (\mathbf{A} \times \{\odot\}) \times (\mathbf{A} \times \{\circ, \odot\})^\omega. \quad (3.7)
\end{aligned}$$

That is,  $\odot$  appears each beginning of a subsequence. Note that by its construction, the first letter is decorated with  $\odot$ , and  $\odot$  appears infinitely many times.

In contrast, for  $a_0 \dots a_n((a_{00}a_{01} \dots a_{0n_0})(a_{10}a_{11} \dots a_{1n_1}) \dots) \in F^+(F^{+\oplus}0) \cong \mathbf{A}^+(\mathbf{A}^+)^\omega$ , we regard that no accepting state is visited in the first part  $a_0 \dots a_n$ , and identify it with the following decorated sequence:

$$\begin{aligned}
& (a_0, \circ) \dots (a_n, \circ)(a_{00}, \odot)(a_{01}, \circ) \dots (a_{0n_0}, \circ)(a_{10}, \odot)(a_{11}, \circ) \dots (a_{1n_1}, \circ) \dots \\
& \in (\mathbf{A} \times \{\circ\}) \times (\mathbf{A} \times \{\circ, \odot\})^\omega. \quad (3.8)
\end{aligned}$$

The first letter is labeled with  $\circ$ , and  $\odot$  appears infinitely many times again. By (3.7) and (3.8), there exists a canonical bijection  $\text{AccRun}(\mathbf{A}) \cong \mathbf{A}^+(\mathbf{A}^+)^\omega + (\mathbf{A}^+)^\omega$ .

**Remark 3.4.4.** As a final coalgebra  $\zeta_0^{F^+}$  is an isomorphism, we can easily see that  $F^+(F^{+\oplus}0) \cong F^{+\oplus}0$ . Indeed, in Example 3.4.3, we have  $(\mathbf{A}^+)^\omega \cong \mathbf{A}^+(\mathbf{A}^+)^\omega$ . However, in this paper, mainly for the sake of simplicity of notations, we explicitly distinguish them and later define categorical decorated trace semantics as a pair  $\text{dtr}_1(c) : X_1 \rightarrow F^+(F^{+\oplus}0)$  and  $\text{dtr}_2(c) : X_2 \rightarrow F^{+\oplus}0$  of Kleisli arrows.

**Remark 3.4.5.** The definition of  $F^+$  is similar to that of the *free monad*  $F^*$  over  $F$ , which is defined as follows: for  $X \in \mathbb{C}$ , the object  $F^*X$  is the carrier of an initial  $(F(\_) + X)$ -algebra. For  $F_{\mathbf{A}} = \mathbf{A} \times (\_)$ , while  $F_{\mathbf{A}}^+ X \cong \mathbf{A}^+ X$ ,  $F_{\mathbf{A}}^* X \cong \mathbf{A}^* X$ . Similarly, the definition of  $F^\oplus$  resembles that of *free completely iterative monad* [80].

### 3.4.2 Lifting $F^+$ and $F^\oplus$ over $\mathcal{Kl}(T)$

In order to take the ‘‘Kleisli approach’’ with the functors  $F^+$  and  $F^\oplus$ , we have to lift them to the Kleisli category.

We show that under certain conditions, a lifting  $\overline{F} : \mathcal{Kl}(T) \rightarrow \mathcal{Kl}(T)$  of  $F$  induces liftings  $\overline{F}^+ : \mathcal{Kl}(T) \rightarrow \mathcal{Kl}(T)$  of  $F^+$  and  $\overline{F}^\oplus : \mathcal{Kl}(T) \rightarrow \mathcal{Kl}(T)$  of  $F^\oplus$ .

#### Definition 3.4.6.

1. Assume  $T$  and  $F(\_ + A)$  constitute a finite trace situation for each  $A \in \mathbb{C}$ . For  $X \in \mathbb{C}$ , we let  $\overline{F}^+X := F^+X$ . For  $f : X \rightarrow Y$ , we define  $\overline{F}^+f : F^+X \rightarrow F^+Y$  as the unique homomorphism from  $\overline{F}(\text{id}_{F^+X} + f) \odot J(\iota_X^F)^{-1}$  to  $J(\iota_Y^F)^{-1}$ .
2. Assume  $T$  and  $F(\_ + A)$  constitute an infinitary trace situation for each  $A \in \mathbb{C}$ . For  $X \in \mathbb{C}$ , we let  $\overline{F}^\oplus X := F^\oplus X$ . For  $f : X \rightarrow Y$ , we define  $\overline{F}^\oplus f : F^\oplus X \rightarrow F^\oplus Y$  as the greatest homomorphism from  $\overline{F}(\text{id}_{F^\oplus X} + f) \odot J\zeta_X^F$  to  $J\zeta_Y^F$ .

$$\begin{array}{ccc}
 F(F^+X + Y) \xrightarrow{\overline{F}(\text{id}_{F^+X} + f)} F(F^+Y + Y) & & F(F^\oplus X + Y) \xrightarrow{\overline{F}(\text{id}_{F^\oplus X} + f)} F(F^\oplus Y + Y) \\
 \uparrow J(\iota_X^F)^{-1} \cong & & \uparrow J\zeta_X^F \cong \\
 F(F^+X + X) & = & F(F^\oplus X + X) \\
 \uparrow J(\iota_X^F)^{-1} \cong & & \uparrow J\zeta_X^F \cong \\
 F^+X & \xrightarrow{\overline{F}^+f} & F^+Y & & F^\oplus X & \xrightarrow{\overline{F}^\oplus f} & F^\oplus Y
 \end{array}$$

**Remark 3.4.7.** It is known that  $J : \mathbb{C} \rightarrow \mathcal{Kl}(T)$  preserves coproducts (see e.g. [60]). This implies that if a lifting  $\overline{F}$  of  $F$  is given then a lifting  $\overline{F}(\_ + A) : \mathcal{Kl}(T) \rightarrow \mathcal{Kl}(T)$  of  $F(\_ + A)$  is given by  $\overline{F}(\_ + A)$ , a composition of  $\overline{F} : \mathcal{Kl}(T) \rightarrow \mathcal{Kl}(T)$  and  $(\_ + A) : \mathcal{Kl}(T) \rightarrow \mathcal{Kl}(T)$  (note that the last  $+$  denotes the coproduct in  $\mathcal{Kl}(T)$ ).

We have to check functoriality of  $\overline{F}^+$  and  $\overline{F}^\oplus$ , and that they are indeed liftings of  $F^+$  and  $F^\oplus$ . Functoriality of  $\overline{F}^+$  is easily proved by the finality. In contrast, functoriality of  $\overline{F}^\oplus$  does not necessarily hold.

**Example 3.4.8.** We define  $F : \mathbf{Sets} \rightarrow \mathbf{Sets}$  by  $F = \{o\} \times (\_) \times (\_)$ . Let  $X = \{x\}$  and  $Y = \{y_1, y_2\}$ , and define  $f : X \rightarrow Y$  and  $g : Y \rightarrow X$  in  $\mathcal{Kl}(\mathcal{D}_s)$  by  $f(x) = [y_1 \mapsto \frac{1}{2}, y_2 \mapsto \frac{1}{2}]$  and  $f(y_1) = f(y_2) = [x \mapsto 1]$ . In a similar manner to Example 3.4.3, we can show that  $F^\oplus X$  is isomorphic to the set of possibly infinite binary trees whose depth is greater than 1, nodes are labeled with  $o$  and leaves are labeled with  $x$ . A set  $F^\oplus Y$  is similar. Let  $t_X \in F^\oplus X$  be an element identified with a tree  $o(x, o(x, o(x, \dots)))$ . For each  $t_Y \in F^\oplus Y$ ,

$$\overline{F}^\oplus f(t_X)(t_Y) = J(\zeta_Y^F)^{-1} \odot \overline{F}(\overline{F}^\oplus f + \text{id}_Y) \odot \overline{F}(\text{id} + f) \odot \zeta_X^F(t_X)(t_Y) = \frac{1}{2} \overline{F}^\oplus f(t_X)(t_Y).$$

This implies  $\overline{F}^\oplus f(t_X)(t_Y) = 0$ , and therefore  $\overline{F}^\oplus g \odot \overline{F}^\oplus f(t_X)(t_X) = 0$ . In contrast,  $\text{id}_X : X \rightarrow X$  is a homomorphism from  $\overline{F}(\text{id} + g) \odot \overline{F}(\text{id} + f) \odot J\zeta_X^F = J\zeta_X^F$  to itself, and  $\text{id}_X(t_X)(t_X) = 1 \neq 0$ . Hence  $\overline{F}^\oplus(g \odot f)(t_X)(t_X) \geq 1$ , and this means that the operation  $\overline{F}^\oplus$  does not satisfy the functoriality.

Hence we need an extra assumption to make  $\overline{F}^\oplus$  a functor. We hereby assume a stronger condition than is needed for the sake of discussions in Section 3.4.

**Definition 3.4.9.** Assume that  $T$  and  $F$  constitute an infinitary trace situation. Let  $\zeta^F : \nu F \rightarrow F(\nu F)$  be a final  $F$ -coalgebra. We say that  $T$  and  $F$  satisfy the *gfp-preserving condition* with respect to an  $\overline{F}$ -algebra  $\sigma : FY \rightarrow Y$  if for each  $X \in \mathbb{C}$  and  $c : X \rightarrow FX$ , if  $l : X \rightarrow Z$  is the greatest homomorphism from  $c$  to  $J\zeta^F$  and the function  $\Phi_{J\zeta^F, \sigma}$  (Definition 2.4.22) has the greatest fixed point  $m : Z \rightarrow Y$ , then  $m \odot l : X \rightarrow Y$  is the greatest fixed point of  $\Phi_{c, \sigma}$ .

$$\begin{array}{ccccc} FX & \xrightarrow{\overline{F}l} & F(\nu F) & \xrightarrow{\overline{F}m} & FY \\ c \uparrow & =_{\nu} & J\zeta^F \uparrow \cong & =_{\nu} & \downarrow \sigma \\ X & \xrightarrow{l} & \nu F & \xrightarrow{m} & Y \end{array}$$

We next check if  $\overline{F}^+$  and  $\overline{F}^\oplus$  are liftings of  $F^+$  and  $F^\oplus$ . We can easily prove  $\overline{F}^+JX = JF^+X$  and  $\overline{F}^\oplus JX = JF^\oplus X$  for each  $X \in \mathbb{C}$  by definition. It remains to prove  $\overline{F}^+Jf = JF^+f$  and  $\overline{F}^\oplus Jf = JF^\oplus f$  for each  $f : X \rightarrow Y$ . The former is immediate by the finality of  $J(\iota_Y^F)^{-1}$ . The latter again requires an assumption.

**Definition 3.4.10.** Assume that  $T$  and  $F$  constitute an infinitary trace situation. Let  $\zeta^F : \nu F \rightarrow F(\nu F)$  be a final  $F$ -coalgebra. We say that  $T$  and  $F$  satisfy the *deterministic-greatest condition* if for  $c : X \rightarrow FX$  in  $\mathbb{C}$ , if  $u : X \rightarrow \nu F$  is the unique homomorphism from  $c$  to  $\zeta^F$  then  $Ju$  is the greatest homomorphism from  $Jc$  to  $J\zeta^F$ .

$$\begin{array}{ccc} FX & \xrightarrow{\overline{F}u} & F(\nu F) \\ Jc \uparrow & =_{\nu} & J\zeta^F \uparrow \cong \\ X & \xrightarrow{Ju} & \nu F \end{array}$$

We now give conditions for the functors in Definition 3.4.6 to be liftings.

**Proposition 3.4.11.**

1. If  $T$  and  $F(\_ + A)$  constitute a finite trace situation for each  $A \in \mathbb{C}$ , the operation  $\overline{F}^+$  in Definition 3.4.6 is a functor and moreover a lifting of  $F^+$ .
2. If  $T$  and  $F(\_ + A)$  constitute an infinitary trace situation, satisfy the *gfp-preserving condition* with respect to an arbitrary algebra and satisfy the *deterministic-greatest condition* for each  $A \in \mathbb{C}$ , then the operation  $\overline{F}^\oplus$  in Definition 3.4.6 is a functor and moreover a lifting of  $F^\oplus$ .

Proposition 3.4.11.2 is proved using the following lemma.

**Lemma 3.4.12.** Assume that  $T$  and  $F$  constitute an infinitary trace situation and satisfy the *gfp-preserving condition* (Definition 3.4.9). For each  $X, A, B \in \mathbb{C}$ ,  $c : X \rightarrow F(X + A)$  and  $f : A \rightarrow B$ , if  $l : X \rightarrow F^\oplus A$  is the greatest homomorphism from  $c$  to  $J\zeta_A^F$ , then  $\overline{F}^\oplus f \odot l : X \rightarrow F^\oplus B$  is the greatest homomorphism from  $\overline{F}(\text{id}_X + f) \odot c$  to  $J\zeta_B^F$ .

$$\begin{array}{ccccc} F(X + B) & \xrightarrow{\overline{F}(l+\text{id})} & F(F^\oplus A + B) & \xrightarrow{\overline{F}(\overline{F}^\oplus f + \text{id})} & F(F^\oplus B + B) \\ \overline{F}(\text{id}+f) \uparrow & & \overline{F}(\text{id}+f) \uparrow & & \uparrow \\ F(X + A) & \xrightarrow{\overline{F}(l+\text{id})} & F(F^\oplus A + A) & =_{\nu} & J\zeta_B^F \cong \\ c \uparrow & =_{\nu} & J\zeta_A^F \uparrow \cong & & \uparrow \\ X & \xrightarrow{l} & F^\oplus A & \xrightarrow{\overline{F}^\oplus f} & F^\oplus B \end{array}$$

**Proof.** By Lemma 2.4.23,  $\overline{F}^\oplus f$  is the greatest fixed point of  $\Phi_{J\zeta_A^F, J(\zeta_B^F)^{-1} \odot \overline{F}(\text{id}+f)}$ . Moreover, the greatest homomorphism from  $\overline{F}(\text{id}_X + f) \odot c$  to  $J\zeta_B^F$  is the greatest

fixed point of  $\Phi_{c, J(\zeta_B^F)^{-1} \circ \overline{F}(\text{id}+f)}$ .

$$\begin{array}{ccccc}
F(X + A) & \xrightarrow{\overline{F}(\text{id}+f)} & F(F^\oplus A + A) & \xrightarrow{\overline{F}(m+\text{id})} & F(F^\oplus B + A) \\
\uparrow c & & \uparrow J\zeta_A^F \cong & & \downarrow \overline{F}(\text{id}+f) \\
& \text{=} \nu & & \text{=} \nu & F(F^\oplus B + B) \\
& & & & \cong \downarrow J(\zeta_B^F)^{-1} \\
X & \xrightarrow{l} & F^\oplus A & \xrightarrow{\overline{F}^\oplus f} & F^\oplus B
\end{array}$$

Hence it is immediate by the gfp-preserving condition.  $\square$

**Proof** (Proposition 3.4.11). Item 1 is immediate by the finality. Item 2 is easily proved by the gfp-preserving condition, the deterministic-greatest condition and Lemma 3.4.12.  $\square$

Hence under appropriate conditions, a lifting  $\overline{F} : \mathcal{Kl}(T) \rightarrow \mathcal{Kl}(T)$  of  $F$  gives rise to those of  $F^+$  and  $F^\oplus$ . By repeating this, we can also define  $\overline{F}^{+\oplus}$ .

We conclude this section by presenting the distributive laws (see Definition 3.1.4) corresponding to the liftings. The proofs are easy.

**Proposition 3.4.13.** *Let  $\lambda : FT \Rightarrow TF$  be a distributive law from  $T$  to  $F$ . For  $A, X \in \mathbb{C}$ , we write  $\dot{\lambda}_{A,X}$  for  $\lambda_{A+X} \circ F[T\kappa_1 \circ \eta_A, T\kappa_2] : F(A+TX) \rightarrow TF(A+X)$ .*

1. *Assume  $T$  and  $F(\_ + A)$  constitute a finite trace situation for each  $A \in \mathbb{C}$ . For  $X \in \mathbb{C}$  we define  $\lambda_{+,X} : F^+TX \rightarrow TF^+X$  as the unique homomorphism from  $\dot{\lambda}_{F^+TX,X} \circ J(\iota_{TX}^F)^{-1}$  to  $J(\iota_X^F)^{-1}$  (see the left diagram below). Then  $\lambda_+ := (\lambda_{+,X})_{X \in \mathbb{C}}$  is a natural transformation  $F^+T \Rightarrow TF^+$  and moreover a distributive law from  $T$  to  $F^+$ .*
2. *Assume  $T$  and  $F(\_ + A)$  constitute an infinitary trace situation and satisfy the gfp-preserving condition and the deterministic-greatest condition for each  $A \in \mathbb{C}$ . For  $X \in \mathbb{C}$ , let  $\lambda_{\oplus,X} : F^\oplus TX \rightarrow TF^\oplus X$  be the greatest homomorphism from  $\dot{\lambda}_{F^\oplus TX,X} \circ J\zeta_{TX}^F$  to  $J\zeta_X^F$  (see the right diagram below). Then  $\lambda_\oplus := (\lambda_{\oplus,X})_{X \in \mathbb{C}}$  is a natural transformation  $F^\oplus T \Rightarrow TF^\oplus$  and moreover a distributive law from  $T$  to  $F^\oplus$ .*

$$\begin{array}{ccc}
F(F^+TX + X) \xrightarrow{\overline{F}(\lambda_{+,X}+\text{id})} F(F^+X + X) & & F(F^\oplus TX + X) \xrightarrow{\overline{F}(\lambda_{\oplus,X}+\text{id})} F(F^\oplus X + X) \\
\uparrow \dot{\lambda}_{F^+TX,X} & & \uparrow \dot{\lambda}_{F^\oplus TX,X} \\
F(F^+TX + TX) & = & J(\iota_X^F)^{-1} \cong & F(F^\oplus TX + TX) & \text{=} \nu & J\zeta_X^F \cong \\
J(\iota_{TX}^F)^{-1} \uparrow \cong & & & J\zeta_{TX}^F \uparrow \cong & & \\
F^+TX & \xrightarrow{\lambda_{+,X}} & F^+X & & F^\oplus TX & \xrightarrow{\lambda_{\oplus,X}} & F^\oplus X & \square
\end{array}$$

### 3.4.3 Decorated Trace Semantics

In this section, we categorically generalize Definition 3.4.1 and define decorated trace semantics for Büchi  $(T, F)$ -systems.

**Assumption 3.4.14.** Throughout this section, let  $T$  be a monad and  $F$  be an endofunctor on  $\mathbb{C}$ , and assume that each homset of  $\mathcal{Kl}(T)$  carries a partial order  $\sqsubseteq$ . We further assume the following conditions for each  $A \in \mathbb{C}$ .

1.  $F^+, F^{+\oplus} : \mathbb{C} \rightarrow \mathbb{C}$  are well-defined and a lifting  $\overline{F}$  is given.
2.  $T$  and  $F(\_ + A)$  satisfy the conditions in Theorem 3.1.13.
3.  $T$  and  $F^+(\_ + A)$  constitute an infinitary trace situation (Definition 3.1.10).

4.  $T$  and  $F^+(\_ + A)$  satisfy the gfp-preserving condition with respect to an arbitrary algebra  $\sigma$  (Definition 3.4.9).
5.  $T$  and  $F^+(\_ + A)$  satisfy the deterministic-greatest condition (Definition 3.4.10).
6. The liftings  $\overline{F^+}(\_ + A)$  and  $\overline{F^{+\oplus}}(\_ + A)$  are obtained from  $\overline{F}(\_ + A)$  and  $\overline{F^+}(\_ + A)$  using the procedure in Definition 3.4.6 (see also Remark 3.4.7).
7. Liftings  $\overline{F^+}(\_ + A)$  and  $\overline{F^{+\oplus}}(\_ + A)$  are locally monotone.

Under the above assumptions, using the datatypes defined in the previous section, we can categorically define *decorated Büchi trace semantics* as follows.

**Definition 3.4.15** ( $\text{dtr}_i(c)$ ). Let  $\mathcal{X} = (X, c, (X_1, X_2))$  be a Büchi  $(T, F)$ -system. The (coalgebraic) decorated trace semantics of  $\mathcal{X}$  is a pair

$$\left( \text{dtr}_1(c): X_1 \rightarrow F^+(F^{+\oplus}0), \text{dtr}_2(c): X_2 \rightarrow F^{+\oplus}0 \right)$$

of arrows that is the greatest fixed point of the following endofunction with respect to the product order induced by  $\sqsubseteq$ .

$$(v_1, v_2) \mapsto \left( J\iota_{F^{+\oplus}0}^F \odot \overline{F}(v_1 + v_2) \odot c_1, J((\zeta_0^{F^+})^{-1} \circ \iota_{F^{+\oplus}0}^F) \odot \overline{F}(v_1 + v_2) \odot c_2 \right) \\ \in \mathcal{Kl}(T)(X_1, F^+(F^{+\oplus}0)) \times \mathcal{Kl}(T)(X_2, F^{+\oplus}0) \quad (3.9)$$

Pictorially,

$$\begin{array}{ccc} \begin{array}{c} \overline{F}(v_1 + v_2) \\ \uparrow \\ F(X_1 + X_2) \end{array} \xrightarrow{\quad} \begin{array}{c} \overline{F}(v_1 + v_2) \\ \uparrow \\ F(F^+(F^{+\oplus}0) + F^{+\oplus}0) \end{array} & \begin{array}{c} \cong \\ \nu \\ \cong \end{array} & \begin{array}{c} \overline{F}(v_1 + v_2) \\ \uparrow \\ F(X_1 + X_2) \end{array} \xrightarrow{\quad} \begin{array}{c} \overline{F}(v_1 + v_2) \\ \uparrow \\ F(F^+(F^{+\oplus}0) + F^{+\oplus}0) \\ \cong \uparrow J(\iota_{F^{+\oplus}0}^F)^{-1} \\ F^+(F^{+\oplus}0) \\ \cong \uparrow J\zeta_0^{F^+} \\ F^{+\oplus}0 \end{array} \\ \begin{array}{c} \uparrow c_1 \\ X_1 \end{array} \xrightarrow{\quad v_1 \quad} \begin{array}{c} \uparrow J(\iota_{F^{+\oplus}0}^F)^{-1} \\ F^+(F^{+\oplus}0) \end{array} & & \begin{array}{c} \uparrow c_2 \\ X_2 \end{array} \xrightarrow{\quad v_2 \quad} \begin{array}{c} \uparrow J\zeta_0^{F^+} \\ F^{+\oplus}0 \end{array} \end{array}$$

We write  $\text{dtr}(c)$  for  $\text{dtr}_1(c) + \text{dtr}_2(c): X_1 + X_2 \rightarrow F^+(F^{+\oplus}0) + F^{+\oplus}0$ .

**Example 3.4.16.** We continue Example 3.2.2. With respect to the bijection  $\text{AccRun}(\mathbf{A}) \cong \mathbf{A}^+(\mathbf{A}^+)^\omega + (\mathbf{A}^+)^\omega$  in Example 3.4.3, we have:  $\text{DecL}_{\mathcal{A}}^{\mathbf{B}} = \text{dtr}_1(c_{\mathcal{A}}) + \text{dtr}_2(c_{\mathcal{A}})$ . See Section 3.6 for a proof that also covers parity automata.

### 3.5 Logical Fixed Point vs. Categorical Fixed Point

In Sections 3.3–3.4, we have introduced two categorical characterizations for languages of Büchi automata. In this section we investigate their relationship. We first explain the intuition using the running example.

**Example 3.5.1.** Let  $\mathcal{A} = (X, \tau, \text{Acc})$  be a nondeterministic Büchi word automaton. Recall that  $L_{\mathcal{A}}^{\mathbf{B}}$  has a type  $X \rightarrow \mathcal{P}(\mathbf{A}^\omega)$  (Definition 2.2.9) while  $\text{DecL}_{\mathcal{A}}^{\mathbf{B}}$  has a type  $X \rightarrow \mathcal{P}(\text{AccRun}(\mathbf{A}))$  where  $\text{AccRun}(\mathbf{A}) \subseteq (\mathbf{A} \times \{\circlearrowleft, \circlearrowright\})^\omega$  (Definition 3.4.1). We define a function  $p: \text{AccRun}(\mathbf{A}) \rightarrow \mathbf{A}^\omega$  by

$$p((a_0, \bullet_0)(a_1, \bullet_1) \dots) := a_0 a_1 \dots$$

Then by the definitions of  $L_{\mathcal{A}}^{\mathbf{B}}$  and  $\text{DecL}_{\mathcal{A}}^{\mathbf{B}}$ , we have  $L_{\mathcal{A}}^{\mathbf{B}}(x) = \mathcal{P}p(\text{DecL}_{\mathcal{A}}^{\mathbf{B}}(x))$  for each  $x \in X$ . With respect to the isomorphism  $\text{AccRun}(\mathbf{A}) \cong (\mathbf{A}^+)^\omega + \mathbf{A}^+(\mathbf{A}^+)^\omega$  in Example 3.4.3,  $p$  has a type  $(\mathbf{A}^+)^\omega + \mathbf{A}^+(\mathbf{A}^+)^\omega \rightarrow \mathbf{A}^\omega$ , and it is given by the canonical “flattening” function.

Categorically, the function  $p$  above is given as follows.

**Definition 3.5.2.** We define  $p_1 : F^+(F^{+\oplus 0}) \rightarrow F^{\oplus 0}$  and  $p_2 : F^{+\oplus 0} \rightarrow F^{\oplus 0}$  so that  $[p_1, p_2]$  is the unique homomorphism from  $(\iota_{F^{+\oplus 0}}^F)^{-1} \circ [\text{id}, \zeta_0^{F^+}]$  to  $\zeta_0^F$ .

$$\begin{array}{ccc}
F(F^+(F^{+\oplus 0}) + F^{+\oplus 0}) & \xrightarrow{F[p_1, p_2]} & F(F^{\oplus 0}) \\
(\iota_{F^{+\oplus 0}}^F)^{-1} \uparrow & & \uparrow \zeta_0^F \\
F^+(F^{+\oplus 0}) & & \text{final} \\
[\text{id}, \zeta_0^{F^+}] \uparrow & & \cong \\
F^+(F^{+\oplus 0}) + F^{+\oplus 0} & \xrightarrow{[p_1, p_2]} & F^{\oplus 0}
\end{array} \quad (3.10)$$

**Example 3.5.3.** Let  $F_A = A \times (\_)$ . According to the characterizations in Example 3.4.3,  $p_1$  and  $p_2$  have types  $A^+(A^+)^\omega \rightarrow A^\omega$  and  $(A^+)^\omega \rightarrow A^\omega$  respectively, and are given by the canonical flattening functions. See also Proposition 3.7.5.

The arrows  $p_1$  and  $p_2$  relate  $\text{tr}^B(c)$  and  $\text{dtr}(c)$  as follows.

**Theorem 3.5.4.** Assume that  $T$  and  $F$  constitute a Büchi trace situation (Definition 3.3.1) and satisfy Assumption 3.4.14. Then for each a Büchi  $(T, F)$ -system  $(X, c, (X_1, X_2))$ , we have  $Jp_1 \odot \text{dtr}_1(c) = \text{tr}_1^B(c)$  and  $Jp_2 \odot \text{dtr}_2(c) = \text{tr}_2^B(c)$ .

$$\begin{array}{ccccc}
F(X_1 + X_2) & \xrightarrow{F(\text{dtr}_1(c) + \text{dtr}_2(c))} & F(F^+(F^{+\oplus 0}) + F^{+\oplus 0}) & \xrightarrow{JF[p_1, p_2]} & F(F^{\oplus 0}) \\
\uparrow c_1 & & \cong \uparrow J(\iota_{F^{+\oplus 0}}^F)^{-1} & = & \cong \uparrow J\zeta_0^F \\
X_1 & \xrightarrow{\text{dtr}_1(c)} & F^+(F^{+\oplus 0}) & \xrightarrow{Jp_1} & F^{\oplus 0} \\
& & \text{tr}_1^B(c) & & 
\end{array} \quad (3.11)$$

$$\begin{array}{ccccc}
F(X_1 + X_2) & \xrightarrow{F(\text{dtr}_1(c) + \text{dtr}_2(c))} & F(F^+(F^{+\oplus 0}) + F^{+\oplus 0}) & \xrightarrow{JF[p_1, p_2]} & F(F^{\oplus 0}) \\
\uparrow c_2 & & \cong \uparrow J(\iota_{F^{+\oplus 0}}^F)^{-1} & = & \cong \uparrow J\zeta_0^F \\
X_2 & \xrightarrow{\text{dtr}_2(c)} & F^{+\oplus 0} & \xrightarrow{Jp_2} & F^{\oplus 0} \\
& & \cong \uparrow J\zeta_0^{F^+} & & 
\end{array} \quad (3.12) \quad \square$$

A full proof covering parity automata is found in Section 3.6. In the rest of this section, we sketch the proof for Büchi  $(T, F)$ -systems to explain the intuition.

We defined  $\text{tr}_1^B(c)$  and  $\text{tr}_2^B(c)$  as a solution of an HES (3.4). Recall from Definition 2.3.6 that when we are solving (3.4), we calculate the following intermediate data:  $l_1^{(1)}$ ,  $f_2^\ddagger$ ,  $l_2^{(2)}$  and  $l_1^{(2)}$ .

In contrast,  $\text{dtr}_1(c)$  and  $\text{dtr}_2(c)$  are simultaneously calculated as the greatest fixed point (Definition 3.4.15). However, their codomains  $F^+(F^{+\oplus 0})$  and  $F^{+\oplus 0}$  were defined in a hierarchized manner (see Section 3.4.1). Because of this, we can give the following ‘‘hierarchized’’ definition of  $\text{dtr}_1(c)$  and  $\text{dtr}_2(c)$  that is in parallel with the procedure to solve the HES (3.4).

**Definition 3.5.5** ( $\tilde{l}_1^{(1)}$ ,  $c_2^\ddagger$ ,  $\tilde{l}_2^{(2)}$ ,  $\tilde{l}_1^{(2)}$ ). We define Kleisli arrows  $\tilde{l}_1^{(1)} : X_1 \rightarrow F^+X_2$ ,  $c_2^\ddagger : X_2 \rightarrow F^+X_2$ ,  $\tilde{l}_2^{(2)} : X_2 \rightarrow F^{+\oplus 0}$  and  $\tilde{l}_1^{(2)} : X_1 \rightarrow F^{+\oplus 0}$  as follows:

- We define  $\tilde{l}_1^{(1)} : X_1 \rightarrow F^+X_2$  as the greatest homomorphism from an  $F(\_ + X_2)$ -coalgebra  $c_1$  to  $J(\iota_{X_2}^F)^{-1}$ .

$$\begin{array}{ccc}
F(X_1 + X_2) & \xrightarrow{\overline{F}(\tilde{l}_1^{(1)} + \text{id})} & F(F^+X_2 + X_2) \\
\uparrow c_1 & \cong \downarrow \tilde{l}_1^{(1)} & \cong \downarrow J(\iota_{X_2}^F)^{-1} \\
X_1 & \xrightarrow{\quad} & F^+X_2
\end{array} \quad (3.13)$$

- We define  $c_2^\ddagger: X_2 \rightarrow F^+X_2$  by:

$$c_2^\ddagger := \left( X_2 \xrightarrow{c_2} F(X_1 + X_2) \xrightarrow{\overline{F}(\tilde{l}_1^{(1)} + \text{id})} F(F^+X_2 + X_2) \xrightarrow{J\iota_{X_2}^F} F^+X_2 \right).$$

- Define  $\tilde{l}_2^{(2)}: X_2 \rightarrow F^{+\oplus 0}$  as the greatest homomorphism from  $c_2^\ddagger$  to  $J\zeta_0^{F^+}$ .

$$\begin{array}{ccc}
F^+X_2 & \xrightarrow{\overline{F^+}(\tilde{l}_2^{(2)})} & F^+(F^{+\oplus 0}) \\
\uparrow c_2^\ddagger & \cong \downarrow \tilde{l}_2^{(2)} & \cong \downarrow J\zeta_0^{F^+} \\
X_2 & \xrightarrow{\quad} & F^{+\oplus 0}
\end{array}$$

- We define  $\tilde{l}_1^{(2)}: X_1 \rightarrow F^+(F^{+\oplus 0})$  as follows:

$$\tilde{l}_1^{(2)} := \left( X_1 \xrightarrow{\tilde{l}_1^{(1)}} F^+X_2 \xrightarrow{\overline{F^+}\tilde{l}_2^{(2)}} F^+(F^{+\oplus 0}) \right).$$

Each of these four data is related to  $\text{dtr}_1(c)$  and  $\text{dtr}_2(c)$  as follows.

**Lemma 3.5.6.** *We have the following:*

1. For  $v_2: X_2 \rightarrow F^+(F^{+\oplus 0})$ ,  $\overline{F^+}v_2 \odot \tilde{l}_1^{(1)}$  is the greatest fixed point of the following function.

$$\left( X_1 \xrightarrow{v_1} F^+(F^{+\oplus 0}) \right) \mapsto \left( \begin{array}{ccc} F(X_1 + X_2) & \xrightarrow{F(v_1+v_2)} & F(F^+(F^{+\oplus 0}) + F^{+\oplus 0}) \\ \uparrow c_1 & & \cong \downarrow J\iota_{F^{+\oplus 0}}^F \\ X_1 & & F^+(F^{+\oplus 0}) \end{array} \right) \quad (3.14)$$

2. For  $v_2: X_2 \rightarrow F^+(F^{+\oplus 0})$ ,  $J(\zeta_0^{F^+})^{-1} \odot \overline{F^+}v_2 \odot c_2^\ddagger = J((\zeta_0^{F^+})^{-1} \odot \iota_{F^{+\oplus 0}}^F) \odot \overline{F}(v_1 + v_2) \odot c_2$ . Pictorially,

$$\left( \begin{array}{ccc} F^+X_2 & \xrightarrow{\overline{F^+}v_2} & F^+(F^{+\oplus 0}) \\ \uparrow c_2^\ddagger & \cong \downarrow J(\zeta_0^{F^+})^{-1} & \\ X_2 & & F^{+\oplus 0} \end{array} \right) = \left( \begin{array}{ccc} F(X_1 + X_2) & \xrightarrow{\overline{F}(v_1+v_2)} & F(F^+(F^{+\oplus 0}) + F^{+\oplus 0}) \\ \uparrow c_2 & & \cong \downarrow J\iota_{F^{+\oplus 0}}^F \\ X_2 & & F^+(F^{+\oplus 0}) \\ & & \cong \downarrow J(\zeta_0^{F^+})^{-1} \\ & & F^{+\oplus 0} \end{array} \right).$$

3.  $\tilde{l}_2^{(2)} = \text{dtr}_2(c)$ .

4.  $\tilde{l}_1^{(2)} = \text{dtr}_1(c)$ . □

We next show the relationship between  $l_1^{(1)}$ ,  $f_2^\ddagger$ ,  $l_2^{(2)}$  and  $l_1^{(2)}$  (intermediate data for the HES (3.4)), and  $\tilde{l}_1^{(1)}$ ,  $c_2^\ddagger$ ,  $\tilde{l}_2^{(2)}$  and  $\tilde{l}_1^{(2)}$  defined above.

The main difference between them is that while  $l_1^{(1)}(u_2)$  is defined as the *least* fixed point,  $\tilde{l}_1^{(1)}$  is the *greatest* homomorphism. The key to filling in this gap is that by Assumption 3.4.14.2 and Theorem 3.1.13, the coalgebra  $J(\iota_{F^{+\oplus 0}}^F)^{-1}$  in

(3.11) is a final coalgebra. This implies that  $\overline{F^+}v_2 \odot \tilde{l}_1^{(1)}$ , which is the greatest fixed point of (3.14), is also the *least* fixed point.

$$\begin{array}{ccccc}
F(X_1 + F^{+\oplus}0) & \xrightarrow{\overline{F}(\tilde{l}_1^{(1)} + \text{id})} & F(F^+X_2 + F^{+\oplus}0) & \xrightarrow{\overline{F}(F^+v_2 + \text{id})} & F(F^+(F^{+\oplus}0) + F^{+\oplus}0) \\
\overline{F}(\text{id} + v_2) \uparrow & & \uparrow \overline{F}(\text{id} + v_2) & & \uparrow J(\iota_{F^{+\oplus}0}^F)^{-1} \\
F(X_1 + X_2) & \xrightarrow{\overline{F}(\tilde{l}_1^{(1)} + \text{id})} & F(F^+X_2 + X_2) & = & \cong \\
c_1 \uparrow & & \cong \uparrow J(\iota_{X_2}^F)^{-1} & & \\
X_1 & \xrightarrow{\tilde{l}_1^{(1)}} & F^+X_2 & \xrightarrow{\overline{F^+}v_2} & F^+(F^{+\oplus}0)
\end{array}$$

Following this observation, we can prove the following lemma, which relate  $l_1^{(1)}$ ,  $f_2^\ddagger$ ,  $l_2^{(2)}$  and  $l_1^{(2)}$ , and  $l_1^{(1)}$ ,  $c_2^\ddagger$ ,  $\tilde{l}_2^{(2)}$  and  $\tilde{l}_1^{(2)}$  one by one.

**Lemma 3.5.7.** Define  $\xi^{F^{+\oplus}} : F^+(F^{\oplus}0) \rightarrow F^{\oplus}0$  by  $\xi^{F^{+\oplus}} := a \circ \kappa_1$ , where  $a$  is the unique homomorphism from  $[(\iota_{F^{\oplus}0}^F)^{-1}, F\kappa_2 \circ \zeta_0^F]$  to a final coalgebra  $\zeta_0^F$ .

$$\begin{array}{ccc}
F(F^+(F^{\oplus}0) + F^{\oplus}0) & \xrightarrow{\overline{F}a} & F(F^{\oplus}0) \\
\uparrow [\text{id}, F\kappa_2] & & \uparrow \text{final} \zeta_0^F \\
F(F^+(F^{\oplus}0) + F^{\oplus}0) + F(F^{\oplus}0) & & \cong \\
\cong \uparrow (\iota_{F^{\oplus}0}^F)^{-1} + \zeta_0^F & & \\
F^+(F^{\oplus}0) \xrightarrow{\kappa_1} F^+(F^{\oplus}0) + F^{\oplus}0 & \xrightarrow{a} & F^{\oplus}0
\end{array}$$

1. For  $u_2: X_2 \rightarrow F^{\oplus}0$ ,

$$\left( X_1 \xrightarrow{l_1^{(1)}(u_2)} F^{\oplus}0 \right) = \left( X_1 \xrightarrow{\tilde{l}_1^{(1)}} F^+X_2 \xrightarrow{F^+u_2} F^+(F^{\oplus}0) \xrightarrow{\xi^{F^{+\oplus}}} F^{\oplus}0 \right).$$

2. For  $u_2: X_2 \rightarrow F^{\oplus}0$ ,  $f_2^\ddagger(u_2) = J\xi^{F^{+\oplus}} \odot \overline{F^+}u_2 \odot c_2^\ddagger$ . Pictorially,

$$\left( X_2 \xrightarrow{f_2^\ddagger(u_2)} F^{\oplus}0 \right) = \left( \begin{array}{ccc} F^+X_2 & \xrightarrow{\overline{F^+}u_2} & F^+(F^{\oplus}0) \\ \uparrow c_2^\ddagger & & \downarrow J\xi^{F^{+\oplus}} \\ X_2 & & F^{+\oplus}0 \end{array} \right).$$

3.  $l_2^{(2)}(*) = Jp_2 \odot \tilde{l}_2^{(2)}$ .

4.  $l_1^{(2)}(*) = Jp_2 \odot \tilde{l}_1^{(2)}$ . □

Lemma 3.5.6.3–4 and Lemma 3.5.7.3–4 together imply Theorem 3.5.4.

## 3.6 Extension to Parity Automata

In this section, we extend the framework in Sections 3.2–3.5 to systems with the *parity acceptance condition*.

### 3.6.1 Categorical Representation of parity Automata

This section generalizes Section 3.2. The following definition corresponds to Definition 3.2.1.

**Definition 3.6.1** (parity  $(T, F)$ -system). Let  $n \in \mathbb{N}$ . A *parity  $(T, F)$ -system* is a triple  $\mathcal{X} = (X, c, (X_1, \dots, X_{2n}))$  of an object  $X \in \mathbb{C}$ , a  $TF$ -coalgebra  $c: X \rightarrow TFX$  and a tuple of objects  $X_1, \dots, X_{2n} \in \mathbb{C}$  such that  $X = X_1 + \dots + X_{2n}$ . For  $i \in \{1, \dots, 2n\}$ , we write  $c_i$  for  $c \circ \kappa_i: X_i \rightarrow \overline{F}X$ .



Intuitively,  $X_i$  is the set of states whose priorities are  $i$ .

**Example 3.6.2.** Let  $F_A := A \times (\_)$ . An  $A$ -labeled nondeterministic parity word automaton (see Remark 2.2.11)  $\mathcal{A} = (X, \tau, \mathbf{p})$  where  $\mathbf{p}: X \rightarrow \{1, \dots, 2n\}$  induces a  $(\mathcal{P}, F_A)$ -system  $\mathcal{X}_{\mathcal{A}} = (X, c_{\mathcal{A}}, (X_1, \dots, X_{2n}))$  defined by  $c_{\mathcal{A}} := \tau$  and  $X_i := \{x \in X \mid \mathbf{p}(x) = i\}$  for each  $i \in \{1, \dots, 2n\}$ .

### 3.6.2 Characterization via Logical Fixed Point: Parity Case

In this section, we extend the discussions in Section 3.3.

The following generalizes Definition 3.3.1.

**Definition 3.6.3** ( $\text{tr}_i^{\mathbf{p}}(c)$ ). Let  $F$  be an endofunctor and  $T$  be a monad on a category  $\mathbb{C}$ . Assume that a  $F$  is equipped with a lifting  $\bar{F}: \mathcal{Kl}(T) \rightarrow \mathcal{Kl}(T)$ , and each homset of  $\mathcal{Kl}(T)$  carries a partial order  $\sqsubseteq$ . We say that  $F$  and  $T$  constitute a *parity trace situation* with respect to  $\sqsubseteq$  if they satisfy the following conditions:

- A final  $F$ -coalgebra  $\zeta^F: \nu F \rightarrow F(\nu F)$  exists.
- For each parity  $(T, F)$ -system  $(X, c, (X_1, \dots, X_{2n}))$ , the following HES has a solution (here  $\mu$  and  $\nu$  appear in the alternating manner).

$$E_c = \begin{cases} u_1 =_{\mu} J\zeta^{-1} \odot \bar{F}[u_1, \dots, u_{2n}] \odot c_1 & \in (\mathcal{Kl}(T)(X_1, \nu F), \sqsubseteq_{X_1, \nu F}) \\ u_2 =_{\nu} J\zeta^{-1} \odot \bar{F}[u_1, \dots, u_{2n}] \odot c_2 & \in (\mathcal{Kl}(T)(X_2, \nu F), \sqsubseteq_{X_2, \nu F}) \\ \vdots \\ u_{2n} =_{\nu} J\zeta^{-1} \odot \bar{F}[u_1, \dots, u_{2n}] \odot c_{2n} & \in (\mathcal{Kl}(T)(X_{2n}, \nu F), \sqsubseteq_{X_{2n}, \nu F}) \end{cases}$$

The solution of  $E_c$  is denoted by  $(\text{tr}_i^{\mathbf{p}}(c): X_i \rightarrow \nu F)_{1 \leq i \leq 2n}$  and is called the (*coalgebraic*) *parity trace semantics* of  $\mathcal{X}$ . Using diagrams, the HES is as follows:

$$\begin{array}{ccccc} \bar{F}[u_1, \dots, u_{2n}] & \bar{F}[u_1, \dots, u_{2n}] & & \bar{F}[u_1, \dots, u_{2n}] & \\ FX \longrightarrow F(\nu F) & FX \longrightarrow F(\nu F) & & FX \longrightarrow F(\nu F) & (3.15) \\ c_1 \uparrow \quad =_{\mu} \quad J\zeta^F \uparrow \cong & c_2 \uparrow \quad =_{\nu} \quad J\zeta^F \uparrow \cong & & c_{2n} \uparrow \quad =_{\nu} \quad J\zeta^F \uparrow \cong & \\ X_1 \xrightarrow{u_1} \nu F, & X_2 \xrightarrow{u_2} \nu F, & \dots, & X_{2n} \xrightarrow{u_{2n}} \nu F. & \end{array}$$

Note that the notions of Büchi  $(T, F)$ -system (Definition 3.2.1) and Büchi trace semantics (Definition 3.3.1) are special cases of parity  $(T, F)$ -system and parity trace semantics.

**Example 3.6.4.** We continue Example 3.6.2. For each  $i \in \{1, \dots, 2n\}$ ,  $\text{tr}_i^{\mathbf{p}}(c_{\mathcal{A}})$  has a type  $X_i \rightarrow A^{\omega}$  (i.e.  $X_i \rightarrow \mathcal{P}(A^{\omega})$ ), and is given by  $\text{tr}_i^{\mathbf{p}}(c_{\mathcal{A}})(x) = L_{\mathcal{A}}^{\mathbf{p}}(x)$  for  $x \in X_i$  (see Section 3.7 for a proof covering parity *tree* automata).

### 3.6.3 Characterization via Categorical Fixed Point: Parity Case

This section generalizes Section 3.4. Recall that we have used objects  $F^+(F^{+\oplus 0})$  and  $F^{+\oplus 0}$ , and coalgebras  $(\iota_{F^{+\oplus 0}}^F)^{-1}$  and  $(\iota_{F^{+\oplus 0}}^F)^{-1} \circ \zeta_0^{F^+}$  in the definition of categorical decorated trace semantics of Büchi  $(T, F)$ -systems  $\mathcal{X} = (X, c, (X_1, X_2))$  (Definition 3.4.15). They were components of the following chain of (co)algebras:

$$F^{+\oplus 0} \xrightarrow{\zeta_0^{F^+} \cong} F^+(F^{+\oplus 0}) \xleftarrow{\iota_{F^{+\oplus 0}}^F \cong} F(F^+(F^{+\oplus 0}) + F^{+\oplus 0}).$$

Similarly, for defining coalgebraic decorated trace semantics of a parity  $(T, F)$ -system  $\mathcal{X} = (X, c, (X_1, \dots, X_{2n}))$  we use a chain whose first component is  $F^{(+\oplus)^n} 0$  (here  $(+\oplus)^n$  denotes  $n$ -repetition of  $+\oplus$ ).

**Definition 3.6.5** ( $F_i^\ddagger, F_j^{(i)}, \alpha_j^{(i)}, \beta_j^{(i)}$ ). For  $i \in \mathbb{N}$ , we define  $F_i^\ddagger : \mathbb{C} \rightarrow \mathbb{C}$  by  $F_i^\ddagger := F^{(+\oplus)^l}$  if  $i = 2l$  and  $F_i^\ddagger := F^{(+\oplus)^{l+1}}$  if  $i = 2l + 1$ . For  $i \in \mathbb{N}$  and  $j \in [0, i]$ , we inductively define  $F_j^{(i)} : \mathbb{C} \rightarrow \mathbb{C}$  as follows:

- $F_i^{(i)} := F_i^\ddagger$ ; and
- $F_j^{(i)} := F_j^\ddagger (\coprod_{k=j+1}^i F_k^{(i)}(\_) + \_)$  for  $j < i$ .

We define an (isomorphic) natural transformation  $\alpha_j^{(i)} : F_j^{(i)} \Rightarrow F_{j-1}^{(i)}$  as follows:

$$\alpha_{j,X}^{(i)} := \begin{cases} \left( \begin{array}{c} F_{j-1}^\ddagger \\ \prod_{k=j+1}^i F_k^{(i)} X+X \end{array} \right)^{-1} & (j \text{ is odd}) \\ \begin{array}{c} F_{j-1}^\ddagger \\ \prod_{k=j+1}^i F_k^{(i)} X+X \end{array} & (j \text{ is even}). \end{cases}$$

Moreover, we define a natural transformation  $\beta_j^{(i)} : F_j^{(i)} \Rightarrow F_0^{(i)}$  by:

$$\beta_{j,X}^{(i)} := \left( F_j^{(i)} X \xrightarrow{\alpha_{j,X}^{(i)}} F_{j-1}^{(i)} X \xrightarrow{\alpha_{j-1,X}^{(i)}} \dots \xrightarrow{\alpha_{1,X}^{(i)}} F_0^{(i)} X \right).$$

The naturality of  $\alpha_j^{(i)}$  is proved by Definition 3.4.2, and it follows the naturality of  $\beta_j^{(i)}$ .

**Example 3.6.6.** We continue Example 3.6.4. Analogously to Example 3.4.3, we can identify  $F_j^{(i)}$  with a set of “decorated runs” as follows:

$$\begin{aligned} F_j^{(i)} X \cong & \{ (a_0, p_0) \dots (a_k, p_k) x \in (\mathbb{A} \times \{1, \dots, i\})^+ \times X \mid p_0 = j \} \\ & \cup \{ (a_0, p_0)(a_1, p_1) \dots \in (\mathbb{A} \times \{1, \dots, i\})^\omega \mid p_0 = j, \limsup_{i \rightarrow \infty} p_i \text{ is even} \}. \end{aligned}$$

We next consider generalizing Definition 3.6.8. For parity  $(T, F)$ -systems, we modify Assumption 3.4.14 as follows.

**Assumption 3.6.7.** Let  $T$  be a monad and  $F$  be an endofunctor on  $\mathbb{C}$ . Assume that each homset of  $\mathcal{Kl}(T)$  carries a partial order  $\sqsubseteq$ . We further assume the following conditions for each  $n \in \mathbb{N}$  and  $A \in \mathbb{C}$ .

1.  $F_n^\ddagger : \mathbb{C} \rightarrow \mathbb{C}$  is well-defined and a lifting  $\overline{F_n^\ddagger} : \mathcal{Kl}(T) \rightarrow \mathcal{Kl}(T)$  of  $F_n^\ddagger$  is given.
2. If  $n$  is even,  $T$  and  $F_n^\ddagger(\_ + A)$  satisfy the conditions in Theorem 3.1.13.
3. If  $n$  is odd,  $T$  and  $F_n^\ddagger(\_ + A)$  constitute an infinitary trace situation.
4. If  $n$  is odd,  $T$  and  $F_n^\ddagger(\_ + A)$  satisfy the gfp-preserving condition for an arbitrary  $\sigma$ .
5. If  $n$  is odd,  $T$  and  $F_n^\ddagger(\_ + A)$  satisfy the deterministic-greatest condition.
6. The lifting  $\overline{F_{n+1}^\ddagger}$  is obtained from  $\overline{F_n^\ddagger}$  using the procedure in Definition 3.4.6.
7. For  $n \in \mathbb{N}$  and  $A \in \mathbb{C}$ ,  $\overline{F_n^\ddagger}(\_ + A)$  is locally monotone.

The following definition is a straight-forward generalization of Definition 3.4.15.

**Definition 3.6.8** ( $\text{dtr}_i(c)$ ). For a parity  $(T, F)$ -system  $(X, c, (X_1, \dots, X_{2n}))$ , we define its *decorated parity trace semantics*  $\text{dtr}_i(c): X_i \rightarrow F_i^{(2n)}0$  as the greatest fixed point of the following function with respect to the product order:

$$\begin{pmatrix} v_1, \\ v_2, \\ \vdots \\ v_{2n} \end{pmatrix} \mapsto \begin{pmatrix} J(\beta_{1,0}^{(2n)})^{-1} \odot \bar{F}(v_1 + \dots + v_{2n}) \odot c_1, \\ J(\beta_{2,0}^{(2n)})^{-1} \odot \bar{F}(v_1 + \dots + v_{2n}) \odot c_2, \\ \vdots \\ J(\beta_{2n,0}^{(2n)})^{-1} \odot \bar{F}(v_1 + \dots + v_{2n}) \odot c_{2n} \end{pmatrix} \in \begin{matrix} (\mathcal{Kl}(T)(X_1, F_1^{(2n)}0), \sqsubseteq) \\ \times (\mathcal{Kl}(T)(X_2, F_2^{(2n)}0), \sqsubseteq) \\ \times \dots \\ \times (\mathcal{Kl}(T)(X_{2n}, F_{2n}^{(2n)}0), \sqsubseteq) \end{matrix} \quad (3.16)$$

Pictorially,

$$\begin{array}{ccccc} \begin{array}{c} \bar{F}(v_1 + \dots + v_{2n}) \\ FX \xrightarrow{\quad} F(\prod F_j^{(2n)}0) \\ c_1 \uparrow \cong \uparrow J\beta_{1,0}^{(2n)} \\ X_1 \xrightarrow{v_1} F_1^{(2n)}0 \end{array} & \begin{array}{c} \bar{F}(v_1 + \dots + v_{2n}) \\ FX \xrightarrow{\quad} F(\prod F_j^{(2n)}0) \\ c_2 \uparrow \cong \uparrow J\beta_{2,0}^{(2n)} \\ X_2 \xrightarrow{v_2} F_2^{(2n)}0 \end{array} & \cdots & \begin{array}{c} \bar{F}(v_1 + \dots + v_{2n}) \\ FX \xrightarrow{\quad} F(\prod F_j^{(2n)}0) \\ c_{2n} \uparrow \cong \uparrow J\beta_{2n,0}^{(2n)} \\ X_{2n} \xrightarrow{v_{2n}} F_{2n}^{(2n)}0 \end{array} & . \end{array}$$

### 3.6.4 Logical Fixed Point vs. Categorical Fixed Point: Parity Case

This section generalizes Section 3.5. We first generalize  $p_1$  and  $p_2$  in Definition 3.5.2. We shall define their generalizations as natural transformations.

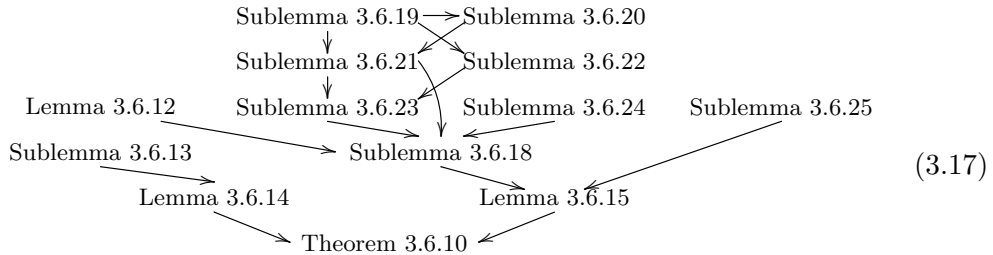
**Definition 3.6.9** ( $p_j^{(i)}$ ). For  $i \in \mathbb{N}$  and  $j \in \{1, \dots, i\}$ , we define  $p_j^{(i)}: F_j^{(i)} \Rightarrow F^\oplus$  so that  $[p_{1,X}^{(i)}, \dots, p_{i,X}^{(i)}]: \prod_{j=1}^i F_j^{(i)}(X) \rightarrow F^\oplus X$  is the unique homomorphism from  $[\beta_{1,X}^{(i)}, \dots, \beta_{i,X}^{(i)}]$  to  $\zeta_X^F$ .

$$\begin{array}{ccc} F(\prod_{j=1}^i F_j^{(i)} X + X) & \xrightarrow{\quad} & F(F^\oplus X + X) \\ \uparrow [p_{1,X}^{(i)}, \dots, p_{i,X}^{(i)}] & & \uparrow \text{final} \\ \prod_{j=1}^i F_j^{(i)} X & \xrightarrow{[\beta_{1,X}^{(i)}, \dots, \beta_{i,X}^{(i)}]} & F^\oplus X \end{array}$$

The following theorem generalizes Theorem 3.5.4.

**Theorem 3.6.10.** For each  $i \in \{1, \dots, 2n\}$ ,  $\text{tr}_i^P(c) = p_{i,0}^{(2n)} \circ \text{dtr}_i(c)$ .

The rest of this section (upto page 57) is devoted to proving the above theorem following the intuition explained in Section 3.5. We have to prove many sublemmas and lemmas. Their dependencies are as follows.



Firstly, we generalize Definition 3.5.5.

**Definition 3.6.11** ( $c_i^\ddagger, \tilde{l}_j^{(i)}$ ). For  $i \in \{1, \dots, 2n\}$  and  $j \in \{1, \dots, i\}$ , using  $\alpha_j^{(i)}$  and  $\beta_j^{(i)}$  defined in Definition 3.6.5, we inductively define  $c_i^\ddagger: X_i \rightarrow F_{i-1}^\ddagger(X_i + \dots + X_{2n})$  and  $\tilde{l}_j^{(i)}: X_j \rightarrow F_j^{(i)}(X_{i+1} + \dots + X_{2n})$  as follows (no need to distinguish the base case from the step case):

- $c_i^\ddagger: X_i \rightarrow F_{i-1}^\ddagger(X_i + \cdots + X_{2n})$  is defined by:

$$c_i^\ddagger := \left( \begin{array}{c} X_i \xrightarrow{c_i} F(X_1 + \cdots + X_{2n}) \xrightarrow{\overline{F}(\tilde{l}_1^{(i-1)} + \cdots + \tilde{l}_{i-1}^{(i-1)} + \text{id}_{X_i} + \cdots + \text{id}_{X_{2n}})} \\ F(\prod_{j=1}^{i-1} F_j^{(i-1)}(X_i + \cdots + X_{2n}) + X_i + \cdots + X_{2n}) \\ = F_0^{(i-1)}(X_i + \cdots + X_{2n}) \xrightarrow{(\beta_{i-1, X_i + \cdots + X_{2n}}^{(i-1)})^{-1}} F_{i-1}^\ddagger(X_i + \cdots + X_{2n}) \end{array} \right).$$

- $\tilde{l}_i^{(i)}: X_i \rightarrow F_i^{(i)}(X_{i+1} + \cdots + X_{2n})$  is defined as follows:

- If  $i = 2k-1$ , we define  $\tilde{l}_i^{(i)}: X_i \rightarrow F_i^{(i)}(X_{i+1} + \cdots + X_{2n}) = F^{(+\oplus)^{k-1}}(X_{i+1} + \cdots + X_{2n})$  as the unique homomorphism from  $c_i^\ddagger$  to  $J\alpha_{i, X_{i+1} + \cdots + X_{2n}}^{(i)} (= J(\iota_{X_{i+1} + \cdots + X_{2n}}^{F^{(+\oplus)^{k-1}}})^{-1})$  (see Assumption 3.6.7.2).

$$\begin{array}{ccc} & \overline{F^{(+\oplus)^k}(\tilde{l}_i^{(i)} + \text{id}_{X_{i+1}} + \cdots + \text{id}_{X_{2n}})} & \\ F^{(+\oplus)^{k-1}} \left( \begin{array}{c} X_i \\ + X_{i+1} + \cdots + X_{2n} \end{array} \right) \xrightarrow{-} & \xrightarrow{-} & F^{(+\oplus)^{k-1}} \left( \begin{array}{c} F^{(+\oplus)^{k-1}}(X_{i+1} + \cdots + X_{2n}) \\ + X_{i+1} + \cdots + X_{2n} \end{array} \right) \\ \uparrow c_i^\ddagger & = & \cong \uparrow J(\iota_{X_{i+1} + \cdots + X_{2n}}^{F^{(+\oplus)^{k-1}}})^{-1} \\ X_i \xrightarrow{-} & \xrightarrow{-} & \xrightarrow{-} F^{(+\oplus)^{k-1}}(X_{i+1} + \cdots + X_{2n}). \end{array}$$

- If  $i = 2k$ , we define  $\tilde{l}_i^{(i)}: X_i \rightarrow F_i^{(i)}(X_{i+1} + \cdots + X_{2n}) = F^{(+\oplus)^k}(X_{i+1} + \cdots + X_{2n})$  as the greatest homomorphism from  $c_i^\ddagger$  to  $J\alpha_{i, X_{i+1} + \cdots + X_{2n}}^{(i)} (= J\zeta_{X_{i+1} + \cdots + X_{2n}}^{F^{(+\oplus)^{k-1}}})$  (see Assumption 3.6.7.3).

$$\begin{array}{ccc} & \overline{F^{(+\oplus)^{k-1}}(\tilde{l}_i^{(i)} + \text{id}_{X_{i+1}} + \cdots + \text{id}_{X_{2n}})} & \\ F^{(+\oplus)^{k-1}} \left( \begin{array}{c} X_i \\ + X_{i+1} + \cdots + X_{2n} \end{array} \right) \xrightarrow{-} & \xrightarrow{-} & F^{(+\oplus)^{k-1}} \left( \begin{array}{c} F^{(+\oplus)^k}(X_{i+1} + \cdots + X_{2n}) \\ + X_{i+1} + \cdots + X_{2n} \end{array} \right) \\ \uparrow c_i^\ddagger & = \nu & \cong \uparrow J\zeta_{X_{i+1} + \cdots + X_{2n}}^{F^{(+\oplus)^{k-1}}} \\ X_i \xrightarrow{-} & \xrightarrow{-} & \xrightarrow{-} F^{(+\oplus)^k}(X_{i+1} + \cdots + X_{2n}). \end{array}$$

- For  $j < i$ ,  $\tilde{l}_j^{(i)}: X_j \rightarrow F_j^{(i)}(X_{i+1} + \cdots + X_{2n})$  is defined by:

$$\tilde{l}_j^{(i)} := \left( \begin{array}{c} X_j \xrightarrow{\tilde{l}_j^{(i-1)}} F_j^{(i-1)}(X_i + X_{i+1} + \cdots + X_{2n}) \xrightarrow{\overline{F_j^{(i-1)}(\tilde{l}_i^{(i)} + J\text{id}_{X_{i+1} + \cdots + X_{2n}})}} \\ F_j^{(i-1)}(F_i^{(i)}(X_{i+1} + \cdots + X_{2n}) + X_{i+1} + \cdots + X_{2n}) \\ = F_j^{(i)}(X_{i+1} + \cdots + X_{2n}) \end{array} \right).$$

Here the last equality is by the following lemma.

**Lemma 3.6.12.** For  $i \in \mathbb{N}$  and  $j \in \{0, \dots, i\}$ ,  $F_j^{(i)}(F_{i+1}^{(i+1)}(\_) + \_) = F_j^{(i+1)}$ .

**Proof.** We prove the statement by the induction on  $j$ .

If  $j = i$  then the statement is immediate by definition.

If  $j = 2l - 1 < i$ , we have:

$$\begin{aligned}
& F_j^{(i)}(F_{i+1}^{(i+1)}(\_)+\_) \\
&= F^{(+\oplus)^{l-1}+} \left( \prod_{k=2l}^i F_k^{(i)}(F_{i+1}^{(i+1)}(\_)+\_) + F_{i+1}^{(i+1)}(\_)+\_ \right) \quad (\text{by definition}) \\
&= F^{(+\oplus)^{l-1}+} \left( \prod_{k=2l}^i F_k^{(i+1)}(\_)+F_{i+1}^{(i+1)}(\_)+\_ \right) \quad (\text{by IH}) \\
&= F^{(+\oplus)^{l-1}+} \left( \prod_{k=2l}^{i+1} F_k^{(i+1)}(\_)+\_ \right) \\
&= F_j^{(i+1)} \quad (\text{by definition}).
\end{aligned}$$

We can similarly prove the statement when  $j$  is even.  $\square$

The sublemma below generalizes Lemma 3.5.6.1–2. That is, it shows that if  $j$  is odd (resp. even), not only  $\tilde{l}_j^{(j)}$  but also  $\tilde{l}_j^{(i)}$  with  $i > j$  is characterized as the unique (resp. greatest) homomorphism.

**Sublemma 3.6.13.** *Let  $i \in \{1, \dots, 2n\}$  and  $j \in \{1, \dots, i\}$ . For simplicity, we write  $X_j^{(i)}$  for  $\prod_{k=j+1}^i F_k^{(i)}(X_{i+1} + \dots + X_{2n}) + X_{i+1} + \dots + X_{2n}$ . Note that by Definition 3.6.5,  $\alpha_{j, X_{i+1} + \dots + X_{2n}}^{(i)}$  has a type  $F_j^{(i)}(X_{i+1} + \dots + X_{2n}) \rightarrow F_{j-1}^{(i)}(X_{i+1} + \dots + X_{2n}) = F_{j-1}^{\ddagger}(F_j^{(i)}(X_{i+1} + \dots + X_{2n}) + X_j^{(i)})$  and hence is an  $F_{j-1}^{\ddagger}(\_ + X_j^{(i)})$ -coalgebra.*

1. *If  $j$  is odd,  $\tilde{l}_j^{(i)} : X_j \rightarrow F_j^{(i)}(X_{i+1} + \dots + X_{2n})$  is the unique homomorphism from an  $F_{j-1}^{\ddagger}(\_ + X_j^{(i)})$ -coalgebra  $\overline{F_{j-1}^{\ddagger}}(\text{id}_{X_j} + \prod_{k=j+1}^i \tilde{l}_k^{(i)} + \text{id}_{X_{i+1} + \dots + X_{2n}}) \odot c_j^{\ddagger}$  to  $\alpha_{j, X_{i+1} + \dots + X_{2n}}^{(i)}$  ( $= J(\iota_{X_j}^{F_{j-1}^{\ddagger}})^{-1}$ ).*
2. *If  $j$  is even,  $\tilde{l}_j^{(i)} : X_j \rightarrow F_j^{(i)}(X_{i+1} + \dots + X_{2n})$  is the greatest homomorphism from an  $F_{j-1}^{\ddagger}(\_ + X_j^{(i)})$ -coalgebra  $\overline{F_{j-1}^{\ddagger}}(\text{id}_{X_j} + \prod_{k=j+1}^i \tilde{l}_k^{(i)} + \text{id}_{X_{i+1} + \dots + X_{2n}}) \odot c_j^{\ddagger}$  to  $\alpha_{j, X_{i+1} + \dots + X_{2n}}^{(i)}$  ( $= J\zeta_{X_j}^{F_{j-1}^{\ddagger}}$ ).*

**Proof.** Item 1 is easily proved by the finality of  $J(\iota_{X_{i+1} + \dots + X_{2n}}^{F_{j-1}^{\ddagger}})^{-1}$ . We prove Item 2 by the induction on  $i$ .

If  $i = j$ , then the statement is immediate by the definition of  $\tilde{l}_j^{(j)} : X_j \rightarrow F_j^{(j)}(X_{j+1} + \dots + X_{2n})$  (Definition 3.6.11).

Let  $i > j$  and assume that  $\tilde{l}_j^{(i-1)} : X_j \rightarrow F_j^{(i-1)}(X_i + \dots + X_{2n})$  is the greatest homomorphism from  $\overline{F_{j-1}^{\ddagger}}(\text{id}_{X_j} + \prod_{k=j+1}^{i-1} \tilde{l}_k^{(i-1)} + \text{id}_{X_i + \dots + X_{2n}}) \odot c_j^{\ddagger}$  to  $J\zeta_{X_j}^{F_{j-1}^{\ddagger}}$ .

By the definition of  $F_j^{(i-1)}$ , we have the following equation.

$$\begin{aligned}
& \overline{F_j^{(i-1)}}(\tilde{l}_i^{(i)} + \text{id}_{X_{i+1} + \dots + X_{2n}}) \\
&= \overline{F_j^{\ddagger}} \left( \prod_{k=j+1}^{i-1} \overline{F_k^{(i-1)}}(\tilde{l}_i^{(i)} + \text{id}_{X_{i+1} + \dots + X_{2n}}) + \tilde{l}_i^{(i)} + \text{id}_{X_{i+1} + \dots + X_{2n}} \right)
\end{aligned}$$

By the definition of a lifting  $\overline{F_j^\ddagger}$  (Definition 3.4.6.2), this means that  $\overline{F_j^{(i-1)}}(\tilde{l}_i^{(i)} + \text{id}_{X_{i+1}+\dots+X_{2n}})$  is the greatest homomorphism from  $\overline{F_{j-1}^\ddagger}(\text{id}_{X_j} + \prod_{k=j+1}^{i-1} F_k^{(i)}(\tilde{l}_i^{(i)} + \text{id}_{X_{i+1}+\dots+X_{2n}})) + \tilde{l}_i^{(i)} + \text{id}_{X_{i+1}+\dots+X_{2n}} \odot J\zeta_{X_j^{(i-1)}}^{F_{j-1}^\ddagger}$  to  $J\zeta_{X_j^{(i)}}^{F_{j-1}^\ddagger}$ . Hence by the gfp-preserving condition,  $\tilde{l}_j^{(i)} = \overline{F_j^{(i-1)}}(\tilde{l}_i^{(i)} + \text{id}) \odot \tilde{l}_j^{(i-1)}$  is the greatest homomorphism from  $\overline{F_{j-1}^\ddagger}(\text{id}_{X_j} + \prod_{k=j+1}^{i-1} F_k^{(i)}(\tilde{l}_i^{(i)} + \text{id}_{X_{i+1}+\dots+X_{2n}})) + \tilde{l}_i^{(i)} + \text{id}_{X_{i+1}+\dots+X_{2n}} \odot \overline{F_{j-1}^\ddagger}(\text{id} + \prod_{k=j+1}^{i-1} \tilde{l}_k^{(i-1)} + \text{id}_{X_{i+1}+\dots+X_{2n}}) \odot c_j^\ddagger$  to  $J\zeta_{X_j^{(i)}}^{F_{j-1}^\ddagger}$ . Here by the definition of  $\tilde{l}_k^{(i)}$ , the former coalgebra can be transformed as follows:

$$\begin{aligned} & \overline{F_{j-1}^\ddagger} \left( \text{id}_{X_j} + \prod_{k=j+1}^{i-1} \overline{F_k^{(i)}}(\tilde{l}_i^{(i)} + \text{id}_{X_{i+1}+\dots+X_{2n}}) + \tilde{l}_i^{(i)} + \text{id}_{X_{i+1}+\dots+X_{2n}} \right) \\ & \quad \odot \overline{F_{j-1}^\ddagger}(\text{id} + \prod_{k=j+1}^{i-1} \tilde{l}_k^{(i-1)} + \text{id}) \odot c_j^\ddagger \\ &= \overline{F_{j-1}^\ddagger} \left( \text{id}_{X_j} + \left( \prod_{k=j+1}^{i-1} \overline{F_k^{(i)}}(\tilde{l}_i^{(i)} + \text{id}_{X_{i+1}+\dots+X_{2n}}) \odot \tilde{l}_k^{(i-1)} \right) + \tilde{l}_i^{(i)} + \text{id}_{X_{i+1}+\dots+X_{2n}} \right) \\ & \quad \odot c_j^\ddagger \\ &= \overline{F_{j-1}^\ddagger}(\text{id}_{X_j} + \prod_{k=j+1}^i \tilde{l}_k^{(i)} + \text{id}_{X_{i+1}+\dots+X_{2n}}) \odot c_j^\ddagger. \end{aligned}$$

Hence the statement is proved. See also Figure 3.1.  $\square$

The above lemma implies the following. It generalizes Lemma 3.5.6.3–4.

**Lemma 3.6.14.** *For each  $i \in \{1, \dots, 2n\}$ ,  $\tilde{l}_i^{(2n)} = \text{dtr}_i(c)$ .*

**Proof.** Assume that  $j$  is odd. By Sublemma 3.6.13.1,  $\tilde{l}_j^{(2n)}$  is the unique homomorphism from  $\overline{F_{j-1}^\ddagger}(\tilde{l}_j^{(i)} + \text{id}_{\prod_{k=j+1}^{2n} X_k}) \odot c_j^\ddagger$  to  $\alpha_{j,0}^{(2n)} = J(\mathcal{L}_{\prod_{k=j+1}^{2n} F_k^{(2n)}_0}^{F_{j-1}^\ddagger})^{-1}$ . This means that it is also the greatest homomorphism.

$$\begin{array}{ccc} \overline{F_{j-1}^\ddagger}(\text{id} + \prod_{k=j+1}^{2n} \tilde{l}_k^{(2n)}) \uparrow & & \overline{F_{j-1}^\ddagger}(\tilde{l}_j^{(2n)} + \text{id}) \xrightarrow{\quad} \overline{F_{j-1}^\ddagger}(\mathcal{L}_{\prod_{k=j+1}^{2n} F_k^{(2n)}_0}^{F_{j-1}^\ddagger})^{-1} \\ \overline{F_{j-1}^\ddagger}(X_j + \dots + X_{2n}) & \stackrel{= \nu}{=} & \uparrow J(\mathcal{L}_{\prod_{k=j+1}^{2n} F_k^{(2n)}_0}^{F_{j-1}^\ddagger})^{-1} \\ \uparrow c_j^\ddagger & & \uparrow \\ X_j & \xrightarrow{\quad \tilde{l}_j^{(2n)} \quad} & F_j^{(2n)}_0 \end{array}$$

Hence by the definition of  $c_j^\ddagger$  (Definition 3.6.11),  $\tilde{l}_j^{(2n)}$  is the greatest fixed point of the following function.

$$f \mapsto \left( \begin{array}{l} J\mathcal{L}_{\prod_{k=j+1}^{2n} F_k^{(2n)}_0}^{F_{j-1}^\ddagger} \odot \overline{F_{j-1}^\ddagger}(f + \tilde{l}_{j+1}^{(2n)} + \dots + \tilde{l}_{2n}^{(2n)}) \\ \odot J(\beta_{j-1, X_i+\dots+X_{2n}}^{(j-1)})^{-1} \odot \overline{F}(\tilde{l}_1^{(i-1)} + \dots + \tilde{l}_{i-1}^{(i-1)} + \text{id}_{\prod_{k=i}^{2n} X_k}) \odot c_j \end{array} \right)$$

Note here that the right-hand side can be transformed as follows:

$$J\mathcal{L}_{\prod_{k=j+1}^{2n} F_k^{(2n)}_0}^{F_{j-1}^\ddagger} \odot \overline{F_{j-1}^\ddagger}(f + \tilde{l}_{j+1}^{(2n)} + \dots + \tilde{l}_{2n}^{(2n)}) \odot J(\beta_{j-1}^{(j-1)} \prod_{k=j}^{2n} X_k)^{-1}$$



$$\begin{aligned}
& \odot \overline{F}(\tilde{l}_1^{(j-1)} + \cdots + \tilde{l}_{j-1}^{(j-1)} + \text{id}_{\prod_{k=i}^{2n} X_k}) \odot c_j \\
= & J \iota_{\prod_{k=j+1}^{2n} F_k^{(2n)} 0}^{F_{j-1}^\dagger} \odot J(\beta_{j-1, \prod_{k=j}^{2n} F_k^{(2n)} 0}^{(j-1)})^{-1} \odot \overline{F_0^{(j-1)}}(f + \tilde{l}_{j+1}^{(2n)} + \cdots + \tilde{l}_{2n}^{(2n)}) \\
& \odot \overline{F}(\tilde{l}_1^{(j-1)} + \cdots + \tilde{l}_{j-1}^{(j-1)} + \text{id}_{\prod_{k=i}^{2n} X_k}) \odot c_j \quad (\text{by naturality of } \beta_{j-1}^{(j-1)}) \\
= & J \iota_{\prod_{k=j+1}^{2n} F_k^{(2n)} 0}^{F_{j-1}^\dagger} \odot J(\beta_{j-1, \prod_{k=j}^{2n} F_k^{(2n)} 0}^{(j-1)})^{-1} \\
& \odot \overline{F}(\tilde{l}_1^{(2n)} + \cdots + \tilde{l}_{j-1}^{(2n)} + f + \tilde{l}_{j+1}^{(2n)} + \cdots + \tilde{l}_{2n}^{(2n)}) \odot c_j \quad (\text{by Definition 3.6.11}) \\
= & J \iota_{\prod_{k=j+1}^{2n} F_k^{(2n)} 0}^{F_{j-1}^\dagger} \odot J(\beta_{j-1,0}^{(2n)})^{-1} \\
& \odot \overline{F}(\tilde{l}_1^{(2n)} + \cdots + \tilde{l}_{j-1}^{(2n)} + f + \tilde{l}_{j+1}^{(2n)} + \cdots + \tilde{l}_{2n}^{(2n)}) \odot c_j \quad (\text{by Definition 3.6.5}) \\
= & J(\beta_{j,0}^{(2n)})^{-1} \odot \overline{F}(\tilde{l}_1^{(2n)} + \cdots + \tilde{l}_{j-1}^{(2n)} + f + \tilde{l}_{j+1}^{(2n)} + \cdots + \tilde{l}_{2n}^{(2n)}) \odot c_j \\
& \quad (\text{by Definition 3.6.5}).
\end{aligned}$$

Hence  $\tilde{l}_j^{(2n)}$  is the greatest fixed point of the following function:

$$f \mapsto J(\beta_{j,0}^{(2n)})^{-1} \odot \overline{F}(\tilde{l}_1^{(2n)} + \cdots + \tilde{l}_{j-1}^{(2n)} + f + \tilde{l}_{j+1}^{(2n)} + \cdots + \tilde{l}_{2n}^{(2n)}) \odot c_j.$$

We can similarly prove the same statement when  $j$  is even. Hence  $(\tilde{l}_1^{(2n)}, \dots, \tilde{l}_{2n}^{(2n)})$  is the greatest fixed point of the function (3.16) in Definition 3.6.8, and this concludes the proof.  $\square$

We have generalized Lemma 3.5.6. We next generalize Lemma 3.5.7.

**Lemma 3.6.15.** *For each  $i \in \{1, \dots, 2n\}$ ,  $\text{tr}_i^p(c) = p_{i,0}^{(2n)} \circ \tilde{l}_i^{(2n)}$ .*

To prove this lemma, we first introduce a natural transformation. As mentioned in Remark 3.4.5,  $F^\oplus$  resembles a monad called the free completely iterative monad. The natural transformation is analogous to its multiplication.

**Definition 3.6.16** ( $\mu^{F^\oplus}$ ). We define a natural transformation  $\mu^{F^\oplus} : F^\oplus F^\oplus \Rightarrow F^\oplus$  by  $\mu^{F^\oplus} := (u_X \circ \kappa_1)_{X \in \mathbb{C}}$ , where  $u_X$  is the unique homomorphism from  $[F[\kappa_1, \kappa_2] \circ \zeta_{F^\oplus X}^F, F[\kappa_2, \kappa_3] \circ \zeta_X^F]$  to  $\zeta_X^F$ .

$$\begin{array}{ccc}
F(F^\oplus F^\oplus X + F^\oplus X + X) & \xrightarrow{F(u_X + \text{id}_X)} & F(F^\oplus X + X) \\
\uparrow [F[\kappa_1, \kappa_2], F[\kappa_2, \kappa_3]] & & \uparrow \text{final} \\
F(F^\oplus F^\oplus X + F^\oplus X) + F(F^\oplus X + X) & & \zeta_X^F \\
\cong \uparrow \zeta_{F^\oplus X}^F + \zeta_X^F & & \uparrow \\
F^\oplus F^\oplus X + F^\oplus X & \xrightarrow{u_X} & F^\oplus X
\end{array}$$

**Example 3.6.17.** For  $F_A = A \times (\_)$ , according to the characterizations in Example 3.4.3,  $\mu_X^{F^\oplus}$  has a type  $(A^+)^+((A^+)^+ X + (A^+)^\omega) + (A^+)^\omega \rightarrow (A^+)^+ X + (A^+)^\omega$ , and is given by the concatenating function that preserves each finite word.

We next prove two sublemmas (Sublemma 3.6.18 and 3.6.25). The following one connects the intermediate solution of the HES in Definition 3.6.3 and  $\tilde{l}_j^{(i)}$  in Definition 3.6.11.

**Sublemma 3.6.18.** *For each  $i \in \{1, \dots, 2n\}$  and  $j \in \{1, \dots, i\}$ , let*

$$l_j^{(i)} : \mathcal{Kl}(T)(X_{i+1}, F^\oplus 0) \times \cdots \times \mathcal{Kl}(T)(X_{2n}, F^\oplus 0) \rightarrow \mathcal{Kl}(T)(X_i, F^\oplus 0)$$



be the intermediate solution of the HES in Definition 3.6.3 (note that  $\nu F$  in Definition 3.6.3 is  $F^\oplus 0$ ). Then for  $(u_k : X_k \rightarrow F^\oplus 0)_{k \in \{i+1, \dots, 2n\}}$ ,

$$l_j^{(i)}(u_{i+1}, \dots, u_{2n}) = J\mu_0^{F^\oplus} \circ Jp_{j, F^\oplus 0}^{(i)} \circ \overline{F_j^{(i)}}[u_{i+1}, \dots, u_{2n}] \circ \tilde{l}_j^{(i)}. \quad (3.18)$$

The proof of the above sublemma is very long. We have to prove six sublemmas (see (3.17)).

**Sublemma 3.6.19.** For  $A \in \mathbb{C}$ , the unique homomorphism from  $[F[\kappa_1, \kappa_2] \circ \zeta_{F^\oplus A}^F, F[\kappa_2, \kappa_3] \circ \zeta_A^F]$  to  $\zeta_A^F$  is given by  $[\mu_A^{F^\oplus}, \text{id}_A]$ .

**Proof.** Let  $u : F^\oplus F^\oplus A + F^\oplus A \rightarrow F^\oplus A$  be the unique homomorphism from  $[F[\kappa_1, \kappa_2] \circ \zeta_{F^\oplus A}^F, F[\kappa_2, \kappa_3] \circ \zeta_A^F]$  to  $\zeta_A^F$ .

Note that  $u = [u \circ \kappa_1, u \circ \kappa_2]$ . By Definition 3.6.16,  $u \circ \kappa_1 = \mu_A^{F^\oplus}$ . It remains to prove  $u \circ \kappa_2 = \text{id}_A$ .

It is easy to see that  $\kappa_2 : F^\oplus A \rightarrow F^\oplus F^\oplus A + F^\oplus A$  is a homomorphism from  $\zeta_A^F$  to  $[F[\kappa_1, \kappa_2] \circ \zeta_{F^\oplus A}^F, F[\kappa_2, \kappa_3] \circ \zeta_A^F]$ . Therefore  $u \circ \kappa_2$  is a homomorphism from  $\zeta_A^F$  to itself, on the one hand. On the other hand,  $\text{id}_A$  is also a homomorphism from  $\zeta_A^F$ . Hence by the finality of  $\zeta_A^F$ , we have  $u \circ \kappa_2 = \text{id}_{F^\oplus A}$ .

$$\begin{array}{ccccc} F(F^\oplus A + A) & \xrightarrow{F(\kappa_2 + \text{id}_A)} & F(F^\oplus F^\oplus A + F^\oplus A + A) & \xrightarrow{F(u + \text{id}_A)} & F(F^\oplus A + A) \\ \cong \uparrow \zeta_A^F & & \uparrow [F[\kappa_1, \kappa_2], F[\kappa_2, \kappa_3]] & & \uparrow \text{final} \zeta_A^F \\ & & F(F^\oplus F^\oplus A + F^\oplus A) + F(F^\oplus A + A) & & \\ & & \cong \uparrow \zeta_{F^\oplus A}^F + \zeta_A^F & & \\ F^\oplus A & \xrightarrow{\kappa_2} & F^\oplus F^\oplus A + F^\oplus A & \xrightarrow{u} & F^\oplus A \quad \square \end{array}$$

**Sublemma 3.6.20.** We define an  $F$ -coalgebra  $\gamma_i : \coprod_{j=1}^{i-1} F_j^{(i)} F^\oplus 0 + F^\oplus 0 \rightarrow F(\coprod_{j=1}^{i-1} F_j^{(i)} F^\oplus 0 + F^\oplus 0)$  as follows:

$$\gamma_i := \left( \begin{array}{c} \coprod_{j=1}^{i-1} F_j^{(i)} F^\oplus 0 + F^\oplus 0 \xrightarrow{[\beta_{1, F^\oplus 0}^{(i)}, \dots, \beta_{i-1, F^\oplus 0}^{(i)}, F\kappa_{i+1} \circ \zeta_0^F]} F(\coprod_{j=1}^i F_j^{(i)} F^\oplus 0 + F^\oplus 0) \\ \xrightarrow{F(\text{id}_{\coprod_{j=1}^i F_j^{(i)} F^\oplus 0} + [\mu_0^{F^\oplus} \circ p_i^{(i)}, \text{id}_{F^\oplus 0}])} F(\coprod_{j=1}^{i-1} F_j^{(i)} F^\oplus 0 + F^\oplus 0) \end{array} \right).$$

Then the unique homomorphism from  $\gamma_i$  to a final coalgebra  $\zeta_0^F : F^\oplus 0 \rightarrow FF^\oplus 0$  is given by the following arrow:

$$[\mu_0^{F^\oplus} \circ p_{1, F^\oplus 0}^{(i)}, \dots, \mu_0^{F^\oplus} \circ p_{i-1, F^\oplus 0}^{(i)}, \text{id}_{F^\oplus 0}] : \coprod_{j=1}^{i-1} F_j^{(i)} 0 + F^\oplus 0 \rightarrow F^\oplus 0.$$

**Proof.** It suffices to show that it is a homomorphism. We have:

$$\begin{aligned} & \zeta_0^F \circ [\mu_0^{F^\oplus} \circ p_{1, F^\oplus 0}^{(i)}, \dots, \mu_0^{F^\oplus} \circ p_{i-1, F^\oplus 0}^{(i)}, \text{id}_{F^\oplus 0}] \\ &= \zeta_0^F \circ [\mu_0^{F^\oplus}, \text{id}_{F^\oplus 0}] \circ ([p_{1, F^\oplus 0}^{(i)}, \dots, p_{i-1, F^\oplus 0}^{(i)}] + \text{id}_{F^\oplus 0}) \\ &= F[\mu_0^{F^\oplus}, \text{id}_{F^\oplus 0}] \circ [\zeta_{F^\oplus 0}^F, F\kappa_2 \circ \zeta_0^F] \circ ([p_{1, F^\oplus 0}^{(i)}, \dots, p_{i-1, F^\oplus 0}^{(i)}] + \text{id}_{F^\oplus 0}) \\ & \hspace{15em} \text{(by Sublemma 3.6.19)} \\ &= F[\mu_0^{F^\oplus}, \text{id}_{F^\oplus 0}] \circ [F([p_{1, F^\oplus 0}^{(i)}, \dots, p_{i-1, F^\oplus 0}^{(i)}] + \text{id}_{F^\oplus 0}) \circ [\beta_{1, F^\oplus 0}^{(i)}, \dots, \beta_{i-1, F^\oplus 0}^{(i)}], \\ & \hspace{15em} F\kappa_2 \circ \zeta_0^F] \hspace{15em} \text{(by Definition 3.6.9)} \\ &= F[\mu_0^{F^\oplus} \circ p_{1, F^\oplus 0}^{(i)}, \dots, \mu_0^{F^\oplus} \circ p_{i-1, F^\oplus 0}^{(i)}, \text{id}_{F^\oplus 0}] \circ [\beta_{1, F^\oplus 0}^{(i)}, \dots, \beta_{i-1, F^\oplus 0}^{(i)}, F\kappa_i \circ \zeta_0^F] \\ &= F[\mu_0^{F^\oplus} \circ p_{1, F^\oplus 0}^{(i)}, \dots, \mu_0^{F^\oplus} \circ p_{i-1, F^\oplus 0}^{(i)}, \text{id}_{F^\oplus 0}] \end{aligned}$$

$$\begin{aligned} & \circ (F(\text{id}_{\prod_{j=1}^i F_j^{(i)} F^{\oplus 0}} + [\mu_0^{F^{\oplus}} \circ p_i^{(i)}, \text{id}_{F^{\oplus 0}}]) \\ & \quad \circ [\beta_{1, F^{\oplus 0}}^{(i)}, \dots, \beta_{i-1, F^{\oplus 0}}^{(i)}, F\kappa_{i+1} \circ \zeta_0^F]). \end{aligned}$$

This concludes the proof.  $\square$

**Sublemma 3.6.21.** *For each  $j \in \{1, \dots, i-1\}$ , we have the following equality.*

$$\mu_0^{F^{\oplus}} \circ p_{j, F^{\oplus 0}}^{(i)} = \mu_0^{F^{\oplus}} \circ p_{j, F^{\oplus 0}}^{(i-1)} \circ F_j^{(i-1)} [\mu_0^{F^{\oplus}} \circ p_{i, F^{\oplus 0}}^{(i)}, \text{id}_{F^{\oplus 0}}].$$

**Proof.** By Sublemma 3.6.20, it suffices to show that the following arrow is a homomorphism from  $\gamma_i$  to  $\zeta_0^F$ .

$$\begin{aligned} & [\mu_0^{F^{\oplus}} \circ p_{1, F^{\oplus 0}}^{(i-1)} \circ F_1^{(i-1)} [\mu_0^{F^{\oplus}} \circ p_{i, F^{\oplus 0}}^{(i)}, \text{id}_{F^{\oplus 0}}], \dots, \\ & \quad \mu_0^{F^{\oplus}} \circ p_{i-1, F^{\oplus 0}}^{(i-1)} \circ F_{i-1}^{(i-1)} [\mu_0^{F^{\oplus}} \circ p_{i, F^{\oplus 0}}^{(i)}, \text{id}_{F^{\oplus 0}}], \text{id}_{F^{\oplus 0}}] \end{aligned}$$

We have:

$$\begin{aligned} & \zeta_0^F \circ [\mu_0^{F^{\oplus}} \circ p_{1, F^{\oplus 0}}^{(i-1)} \circ F_1^{(i-1)} [\mu_0^{F^{\oplus}} \circ p_{i, F^{\oplus 0}}^{(i)}, \text{id}_{F^{\oplus 0}}], \dots, \\ & \quad \mu_0^{F^{\oplus}} \circ p_{i-1, F^{\oplus 0}}^{(i-1)} \circ F_{i-1}^{(i-1)} [\mu_0^{F^{\oplus}} \circ p_{i, F^{\oplus 0}}^{(i)}, \text{id}_{F^{\oplus 0}}], \text{id}_{F^{\oplus 0}}] \\ & = \zeta_0^F \circ [\mu_0^{F^{\oplus}}, \text{id}_{F^{\oplus 0}}] \circ \left( [p_{1, F^{\oplus 0}}^{(i-1)} \circ F_1^{(i-1)} [\mu_0^{F^{\oplus}} \circ p_{i, F^{\oplus 0}}^{(i)}, \text{id}_{F^{\oplus 0}}], \dots, \right. \\ & \quad \left. p_{i-1, F^{\oplus 0}}^{(i-1)} \circ F_{i-1}^{(i-1)} [\mu_0^{F^{\oplus}} \circ p_{i, F^{\oplus 0}}^{(i)}, \text{id}_{F^{\oplus 0}}]] + \text{id}_{F^{\oplus 0}} \right) \\ & = F[\mu_0^{F^{\oplus}}, \text{id}_{F^{\oplus 0}}] \circ [\zeta_{F^{\oplus 0}}^F, F\kappa_2 \circ \zeta_0^F] \\ & \quad \circ \left( [p_{1, F^{\oplus 0}}^{(i-1)} \circ F_1^{(i-1)} [\mu_0^{F^{\oplus}} \circ p_{i, F^{\oplus 0}}^{(i)}, \text{id}_{F^{\oplus 0}}], \dots, \right. \\ & \quad \left. p_{i-1, F^{\oplus 0}}^{(i-1)} \circ F_{i-1}^{(i-1)} [\mu_0^{F^{\oplus}} \circ p_{i, F^{\oplus 0}}^{(i)}, \text{id}_{F^{\oplus 0}}]] + \text{id}_{F^{\oplus 0}} \right) \quad (\text{by Sublemma 3.6.19}) \\ & = F[\mu_0^{F^{\oplus}}, \text{id}_{F^{\oplus 0}}] \\ & \quad \circ \left[ \zeta_{F^{\oplus 0}}^F \circ [p_{1, A}^{(i-1)}, \dots, p_{i-1, A}^{(i-1)}] \right. \\ & \quad \left. \circ (F_1^{(i-1)} [\mu_{F^{\oplus 0}}^{F^{\oplus}} \circ p_{i, F^{\oplus 0}}^{(i)}, \text{id}_{F^{\oplus 0}}] + \dots + F_{i-1}^{(i-1)} [\mu_{F^{\oplus 0}}^{F^{\oplus}} \circ p_{i, F^{\oplus 0}}^{(i)}, \text{id}_{F^{\oplus 0}}]), \right. \\ & \quad \left. F\kappa_2 \circ \zeta_0^F \right] \\ & = F[\mu_0^{F^{\oplus}}, \text{id}_{F^{\oplus 0}}] \\ & \quad \circ \left[ F([p_{1, F^{\oplus 0}}^{(i-1)}, \dots, p_{i-1, F^{\oplus 0}}^{(i-1)}] + \text{id}_{F^{\oplus 0}}) \circ [\beta_{1, F^{\oplus 0}}^{(i-1)}, \dots, \beta_{i-1, F^{\oplus 0}}^{(i-1)}] \right. \\ & \quad \left. \circ (F_1^{(i-1)} [\mu_{F^{\oplus 0}}^{F^{\oplus}} \circ p_{i, F^{\oplus 0}}^{(i)}, \text{id}_{F^{\oplus 0}}] + \dots + F_{i-1}^{(i-1)} [\mu_{F^{\oplus 0}}^{F^{\oplus}} \circ p_{i, F^{\oplus 0}}^{(i)}, \text{id}_{F^{\oplus 0}}]), \right. \\ & \quad \left. F\kappa_2 \circ \zeta_0^F \right] \quad (\text{by Definition 3.6.9}) \\ & = F[\mu_0^{F^{\oplus}}, \text{id}_{F^{\oplus 0}}] \\ & \quad \circ \left[ F([p_{1, F^{\oplus 0}}^{(i-1)}, \dots, p_{i-1, F^{\oplus 0}}^{(i-1)}] + \text{id}_{F^{\oplus 0}}) \right. \\ & \quad \left. \circ F(\prod_{j=1}^{i-1} F_j^{(i-1)} [\mu_{F^{\oplus 0}}^{F^{\oplus}} \circ p_{i, F^{\oplus 0}}^{(i)}, \text{id}_{F^{\oplus 0}}] + [\mu_{F^{\oplus 0}}^{F^{\oplus}} \circ p_{i, F^{\oplus 0}}^{(i)}, \text{id}_{F^{\oplus 0}}]) \right. \\ & \quad \left. \circ [\beta_{1, F_i^{(i)} F^{\oplus 0} + F^{\oplus 0}}^{(i-1)}, \dots, \beta_{i-1, F_i^{(i)} F^{\oplus 0} + F^{\oplus 0}}^{(i-1)}], \right. \\ & \quad \left. F\kappa_2 \circ \zeta_0^F \right] \quad (\text{by naturality}) \\ & = F[\mu_0^{F^{\oplus}}, \text{id}_{F^{\oplus 0}}] \\ & \quad \circ \left[ F([p_{1, F^{\oplus 0}}^{(i-1)}, \dots, p_{i-1, F^{\oplus 0}}^{(i-1)}] + \text{id}_{F^{\oplus 0}}) \right. \end{aligned}$$

$$\begin{aligned}
& \circ F\left(\prod_{j=1}^{i-1} F_j^{(i-1)} [\mu_{F^\oplus 0}^{F^\oplus} \circ p_{i, F^\oplus 0}^{(i)}, \text{id}_{F^\oplus 0}] + [\mu_{F^\oplus 0}^{F^\oplus} \circ p_{i, F^\oplus 0}^{(i)}, \text{id}_{F^\oplus 0}]\right) \\
& \circ [\beta_{1, F^\oplus 0}^{(i)}, \dots, \beta_{i-1, F^\oplus 0}^{(i)}], \\
& F\kappa_2 \circ \zeta_0^F \Big] \quad \text{(by Definition 3.6.5)} \\
= & F[\mu_0^{F^\oplus}, \text{id}_{F^\oplus 0}] \circ F([p_{1, F^\oplus 0}^{(i-1)}, \dots, p_{i-1, F^\oplus 0}^{(i-1)}] + \text{id}_{F^\oplus 0}) \\
& \circ F\left(\prod_{j=1}^{i-1} F_j^{(i-1)} [\mu_{F^\oplus 0}^{F^\oplus} \circ p_{i, F^\oplus 0}^{(i)}, \text{id}_{F^\oplus 0}] + [\mu_{F^\oplus 0}^{F^\oplus} \circ p_{i, F^\oplus 0}^{(i)}, \text{id}_{F^\oplus 0}]\right) \\
& \circ [\beta_{1, F^\oplus 0}^{(i)}, \dots, \beta_{i-1, F^\oplus 0}^{(i)}, F\kappa_{i+1} \circ \zeta_0^F] \\
= & F[\mu_{F^\oplus 0}^{F^\oplus} \circ p_{1, F^\oplus 0}^{(i-1)} \circ F_1^{(i-1)} [\mu_A^{F^\oplus} \circ p_{i, F^\oplus 0}^{(i)}, \text{id}_{F^\oplus 0}], \dots, \\
& \mu_{F^\oplus 0}^{F^\oplus} \circ p_{i-1, F^\oplus 0}^{(i-1)} \circ F_{i-1}^{(i-1)} [\mu_{F^\oplus 0}^{F^\oplus} \circ p_{i, F^\oplus 0}^{(i)}, \text{id}_{F^\oplus 0}]] \\
& \circ (F(\text{id}_{F_1^{(i)} F^\oplus 0} + \dots + \text{id}_{F_{i-1}^{(i)} F^\oplus 0} + [\mu_0^{F^\oplus} \circ p_i^{(i)}, \text{id}_{F^\oplus 0}]) \\
& \circ [\beta_{1, F^\oplus 0}^{(i)}, \dots, \beta_{i-1, F^\oplus 0}^{(i)}, F\kappa_{i+1} \circ \zeta_0^F]).
\end{aligned}$$

This concludes the proof.  $\square$

**Sublemma 3.6.22.** *The unique homomorphism from an  $F$ -coalgebra*

$$[\beta_{1, F^\oplus 0}^{(i)}, \dots, \beta_{i, F^\oplus 0}^{(i)}, F\kappa_{i+1} \circ \zeta_0^F] : \prod_{j=1}^i F_j^{(i)} F^\oplus 0 + F^\oplus 0 \rightarrow F\left(\prod_{j=1}^i F_j^{(i)} F^\oplus 0 + F^\oplus 0\right)$$

to a final coalgebra  $\zeta_0^F : F^\oplus 0 \rightarrow FF^\oplus 0$  is given by the following arrow:

$$[\mu_0^{F^\oplus} \circ p_{1, F^\oplus 0}^{(i)}, \dots, \mu_0^{F^\oplus} \circ p_{i, F^\oplus 0}^{(i)}, \text{id}_{F^\oplus 0}] : \prod_{j=1}^i F_j^{(i)} 0 + F^\oplus 0 \rightarrow F^\oplus 0.$$

**Proof.** It suffices to show that the arrow is a homomorphism. We have:

$$\begin{aligned}
& \zeta_0^F \circ [\mu_0^{F^\oplus} \circ p_{1, F^\oplus 0}^{(i)}, \dots, \mu_0^{F^\oplus} \circ p_{i, F^\oplus 0}^{(i)}, \text{id}_{F^\oplus 0}] \\
= & \zeta_0^F \circ [\mu_0^{F^\oplus}, \text{id}_{F^\oplus 0}] \circ ([p_{1, F^\oplus 0}^{(i)}, \dots, p_{i, F^\oplus 0}^{(i)}] + \text{id}_{F^\oplus 0}) \\
= & F[\mu_0^{F^\oplus}, \text{id}_{F^\oplus 0}] \circ [\zeta_{F^\oplus 0}^F, F\kappa_2 \circ \zeta_0^F] \circ ([p_{1, F^\oplus 0}^{(i)}, \dots, p_{i, F^\oplus 0}^{(i)}] + \text{id}_{F^\oplus 0}) \\
& \quad \text{(by Sublemma 3.6.19)} \\
= & F[\mu_0^{F^\oplus}, \text{id}_{F^\oplus 0}] \circ F([p_{1, F^\oplus 0}^{(i)}, \dots, p_{i, F^\oplus 0}^{(i)}] + \text{id}_{F^\oplus 0}) \circ [\beta_{1, F^\oplus 0}^{(i)}, \dots, \beta_{i, F^\oplus 0}^{(i)}], \\
& \quad F\kappa_2 \circ \zeta_0^F \Big] \quad \text{(by Definition 3.6.9)} \\
= & F[\mu_0^{F^\oplus} \circ p_{1, F^\oplus 0}^{(i)}, \dots, \mu_0^{F^\oplus} \circ p_{i, F^\oplus 0}^{(i)}, \text{id}_{F^\oplus 0}] \circ [\beta_{1, F^\oplus 0}^{(i)}, \dots, \beta_{i, F^\oplus 0}^{(i)}, F\kappa_{i+1} \circ \zeta_0^F].
\end{aligned}$$

This concludes the proof.  $\square$

**Sublemma 3.6.23.** *Let  $i > 0$  and  $A \in \mathbb{C}$ . We define an  $F_{i-1}^\ddagger(\_ + F^\oplus 0)$ -algebra  $\sigma_i : F_{i-1}^\ddagger(F^\oplus 0 + F^\oplus 0) \rightarrow F^\oplus 0$  as follows:*

$$\sigma_i := \left( \begin{array}{c} F_{i-1}^\ddagger(F^\oplus 0 + F^\oplus 0) \xrightarrow{F_{i-1}^\ddagger[\text{id}_{F^\oplus 0}, \text{id}_{F^\oplus 0}]} F_{i-1}^\ddagger F^\oplus 0 \xrightarrow{\beta_{i-1, F^\oplus 0}^{(i-1)}} \\ F\left(\prod_{j=1}^{i-1} F_j^{(i-1)} F^\oplus 0 + F^\oplus 0\right) \xrightarrow{F[\mu_0^{F^\oplus} \circ p_{1, F^\oplus 0}^{(i-1)}, \dots, \mu_0^{F^\oplus} \circ p_{i-1, F^\oplus 0}^{(i-1)}, \text{id}_{F^\oplus 0}]} \\ FF^\oplus 0 \xrightarrow{(\zeta_0^F)^{-1}} F^\oplus 0 \end{array} \right).$$

Then if  $i$  is even (resp. odd),  $J\mu_0^{F^\oplus} \circ Jp_{i, F^\oplus 0}^{(i)} : F_i^\ddagger F^\oplus 0 \rightarrow F^\oplus 0$  is the greatest fixed point of  $\Phi_{J\zeta_{F^\oplus 0}^{F^\oplus}, J\sigma_i}^{F_{i-1}^\ddagger}$  (resp.  $\Phi_{J(\zeta_{F^\oplus 0}^{F^\oplus})^{-1}, J\sigma_i}^{F_{i-1}^\ddagger}$ ), see Definition 2.4.22).

**Proof.** Assume that  $i$  is even. We write  $\Phi$  for  $\Phi_{J\zeta_{F^{\oplus 0}}^{F^{\ddagger}_{i-1}}, J\sigma_i}$  for simplicity. For  $f : F_i^{\ddagger}F^{\oplus 0} \rightarrow F^{\oplus 0}$ , we have:

$$\begin{aligned}
& \Phi(f) \\
&= J\sigma_i \odot \overline{F_{i-1}^{\ddagger}}(f + \text{id}_{F^{\oplus 0}}) \odot J\zeta_{F^{\oplus 0}}^{F^{\ddagger}_{i-1}} \quad (\text{by definition}) \\
&= J(\zeta_0^F)^{-1} \odot JF[\mu_0^{F^{\oplus}} \circ p_{1, F^{\oplus 0}}^{(i-1)}, \dots, \mu_0^{F^{\oplus}} \circ p_{i-1, F^{\oplus 0}}^{(i-1)}, \text{id}_{F^{\oplus 0}}] \odot J\beta_{i-1, F^{\oplus 0}}^{(i-1)} \\
&\quad \odot \overline{F_{i-1}^{\ddagger}}[\text{id}_{F^{\oplus 0}}, \text{id}_{F^{\oplus 0}}] \odot \overline{F_{i-1}^{\ddagger}}(f + \text{id}_{F^{\oplus 0}}) \odot J\zeta_{F^{\oplus 0}}^{F^{\ddagger}_{i-1}} \quad (\text{by definition}) \\
&= J(\zeta_0^F)^{-1} \odot JF[\mu_0^{F^{\oplus}} \circ p_{1, F^{\oplus 0}}^{(i-1)}, \dots, \mu_0^{F^{\oplus}} \circ p_{i-1, F^{\oplus 0}}^{(i-1)}, \text{id}_{F^{\oplus 0}}] \odot \overline{F_0^{(i-1)}}[f, \text{id}_{F^{\oplus 0}}] \\
&\quad \odot J\beta_{i-1, F_i^{\ddagger}F^{\oplus 0} + F^{\oplus 0}}^{(i-1)} \odot J\zeta_{F^{\oplus 0}}^{F^{\ddagger}_{i-1}} \quad (\text{by the naturality of } \beta_{i-1}^{(i-1)}) \\
&= J(\zeta_0^F)^{-1} \odot JF[\mu_0^{F^{\oplus}} \circ p_{1, F^{\oplus 0}}^{(i-1)}, \dots, \mu_0^{F^{\oplus}} \circ p_{i-1, F^{\oplus 0}}^{(i-1)}, \text{id}_{F^{\oplus 0}}] \odot \overline{F_0^{(i-1)}}[f, \text{id}_{F^{\oplus 0}}] \\
&\quad \odot J\beta_{i, F^{\oplus 0}}^{(i)} \quad (\text{by Definition 3.6.5}) \\
&= J(\zeta_0^F)^{-1} \odot \overline{F}[J\mu_0^{F^{\oplus}} \odot Jp_{1, F^{\oplus 0}}^{(i-1)} \odot \overline{F_1^{(i-1)}}[f, \text{id}_{F^{\oplus 0}}], \dots, \\
&\quad J\mu_0^{F^{\oplus}} \odot Jp_{i-1, F^{\oplus 0}}^{(i-1)} \odot \overline{F_{i-1}^{(i-1)}}[f, \text{id}_{F^{\oplus 0}}], f, \text{id}_{F^{\oplus 0}}] \odot J\beta_{i, F^{\oplus 0}}^{(i)} \\
&\quad (\text{by Definition 3.6.5}). \quad (3.19)
\end{aligned}$$

We now show that  $J\mu_0^{F^{\oplus}} \odot p_{i, F^{\oplus 0}}^{(i)}$  is a fixed point of  $\Phi$ . By the equation (3.19) above, we have:

$$\begin{aligned}
& \Phi(J\mu_0^{F^{\oplus}} \odot p_{i, F^{\oplus 0}}^{(i)}) \\
&= J(\zeta_0^F)^{-1} \odot \overline{F}[J\mu_0^{F^{\oplus}} \odot Jp_{1, F^{\oplus 0}}^{(i-1)} \odot \overline{F_1^{(i-1)}}[J\mu_0^{F^{\oplus}} \odot p_{i, F^{\oplus 0}}^{(i)}, \text{id}_{F^{\oplus 0}}], \dots, \\
&\quad J\mu_0^{F^{\oplus}} \odot Jp_{i-1, F^{\oplus 0}}^{(i-1)} \odot \overline{F_{i-1}^{(i-1)}}[J\mu_0^{F^{\oplus}} \odot p_{i, F^{\oplus 0}}^{(i)}, \text{id}_{F^{\oplus 0}}], p_{i, F^{\oplus 0}}^{(i)}, \text{id}_{F^{\oplus 0}}] \\
&\quad \odot J\beta_{i, F^{\oplus 0}}^{(i)} \\
&= J(\zeta_0^F)^{-1} \odot \overline{F}[J\mu_0^{F^{\oplus}} \odot Jp_{1, F^{\oplus 0}}^{(i)}, \dots, J\mu_0^{F^{\oplus}} \odot Jp_{1, F^{\oplus 0}}^{(i)}, J\mu_0^{F^{\oplus}} \odot p_{i, F^{\oplus 0}}^{(i)}, \text{id}_{F^{\oplus 0}}] \\
&\quad \odot J\beta_{i, F^{\oplus 0}}^{(i)} \quad (\text{by Sublemma 3.6.21}) \\
&= J\mu_0^{F^{\oplus}} \odot Jp_{j, F^{\oplus 0}}^{(i)} \quad (\text{by Sublemma 3.6.22}).
\end{aligned}$$

Hence  $J\mu_0^{F^{\oplus}} \odot Jp_{i, F^{\oplus 0}}^{(i)}$  is a fixed point of  $\Phi$ .

It remains to show that it is the greatest fixed point. Let  $f$  be a fixed point of  $\Phi$ . For each  $j \in \{1, \dots, i-1\}$ , we have:

$$\begin{aligned}
& J\zeta_0^F \odot (J\mu_0^{F^{\oplus}} \odot Jp_j^{(i-1)} \odot \overline{F_j^{(i-1)}}[f, \text{id}_{F^{\oplus 0}}]) \\
&= JF[\mu_0^{F^{\oplus}} \circ p_{1, F^{\oplus 0}}^{(i-1)}, \dots, \mu_0^{F^{\oplus}} \circ p_{j-1, F^{\oplus 0}}^{(i-1)}, \text{id}_{F^{\oplus 0}}] \odot J\beta_{j, F^{\oplus 0}}^{(i-1)} \odot \overline{F_j^{(i-1)}}[f, \text{id}_{F^{\oplus 0}}] \\
&\quad (\text{by Sublemma 3.6.22}) \\
&= JF[\mu_0^{F^{\oplus}} \circ p_{1, F^{\oplus 0}}^{(i-1)}, \dots, \mu_0^{F^{\oplus}} \circ p_{j-1, F^{\oplus 0}}^{(i-1)}, \text{id}_{F^{\oplus 0}}] \odot \overline{F_0^{(i-1)}}[f, \text{id}_{F^{\oplus 0}}] \\
&\quad \odot J\beta_{j, F_i^{\ddagger}F^{\oplus 0} + F^{\oplus 0}}^{(i-1)} \quad (\text{by the naturality of } \beta_j^{(i-1)}) \\
&= JF[\mu_0^{F^{\oplus}} \circ p_{1, F^{\oplus 0}}^{(i-1)}, \dots, \mu_0^{F^{\oplus}} \circ p_{i-1, F^{\oplus 0}}^{(i-1)}, \text{id}_{F^{\oplus 0}}] \odot \overline{F_0^{(i-1)}}[f, \text{id}_{F^{\oplus 0}}] \odot J\beta_{j, F^{\oplus 0}}^{(i)} \\
&\quad (\text{by Definition 3.6.5})
\end{aligned}$$

$$\begin{aligned}
&= \overline{F}[J\mu_0^{F^\oplus} \odot Jp_{1,F^\oplus}^{(i-1)} \odot \overline{F_1^{(i-1)}}[f, \text{id}_{F^\oplus}], \dots, \\
&\quad J\mu_0^{F^\oplus} \odot Jp_{i-1,F^\oplus}^{(i-1)} \odot \overline{F_{i-1}^{(i-1)}}[f, \text{id}_{F^\oplus}], f, \text{id}_{F^\oplus}] \odot J\beta_{j,F^\oplus}^{(i)} \\
&\hspace{15em} \text{(by Definition 3.6.5)}.
\end{aligned}$$

Therefore, together with the equation (3.19) proved above, we can see that the following arrow is a homomorphism from  $[\beta_{1,F^\oplus}^{(i)}, \dots, \beta_{i,F^\oplus}^{(i)}, F\kappa_{i+1} \circ \zeta_0^F]$  to  $\zeta_0^F$ .

$$\begin{aligned}
&[J\mu_0^{F^\oplus} \odot Jp_1^{(i-1)} \odot \overline{F_1^{(i-1)}}[f, \text{id}_{F^\oplus}], \dots, \\
&\quad J\mu_0^{F^\oplus} \odot Jp_{i-1}^{(i-1)} \odot \overline{F_{i-1}^{(i-1)}}[f, \text{id}_{F^\oplus}], f, \text{id}_{F^\oplus}]
\end{aligned}$$

Hence by Sublemma 3.6.22 and the deterministic-greatest condition (Assumption 3.6.7.5), we have:

$$\begin{aligned}
&[J\mu_0^{F^\oplus} \odot Jp_{1,F^\oplus}^{(i-1)} \odot \overline{F_1^{(i-1)}}[f, \text{id}_{F^\oplus}], \dots, \\
&\quad J\mu_0^{F^\oplus} \odot Jp_{i-1,F^\oplus}^{(i-1)} \odot \overline{F_{i-1}^{(i-1)}}[f, \text{id}_{F^\oplus}], f, \text{id}_{F^\oplus}] \\
&\quad \sqsubseteq J[\mu_0^{F^\oplus} \circ p_{1,F^\oplus}^{(i)}, \dots, \mu_0^{F^\oplus} \circ p_{i,F^\oplus}^{(i)}, \text{id}_{F^\oplus}].
\end{aligned}$$

This immediately implies  $f \sqsubseteq J\mu_0^{F^\oplus} \odot Jp_{i,F^\oplus}^{(i)}$ . Hence  $J\mu_0^{F^\oplus} \odot Jp_{i,F^\oplus}^{(i)}$  is the greatest fixed point of  $\Phi$ .

The proof when  $i$  is odd is similar.  $\square$

**Sublemma 3.6.24.** Let  $T$  be a monad and  $F$  be an endofunctor on  $\mathbb{C}$ . Assume that a lifting  $\overline{F} : \mathcal{Kl}(T) \rightarrow \mathcal{Kl}(T)$  of  $F$  is given and each homset of  $\mathcal{Kl}(T)$  carries a partial order  $\sqsubseteq$ . Assume further that they satisfy the conditions in Theorem 3.1.13. Let  $\iota^F : FA \rightarrow A$  be an initial  $F$ -algebra. For each  $X, Y \in \mathbb{C}$ ,  $c : X \rightarrow FX$  and  $\sigma : FY \rightarrow Y$ , if  $u : X \rightarrow A$  is the unique homomorphism from  $c$  to  $J(\iota^F)^{-1}$  and a function  $\Phi_{J(\iota^F), \sigma}$  (see Definition 2.4.22) has a fixed point  $m : A \rightarrow Y$ , then  $m \odot u : X \rightarrow Y$  is the least fixed point of  $\Phi_{c, \sigma}$ .

$$\begin{array}{ccc}
FX & \xrightarrow{\overline{F}u} & FA & \xrightarrow{\overline{F}m} & FY \\
c \uparrow & & J(\iota^F)^{-1} \uparrow \cong & = & \downarrow \sigma \\
X & \xrightarrow{u} & A & \xrightarrow{m} & Y
\end{array}$$

**Proof.** It is easy to see that  $m \odot u$  is a fixed point of  $\Phi_{c, \sigma}$ . We shall show that it is the least fixed point. Let  $f : X \rightarrow Y$  be a fixed point of  $\Phi_{c, \sigma}$ .

By the conditions in Theorem 3.1.13, a homset  $\mathcal{Kl}(T)(X, A)$  is  $\omega$ -complete and has the least element  $\perp$ , and a function  $\Phi_{c, J\iota^F} : \mathcal{Kl}(T)(X, A) \rightarrow \mathcal{Kl}(T)(X, A)$  is monotone and  $\omega$ -continuous. Hence by the Kleene fixed point theorem (Theorem 2.2.3.2), the unique fixed point  $u : X \rightarrow A$  of  $\Phi_{c, J\iota^F}$  is given by  $\bigsqcup_{i \in \omega} \Phi_{c, J\iota^F}^i(\perp)$ .

We now prove  $m \odot \Phi_{c, J\iota^F}^i(\perp) \sqsubseteq f$  by the induction on  $i$ .

- If  $i = 0$ , by the conditions in Theorem 3.1.13, we have:

$$m \odot \Phi_{c, J\iota^F}^0(\perp) = m \odot \perp = \perp \sqsubseteq f.$$

- Assume that  $m \odot \Phi_{c, J\iota^F}^i(\perp) \sqsubseteq f$ . Then we have:

$$\begin{aligned}
&m \odot \Phi_{c, J\iota^F}^{i+1}(\perp) \\
&= m \odot J\iota^F \odot F(\Phi_{c, J\iota^F}^i(\perp)) \odot c \hspace{10em} \text{(by definition)}
\end{aligned}$$

$$\begin{aligned}
&= \sigma \odot Fm \odot J(\iota^F)^{-1} \odot J\iota^F \odot F(\Phi_{c, J\iota^F}^i(\perp)) \odot c \quad (m \text{ is a fixed point}) \\
&= \sigma \odot F(m \odot \Phi_{c, J\iota^F}^i(\perp)) \odot c \\
&\sqsubseteq \sigma \odot Ff \odot c \quad (\text{by IH}) \\
&= f \quad (f \text{ is a fixed point}) .
\end{aligned}$$

Hence we have  $m \odot \Phi_{c, J\iota^F}^i(\perp) \sqsubseteq f$  for each  $i \in \omega$ . Therefore we have:

$$m \odot u = m \odot \bigsqcup_{i \in \omega} \Phi_{c, J\iota^F}^i(\perp) = \bigsqcup_{i \in \omega} (m \odot \Phi_{c, J\iota^F}^i(\perp)) \sqsubseteq f .$$

Hence  $m \odot u$  is the least fixed point.  $\square$

**Proof** (Sublemma 3.6.18). We prove Equation (3.18) by the induction on  $i$ . We do not have to distinguish the base case from the step case.

We first prove Equation (3.18) for  $j = i$ . Assume that  $i$  is even. By the definition of intermediate solutions (Definition 2.3.6), it suffices to show that  $J\mu_0^{F^\oplus} \odot Jp_{i, F^\oplus}^{(i)} \odot \overline{F_i^{(i)}}[u_{i+1}, \dots, u_{2n}] \odot \tilde{l}_i^{(i)}$  is the greatest fixed point of the following function:

$$f \mapsto J(\zeta_0^F)^{-1} \odot \overline{F}[l_1^{(i-1)}(f, u_{i+1}, \dots, u_{2n}), \dots, l_{i-1}^{(i-1)}(f, u_{i+1}, \dots, u_{2n}), f, u_{i+1}, \dots, u_{2n}] \odot c_i .$$

Here the right-hand side can be deformed as follows:

$$\begin{aligned}
&J(\zeta_0^F)^{-1} \odot \overline{F}[l_1^{(i-1)}(f, u_{i+1}, \dots, u_{2n}), \dots, l_{i-1}^{(i-1)}(f, u_{i+1}, \dots, u_{2n}), \\
&\quad f, u_{i+1}, \dots, u_{2n}] \odot c_i \\
&= J(\zeta_0^F)^{-1} \odot \overline{F}[J\mu_0^{F^\oplus} \odot Jp_1^{(i-1)} \odot \overline{F_1^{(i-1)}}[f, u_{i+1}, \dots, u_{2n}] \odot \tilde{l}_1^{(i-1)}, \dots, \\
&\quad J\mu_0^{F^\oplus} \odot Jp_{i-1}^{(i-1)} \odot \overline{F_{i-1}^{(i-1)}}[f, u_{i+1}, \dots, u_{2n}] \odot \tilde{l}_{i-1}^{(i-1)}, \\
&\quad f, u_{i+1}, \dots, u_{2n}] \odot c_i \quad (\text{by IH}) \\
&= J(\zeta_0^F)^{-1} \odot \overline{F}[J\mu_0^{F^\oplus} \odot Jp_1^{(i-1)} \odot \overline{F_1^{(i-1)}}[f, u_{i+1}, \dots, u_{2n}], \dots, \\
&\quad J\mu_0^{F^\oplus} \odot Jp_{i-1}^{(i-1)} \odot \overline{F_{i-1}^{(i-1)}}[f, u_{i+1}, \dots, u_{2n}], \\
&\quad f, u_{i+1}, \dots, u_{2n}] \odot J\beta_{i-1, X_i + \dots + X_{2n}}^{(i-1)} \odot c_i^\ddagger \\
&\quad (\text{by the definition of } c_i^\ddagger) \\
&= J(\zeta_0^F)^{-1} \odot \overline{F}[Jp_1^{(i-1)}, \dots, Jp_{i-1}^{(i-1)}, \text{id}_{F^\oplus}] \\
&\quad \odot \overline{F}(\bigsqcup_{j=1}^{i-1} \overline{F_j^{(i-1)}}[f, u_{i+1}, \dots, u_{2n}] + [f, u_{i+1}, \dots, u_{2n}]) \\
&\quad \odot J\beta_{i-1, X_i + \dots + X_{2n}}^{(i-1)} \odot c_i^\ddagger \\
&= J(\zeta_0^F)^{-1} \odot \overline{F}[Jp_1^{(i-1)}, \dots, Jp_{i-1}^{(i-1)}, \text{id}_{F^\oplus}] \odot \overline{F_0^{(i-1)}}[f, u_{i+1}, \dots, u_{2n}] \\
&\quad \odot J\beta_{i-1, X_i + \dots + X_{2n}}^{(i-1)} \odot c_i^\ddagger \quad (\text{by the definition of } F_0^{(i-1)}) \\
&= J(\zeta_0^F)^{-1} \odot JF[\mu_0^{F^\oplus} \circ p_1^{(i-1)}, \dots, \mu_0^{F^\oplus} \circ p_{i-1}^{(i-1)}, \text{id}_{F^\oplus}] \odot J\beta_{i-1, F^\oplus}^{(i-1)} \\
&\quad \odot \overline{F_{i-1}^\ddagger}[f, u_{i+1}, \dots, u_{2n}] \odot c_i^\ddagger \quad (\text{by naturality of } \beta_{i-1}^{(i-1)}) \\
&= J(\zeta_0^F)^{-1} \odot JF[\mu_0^{F^\oplus} \circ p_1^{(i-1)}, \dots, \mu_0^{F^\oplus} \circ p_{i-1}^{(i-1)}, \text{id}_{F^\oplus}] \odot J\beta_{i-1, F^\oplus}^{(i-1)} \\
&\quad \odot JF_{i-1}^\ddagger[\text{id}_{F^\oplus}, \text{id}_{F^\oplus}] \odot \overline{F_{i-1}^\ddagger}(f + \text{id}_{F^\oplus}) \odot \overline{F_{i-1}^\ddagger}(\text{id}_{X_i} + [u_{i+1}, \dots, u_{2n}]) \odot c_i^\ddagger
\end{aligned}$$

$$= J\sigma_i \odot \overline{F_{i-1}^\ddagger}(f + \text{id}_{F^\oplus 0}) \odot \overline{F_{i-1}^\ddagger}(\text{id}_{X_i} + [u_{i+1}, \dots, u_{2n}]) \odot c_i^\ddagger$$

(by the definition of  $\sigma_i$ ).

Therefore we have to show that  $J\mu_0^{F^\oplus} \odot Jp_{i,F^\oplus 0}^{(i)} \odot \overline{F_i^{(i)}}[u_{i+1}, \dots, u_{2n}] \odot \tilde{l}_i^{(i)}$  is the greatest fixed point of  $\Phi_{\overline{F_{i-1}^\ddagger}(\text{id}_{X_i} + [u_{i+1}, \dots, u_{2n}]) \odot c_i^\ddagger, J\sigma_i}$  (see also Figure 3.2).

By definition,  $\tilde{l}_i^{(i)}$  is the greatest homomorphism from  $c_i^\ddagger$  to  $J\zeta_{X_{i+1} + \dots + X_{2n}}^F$ , and  $\overline{F_i^{(i)}}[u_{i+1}, \dots, u_{2n}]$  is the greatest homomorphism from  $\overline{F_{i-1}^\ddagger}(\text{id}_{X_{i+1} + \dots + X_{2n}} + [u_{i+1}, \dots, u_{2n}]) \odot J\zeta_{X_{i+1} + \dots + X_{2n}}^{F_{i-1}^\ddagger}$  to  $J\zeta_{F^\oplus 0}^{F_{i-1}^\ddagger}$ . Therefore by Lemma 3.4.12,  $\overline{F_i^{(i)}}[u_{i+1}, \dots, u_{2n}] \odot \tilde{l}_i^{(i)}$  is the greatest homomorphism from  $\overline{F_{i-1}^\ddagger}(\text{id}_{X_i} + [u_{i+1}, \dots, u_{2n}]) \odot c_i^\ddagger$  to  $J\zeta_{F^\oplus 0}^{F_{i-1}^\ddagger}$ .

By Sublemma 3.6.23,  $J\mu_0^{F^\oplus} \odot Jp_{i,F^\oplus 0}^{(i)}$  is the greatest fixed point of  $\Phi_{J\zeta_{F^\oplus 0}^{F_{i-1}^\ddagger}, J\sigma_i}$ .

Therefore by the gfp-preserving condition (Assumption 3.6.7.4),  $J\mu_0^{F^\oplus} \odot Jp_{i,F^\oplus 0}^{(i)} \odot \overline{F_i^{(i)}}[u_{i+1}, \dots, u_{2n}] \odot \tilde{l}_i^{(i)}$  is the greatest fixed point of  $\Phi_{\overline{F_{i-1}^\ddagger}(\text{id}_{X_i} + [u_{i+1}, \dots, u_{2n}]) \odot c_i^\ddagger, J\sigma_i}$ .

Equation (3.18) is similarly proved when  $i$  is odd, except that we use the finality instead of Lemma 3.4.12, and Sublemma 3.6.24 instead of the gfp-preserving condition respectively.

It remains to prove Equation (3.18) for  $j < i$ . We have:

$$\begin{aligned} & l_j^{(i)}(u_{i+1}, \dots, u_{2n}) \\ &= l_j^{(i-1)}(l_i^{(i)}(u_{i+1}, \dots, u_{2n}), u_{i+1}, \dots, u_{2n}) && \text{(by definition)} \\ &= l_j^{(i-1)}(J\mu_0^{F^\oplus} \odot Jp_{i,F^\oplus 0}^{(i)} \odot \overline{F_i^{(i)}}[u_{i+1}, \dots, u_{2n}] \odot \tilde{l}_i^{(i)}, u_{i+1}, \dots, u_{2n}) \\ & && \text{(by the discussion above)} \\ &= J\mu_0^{F^\oplus} \odot Jp_{j,F^\oplus 0}^{(i-1)} \\ & \quad \odot \overline{F_j^{(i-1)}}[J\mu_0^{F^\oplus} \odot Jp_{i,F^\oplus 0}^{(i)} \odot \overline{F_i^{(i)}}[u_{i+1}, \dots, u_{2n}] \odot \tilde{l}_i^{(i)}, u_{i+1}, \dots, u_{2n}] \odot \tilde{l}_j^{(i-1)} \\ & && \text{(by IH)} \\ &= J\mu_0^{F^\oplus} \odot Jp_{j,F^\oplus 0}^{(i-1)} \\ & \quad \odot \overline{F_j^{(i-1)}}[J\mu_0^{F^\oplus} \odot Jp_{i,F^\oplus 0}^{(i)} \odot \overline{F_i^{(i)}}[u_{i+1}, \dots, u_{2n}], u_{i+1}, \dots, u_{2n}] \odot \tilde{l}_j^{(i)} \\ & && \text{(by Definition 3.6.11)} \\ &= J\mu_0^{F^\oplus} \odot Jp_{j,F^\oplus 0}^{(i-1)} \odot \overline{F_j^{(i-1)}}[J\mu_0^{F^\oplus} \odot Jp_{i,F^\oplus 0}^{(i)}, \text{id}_{F^\oplus 0}] \odot \overline{F_j^{(i)}}[u_{i+1}, \dots, u_{2n}] \odot \tilde{l}_j^{(i)} \\ & && \text{(by Lemma 3.6.12)} \\ &= J\mu_0^{F^\oplus} \odot Jp_{j,F^\oplus 0}^{(i)} \odot \overline{F_j^{(i)}}[u_{i+1}, \dots, u_{2n}] \odot \tilde{l}_j^{(i)} && \text{(by Sublemma 3.6.21).} \end{aligned}$$

This concludes the proof.  $\square$

The following is the second sublemma for proving Lemma 3.6.15. It is about a property on  $p_j^{(i)} : F_j^{(i)} \Rightarrow F^\oplus$ .

**Sublemma 3.6.25.** *For each  $j \in \{1, \dots, i\}$ ,  $\mu_0^{F^\oplus} \circ p_{j,F^\oplus 0}^{(i)} \circ F_j^{(i)} \mathbf{i}_{F^\oplus 0} = p_{j,0}^{(i)}$ .*

**Proof.** By definition, it suffices to prove that the following arrow is a homomorphism from  $[\beta_{1,0}^{(i)}, \dots, \beta_{i,0}^{(i)}]$  to  $\zeta_0^F$ .

$$[\mu_0^{F^\oplus} \circ p_{1,F^\oplus 0}^{(i)} \circ F_1^{(i)} \mathbf{i}_{F^\oplus 0}, \dots, \mu_0^{F^\oplus} \circ p_{i,F^\oplus 0}^{(i)} \circ F_i^{(i)} \mathbf{i}_{F^\oplus 0}] : \prod_{j=1}^i F_j^{(i)} \mathbf{0} \rightarrow F^\oplus \mathbf{0}$$





We have:

$$\begin{aligned}
& \zeta_0^F \circ [\mu_0^{F^\oplus} \circ p_{1,F^\oplus}^{(i)} \circ F_1^{(i)} i_{F^\oplus 0}, \dots, \mu_0^{F^\oplus} \circ p_{i,F^\oplus}^{(i)} \circ F_i^{(i)} i_{F^\oplus 0}] \\
&= F[\mu_0^{F^\oplus}, \text{id}_{F^\oplus 0}] \circ \zeta_{F^\oplus 0}^F \circ [p_{1,F^\oplus 0}^{(i)} \circ F_1^{(i)} i_{F^\oplus 0}, \dots, p_{i,F^\oplus 0}^{(i)} \circ F_i^{(i)} i_{F^\oplus 0}] \\
&\hspace{20em} \text{(by Definition 3.6.16)} \\
&= F[\mu_0^{F^\oplus}, \text{id}_{F^\oplus 0}] \circ F([p_{1,F^\oplus 0}^{(i)}, \dots, p_{i,F^\oplus 0}^{(i)}] + \text{id}_{F^\oplus 0}) \\
&\quad \circ [\beta_{1,F^\oplus 0}^{(i)} \circ F_1^{(i)} i_{F^\oplus 0}, \dots, \beta_{i,F^\oplus 0}^{(i)} \circ F_i^{(i)} i_{F^\oplus 0}] \hspace{5em} \text{(by Definition 3.6.9)} \\
&= F[\mu_0^{F^\oplus}, \text{id}_{F^\oplus 0}] \circ F([p_{1,F^\oplus 0}^{(i)}, \dots, p_{i,F^\oplus 0}^{(i)}] + \text{id}_{F^\oplus 0}) \circ F_0^{(i)} i_{F^\oplus 0} \circ [\beta_{1,0}^{(i)}, \dots, \beta_{i,0}^{(i)}] \\
&\hspace{20em} \text{(by naturality)} \\
&= F\left([\mu_0^{F^\oplus} \circ p_{1,F^\oplus 0}^{(i)} \circ F_1^{(i)} i_{F^\oplus 0}, \dots, \mu_0^{F^\oplus} \circ p_{i,F^\oplus 0}^{(i)} \circ F_i^{(i)} i_{F^\oplus 0}]\right) \circ [\beta_{1,0}^{(i)}, \dots, \beta_{i,0}^{(i)}].
\end{aligned}$$

This concludes the proof.  $\square$

**Proof** (Lemma 3.6.15). Immediate by Sublemma 3.6.18 and 3.6.25  $\square$

**Proof** (Theorem 3.6.10). Immediate by Lemma 3.6.14 and Lemma 3.6.15.  $\square$

We note that Lemma 3.6.14 implies the existence of a solution of the HES in Definition 3.6.8.

### 3.7 Extension to Nondeterministic Parity Tree Automata

We used nondeterministic Büchi and parity *word* automata as a running example in the previous sections. In this section, we apply our framework to nondeterministic parity *tree* automata (NPTA, Definition 2.2.4). Categorically, this means that we extend  $F$  from  $F_A = A \times (-)$  to  $F_\Sigma = \coprod_{i \in \omega} \Sigma_i \times (-)^i$ .

#### 3.7.1 Trace Semantics of NPTA via Logical Fixed Point

A coalgebraic parity trace semantics  $\text{tr}_i^{\mathcal{P}}(c)$  (Section 3.6.2) characterizes  $L_A^{\mathcal{P}}$ .

**Proposition 3.7.1.** *We define a partial order  $\sqsubseteq$  for each homset of  $\mathcal{Kl}(\mathcal{P})$  as in Example 3.1.11, and define a lifting  $\overline{F}_\Sigma : \mathcal{Kl}(\mathcal{P}) \rightarrow \mathcal{Kl}(\mathcal{P})$  of  $F_\Sigma$  as in Example 3.1.6. Then we have:*

1.  $\mathcal{P}$  and  $F_\Sigma$  constitute a parity trace situation (Definition 3.6.3).
2. The carrier set of a final  $F_\Sigma$ -coalgebra is isomorphic to  $\text{Tree}_\Sigma^\infty$ .
3. For a  $\Sigma$ -labeled NPTA  $\mathcal{A} = (X, \tau, \mathbf{p})$ , we define a parity  $(\mathcal{P}, F_\Sigma)$ -system  $(X, c, (X_1, \dots, X_{2n}))$  by  $c := \tau$  and  $X_i := \{x \mid \mathbf{p}(x) = i\}$ . Then we have:  $[\text{tr}_1^{\mathcal{P}}(c), \dots, \text{tr}_{2n}^{\mathcal{P}}(c)] = L_A^{\mathcal{P}}$ .  $\square$

Item 1 is immediate as each homset of  $\mathcal{Kl}(\mathcal{P})$  constitutes a complete lattice. Item 2 is fundamental (see Example 2.4.18). Item 3 will be proved later, using Theorem 3.6.10 (see page 61).

#### 3.7.2 Trace Semantics of NPTA via Categorical Fixed Point

We next focus on  $\text{dtr}_i(c)$ . We first describe a datatype  $(F_\Sigma)_j^{(i)} A$  and an arrow  $\beta_{j,A}^{(i)} : (F_\Sigma)_j^{(i)} A \rightarrow (F_\Sigma)_0^{(i)} A$  concretely.

To explain the intuition, we first focus on *Büchi* tree automata and describe  $(F_\Sigma)_2^\dagger 0 = F_\Sigma^{+\oplus} 0$  referring to the construction of a final coalgebra in Theorem 2.4.17. We can easily see that  $F_\Sigma^+ A \cong \text{Tree}_\Sigma^+(A)$  where  $\text{Tree}_\Sigma^+(A) := \text{Tree}_{\Sigma+A}^* \setminus \{(x) \mid x \in A\}$  (see Definition 2.2.1). Hence for each  $i \in \omega$ ,

$$\begin{aligned} & (F_\Sigma^+(-+0))^i 1 \\ & \cong \underbrace{\text{Tree}_\Sigma^+(\text{Tree}_\Sigma^+(\dots \text{Tree}_\Sigma^+(\{*\}) \dots))}_i \\ & \cong \left\{ \xi \in \text{Tree}_{\Sigma \times \{\circ, \odot\} + \{*\}}^* \mid \begin{array}{l} \text{the root node is labeled by } \odot, \text{ and for each branch} \\ \text{whose last component is } *, \odot \text{ appears exactly } i\text{-times} \end{array} \right\}. \end{aligned}$$

By Theorem 2.4.17,  $F_\Sigma^{+\oplus}$  is a limit of the following sequence:  $1 \xleftarrow{!} F_\Sigma^+ 1 \xleftarrow{F_\Sigma^+!} (F_\Sigma^+)^2 1 \xleftarrow{(F_\Sigma^+)^2!} \dots$ . Hence  $F_\Sigma^{+\oplus} A$  is characterized as follows:

$$F_\Sigma^{+\oplus} A \cong \left\{ \xi \in \text{Tree}_{\Sigma \times \{\circ, \odot\} + A}^\infty \mid \begin{array}{l} \text{the root node is labeled by } \odot, \text{ and for each} \\ \text{infinite branch } \odot \text{ appears infinitely often} \end{array} \right\}.$$

In general, we have the following characterization (the proof is similar to the Büchi case).

**Proposition 3.7.2.** *For  $i \in \mathbb{N}$ ,  $j \in \{1, \dots, i\}$  and a set  $A$ , we define a set  $\text{AccTree}_j^{(i)}(\Sigma, A) \subseteq \text{Tree}_{\Sigma \times \{1, \dots, i\} + A}^\infty$  by:*

$$\text{AccTree}_j^{(i)}(\Sigma, A) := \left\{ \xi \in \text{Tree}_{\Sigma \times \{1, \dots, i\} + A}^\infty \mid \begin{array}{l} \text{the root node is labeled by } j, \text{ and for} \\ \text{each infinite branch, the maximum} \\ \text{priority appearing infinitely is even} \end{array} \right\}.$$

Moreover, we let

$$\text{AccTree}_0^{(i)}(\Sigma, A) := \left\{ \xi \in \text{Tree}_{\Sigma \times \{0, \dots, i\} + A}^\infty \mid \begin{array}{l} \text{only the root node is labeled by } 0, \text{ and for} \\ \text{each infinite branch, the maximum} \\ \text{priority appearing infinitely is even} \end{array} \right\}.$$

We define a function

$$\text{decomp}_j^{(i)} : \text{AccTree}_j^{(i)}(\Sigma, A) \rightarrow \text{AccTree}_{j-1}^{(i)}(\Sigma, A)$$

by  $\text{decomp}_j^{(i)}(D, l) := (D, l')$  where

$$l'(w) := \begin{cases} (a, j-1) & (w = \langle \rangle \text{ and } l(\langle \rangle) = (a, j)) \\ l(w) & (\text{otherwise}). \end{cases}$$

Then  $\text{AccTree}_j^{(i)}(\Sigma, A) \cong (F_\Sigma)_j^{(i)} A$ , and

$$\begin{aligned} & \left( \text{decomp}_j^{(i)} : \text{AccTree}_j^{(i)}(\Sigma, A) \rightarrow \text{AccTree}_{j-1}^{(i)}(\Sigma, A) \right) \\ & \cong \left( \alpha_{j,A}^{(i)} : (F_\Sigma)_j^{(i)} A \rightarrow (F_\Sigma)_{j-1}^{(i)} A \right). \quad \square \end{aligned}$$

By using the characterizations above, we can concretely prove that the assumptions required in the previous sections are satisfied by  $\mathcal{P}$  and  $F_\Sigma$ .

**Proposition 3.7.3.** *Assumption 3.6.7 is satisfied by  $(T, F) = (\mathcal{P}, F_\Sigma)$ .*

**Proof.** Condition 2 is proved in a similar manner to [47]. Condition 3 is proved in a similar manner to [113, Theorem 4.3] using Proposition 3.7.2.

We prove Condition 4. For notational simplicity, let  $G := (F_\Sigma)_i^\dagger(-+A)$ ,  $Z := ((F_\Sigma)_i^\dagger)^\oplus A$  and  $\zeta := \zeta_A^{(F_\Sigma)_i^\dagger}$  for simplicity. By Proposition 3.7.2,  $GX \cong \text{AccTree}_i^{(i)}(\Sigma, X+A)$ .

Let  $c: X \rightarrow \overline{GX}$  and  $\sigma: GY \rightarrow Y$ . Let  $l: X \rightarrow Z$  be the greatest homomorphism from  $c$  to  $J\zeta$ , and  $m: Z \rightarrow Y$  be the greatest fixed point of  $\Phi_{J\zeta, \sigma}$ . It is easy to see that  $m \odot l$  is a fixed point of  $\Phi_{c, \sigma}$ . We show that it is the greatest fixed point. Let  $t: X \rightarrow Y$  be a fixed point of  $\Phi_{c, \sigma}$  and let  $x \in X$ . Assume  $y \in t(x)$ . We prove  $y \in m \odot l(x)$ .

For each  $k \in \omega$ , we inductively define  $\kappa_k: Z \rightarrow G^k 1$  as follows:  $\kappa_0 := !_Z$  and  $\kappa_{k+1} := G\kappa_k \circ \zeta$ . By Theorem 2.4.17,  $(Z, (\kappa_k)_{k \in \omega})$  is a limit over a final sequence  $1 \xleftarrow{!} G1 \xleftarrow{G!} G^2 1 \xleftarrow{G^2!} \dots$ . Concretely, this limit is given as follows:

$$Z \cong \left\{ (b_k \in G^k 1)_{k \in \omega} \mid \forall k \in \omega. G^{k!}(b_{k+1}) = b_k \right\}.$$

For each  $k \in \omega$ , we inductively define  $x_k \in G^k X$  and  $y_k \in G^k Y$  so that  $y_k \in \overline{G^k} t(x_k)$  as follows.

- When  $k = 0$ ,  $x_k := x$  and  $y_k := y$ .
- Assume that  $x_k, y_k$  and  $b_k$  are defined so that  $x_k \in G^k X$ ,  $y_k \in G^k Y$  and  $y_k \in \overline{G^k} t(x_k)$ . As  $t = \sigma \odot \overline{G} t \odot c$ , there exist  $x' \in G^{k+1} X$  and  $y' \in G^{k+1} Y$  such that  $x' \in \overline{G^k} c(x_k)$ ,  $y' \in \overline{G^{k+1}} t(x')$  and  $y_k \in \overline{G^k} \sigma(y')$ . We choose one of such a pair and let  $x_{k+1} := x'$  and  $y_{k+1} := y'$ .

Moreover, for each  $l \in \omega$ , we define  $b_l \in G^l 1$  by  $b_l := G^{l!}_X(x_l)$  and let  $z_k := (b_{l+k})_{l \in \omega} \in G^k Z$  for each  $k \in \omega$ .

We prove  $z_0 \in l(x)$  and  $y \in m(z_0)$ . We first prove the former. Let  $\mathbf{m}$  be an ordinal such that  $\mathcal{Kl}(\mathcal{P})(X, Z) < |\mathbf{m}|$ . By the dual of Theorem 2.3.2.3,  $l = \Phi_{c, (J\zeta)^{-1}}^{\mathbf{m}}(\top)$ . By the transfinite induction on  $\mathbf{a}$ , we prove the following:

$$\forall \mathbf{a}. \forall k \in \omega. z_k \in \overline{G^k}(\Phi_{c, J\zeta^{-1}}^{\mathbf{a}}(\top))(x_k). \quad (3.20)$$

- Let  $\mathbf{a} = 0$ . By the characterization  $GX \cong \text{AccTree}_i^{(i)}(\Sigma, X + A)$ ,  $z_k \in \overline{G^k}(\top)(x_k)$  for each  $k$ . Hence it is proved.
- Let  $\mathbf{a}$  be a successor ordinal. Then we have:

$$\begin{aligned} z_k \in \overline{G^k}(\Phi_{c, J\zeta^{-1}}^{\mathbf{a}}(\top))(x_k) &\Leftrightarrow z_k \in \overline{G^k} J\zeta^{-1} \odot \overline{G^{k+1}}(\Phi^{\mathbf{a}-1}(\top)) \odot \overline{G^k} c(x_k) \\ &\Leftarrow z_{k+1} \in \overline{G^{k+1}}(\Phi^{\mathbf{a}-1}(\top))(x_{k+1}). \end{aligned}$$

The last statement holds by the induction hypothesis.

- Let  $\mathbf{a}$  be a limit ordinal. If  $z_k \in \overline{G^k}(\Phi_{c, J\zeta^{-1}}^{\mathbf{a}'}(\top))(x_k)$  for each  $\mathbf{a}' < \mathbf{a}$ , then  $z_k \in \bigcap_{\mathbf{a}' < \mathbf{a}} \overline{G^k}(\Phi_{c, J\zeta^{-1}}^{\mathbf{a}'}(\top))(x_k)$ . Hence it is proved.

Hence we have  $z_k \in \Phi_{c, J\zeta^{-1}}^{\mathbf{a}}(\top)(x_k)$  for each  $\mathbf{a}$  and  $k$ . By letting  $\mathbf{a} = \mathbf{m}$  and  $k = 0$ , we have  $z_0 \in l(x)$ . We can similarly prove  $y \in m(z_0)$ . Hence Condition 4 is satisfied.

We prove that Condition 5 is satisfied. Let  $c: X \rightarrow (F_\Sigma)_n^\ddagger(X + A)$  be an  $(F_\Sigma)_n^\ddagger(\_ + A)$ -coalgebra and  $u: X \rightarrow ((F_\Sigma)_n^\ddagger)^\oplus A$  be the unique homomorphism from  $c$  to  $\zeta_A^{(F_\Sigma)_n^\ddagger}$ . Let  $f: X \rightarrow ((F_\Sigma)_n^\ddagger)^\oplus A$  be a homomorphism from  $Jc$  to  $J\zeta_A^{(F_\Sigma)_n^\ddagger}$ .

$$\begin{array}{ccc} (F_\Sigma)_n^\ddagger(X + A) & \xrightarrow{\overline{F}l} & (F_\Sigma)_n^\ddagger(((F_\Sigma)_n^\ddagger)^\oplus A + A) \\ Jc \uparrow & & J\zeta_A^{(F_\Sigma)_n^\ddagger} \uparrow \cong \\ X & \xrightarrow{Ju} & ((F_\Sigma)_n^\ddagger)^\oplus A \\ & \searrow f & \uparrow \\ & & J\zeta_A^{(F_\Sigma)_n^\ddagger} \end{array}$$

and  $x \in X$ , and assume that  $s \in f(x)$ . Then as  $f$  and  $Ju$  are homomorphisms from  $Jc$  to  $J\zeta_A^{(F_\Sigma)^\ddagger}$ , we can prove  $s = u(x)$  by the induction on the structure of  $s$ . Hence  $Ju$  is the greatest homomorphism.

By Conditions 4–5, we can inductively define a lifting  $\overline{(F_\Sigma)^\ddagger} : \mathcal{Kl}(T) \rightarrow \mathcal{Kl}(T)$  for each  $n \in \mathbb{N}$ . Hence Condition 1 and Condition 6 are satisfied.

Using Proposition 3.7.2, we can prove that Condition 7 is satisfied in a similar manner to [47, 113].  $\square$

We now explain  $\text{dtr}_i(c)$  for NPTAs. For simplicity, we write  $\text{AccTree}_j^{(i)}(\Sigma)$  for  $\text{AccTree}_j^{(i)}(\Sigma, \emptyset)$ . By Proposition 3.7.2,  $\text{dtr}_i(c)$  has the following type:

$$\text{dtr}_i(c) : X_i \rightarrow \mathcal{P}(\text{AccTree}_i^{(2n)}(\Sigma)).$$

**Proposition 3.7.4.** *Let  $\mathcal{A} = (X, \tau, \mathbf{p})$  be a  $\Sigma$ -labeled NPTA where  $\mathbf{p} : X \rightarrow \{1, \dots, 2n\}$ . We define  $\mathbf{p} : \text{Run}_{\mathcal{A}}^\infty(X) \rightarrow \text{Tree}_{\Sigma \times \{1, \dots, 2n\}}^\infty$  by  $\mathbf{p}(D, l) := (D, l')$  where  $l'(w) := (a, \mathbf{p}(x))$  if  $l(w) = (a, x)$ . We define a parity  $(\mathcal{P}, F_\Sigma)$ -system  $(X, c, (X_1, \dots, X_{2n}))$  as in Proposition 3.7.1.3. Then for each  $i \in \{1, \dots, 2n\}$  and  $x \in X_i$ ,*

$$\text{dtr}_i(c)(x) = \{\mathbf{p}(\rho) \in \text{AccTree}_i^{(2n)}(\Sigma) \mid \rho \in \text{Run}_{\mathcal{A}}^{\text{Acc}}(x)\}.$$

**Proof.** By Proposition 3.7.2, the type of  $\beta_{j,A}^{(i)}$  is isomorphic to the following:

$$\text{AccTree}_j^{(i)}(\Sigma, A) \rightarrow \prod_{n \in \omega} \Sigma_n \times \left( \prod_{k=1}^i \text{AccTree}_k^{(i)}(\Sigma, A) + A \right)^n.$$

Moreover, it is given by  $\beta_{j,A}^{(i)}(\xi) = (a, \xi_0, \dots, \xi_{n-1})$  for  $\xi = ((a, j), \xi_1, \dots, x_{n-1}) \in \text{AccTree}_j^{(i)}(\Sigma, A)$ .

By Proposition 3.7.2 and the definition of  $\text{Run}_{\mathcal{A}}^{\text{Acc}}(x)$ , we can see that  $\{\mathbf{p}(\rho) \mid \rho \in \text{Run}_{\mathcal{A}}^{\text{Acc}}(x)\} \subseteq F_i^{(2n)}0$ . For each  $i \in \{1, \dots, 2n\}$ , we define  $f_i : X_i \rightarrow \mathcal{P}(\text{AccTree}_i^{(2n)}(\Sigma))$  by  $f_i(x) := \{\mathbf{p}(\rho) \mid \rho \in \text{Run}_{\mathcal{A}}^{\text{Acc}}(x)\}$ . We show that a tuple  $(f_i)_{i \in \{1, \dots, 2n\}}$  is the greatest fixed point of the function (3.16) in Definition 3.6.8.

We first prove that it is a fixed point. For each  $x \in X_i$ , we have:

$$\begin{aligned} & J(\beta_{i,0}^{(2n)})^{-1} \odot \overline{F_\Sigma}(f_1 + \dots + f_{2n}) \odot c_i(x) \\ &= J(\beta_{i,0}^{(2n)})^{-1} \odot \overline{F_\Sigma}(f_1 + \dots + f_{2n}) (\{(a, x_0, \dots, x_{m-1}) \in \tau(x)\}) \\ &= J(\beta_{i,0}^{(2n)})^{-1} \left( \{(a, \mathbf{p}(\rho_0), \dots, \mathbf{p}(\rho_{m-1})) \in \Sigma_m \times (\prod_{k=1}^{2n} \text{AccTree}_k^{(2n)}(\Sigma))^m \right. \\ &\quad \left. \mid (a, x_0, \dots, x_{m-1}) \in \tau(x), \forall t \in \{0, \dots, m-1\}. \rho_t \in \text{Run}_{\mathcal{A}}^{\text{Acc}}(x_t)\} \right) \\ &= \{((a, i), \mathbf{p}(\rho_0), \dots, \mathbf{p}(\rho_{m-1})) \in \text{AccTree}_i^{(2n)}(\Sigma) \\ &\quad \mid (a, x_0, \dots, x_{m-1}) \in \tau(x), \forall t \in \{0, \dots, m-1\}. \rho_t \in \text{Run}_{\mathcal{A}}^{\text{Acc}}(x_t)\} \\ &= \{\mathbf{p}(\rho) \in \text{AccTree}_i^{(2n)}(\Sigma) \mid \rho \in \text{Run}_{\mathcal{A}}^{\text{Acc}}(x)\} \\ &= f_i(x). \end{aligned}$$

Hence  $(f_i)_{1 \leq i \leq 2n}$  is a fixed point of the function (3.16) in Definition 3.6.8.

We next show that  $(f_i)_{1 \leq i \leq 2n}$  is the greatest fixed point. Let  $(g_i : X_i \rightarrow \mathcal{P}\text{AccTree}_i^{(2n)}(\Sigma))_{1 \leq i \leq 2n}$  be a tuple of functions such that  $g_i = J(\beta_{i,0}^{(2n)})^{-1} \odot \overline{F_\Sigma}(g_1 + \dots + g_{2n}) \odot c_i$  for each  $i$ . It suffices to show that  $g_i(x) \subseteq f_i(x)$  for each  $i \in \{1, \dots, 2n\}$  and  $x \in X_i$ .

Let  $i \in \{1, \dots, 2n\}$  and  $x \in X_i$ , and assume that  $\xi = (D, l) \in g_i(x)$ . We write  $l_1(w)$  and  $l_2(w)$  for  $\pi_1(l(w)) \in \Sigma$  and  $\pi_2(l(w)) \in \{1, \dots, 2n\}$  respectively. We define a function  $l' : D \rightarrow \Sigma \times X$  by  $l'(w) := (l_1(w), l'_2(w))$ , where  $l'_2(w) \in X$  is inductively defined as follows so that the following condition is satisfied: for each  $w \in D$ ,  $\xi_w \in [g_1, \dots, g_{2n}](l'_2(w))$  (recall that  $\xi_w$  denotes the  $w$ -th subtree of  $\xi$ ).

- For  $w = \langle \rangle$ , we let  $l'_2(\langle \rangle) := x$ .
- Let  $w \in D$  and assume  $|l_1(w)| = m$ . Assume that we have fixed  $l'_2(w)$  so that the condition above is satisfied. Assume further that  $l'_2(w) \in X_i$  and that the root node of the  $w$ -th subtree  $\xi_w$  of  $\xi$  is labeled by  $a \in \Sigma_m$ . Then  $\xi_w$  has a shape  $((a, i), \xi_{w0}, \dots, \xi_{w(m-1)})$ . We have:

$$\begin{aligned}
\xi_w &= ((a, i), \xi_{w0}, \dots, \xi_{w(m-1)}) \\
&\in g_i(l'_2(w)) \\
&= J(\beta_{i,0}^{(2n)})^{-1} \odot \overline{F_\Sigma}(g_1 + \dots + g_{2n}) \odot c_i(x') \\
&= J(\beta_{i,0}^{(2n)})^{-1} \odot \overline{F_\Sigma}(g_1 + \dots + g_{2n}) \left( \{(a', x_0, \dots, x_{m'-1}) \in \tau(l'_2(w))\} \right) \\
&= J(\beta_{i,0}^{(2n)})^{-1} \left( \{(a', \xi'_0, \dots, \xi'_{m'-1}) \in \Sigma_{m'} \times \left( \prod_{k=1}^{2n} \text{AccTree}_k^{(2n)}(\Sigma) \right)^{m'} \right. \\
&\quad \left. | (a', x_0, \dots, x_{m'-1}) \in \tau(l'_2(w)), \right. \\
&\quad \left. \xi'_t \in [g_1, \dots, g_{2n}](x_t) \text{ for each } t \in \{0, \dots, m' - 1\} \} \right) \\
&= \{((a', i), \xi'_0, \dots, \xi'_{m'-1}) \in \text{AccTree}_i^{(2n)}(\Sigma) \\
&\quad | (a', x_0, \dots, x_{m'-1}) \in \tau(l'_2(w)), \\
&\quad \xi'_t \in [g_1, \dots, g_{2n}](x_t) \text{ for each } t \in \{0, \dots, m' - 1\} \}.
\end{aligned}$$

This means that there exists  $(a, x_0, \dots, x_{m-1}) \in \tau(l'_2(w))$  such that  $\xi_{wt} \in [g_1, \dots, g_{2n}](x_t)$  for each  $t \in \{0, \dots, m - 1\}$ . We let  $l'_2(wt) := x_t$  for each  $t$ .

Let  $\rho = (D, l')$ . By its construction, we can easily see that  $\mathbf{p}(\rho) = \xi$ .

By the construction,  $\rho$  is a run tree over  $\mathcal{A}$ , and moreover, as  $\xi \in \text{AccTree}_i^{(2n)}(\Sigma)$ ,  $\rho$  is accepting. Therefore by the definition of  $f_i$ , we have  $\mathbf{p}(\rho) \in f_i(x)$ . This concludes the proof.  $\square$

Finally, we check what  $p_i$  in Definition 3.6.9 characterizes.

**Proposition 3.7.5.** *We define  $\text{DelSt}_j^{(i)} : \text{AccTree}_j^{(i)}(\Sigma, A) \rightarrow \text{Tree}_{\Sigma+A}^\infty$  by  $\text{DelSt}_j^{(i)}(D, l) := (D, l')$  where  $l'(w) := \pi_1(l(w))$ . Then with respect to the isomorphism in Proposition 3.7.2,  $\text{DelSt}_j^{(i)}(\xi) = p_{j,A}^{(i)}(\xi)$ .*

**Proof.** It is easy to see that  $\text{DelSt}_j^{(i)} : \text{AccTree}_j^{(i)}(\Sigma, A) \rightarrow \text{Tree}_{\Sigma+A}^\infty$  satisfies the following equality for each  $\xi = ((a, i), (\xi_0, \dots, \xi_{m-1})) \in \text{AccTree}_j^{(i)}(\Sigma, A)$ .

$$\text{DelSt}_j^{(i)}(\xi) = \left( a, ([\text{DelSt}_1^{(i)}, \dots, \text{DelSt}_i^{(i)}](\xi_0), \dots, [\text{DelSt}_1^{(i)}, \dots, \text{DelSt}_i^{(i)}](\xi_{m-1})) \right).$$

By Proposition 3.7.2, this means that  $[\text{DelSt}_1^{(i)}, \dots, \text{DelSt}_i^{(i)}]$  is a homomorphism from  $[\beta_{1,A}^{(i)}, \dots, \beta_{i,A}^{(i)}]$  to  $\zeta_A^F$ . Therefore immediate by the definition of  $p_j^{(i)}$ .  $\square$

We can now prove Proposition 3.7.1.3 as follows: for  $x \in X_i$ ,

$$\begin{aligned}
L_{\mathcal{A}}^P(x) &= \{ \text{DelSt}_i^{(2n)}(\mathbf{p}(\rho)) \mid \rho \in \text{Run}_{\mathcal{A}}^{\text{Acc}}(x) \} && \text{(by definition)} \\
&= Jp_1 \odot \text{dtr}_i(c)(x) && \text{(by Proposition 3.7.4 and Proposition 3.7.5)} \\
&= \text{tr}_i^P(c)(x) && \text{(by Theorem 3.6.10).}
\end{aligned}$$

### 3.8 Extension to Probabilistic Automata

In this section, we fix  $T$  to the sub-Giry monad  $\mathcal{G}_s$ . A parity  $(\mathcal{G}_s, F_\Sigma)$ -system represents a *probabilistic parity tree automaton* (PPTA).

#### 3.8.1 Trace Semantics of PPTA via Logical Fixed Point

We apply the framework in Section 3.3 to PPTAs. The proposition below is proved in a similar manner to Proposition 3.7.1.

**Proposition 3.8.1.** *Define an order on each homset of  $\mathcal{Kl}(\mathcal{G}_s)$  as Example 3.1.11, and define a lifting  $\overline{F}_\Sigma : \mathcal{Kl}(\mathcal{G}_s) \rightarrow \mathcal{Kl}(\mathcal{G}_s)$  of  $F_\Sigma$  as Example 3.1.6. Then:*

1.  $\mathcal{G}_s$  and  $F_\Sigma$  constitute a parity trace situation (Definition 3.6.3).
2. The carrier set of a final  $F_\Sigma$ -coalgebra is isomorphic to  $(\text{Tree}_\Sigma^\infty, \mathfrak{F}_{\text{Tree}_\Sigma^\infty})$ .
3. For a  $\Sigma$ -labeled PPTA  $\mathcal{A} = ((X, \mathfrak{F}_X), \xi, \mathbf{p})$  where  $\mathbf{p}: X \rightarrow \{1, \dots, 2n\}$ , we define a parity  $(\mathcal{G}_s, F_\Sigma)$ -system  $((X, \mathfrak{F}_X), c_{\mathcal{A}}, ((X_1, \mathfrak{F}_{X_1}), \dots, (X_{2n}, \mathfrak{F}_{X_{2n}})))$  as follows:  $c_{\mathcal{A}} = \xi$ ,  $X_i := \{x \in X \mid \mathbf{p}(x) = i\}$  and  $\mathfrak{F}_{X_i} = \{A \cap X_i \mid A \in \mathfrak{F}_X\}$ . Then for each  $x \in X$ , we have:  $\text{tr}^{\mathbf{p}}(c)(x) = L_{\mathcal{A}}^{\mathbf{p}}(x)$ .  $\square$

**Remark 3.8.2.** When  $T = \mathcal{G}_s$ , the HES in Definition 3.6.3 is over a poset that is not a complete lattice. Let  $\Sigma = \{*\}$  where  $|*| = 0$  for example. Then  $F_\Sigma X = \{*\}$  for each  $X$ , and the carrier  $\nu F_\Sigma$  of a final  $F_\Sigma$ -coalgebra is given by  $\{*\}$ . It is not so hard to see that a measurable space  $\mathcal{G}_s(\nu F_\Sigma)$  is isomorphic to  $[0, 1]$  equipped with the standard  $\sigma$ -algebra. Suppose that we are given a Büchi  $(\mathcal{G}_s, F_\Sigma)$ -system  $(X, c, (X_1, X_2))$  such that  $X = X_1 = \mathbb{R}$  and  $X_2 = \emptyset$ , both of which are equipped with the standard  $\sigma$ -algebras. Then a poset  $\mathcal{Kl}(\mathcal{G}_s)(X, \nu F_\Sigma)$  with the order in Example 3.1.11 is not a complete lattice (nor an upward directed set). Indeed, for an arbitrary unmeasurable subset  $A \subseteq X = \mathbb{R}$  of  $\mathbb{R}$ , a directed subset  $\{\chi_a : X \rightarrow [0, 1] \mid a \in A\} \subseteq \mathcal{Kl}(\mathcal{G}_s)(X, \nu F_\Sigma) \cong \mathbf{SB}(X, [0, 1])$ , where  $\chi_a(x) = 1$  if  $x = a$  and 0 otherwise, does not have the supremum.

#### 3.8.2 Trace Semantics of PPTA via Categorical Fixed Point

In fact, it is still open if  $T = \mathcal{G}_s$  and  $F = F_\Sigma$  satisfy Assumption 3.4.14. The challenging part is the gfp-preserving condition (Assumption 3.4.14.4). However, by carefully checking the proofs of the lemmas and the propositions where the gfp-preserving condition is used (concretely, Proposition 3.4.11, Sublemma 3.6.13 and Lemma 3.6.15), we can show that Assumption 3.4.14.4 can be relaxed to the following weaker but more complicated conditions.

**Assumption 3.8.3.** When  $n$  is odd, the following conditions are satisfied.

- 4'-1.  $T$  and  $F_n^\dagger(\_ + A)$  satisfy the gfp-preserving condition with respect to an algebra  $\overline{F_n^\dagger}((F_i^\dagger)^\oplus B + A) \xrightarrow{\overline{F_n^\dagger}(\text{id}+f)} \overline{F_n^\dagger}((F_n^\dagger)^\oplus B + B) \xrightarrow{J(\zeta_B^{F_n^\dagger})^{-1}} (F_n^\dagger)^\oplus B$  for each  $f : A \rightarrow B$ ;
- 4'-2.  $T$  and  $F^+(\_ + A)$  satisfy the gfp-preserving condition with respect to an algebra  $F^+(F^{\oplus\oplus} A + A) \xrightarrow{J_\tau} F^\oplus(F^{\oplus\oplus} A + A) \xrightarrow{J(\zeta_A^{F^\oplus})^{-1}} F^{\oplus\oplus} A$  where  $\tau$  is the unique homomorphism from  $(\iota_{F^{\oplus\oplus} A + A}^F)^{-1}$  to  $\zeta_{F^{\oplus\oplus} A + A}^F$ ; and

4'-3.  $T$  and  $F(\_ + A)$  satisfy the gfp-preserving condition with respect to an algebra  $F(F^\oplus A + F^\oplus A + A) \xrightarrow{JF(\text{id}, \text{id}) + \text{id}} F(F^\oplus A + A) \xrightarrow{J(\zeta_A^F)^{-1}} F^\oplus A$ .

**Proposition 3.8.4.** *Theorem 3.6.10 still holds even if we replace Condition 4 of Assumption 3.6.7 with the conditions in Assumption 3.8.3.  $\square$*

The weakened conditions above are satisfied by  $T = \mathcal{G}_s$  and  $F = F_\Sigma$  on **SB**.

**Proposition 3.8.5.** *We define a partial order  $\sqsubseteq$  for each homset of  $\mathcal{Kl}(\mathcal{G}_s)$  as in Example 3.1.11, and define a lifting  $\overline{F}_\Sigma : \mathcal{Kl}(\mathcal{G}_s) \rightarrow \mathcal{Kl}(\mathcal{G}_s)$  of  $F_\Sigma$  as in Example 3.1.6. Then Conditions 1–3 and 5–7 of Assumption 3.6.7 and Conditions 4'-1–3 of Assumption 3.8.3 are satisfied by  $(T, F) = (\mathcal{G}_s, F_\Sigma)$ .*

**Proof.** Conditions 1, 2, 3, 6 and 7 are proved in a similar manner to Proposition 3.7.3.

We prove that Condition 4'-1 is satisfied. Let  $c: X \rightarrow (F_\Sigma)_i^\ddagger(X + A)$ . Let  $l: X \rightarrow ((F_\Sigma)_i^\ddagger)^\oplus A$  be the greatest homomorphism from  $c$  to  $J\zeta_A^{(F_\Sigma)_i^\ddagger}$ . Let  $m: ((F_\Sigma)_i^\ddagger)^\oplus A \rightarrow ((F_\Sigma)_i^\ddagger)^\oplus B$  be the greatest fixed point of  $g \mapsto (J(\zeta_B^{(F_\Sigma)_i^\ddagger})^{-1} \odot (F_\Sigma)_i^\ddagger(\text{id} + f)) \odot \overline{(F_\Sigma)_i^\ddagger}(g + \text{id}) \odot \zeta_A^{(F_\Sigma)_i^\ddagger}$ . Let  $t: X \rightarrow ((F_\Sigma)_i^\ddagger)^\oplus B$  be the greatest fixed point of  $h \mapsto (J(\zeta_B^{(F_\Sigma)_i^\ddagger})^{-1} \odot (F_\Sigma)_i^\ddagger(\text{id} + f)) \odot \overline{(F_\Sigma)_i^\ddagger}(h + \text{id}) \odot c$ . Pictorially,

$$\begin{array}{ccccc}
(F_\Sigma)_i^\ddagger(X + A) & \xrightarrow{\quad} & (F_\Sigma)_i^\ddagger(((F_\Sigma)_i^\ddagger)^\oplus A + A) & \xrightarrow{\quad} & (F_\Sigma)_i^\ddagger(((F_\Sigma)_i^\ddagger)^\oplus B + A) \\
\uparrow c & & \uparrow J\zeta_A^{(F_\Sigma)_i^\ddagger} \cong & & \downarrow \overline{(F_\Sigma)_i^\ddagger}(\text{id} + f) \\
& \xrightarrow{\quad} & & \xrightarrow{\quad} & (F_\Sigma)_i^\ddagger(((F_\Sigma)_i^\ddagger)^\oplus B + B) \\
& & & & \downarrow J(\zeta_B^{(F_\Sigma)_i^\ddagger})^{-1} \\
X & \xrightarrow{\quad l \quad} & ((F_\Sigma)_i^\ddagger)^\oplus A & \xrightarrow{\quad m \quad} & ((F_\Sigma)_i^\ddagger)^\oplus B \\
& & \uparrow & & \uparrow t \\
& & X & & 
\end{array}$$

It is easy to see that  $m \odot l$  is a fixed point of  $h \mapsto (J(\zeta_B^{(F_\Sigma)_i^\ddagger})^{-1} \odot \overline{(F_\Sigma)_i^\ddagger}(\text{id} + f)) \odot \overline{(F_\Sigma)_i^\ddagger}(h + \text{id}) \odot c$  and hence we have  $m \odot l \sqsubseteq t$ . We prove the opposite direction.

For each  $k \in \omega$ , we inductively define  $\pi_k : ((F_\Sigma)_i^\ddagger)^\oplus A \rightarrow ((F_\Sigma)_i^\ddagger(\_ + A))^k 1$  as follows:  $\pi_0 := !_{((F_\Sigma)_i^\ddagger)^\oplus A}$  and  $\pi_{k+1} := ((F_\Sigma)_i^\ddagger(\_ + \text{id}_A))^k ! \circ (F_\Sigma)_i^\ddagger(\pi_k + \text{id}_A) \circ \zeta_A^{(F_\Sigma)_i^\ddagger}$ . We define  $\pi'_k : ((F_\Sigma)_i^\ddagger)^\oplus B \rightarrow ((F_\Sigma)_i^\ddagger(\_ + B))^k 1$  in a similar manner.

By Theorem 2.4.17,  $(((F_\Sigma)_i^\ddagger)^\oplus A, (\pi_k)_{k \in \omega})$  is a limit over a final sequence  $1 \xleftarrow{!} (F_\Sigma)_i^\ddagger(1 + A) \xleftarrow{(F_\Sigma)_i^\ddagger(1 + \text{id}_A)} (F_\Sigma)_i^\ddagger(((F_\Sigma)_i^\ddagger(1 + A) + A)) \xleftarrow{(F_\Sigma)_i^\ddagger((F_\Sigma)_i^\ddagger(1 + \text{id}_A) + \text{id}_A)} \dots$ . Similarly,  $(((F_\Sigma)_i^\ddagger)^\oplus B, (\pi'_k)_{k \in \omega})$  is a limit over  $1 \xleftarrow{!} (F_\Sigma)_i^\ddagger(1 + B) \xleftarrow{(F_\Sigma)_i^\ddagger(1 + \text{id}_B)} (F_\Sigma)_i^\ddagger(((F_\Sigma)_i^\ddagger(1 + B) + B)) \xleftarrow{(F_\Sigma)_i^\ddagger((F_\Sigma)_i^\ddagger(1 + \text{id}_B) + \text{id}_B)} \dots$ .

It is known that  $\mathcal{G}_s : \mathbf{SB} \rightarrow \mathbf{SB}$  preserves a limit of an  $\omega^{\text{op}}$ -sequence consisting of standard Borel sets [96]. This means that  $(((F_\Sigma)_i^\ddagger)^\oplus A, (J\pi_k)_{k \in \omega})$  is a limit over a sequence  $1 \xleftarrow{J!} (F_\Sigma)_i^\ddagger(1 + A) \xleftarrow{J(F_\Sigma)_i^\ddagger(1 + \text{id}_A)} (F_\Sigma)_i^\ddagger(((F_\Sigma)_i^\ddagger(1 + A) + A)) \xleftarrow{J(F_\Sigma)_i^\ddagger((F_\Sigma)_i^\ddagger(1 + \text{id}_A) + \text{id}_A)} \dots$ . It is easy to see that it is also a 2-limit. That is, for two cones  $(X, (\gamma_k^1)_{k \in \omega})$  and  $(X, (\gamma_k^2)_{k \in \omega})$  over  $1 \xleftarrow{J!} (F_\Sigma)_i^\ddagger(1 + A) \xleftarrow{J(F_\Sigma)_i^\ddagger(1 + \text{id}_A)} (F_\Sigma)_i^\ddagger(((F_\Sigma)_i^\ddagger(1 + A) + A)) \xleftarrow{J(F_\Sigma)_i^\ddagger((F_\Sigma)_i^\ddagger(1 + \text{id}_A) + \text{id}_A)} \dots$  such that  $\gamma_k^1 \sqsubseteq \gamma_k^2$  for each  $k \in \omega$ , if we write  $l^1$  (resp.  $l^2$ ) for the mediating arrow from  $(X, (\gamma_k^1)_{k \in \omega})$

(resp.  $(X, (\gamma_k^2)_{k \in \omega})$ ) to  $((F_\Sigma)_i^\ddagger)^\oplus A, (J\pi_k)_{k \in \omega}$ , then we have  $l^1 \sqsubseteq l^2$ . Similarly,  $((F_\Sigma)_i^\ddagger)^\oplus B, (J\pi'_k)_{k \in \omega}$  is a 2-limit over  $1 \xleftarrow{J!} (F_\Sigma)_i^\ddagger(1+B) \xleftarrow{J(F_\Sigma)_i^\ddagger(1+\text{id}_B)} (F_\Sigma)_i^\ddagger((F_\Sigma)_i^\ddagger(1+B)+B) \xleftarrow{J(F_\Sigma)_i^\ddagger((F_\Sigma)_i^\ddagger(1+\text{id}_B)+\text{id}_B)} \dots$

We inductively define a cone  $(X, (\gamma_k : X \rightarrow ((F_\Sigma)_i^\ddagger(\_ + A))^k 1)_{k \in \omega})$  over  $1 \xleftarrow{J!} (F_\Sigma)_i^\ddagger(1+A) \xleftarrow{J(F_\Sigma)_i^\ddagger(1+\text{id}_A)} (F_\Sigma)_i^\ddagger((F_\Sigma)_i^\ddagger(1+A)+A) \xleftarrow{J(F_\Sigma)_i^\ddagger((F_\Sigma)_i^\ddagger(1+\text{id}_A)+\text{id}_A)} \dots$  as follows:

- For each  $s \in \omega$ , we define an arrow  $f_s : X \rightarrow 1$  as follows: i)  $f_0 := J!_X$  and ii)  $f_{s+1} := J! \odot (F_\Sigma)_i^\ddagger(f_s + \text{id}_A) \odot c$ . It is easy to see that  $f_0 \sqsupseteq f_1 \sqsupseteq \dots$ . We define  $f_\omega : X \rightarrow 1$  by  $f_\omega := \prod_{s \in \omega} f_s$ . As the composition  $\odot$  in  $\mathcal{Kl}(\mathcal{G}_s)$  is  $\omega^{\text{op}}$ -continuous, by the Kleene fixed point theorem,  $f_\omega$  is the greatest fixed point of a function  $h \mapsto J! \odot (F_\Sigma)_i^\ddagger(h + \text{id}_A) \odot c$ . We let  $\gamma_0 := f_\omega$ .
- $\gamma_{k+1} := J(F_\Sigma)_i^\ddagger(\gamma_k + \text{id}_A) \odot c$ .

$$\begin{array}{ccccccc}
X & \xrightarrow{c} & \overline{G}X & \xrightarrow{\overline{G}c} & \overline{G}^2X & \xrightarrow{\overline{G}^2c} & \overline{G}^3X & \xrightarrow{\overline{G}^3c} & \dots \\
\downarrow \gamma_0 & \text{=} & \downarrow \overline{G}\gamma_0 & \text{=} & \downarrow \overline{G}^2\gamma_0 & \text{=} & \downarrow \overline{G}^3\gamma_0 & & \\
1 & \xleftarrow{J!} & \overline{G}1 & \xleftarrow{JG!} & \overline{G}^21 & \xleftarrow{JG^2!} & \overline{G}^31 & \xleftarrow{JG^3!} & \dots \\
\uparrow J! & \text{=} & \uparrow JG! & \text{=} & \uparrow JG^2! & \text{=} & \uparrow JG^3! & & \\
Z & \xrightarrow{J\zeta_A^{(F_\Sigma)_i^\ddagger}} & \overline{G}Z & \xrightarrow{JG\zeta_A^{(F_\Sigma)_i^\ddagger}} & \overline{G}^2Z & \xrightarrow{JG^2\zeta_A^{(F_\Sigma)_i^\ddagger}} & \overline{G}^3Z & \xrightarrow{JG^3\zeta_A^{(F_\Sigma)_i^\ddagger}} & \dots
\end{array}$$

(Here  $G := (F_\Sigma)_i^\ddagger(\_ + A)$  and  $Z := ((F_\Sigma)_i^\ddagger)^\oplus A$ .)

Let  $l' : X \rightarrow ((F_\Sigma)_i^\ddagger)^\oplus A$  be the unique mediating arrow from  $(X, (\gamma_k)_{k \in \omega})$  to a 2-limit  $((F_\Sigma)_i^\ddagger)^\oplus A, (J\pi_k)_{k \in \omega}$ . By its definition, we can easily see that  $l'$  is a homomorphism from  $c$  to  $J\zeta_A^{(F_\Sigma)_i^\ddagger}$ . Moreover, for an arbitrary homomorphism  $l'' : X \rightarrow ((F_\Sigma)_i^\ddagger)^\oplus A$  from  $c$  to  $J\zeta_A^{(F_\Sigma)_i^\ddagger}$ , if we define  $\gamma'_k : X \rightarrow ((F_\Sigma)_i^\ddagger(\_ + A))^k 1$  by  $\gamma'_k := J\pi_k \odot l''$  for each  $k \in \omega$ , then  $l''$  is a mediating arrow from  $(X, (\gamma'_k)_{k \in \omega})$  to  $((F_\Sigma)_i^\ddagger)^\oplus A, (J\pi_k)_{k \in \omega}$ . By definition, we have  $\gamma'_k \sqsubseteq \gamma_k$  for each  $k$ . Therefore as  $((F_\Sigma)_i^\ddagger)^\oplus A, (J\pi_k)_{k \in \omega}$  is a 2-limit, we have  $l'' \leq l'$ . This means that the mediating arrow from  $(X, (\gamma_k)_{k \in \omega})$  to  $((F_\Sigma)_i^\ddagger)^\oplus A, (J\pi_k)_{k \in \omega}$  is given by  $l$ .

Note that  $m$  (resp.  $t$ ) is the greatest homomorphism from  $(F_\Sigma)_i^\ddagger(\text{id}_A + f) \odot J\zeta_A^{(F_\Sigma)_i^\ddagger}$  (resp.  $(F_\Sigma)_i^\ddagger(\text{id}_X + f) \odot c$ ) to  $J\zeta_B^{(F_\Sigma)_i^\ddagger}$ . Hence in a similar manner to  $(X, (\gamma_k)_{k \in \omega})$ , we can define a cone  $((F_\Sigma)_i^\ddagger)^\oplus A, (\varepsilon_k : ((F_\Sigma)_i^\ddagger)^\oplus A \rightarrow ((F_\Sigma)_i^\ddagger(\_ + B))^k 1)_{k \in \omega}$  (resp.  $(X, (\delta_k : X \rightarrow ((F_\Sigma)_i^\ddagger(\_ + B))^k 1)_{k \in \omega})$ ) over  $1 \xleftarrow{J!} (F_\Sigma)_i^\ddagger(1+B) \xleftarrow{J(F_\Sigma)_i^\ddagger(1+\text{id}_B)} (F_\Sigma)_i^\ddagger((F_\Sigma)_i^\ddagger(1+B)+B) \xleftarrow{J(F_\Sigma)_i^\ddagger((F_\Sigma)_i^\ddagger(1+\text{id}_B)+\text{id}_B)} \dots$ , and  $m$  (resp.  $t$ ) is the mediating arrow from the cone to  $((F_\Sigma)_i^\ddagger)^\oplus B, (J\pi'_k)_{k \in \omega}$ .

$$\begin{array}{ccc}
X & \xrightarrow{l} & ((F_\Sigma)_i^\ddagger)^\oplus A & \xrightarrow{m} & ((F_\Sigma)_i^\ddagger)^\oplus B & \xrightarrow{J\pi'_k} & ((F_\Sigma)_i^\ddagger(\_ + B))^k 1 \\
& \searrow & \downarrow J\pi_k & \searrow \varepsilon_k & \downarrow J\pi_k & \searrow & \\
& \searrow & ((F_\Sigma)_i^\ddagger(\_ + A))^k 1 & & ((F_\Sigma)_i^\ddagger(\_ + B))^k 1 & & \\
& \searrow & \downarrow \gamma_k & & \downarrow \delta_k & & \\
& \searrow & & & & & 
\end{array}$$

By definition and that  $((F_\Sigma)_i^\ddagger)^\oplus B, (J\pi'_k)_{k \in \omega}$  is a 2-limit, to prove  $t \sqsubseteq m \odot l$ , it suffices to prove  $\delta_k \sqsubseteq \varepsilon_k \odot l$  for each  $k \in \omega$ . We prove it by an induction on  $k$ .



- Let  $k = 0$ . Note that for each  $X \in \mathbf{SB}$ ,  $\mathcal{Kl}(\mathcal{G}_s)(X, 1)$  has the greatest element  $\top_{X,1}$ . We first prove  $\delta_0 \sqsubseteq \top_{((F_\Sigma)_i^\ddagger)^\oplus_{A,1}} \odot l$ . Note that the right-hand side is equivalent to  $\gamma_0$ . Recall that  $\delta_0$  is the greatest fixed point of  $h'' \mapsto J! \odot \overline{(F_\Sigma)_i^\ddagger}(h'' + f) \odot c$  while  $\gamma_0$  is the greatest fixed point of  $h \mapsto J! \odot (F_\Sigma)_i^\ddagger(h + \text{id}_A) \odot c$ . Hence by an easy induction, we can prove that  $\delta_0 \sqsubseteq \gamma_0$  and therefore  $\delta_0 \sqsubseteq \top_{((F_\Sigma)_i^\ddagger)^\oplus_{A,1}} \odot l$  is proved.

Note that  $\varepsilon_0$  is the greatest fixed point of  $h' \mapsto J! \odot F(h' + \text{id}_B) \odot \overline{(F_\Sigma)_i^\ddagger}(\text{id}_A + f) \odot J\zeta_A^{(F_\Sigma)_i^\ddagger}$ . Again by an inductive manner, we can prove  $\delta_0 \sqsubseteq \varepsilon_0 \odot l$ .

- When  $k > 0$ , we can prove  $\delta_k \sqsubseteq \varepsilon_k \odot l$  by the definitions of  $\delta_k$  and  $\varepsilon_k$ , and the induction hypothesis.

Hence we have  $t = m \odot l$ .

Condition 4'-2 and Condition 4'-3 are similarly proved.

It is easy to prove  $Jf \sqsubseteq g$  implies  $g = Jf$ . Hence Condition 5 holds.  $\square$

Hence we can consider decorated trace semantics  $\text{dtr}_1(c), \dots, \text{dtr}_{2n}(c)$  for parity  $(\mathcal{G}_s, F_\Sigma)$ -systems. The datatype  $(F_\Sigma)_j^{(i)}$  and its accompanying coalgebraic structure  $\beta_j^{(i)}$  are given as follows.

**Lemma 3.8.6.** For  $i \in \mathbb{N}$  and  $j \in \{1, \dots, i\}$ ,

$$(F_\Sigma)_j^{(i)} A \cong (\text{AccTree}_j^{(i)}(\Sigma, A), \mathfrak{F}_{\text{AccTree}_j^{(i)}(\Sigma, A)})$$

where  $\mathfrak{F}_{\text{AccTree}_j^{(i)}(\Sigma, A)} := \mathfrak{F}_{\text{Tree}_{(\Sigma+A) \times \{1, \dots, i\}}^\infty} \cap \text{AccTree}_j^{(i)}(\Sigma, A)$ . Moreover, if  $i$  is odd then the function  $\text{decomp}_j^{(i)}$  in Proposition 3.7.2 coincides with  $\alpha_{j,A}^{(i)}$ .  $\square$

We can now characterize  $\text{dtr}_i(c)$ .

**Proposition 3.8.7.** Let  $\mathcal{A} = ((X, \mathfrak{F}_X), \xi, \mathfrak{p})$  be a  $\Sigma$ -labeled PPTA such that  $\mathfrak{p}: X \rightarrow \{1, \dots, 2n\}$ . We define a parity  $(\mathcal{G}_s, F_\Sigma)$ -system

$$((X, \mathfrak{F}_X), c_{\mathcal{A}}, ((X_1, \mathfrak{F}_{X_1}), \dots, (X_{2n}, \mathfrak{F}_{X_{2n}})))$$

as in Proposition 3.8.1. We define  $\mathfrak{p}: \text{Run}_{\mathcal{A}}^\infty \rightarrow \text{Tree}_{\Sigma \times \{1, \dots, 2n\}}^\infty$  as in Proposition 3.7.4. Then for each  $i \in \{1, \dots, 2n\}$ ,  $x \in X_i$  and  $A \in \mathfrak{F}_{\text{AccTree}_i^{(2n)}(\Sigma, A)}$ , with respect to the isomorphism in Lemma 3.8.6,

$$\text{dtr}_i(c_{\mathcal{A}})(x)(A) = L_{\mathcal{A}}^{\text{Run}}(x) \left( \left\{ \rho \in \text{Run}_{\mathcal{A}}^\infty \mid \mathfrak{p}(\rho) \in A \right\} \right).$$

**Proof.** Proved in a similar manner to [113, Theorem A.13].  $\square$

We can characterize  $p_i$  in a similar manner to the nondeterministic case (Proposition 3.7.5).

As in the previous section, the Proposition 3.8.7 implies Proposition 3.8.1.3.

### 3.9 Conclusion and Related Work

We have introduced two categorical characterizations for languages of NBTAs (Definition 3.3.1 and 3.4.15) and NPTAs (Definition 3.6.3 and 3.6.8). One of them considers logical fixed points (i.e. fixed points in homsets) while the other considers categorical fixed points (i.e. fixed points in categories). We have proved that the latter characterization induces the former one (Theorem 3.5.4 and 3.6.10) and hence they can be thought of as essentially the same characterization.

**Related Work** A categorical characterization of the Büchi and parity condition is also found in some existing work. In [24], using the *lasso-characterization* of the Büchi condition, a Büchi automaton is modeled as a coalgebra in the product category  $\mathbf{Sets} \times \mathbf{Sets}$ . Because of the use of the lasso characterization, dealing with infinite-state automata seems to be difficult for their framework. Extension to probabilistic systems also seems to be hard in their framework. However, in contrast, there exist notions that are well-characterized in their framework but seems to be difficult to be characterized in our framework, like *bisimilarity*. In [116], a notion of *coalgebra automaton* is introduced. It is an automaton that takes coalgebras as inputs and classifies them with respect to the Büchi, parity or Muller acceptance condition.

We have used the notion of “alternating fixed point of functors” in Chapter 3. The same notion is also used in [38, 4]. In [38] the authors characterize the set of continuous functions from  $A^\omega$  to  $B^\omega$  as an alternating fixed point  $\nu X. \mu Y. (B \times X) + Y^A$  of a functor. Although the data type is not exactly the same as the one used by us, it also has a Büchi-like flavor: for a continuous function  $f : A^\omega \rightarrow B^\omega$ , if  $f(a_0a_1 \dots) = b_0b_1 \dots$  then each  $b_i$  should be determined by some *finite* prefix of  $a_0a_1 \dots$ . Hence a continuous function  $f$  can be regarded as an *infinite* repetition of such assignments determined by finite prefix. In [4, Section 7] a sufficient condition for the existence of such an alternating fixed point is discussed.

## Chapter 4

# Categorical Fair Simulation

Using the categorical framework developed in the previous chapter, we categorically generalize *fair simulation*.

*Simulation* is a notion often used to prove behavioral inclusion between transition systems. Simulation notions are defined for various systems [66, 55]. Fair simulation is one of them, which was originally introduced for nondeterministic Büchi word automata [50, 32]. It was generalized for nondeterministic Büchi tree automata (NBTAs) in [117].

Other well-known simulation notions are *forward* and *backward simulation* [74] for nondeterministic automata. Categorical generalizations of those simulation notions are known as *Kleisli simulation* [43]. In this chapter, we categorically generalize fair simulation by extending the framework of Kleisli simulation. We then concretize it for probabilistic Büchi tree automata (PBTAs), quantitative variants of nondeterministic Büchi tree automata, to induce a new simulation notion.

This chapter is organized as follows. We review notions of forward and fair simulation in Section 4.1 and that of Kleisli simulation in Section 4.2. A categorical generalization of fair simulation is in Section 4.3. Section 4.4 is devoted to a “sanity check”: we concretize the obtained categorical framework for NBTAs and observe that we rediscover the conventional notion of fair simulation. We concretize the framework for PBTAs and induce a new simulation notion in Section 4.5.

This chapter is based on [110].

### 4.1 Simulation

In this section, we review two notions of *simulation*. Simulation is often used to verify *behavioral inclusion* between transition systems. For nondeterministic word automata, simulation notions called *forward simulation* and *backward simulation* are well-known. We hereby review the forward one. It is defined in terms of a parity game (Definition 2.2.25).

**Definition 4.1.1** (forward simulation, [74]). Let  $\Sigma$  be a ranked alphabet and  $\mathcal{A} = (X, \tau)$  and  $\mathcal{B} = (Y, \sigma)$  be  $\Sigma$ -labeled NTAs (Definition 2.2.4). Let  $R \subseteq X \times Y$ . We define a parity game  $G_{\mathcal{A}, \mathcal{B}, R}^{\text{fwd}} = (X_{\mathcal{A}, \mathcal{B}, R}^{\text{Max}}, X_{\mathcal{A}, \mathcal{B}, R}^{\text{Min}}, E_{\mathcal{A}, \mathcal{B}, R}^{\text{Max}}, E_{\mathcal{A}, \mathcal{B}, R}^{\text{Min}}, \mathbf{P}_{\mathcal{A}, \mathcal{B}})$  as follows:

$$X_{\mathcal{A}, \mathcal{B}, R}^{\text{Max}} := R + \left( \prod_{i \in \omega} \Sigma_i \times X^i \right) \times Y$$
$$X_{\mathcal{A}, \mathcal{B}, R}^{\text{Min}} := R + R^*$$

$$\begin{aligned}
E_{\mathcal{A},\mathcal{B},R}^{\text{Max}} &:= \{(r, r) \in R \times R \mid r \in R\} \\
&\quad + \left\{ \left( \begin{array}{l} ((a, x_1, \dots, x_i), y), \\ ((x_1, y_1), \dots, (x_i, y_i)) \end{array} \in \prod_i \Sigma_i \times X^i \times R^* \mid \begin{array}{l} (a, y_1, \dots, y_n) \in \sigma(y), \\ \forall j. (x_j, y_j) \in R \end{array} \right) \right\} \\
E_{\mathcal{A},\mathcal{B},R}^{\text{Min}} &:= \left\{ \left( \begin{array}{l} ((x, y), \\ ((a, x_1, \dots, x_i), y)) \end{array} \in R \times \prod_i \Sigma_i \times X^i \mid \begin{array}{l} (a, x_1, \dots, x_n) \in \tau(x) \end{array} \right) \right\} \\
&\quad + \left\{ \left( ((x_1, y_1), \dots, (x_i, y_i)), (x_j, y_j) \right) \in R^* \times R \mid j \in \{1, \dots, i\} \right\} \\
\rho_{\mathcal{A},\mathcal{B},R}(t) &:= 0.
\end{aligned}$$

We call  $R$  a *forward simulation* from  $\mathcal{A}$  to  $\mathcal{B}$  if Player Max is winning in the game  $G_{\mathcal{A},\mathcal{B},R}^{\text{fwd}}$  from each state  $(x, x')$  in  $R$ .

Note that in the parity game  $G_{\mathcal{A},\mathcal{B},R}^{\text{fwd}}$ , all the states are assigned a priority 0. Hence Player Max wins if the play continues infinitely or Player Min gets stuck.

Using a forward simulation, we can check both finite and infinitary language inclusions. This property is called *soundness* of simulation.

**Theorem 4.1.2** ([74]). *If  $R$  is a forward simulation from  $\mathcal{A}$  to  $\mathcal{B}$ ,  $(x, y) \in R$  implies  $L_{\mathcal{A}}^*(x) \subseteq L_{\mathcal{B}}^*(y)$  and  $L_{\mathcal{A}}^\infty(x) \subseteq L_{\mathcal{B}}^\infty(y)$ .*  $\square$

A simulation notion was also defined for nondeterministic Büchi word automata [50, 32] and named *fair simulation*. It was extended for nondeterministic Büchi tree automata in [117].

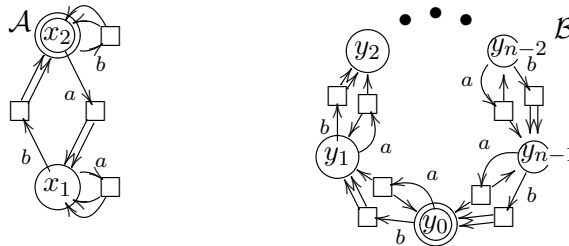
**Definition 4.1.3** (fair simulation, [50, 32, 117]). Let  $\mathcal{A} = (X, \tau, \text{Acc}_{\mathcal{A}})$  and  $\mathcal{B} = (Y, \sigma, \text{Acc}_{\mathcal{B}})$  be finite-state  $\Sigma$ -labeled NBTAs. Let  $R \subseteq X \times Y$ . We define a parity game  $G_{\mathcal{A},\mathcal{B},R}^{\text{fair}} = (X_{\mathcal{A},\mathcal{B},R}^{\text{Max}}, X_{\mathcal{A},\mathcal{B},R}^{\text{Min}}, E_{\mathcal{A},\mathcal{B},R}^{\text{Max}}, E_{\mathcal{A},\mathcal{B},R}^{\text{Min}}, \rho_{\mathcal{A},\mathcal{B}})$  so that  $X_{\mathcal{A},\mathcal{B},R}^{\text{Max}}$ ,  $X_{\mathcal{A},\mathcal{B},R}^{\text{Min}}$ ,  $E_{\mathcal{A},\mathcal{B},R}^{\text{Max}}$ , and  $E_{\mathcal{A},\mathcal{B},R}^{\text{Min}}$  are the same as Definition 4.1.1 and  $\rho_{\mathcal{A},\mathcal{B}}$  is defined as follows:

$$\rho_{\mathcal{A},\mathcal{B},R}(x, y) := \begin{cases} 0 & (x \notin \text{Acc}_{\mathcal{A}}, y \notin \text{Acc}_{\mathcal{B}}) \\ 1 & (x \in \text{Acc}_{\mathcal{A}}, y \notin \text{Acc}_{\mathcal{B}}) \\ 2 & (y \in \text{Acc}_{\mathcal{B}}) \end{cases} \quad \text{and} \quad \rho_{\mathcal{A},\mathcal{B},R}((a, x_1, \dots, x_i), y) := 0$$

We call  $R$  a *fair simulation* from  $\mathcal{A}$  to  $\mathcal{B}$  if Player Max is winning in the game  $G_{\mathcal{A},\mathcal{B},R}^{\text{fair}}$  from each state  $(x, x')$  in  $R$ .

**Theorem 4.1.4** (soundness of fair simulation, [50, 32, 117]). *If  $R$  is a fair simulation from  $\mathcal{A}$  to  $\mathcal{B}$ ,  $(x, y) \in R$  implies  $L_{\mathcal{A}}^{\text{B}}(x) \subseteq L_{\mathcal{B}}^{\text{B}}(y)$ .*  $\square$

**Example 4.1.5.** Let  $\Sigma$  be a ranked alphabet defined by  $\Sigma = \{a, b\}$  and  $|a| = |b| = 2$ . Let  $\mathcal{A} = (X, \tau, \text{Acc}_{\mathcal{A}})$  and  $\mathcal{B} = (Y, \sigma, \text{Acc}_{\mathcal{B}})$  be  $\Sigma$ -labeled NBTAs illustrated below, where we write  $z \xrightarrow{\sigma} \square \rightrightarrows z_1^1 z_2^2$  when  $(c, z_1, z_2) \in u(z)$  for  $u \in \{\tau, \sigma\}$ .



If we let  $R := X \times Y$ , then  $R$  is a fair simulation from  $\mathcal{A}$  to  $\mathcal{B}$ .

## 4.2 Kleisli Simulation

In this section, we review the notion of *Kleisli simulation*. It is a categorical generalization of forward (and backward) simulation (Definition 4.1.1). We hereby review the forward one.

**Definition 4.2.1** (forward Kleisli simulation, [43]). Let  $\mathcal{X} = (X, c)$  and  $\mathcal{Y} = (Y, d)$  be  $(T, F)$ -systems. Assume that each homset of  $\mathcal{Kl}(T)$  carries a partial order  $\sqsubseteq$ . A *forward (Kleisli) simulation* from  $\mathcal{X}$  to  $\mathcal{Y}$  is a Kleisli arrow  $f : Y \rightarrow X$  such that  $c \odot f \sqsubseteq \overline{F}f \odot d$ .

**Theorem 4.2.2** (soundness of forward Kleisli simulation). *Let  $\mathcal{X} = (X, c)$  and  $\mathcal{Y} = (Y, d)$  be  $(T, F)$ -systems. Assume that each homset of  $\mathcal{Kl}(T)$  carries a partial order  $\sqsubseteq$ , and the assumptions in Theorem 3.1.13 are satisfied. Assume further that there exists a forward Kleisli simulation from  $\mathcal{X}$  to  $\mathcal{Y}$ .*

1. ([43]) *By Theorem 3.1.13,  $T$  and  $F$  constitute a finite trace semantics. We have:  $\text{tr}(c) \sqsubseteq \text{tr}(d) \odot f : X \rightarrow \mu F$ .*
2. *With respect to  $\text{tr}^*(c)$  and  $\text{tr}^*(d)$  in Proposition 3.1.14, we have:  $\text{tr}^*(c) \sqsubseteq \text{tr}^*(d) \odot f : X \rightarrow \nu F$ .*
3. ([113]) *Assume that  $T$  and  $F$  constitute an infinitary trace situation. We have:  $\text{tr}^\infty(c) \sqsubseteq \text{tr}^\infty(d) \odot f : X \rightarrow \nu F$ .  $\square$*

Item 2 is immediate by Item 1 and Proposition 3.1.14.

## 4.3 Kleisli Fair Simulation

In this section, we extend the notion of Kleisli simulation (Definition 4.2.1) so that it generalizes fair simulation (Definition 4.1.3). To this end, we employ two existing notions—*partially additive monad* and *lattice-theoretic progress measure*.

### 4.3.1 Partially Additive Monad

Let  $T$  be a monad and  $g_1 : V \rightarrow X_1$  and  $g_2 : V \rightarrow X_2$  be Kleisli arrows in  $\mathcal{Kl}(T)$ . In general, we cannot canonically define an arrow  $g : V \rightarrow X_1 + X_2$  from  $g_1$  and  $g_2$ . A *partially additive monad* is a monad with such a “codomain join” operation.

For example, when  $T = \mathcal{P}$ , we can define  $g : V \rightarrow X_1 + X_2$  by  $g(v) := g_1(v) \cup g_2(v)$ . When  $T = \mathcal{G}_s$ , if  $g_1(v)(X_1) + g_2(v)(X_2) \leq 1$  then we can define  $g : V \rightarrow X_1 + X_2$  by  $g(v)(A) := g_1(v)(A) + g_2(v)(A)$ . Note that in the latter case, the “codomain join” is not always defined, i.e. it is a partial operation.

**Definition 4.3.1** (partially additive monad, [59, 25]). Let  $\mathbb{C}$  be a category with an initial object  $0$ , binary products and binary coproducts. A monad  $T$  on  $\mathbb{C}$  is called a *partially additive monad* if it satisfies the following conditions:

1. The object  $T0$  is a final object in  $\mathbb{C}$ .<sup>1</sup>
2. Let  $X_1, X_2 \in \mathbb{C}$ . We define  $p_1 : X_1 + X_2 \rightarrow TX_1$  and  $p_2 : X_1 + X_2 \rightarrow TX_2$  by  $p_1 := [\eta_{X_1}, \perp_{X_2, X_1}]$  and  $p_2 := [\perp_{X_1, X_2}, \eta_{X_2}]$ , where for  $X, Y \in \mathbb{C}$ ,

<sup>1</sup>This implies that  $0$  is both an initial and final object in  $\mathcal{Kl}(T)$ . Such an object is called a *zero object*.

$\perp_{X,Y}: X \rightarrow TY$  is defined by  $\perp_{X,Y} := X \xrightarrow{!X} T0 \xrightarrow{T!Y} TY$ . (See also Remark 4.3.4.) Then the following arrow is a monomorphism<sup>2</sup>.

$$T(X_1 + X_2) \xrightarrow{\langle \mu_{X_1} \circ T p_1, \mu_{X_2} \circ T p_2 \rangle} TX_1 \times TX_2$$

The above assumptions allow us to define a ‘‘codomain join’’ and its inverse operation as follows.

**Definition 4.3.2** (codomain restriction and codomain join, [25, 59]). Let  $T$  be a partially additive monad. For each  $V, X_1, X_2 \in \mathbb{C}$  and  $i \in \{1, 2\}$ , we define a function called a *codomain restriction*  $\_ \uparrow^{X_i}: \mathcal{Kl}(T)(V, X_1 \times X_2) \rightarrow \mathcal{Kl}(T)(V, X_i)$  and a partial function called a *codomain join*  $\langle\langle \_ , \_ \rangle\rangle: \mathcal{Kl}(T)(V, X_1) \times \mathcal{Kl}(T)(V, X_2) \rightarrow \mathcal{Kl}(T)(V, X_1 \times X_2)$  as follows.

$$g \uparrow^{X_i} := \left( V \xrightarrow{g} T(X_1 + X_2) \xrightarrow{\langle \mu_{X_1} \circ T p_1, \mu_{X_2} \circ T p_2 \rangle} TX_1 \times TX_2 \xrightarrow{\pi_i} TX_i \right) \quad \text{and}$$

$$\langle\langle g_1, g_2 \rangle\rangle := \begin{cases} g & (\exists g: V \rightarrow X_1 \times X_2. \forall i \in \{1, 2\}. g_i = g \uparrow^{X_i}) \\ \text{undefined} & (\text{otherwise}). \end{cases}$$

Note that by Condition 2 of Definition 4.3.1,  $\langle\langle g_1, g_2 \rangle\rangle$  is unique if it exists. See also the following diagram.

$$\begin{array}{ccc} T(X_1 + X_2) & \xrightarrow{\langle \mu_{X_1} \circ T p_1, \mu_{X_2} \circ T p_2 \rangle} & TX_1 \times TX_2 \\ & \swarrow \langle\langle g_1, g_2 \rangle\rangle & \uparrow \langle\langle g_1, g_2 \rangle\rangle \\ & & V \end{array}$$

**Example 4.3.3.** As we have already mentioned, the powerset monad  $\mathcal{P}$  is partially additive. The operations in Definition 4.3.2 is given as follows:

$$g \uparrow^{X_i}(v) = \{x \in X_i \mid x \in g(v)\} \quad \text{and} \quad \langle\langle g_1, g_2 \rangle\rangle(v) = g_1(v) \cup g_2(v).$$

The sub-Giry monad  $\mathcal{G}_s$  is also partially additive, and we have:

$$g \uparrow^{X_i}(v)(A) = g(v)(A), \quad \text{and}$$

$$\langle\langle g_1, g_2 \rangle\rangle(v)(A) = \begin{cases} g_1(v)(A \cap X_1) + g_2(v)(A \cap X_2) & (g_1(v)(A \cap X_1) + g_2(v)(A \cap X_2) \leq 1) \\ \text{undefined} & (\text{otherwise}). \end{cases}$$

**Remark 4.3.4.** Let  $T$  be a partially additive monad such that  $\mathcal{Kl}(T)$  is a **Cppo**-enriched category. We have used the same symbol  $\perp_{X,Y}$  for (i) the least element in a homset of a **Cppo**-enriched category (Definition 3.1.12) and (ii) an arrow  $T!Y \circ !X$  in the Kleisli category of a partially additive monad. A confusion is unlikely because of the following reason: in the next section (see Theorem 4.3.14) we assume that composition in  $\mathcal{Kl}(T)$  is left-strict, i.e.  $\perp_{X,Y} \odot g = \perp_{Z,Y}$  for each  $g: Z \rightarrow Y$ , where  $\perp_{X,Y}$  and  $\perp_{Z,Y}$  are defined in the sense of (i). It is easy to see that under this assumption, (i) and (ii) coincide.

We conclude this section by presenting some properties of codomain restrictions and joins. Their proofs are easy.

**Lemma 4.3.5.** *Let  $T$  be a partially additive monad.*

<sup>2</sup>An arrow  $m: X \rightarrow Y$  is a *monomorphism* if  $m \circ f = m \circ g$  implies  $f = g$

1. For  $g: V \rightarrow X_1 + X_2$ , the codomain join  $\langle\langle g \upharpoonright^{X_1}, g \upharpoonright^{X_2} \rangle\rangle$  is always defined and given by  $g$ . Conversely, if  $\langle\langle g_1, g_2 \rangle\rangle$  is defined then  $(\langle\langle g_1, g_2 \rangle\rangle) \upharpoonright^{X_i} = g_i$  for  $i \in \{1, 2\}$ .

2. For  $f: W \rightarrow V$ ,  $g_1: V \rightarrow X_1$ ,  $g_2: V \rightarrow X_2$ ,  $h_1: X_1 \rightarrow Y_1$  and  $h_2: X_2 \rightarrow Y_2$  such that  $\langle\langle g_1, g_2 \rangle\rangle$  is defined, we have:

$$\langle\langle g_1, g_2 \rangle\rangle \odot f = \langle\langle g_1 \circ f, g_2 \circ f \rangle\rangle \quad \text{and} \quad (h_1 + h_2) \circ \langle\langle g_1, g_2 \rangle\rangle = \langle\langle h_1 \circ g_1, h_2 \circ g_2 \rangle\rangle.$$

3. For  $g: V \rightarrow X$ ,  $\langle\langle g, \perp_{V,X} \rangle\rangle$  and  $\langle\langle \perp_{V,X}, g \rangle\rangle$  are always defined and we have

$$[\text{id}_X, \text{id}_X] \odot \langle\langle g, \perp_{V,X} \rangle\rangle = [\text{id}_X, \text{id}_X] \odot \langle\langle \perp_{V,X}, g \rangle\rangle = g. \quad \square$$

### 4.3.2 Lattice-Theoretic Progress Measures

As its name suggests, the notion of lattice-theoretic progress measure stems from that of progress measure. We first review the latter. As parity games (Definition 2.2.25) have many applications, methods for calculating the winning region of a parity game are extensively studied. Progress measure is one of them.

A progress measure is defined as a function that assigns each state of a parity game a tuple  $(\mathbf{a}_1, \dots, \mathbf{a}_n)$  of ordinals. Intuitively, an ordinal  $\mathbf{a}_k$  ‘‘counts’’ the number of states with an odd priority  $2k - 1$  so that such states are not visited infinitely many times without visiting states with greater priorities. Ranking function discussed in the next chapter (Definition 5.1.1) can be thought of as its special case.

**Definition 4.3.6** ( $\leq_i$ ). Let  $\mathcal{S} = (X^{\text{Max}}, X^{\text{Min}}, E^{\text{Max}}, E^{\text{Min}}, \mathbf{p})$  be a parity game such that  $\mathbf{p}: X^{\text{Max}} \rightarrow \{1, \dots, 2n\}$ . A *prioritized ordinal* for  $\mathcal{S}$  is an  $n$ -tuple  $(\mathbf{a}_1, \dots, \mathbf{a}_n)$  of ordinals. Moreover, for each  $i \in \{1, \dots, 2n\}$  the  *$i$ -th truncated (pointwise) order* is a preorder  $\leq_i$  between prioritized ordinals defined as follows:  $(\mathbf{a}_1, \dots, \mathbf{a}_n) \leq_i (\mathbf{a}'_1, \dots, \mathbf{a}'_n)$  holds if (i)  $i = 2n$ ; or (ii)  $i \leq 2n - 1$  and  $\mathbf{a}_j \leq \mathbf{a}'_j$  for each  $j \in \{a, \dots, n\}$  where  $a = \frac{i+1}{2}$  if  $i$  is odd and  $\frac{i}{2} + 1$  if  $i$  is even. We write  $(\mathbf{a}_1, \dots, \mathbf{a}_n) <_i (\mathbf{a}'_1, \dots, \mathbf{a}'_n)$  if  $(\mathbf{a}_1, \dots, \mathbf{a}_n) \leq_i (\mathbf{a}'_1, \dots, \mathbf{a}'_n)$  and  $(\mathbf{a}_1, \dots, \mathbf{a}_n) \not\geq_i (\mathbf{a}'_1, \dots, \mathbf{a}'_n)$ .

**Definition 4.3.7** (progress measure for two-player games, [67]). Let  $\mathcal{S} = (X^{\text{Max}}, X^{\text{Min}}, E^{\text{Max}}, E^{\text{Min}}, \mathbf{p})$  be a parity game such that  $\mathbf{p}: X^{\text{Max}} \rightarrow \{1, \dots, 2n\}$ . We write  $[\mathbf{a}]$  for  $\mathbf{a} + 1 (= \{\mathbf{a}' \leq \mathbf{a}\})$ . A *progress measure* for  $E$  is a pair

$$p = ((\overline{\mathbf{a}}_1, \dots, \overline{\mathbf{a}}_n), p: X^{\text{Max}} \rightarrow [\overline{\mathbf{a}}_1] \times \dots \times [\overline{\mathbf{a}}_n] + \{\top\})$$

of a prioritized ordinal and a function that satisfy the following conditions for each  $x \in X^{\text{Max}}$  such that  $\mathbf{p}(x) = i$ .

1. If  $i$  is odd, then

$$\exists y \in X^{\text{Min}} \text{ s.t. } (x, y) \in E^{\text{Max}}. \forall x' \in X^{\text{Max}} \text{ s.t. } (y, x') \in E^{\text{Min}}. p(x') <_i p(x).$$

2. If  $i$  is even, then

$$\exists y \in X^{\text{Min}} \text{ s.t. } (x, y) \in E^{\text{Max}}. \forall x' \in X^{\text{Max}} \text{ s.t. } (y, x') \in E^{\text{Min}}. p(x') \leq_i p(x).$$

Here we regard  $\top$  as the greatest element with respect to  $\leq_i$  for each  $i$ .

Progress measure provides us with a sound and complete method for deciding the winner of a parity game.

**Theorem 4.3.8** ([67]). *Let  $\mathcal{S} = (X^{\text{Max}}, X^{\text{Min}}, E^{\text{Max}}, E^{\text{Min}}, p)$  be a parity game.*

1. (**Soundness**) *If  $p = ((\bar{\mathbf{a}}_1, \dots, \bar{\mathbf{a}}_n), p)$  is a progress measure for  $\mathcal{S}$ , then for each  $x \in X^{\text{Max}}$ ,  $p(x) \neq \top$  implies  $x \in \text{Win}_{\mathcal{S}}$ .*
2. (**Completeness**) *There exists a progress measure  $p = ((\bar{\mathbf{a}}_1, \dots, \bar{\mathbf{a}}_n), p)$  such that  $p(x) \neq \top$  for each  $x \in \text{Win}_{\mathcal{S}}$ .  $\square$*

Theorem 4.3.8 says that we can *underapproximate* the winning region  $\text{Win}_{\mathcal{S}}$  of a parity game using a progress measure. *Lattice-theoretic progress measure* [49] is a generalization of progress measure, which we can use for underapproximating solutions of an HES. Because  $\mu$  and  $\nu$  do not necessarily appear in the alternating manner in our definition of HES, we have to first modify Definition 4.3.6.

**Definition 4.3.9** ( $\leq_i$  for HESs). Let  $E$  be an HES as in Definition 2.3.5. Let  $k := |\{i \mid \eta_i = \mu\}|$  and define  $i_1, \dots, i_k \in \{1, \dots, m\}$  so that  $i_1 < i_2 < \dots < i_k$  and  $\eta_{i_a} = \mu$  for each  $a \in \{1, \dots, k\}$ . A *prioritized ordinal* for  $E$  is a  $k$ -tuple  $(\mathbf{a}_1, \dots, \mathbf{a}_k)$  of ordinals. Moreover, for each  $i \in \{1, \dots, m\}$  the  *$i$ -th truncated (pointwise) order* is a preorder  $\leq_i$  between prioritized ordinals defined as follows:  $(\mathbf{a}_1, \dots, \mathbf{a}_k) \leq_i (\mathbf{a}'_1, \dots, \mathbf{a}'_k)$  holds if (i)  $i_k < i$ , or (ii)  $i \leq i_k$  and  $\mathbf{a}_j \leq \mathbf{a}'_j$  for each  $j \in \{a, \dots, k\}$  where we define  $a \in \{1, \dots, k\}$  so that  $i_1 < \dots < i_{a-1} < i \leq i_a < \dots < i_k$ .

**Definition 4.3.10** (lattice-theoretic progress measure, [49]). Let  $E$  be an HES as in Definition 2.3.5, and define  $\{i_1, \dots, i_k\}$  as in Definition 4.3.9. We assume that for each  $i \in \{1, \dots, m\}$ ,  $L_i$  has the smallest element  $\perp$ . A *(lattice-theoretic) progress measure* for  $E$  is a pair

$$p = ((\bar{\mathbf{a}}_1, \dots, \bar{\mathbf{a}}_k), (p_i(\mathbf{a}_1, \dots, \mathbf{a}_k))_{i \in \{1, \dots, m\}, 0 \leq \mathbf{a}_1 \leq \bar{\mathbf{a}}_1, \dots, 0 \leq \mathbf{a}_k \leq \bar{\mathbf{a}}_k})$$

of a prioritized ordinal and a family of elements  $p_i(\mathbf{a}_1, \dots, \mathbf{a}_k) \in L_i$  that satisfy the following conditions for each prioritized ordinal  $(\mathbf{a}_1, \dots, \mathbf{a}_k)$ .

1. (**Monotonicity**) For each  $i \in \{1, \dots, m\}$ ,  $(\mathbf{a}_1, \dots, \mathbf{a}_k) \leq_i (\mathbf{a}'_1, \dots, \mathbf{a}'_k)$  implies  $p_i(\mathbf{a}_1, \dots, \mathbf{a}_k) \sqsubseteq p_i(\mathbf{a}'_1, \dots, \mathbf{a}'_k)$ .
2. ( **$\mu$ -variables, base case**) Let  $a \in \{1, \dots, k\}$ . If  $\mathbf{a}_a = 0$  then  $p_{i_a}(\mathbf{a}_1, \dots, \mathbf{a}_a, \dots, \mathbf{a}_k) = \perp$ .
3. ( **$\mu$ -variables, step case**) Let  $a \in \{1, \dots, k\}$ . If  $\mathbf{a}$  is a successor ordinal, there exist ordinals  $\mathbf{b}_1, \dots, \mathbf{b}_{a-1}$  such that  $\mathbf{b}_1 \leq \bar{\mathbf{a}}_1, \dots, \mathbf{b}_{a-1} \leq \bar{\mathbf{a}}_{a-1}$  and

$$p_{i_a}(\mathbf{a}_1, \dots, \mathbf{a}_{a-1}, \mathbf{a}_a, \mathbf{a}_{a+1}, \dots, \mathbf{a}_k) \sqsubseteq f_{i_a} \left( \begin{array}{c} p_1(\mathbf{b}_1, \dots, \mathbf{b}_{a-1}, \mathbf{a}_a - 1, \mathbf{a}_{a+1}, \dots, \mathbf{a}_k), \\ \dots, \\ p_m(\mathbf{b}_1, \dots, \mathbf{b}_{a-1}, \mathbf{a}_a - 1, \mathbf{a}_{a+1}, \dots, \mathbf{a}_k) \end{array} \right). \quad (4.1)$$

4. ( **$\mu$ -variables, limit case**) Let  $a \in \{1, \dots, k\}$ . If  $\mathbf{a}_a$  is a limit ordinal, then the supremum  $\bigsqcup_{\mathbf{b} < \mathbf{a}_a} p_{i_a}(\mathbf{a}_1, \dots, \mathbf{a}_{a-1}, \mathbf{b}, \mathbf{a}_{a+1}, \dots, \mathbf{a}_k) \in L_{i_a}$  exists and

$$p_{i_a}(\mathbf{a}_1, \dots, \mathbf{a}_{a-1}, \mathbf{a}_a, \mathbf{a}_{a+1}, \dots, \mathbf{a}_k) \sqsubseteq \bigsqcup_{\mathbf{b} < \mathbf{a}_a} p_{i_a}(\mathbf{a}_1, \dots, \mathbf{a}_{a-1}, \mathbf{b}, \mathbf{a}_{a+1}, \dots, \mathbf{a}_k). \quad (4.2)$$



5. ( $\nu$ -variables) Let  $i \in \{1, \dots, m\} \setminus \{i_1, \dots, i_k\}$ . Define  $a \in \{1, \dots, k+1\}$  be so that  $i_1 < \dots < i_{a-1} < i < i_a < \dots < i_k$  (if  $i_k < i$  then  $a := k+1$ ). There exist ordinals  $\mathbf{b}_1, \dots, \mathbf{b}_{a-1}$  such that  $\mathbf{b}_1 \leq \overline{\mathbf{a}}_1, \dots, \mathbf{b}_{a-1} \leq \overline{\mathbf{a}}_{a-1}$  and

$$p_i(\mathbf{a}_1, \dots, \mathbf{a}_{a-1}, \mathbf{a}_a, \dots, \mathbf{a}_k) \sqsubseteq f_i \left( \begin{array}{c} p_1(\mathbf{b}_1, \dots, \mathbf{b}_{a-1}, \mathbf{a}_a, \dots, \mathbf{a}_k), \\ \dots, \\ p_m(\mathbf{b}_1, \dots, \mathbf{b}_{a-1}, \mathbf{a}_a, \dots, \mathbf{a}_k) \end{array} \right). \quad (4.3)$$

We call each  $p_i(\mathbf{a}_1, \dots, \mathbf{a}_k)$  an *approximant*.

The following result corresponds to Theorem 4.3.8, which is proved using Theorem 2.3.2 and Corollary 2.3.3.

**Theorem 4.3.11** (cf. [49]). *Let  $E$  be an HES as in Definition 2.3.5, and assume that  $E$  has a solution. We further assume that for each  $i \in \{1, \dots, m\}$ , the poset  $(L_i, \sqsubseteq_i)$  has the least element and either of the following conditions is satisfied:*

- (i)  $(L_i, \sqsubseteq_i)$  is  $\omega$ -complete, and for each  $l_{i+1} \in L_{i+1}, \dots, l_m \in L_m$ , the function  $f_i^\ddagger(\_, l_{i+1}, \dots, l_m) : L_i \rightarrow L_i$  (Definition 2.3.6) is  $\omega$ -continuous; or
- (ii)  $(L_i, \sqsubseteq_i)$  is directed complete.

Then we have the following.

1. (**Soundness**) If  $p = ((\overline{\mathbf{a}}_1, \dots, \overline{\mathbf{a}}_k), (p_i(\mathbf{a}_1, \dots, \mathbf{a}_k))_{i, \mathbf{a}_1, \dots, \mathbf{a}_k})$  is a progress measure for  $E$  then we have  $p_i(\overline{\mathbf{a}}_1, \dots, \overline{\mathbf{a}}_k) \sqsubseteq l_i^{\text{sol}}$  for each  $i \in \{1, \dots, m\}$ .
2. (**Completeness**) There exists a progress measure  $p = ((\overline{\mathbf{a}}_1, \dots, \overline{\mathbf{a}}_k), (p_i(\mathbf{a}_1, \dots, \mathbf{a}_k))_{i, \mathbf{a}_1, \dots, \mathbf{a}_k})$  such that  $p_i(\overline{\mathbf{a}}_1, \dots, \overline{\mathbf{a}}_k) = u_i^{\text{sol}}$  for each  $i \in \{1, \dots, m\}$ . Especially if Assumption (i) is satisfied above, there exists a progress measure such that  $\overline{\mathbf{a}}_i \leq \omega$  for each  $i \in \{1, \dots, m\}$ .  $\square$

Definition 4.3.7 and Definition 4.3.10 might seem rather different, but in fact the latter generalizes the former in the following sense.

**Proposition 4.3.12** ([49]). *Let  $\mathcal{S} = (X^{\text{Max}}, X^{\text{Min}}, E^{\text{Max}}, E^{\text{Min}}, \mathbf{p})$  be a parity game such that  $\mathbf{p} : X^{\text{Max}} \rightarrow \{1, \dots, 2n\}$ . For each  $i \in \{1, \dots, 2n\}$ , we let  $X_i^{\text{Max}} := \{x \in X^{\text{Max}} \mid \mathbf{p}(x) = i\}$ . We define functions  $\square_{\mathcal{S}} : \mathcal{P}X^{\text{Max}} \rightarrow \mathcal{P}X^{\text{Min}}$  and  $\diamond_{\mathcal{S}} : \mathcal{P}X^{\text{Min}} \rightarrow \mathcal{P}X^{\text{Max}}$  as follows:*

$$\begin{aligned} \square_{\mathcal{S}}(S) &:= \{y \in X^{\text{Min}} \mid \forall x' \in X^{\text{Max}} \text{ s.t. } (y, x') \in E^{\text{Min}}. x' \in S\} \quad \text{and} \\ \diamond_{\mathcal{S}}(T) &:= \{x \in X^{\text{Max}} \mid \exists y \in X^{\text{Min}} \text{ s.t. } (x, y) \in E^{\text{Max}}. y \in T\}. \end{aligned}$$

For  $i \in \{1, \dots, 2n\}$ , we define  $\diamond_{\mathcal{S}, i} : \mathcal{P}X^{\text{Min}} \rightarrow \mathcal{P}X_i^{\text{Max}}$  by  $\diamond_{\mathcal{S}, i}(T) := \diamond_{\mathcal{S}}(T) \cap X_i$ . We define an HES as follows (here  $\mu$  and  $\nu$  appear in an alternating manner):

$$\left\{ \begin{array}{l} u_1 =_{\mu} \diamond_{\mathcal{S}, 1}(\square_{\mathcal{S}}(u_1 + u_2 + \dots + u_{2n})) \in (\mathcal{P}X_1^{\text{Max}}, \subseteq) \\ u_2 =_{\nu} \diamond_{\mathcal{S}, 2}(\square_{\mathcal{S}}(u_1 + u_2 + \dots + u_{2n})) \in (\mathcal{P}X_2^{\text{Max}}, \subseteq) \\ \vdots \\ u_{2n} =_{\nu} \diamond_{\mathcal{S}, 2n}(\square_{\mathcal{S}}(u_1 + u_2 + \dots + u_{2n})) \in (\mathcal{P}X_{2n}^{\text{Max}}, \subseteq) \end{array} \right. \quad (4.4)$$

1. Let  $p = ((\overline{\mathbf{a}}_1, \dots, \overline{\mathbf{a}}_n), p)$  be a progress measure for  $\mathcal{S}$ . For  $i \in \{1, \dots, 2n\}$  and ordinals  $\mathbf{a}_1 \in [\overline{\mathbf{a}}_1], \dots, \mathbf{a}_n \in [\overline{\mathbf{a}}_n]$ , we define a subset  $p_i(\mathbf{a}_1, \dots, \mathbf{a}_n) \in \mathcal{P}X_i^{\text{Max}}$  as follows:

$$p_i(\mathbf{a}_1, \dots, \mathbf{a}_n) := \left\{ x \in X^{\text{Max}} \mid p(x) \leq_i (\mathbf{a}_1, \dots, \mathbf{a}_n) \right\}.$$

Then a pair  $((\overline{\mathbf{a}}_1, \dots, \overline{\mathbf{a}}_n), (p_i(\mathbf{a}_1, \dots, \mathbf{a}_n))_{i, \mathbf{a}_1, \dots, \mathbf{a}_n})$  is a lattice-theoretic progress measure for the HES (4.4).

2. Let  $p = ((\bar{\mathbf{a}}_1, \dots, \bar{\mathbf{a}}_n), (p_i(\mathbf{a}_1, \dots, \mathbf{a}_n))_{i, \mathbf{a}_1, \dots, \mathbf{a}_n})$  be a lattice-theoretic progress measure for the HES (4.4). We define a function  $p : X^{\text{Max}} \rightarrow [\bar{\mathbf{a}}_1] \times \dots \times [\bar{\mathbf{a}}_n] + \{\top\}$  so that  $p(x)$  is one of the minimum elements of the following set with respect to the preorder  $\leq_i$ :

$$\{(\mathbf{a}_1, \dots, \mathbf{a}_n) \mid x \in p_i(\mathbf{a}_1, \dots, \mathbf{a}_n)\}.$$

Then a pair  $((\bar{\mathbf{a}}_1, \dots, \bar{\mathbf{a}}_n), p)$  is a progress measure for  $\mathcal{S}$ .  $\square$

The proposition above also shows that the winning region of a parity game can be calculated as the solution of an HES (4.4).

### 4.3.3 Kleisli Fair Simulation with Dividing

We are now ready to present our categorical definition of fair simulation and its soundness theorem.

**Definition 4.3.13** (fair simulation with dividing). Assume that  $T$  and  $F$  constitute a Büchi trace situation (Definition 3.3.1) with respect to  $\sqsubseteq$ . We further assume that the monad  $T$  is partially additive. Let  $\mathcal{X} = (X, c, (X_1, X_2))$  and  $\mathcal{Y} = (Y, d, (Y_1, Y_2))$  be Büchi  $(T, F)$ -systems, and  $\bar{\mathbf{a}}$  be an ordinal. A (Kleisli,  $\bar{\mathbf{a}}$ -bounded) fair simulation with dividing from  $\mathcal{X}$  to  $\mathcal{Y}$  is an arrow  $f : Y \rightarrow X$  that satisfies the following conditions (for simplicity, we write  $f_{ji}$  for  $(f \odot \kappa_j) \uparrow^{X_i} : Y_j \rightarrow X_i$  (see Definition 4.3.2)).

- A. The arrow  $f : Y \rightarrow X$  is a forward Kleisli simulation from  $\mathcal{X}$  to  $\mathcal{Y}$ .
- B. There exist a pair  $d_{11}, d_{12} : Y_1 \rightarrow \bar{F}Y$  of arrows such that  $[\text{id}_{\bar{F}Y}, \text{id}_{\bar{F}Y}] \odot \langle\langle d_{11}, d_{12} \rangle\rangle = d_1$  and a pair of increasing transfinite sequences

$$f_{11}^{(0)} \sqsubseteq f_{11}^{(1)} \sqsubseteq \dots \sqsubseteq f_{11}^{(\bar{\mathbf{a}})} : Y_1 \rightarrow X_1 \text{ and } f_{12}^{(0)} \sqsubseteq f_{12}^{(1)} \sqsubseteq \dots \sqsubseteq f_{12}^{(\bar{\mathbf{a}})} : Y_1 \rightarrow X_2,$$

such that a codomain join  $\langle\langle f_{11}^{(\mathbf{a})}, f_{12}^{(\mathbf{a})} \rangle\rangle$  exists for each  $\mathbf{a} \leq \bar{\mathbf{a}}$ , and the following conditions are satisfied:

- (a) (**Approximate  $f_{11}$  and  $f_{12}$** ) We have  $f_{11}^{(\bar{\mathbf{a}})} = f_{11}$  and  $f_{12}^{(\bar{\mathbf{a}})} = f_{12}$ .
- (b) ( $f_{11}^{(\mathbf{a})}$ ) For each  $\mathbf{a}$ ,  $c_1 \odot f_{11}^{(\mathbf{a})} \sqsubseteq \bar{F}[\langle\langle f_{11}^{(\mathbf{a})}, f_{12}^{(\mathbf{a})} \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle] \odot d_{11}$ .
- (c) ( $f_{12}^{(\mathbf{a})}$ , **the base case**) If  $\mathbf{a} = 0$ , then  $f_{12}^{(\mathbf{a})} = \perp$ .
- (d) ( $f_{12}^{(\mathbf{a})}$ , **the step case**) If  $\mathbf{a}$  is a successor ordinal, then  $c_2 \odot f_{12}^{(\mathbf{a})} \sqsubseteq \bar{F}[\langle\langle f_{11}^{(\mathbf{a}-1)}, f_{12}^{(\mathbf{a}-1)} \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle] \odot d_{12}$ .
- (e) ( $f_{12}^{(\mathbf{a})}$ , **the limit case**) If  $\mathbf{a}$  is a limit ordinal, then the supremum  $\bigsqcup_{\mathbf{a}' < \mathbf{a}} f_{12}^{(\mathbf{a}' )}$  exists and  $f_{12}^{(\mathbf{a})} \sqsubseteq \bigsqcup_{\mathbf{a}' < \mathbf{a}} f_{12}^{(\mathbf{a}' )}$ .

We call the pair  $d_{11}, d_{12}$  of arrows a *dividing* of  $d_1$ , and the sequences  $f_{11}^{(0)} \sqsubseteq \dots \sqsubseteq f_{11}^{(\bar{\mathbf{a}})}$  and  $f_{12}^{(0)} \sqsubseteq \dots \sqsubseteq f_{12}^{(\bar{\mathbf{a}})}$  *approximating sequences*.

$$\begin{array}{ccc} FY & \xrightarrow{\bar{F}[\langle\langle f_{11}^{(\mathbf{a})}, f_{12}^{(\mathbf{a})} \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle]} & FX \\ d_{11} \uparrow & \sqsupseteq & \uparrow c_1 \\ Y_1 & \xrightarrow{f_{11}^{(\mathbf{a})}} & X_1 \end{array} \quad \begin{array}{ccc} FY & \xrightarrow{\bar{F}[\langle\langle f_{11}^{(\mathbf{a})}, f_{12}^{(\mathbf{a})} \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle]} & FX \\ d_{12} \uparrow & \sqsupseteq & \uparrow c_2 \\ Y_1 & \xrightarrow{f_{12}^{(\mathbf{a}+1)}} & X_2 \end{array}$$

**Theorem 4.3.14** (soundness). Let  $\bar{\mathbf{a}}$  be an ordinal. Assume the following.

1. The Kleisli category  $\mathcal{Kl}(T)$  and the lifting  $\overline{F} : \mathcal{Kl}(T) \rightarrow \mathcal{Kl}(T)$  of  $F$  are **Cppo**-enriched with respect to  $\sqsubseteq$  (see Definition 3.1.12).
2. The codomain restriction  $(\_) \upharpoonright^{X_i}$ , the codomain join  $\langle\langle \_, \_ \rangle\rangle$  and the co-coupling  $[\_, \_]$  of Kleisli arrows are all monotone and  $\omega$ -continuous with respect to  $\sqsubseteq$ .
3. The codomain join is downward closed, i.e. for  $f_1, g_1 : V \rightarrow X_1$  and  $f_2, g_2 : V \rightarrow X_2$  such that  $f_1 \sqsubseteq g_1$  and  $f_2 \sqsubseteq g_2$ , if  $\langle\langle g_1, g_2 \rangle\rangle$  is defined then  $\langle\langle f_1, f_2 \rangle\rangle$  is also defined.
4. The Kleisli composition  $\odot$  is left and right-strict, i.e.  $\perp \odot f = \perp$  and  $f \odot \perp = \perp$ .
5. For each limit ordinal  $\mathfrak{a} \leq \overline{\mathfrak{a}}$ , post-composition in  $\mathcal{Kl}(T)$  is  $\mathfrak{a}$ -continuous, i.e. if the supremum  $\bigsqcup_{i < \mathfrak{a}} f_i$  exists then  $\bigsqcup_{i < \omega} (g \odot f_i)$  also exists and  $g \odot (\bigsqcup_{i < \mathfrak{a}} f_i) = \bigsqcup_{i < \omega} (g \odot f_i)$ .

If there exists a fair simulation with dividing  $f : Y \rightarrow X$  from  $\mathcal{X} = (X, c, (X_1, X_2))$  to  $\mathcal{Y} = (Y, c, (Y_1, Y_2))$ , then we have  $\text{tr}^B(c) \odot f \sqsubseteq \text{tr}^B(d) : Y \rightarrow \nu F$ .

By Assumption 1, we can apply the correctness results of progress measures in Theorem 4.3.11 for the HES (3.4) in Definition 3.3.1. Note also that by Assumption 3, the codomain join  $\langle\langle f_{11}^{(\mathfrak{a})}, f_{12}^{(\mathfrak{a})} \rangle\rangle$  in Definition 4.3.13 always exists.

We prove the theorem using two lemmas.

**Lemma 4.3.15.** *Recall that  $\text{tr}_1^B(c) : X_1 \rightarrow \nu F$  and  $\text{tr}_2^B(c) : X_2 \rightarrow \nu F$  are given as the solutions  $u_1^{\text{sol}}$  and  $u_2^{\text{sol}}$  of the HES (3.4) in Definition 3.3.1. By Assumption 1 of Theorem 4.3.14 and the completeness theorem of progress measure (Theorem 4.3.11.2), there exists a progress measure*

$$p_{\mathcal{X}} = \left( (\overline{\mathfrak{b}}_1), (u_1(\mathfrak{b}_1) : X_1 \rightarrow \nu F, u_2(\mathfrak{b}_1) : X_2 \rightarrow \nu F)_{\mathfrak{b}_1 \leq \overline{\mathfrak{b}}_1} \right)$$

for the HES (3.4) such that  $\overline{\mathfrak{b}}_1 \leq \omega$ ,  $u_1(\overline{\mathfrak{b}}_1) = \text{tr}_1^B(c)$  and  $u_2(\overline{\mathfrak{b}}_1) = \text{tr}_2^B(c)$ . We define a pair

$$p = \left( (\overline{\mathfrak{c}}_1, \overline{\mathfrak{c}}_2), (h_1(\mathfrak{c}_1, \mathfrak{c}_2), h_2(\mathfrak{c}_1, \mathfrak{c}_2), h_3(\mathfrak{c}_1, \mathfrak{c}_2))_{\mathfrak{c}_1 \leq \overline{\mathfrak{c}}_1, \mathfrak{c}_2 \leq \overline{\mathfrak{c}}_2} \right)$$

of a pair of ordinals and a family of ordinal-indexed Kleisli arrows  $h_1(\mathfrak{c}_1, \mathfrak{c}_2) : Y_1 \rightarrow \nu F$ ,  $h_2(\mathfrak{c}_1, \mathfrak{c}_2) : Y_1 \rightarrow \nu F$  and  $h_3(\mathfrak{c}_1, \mathfrak{c}_2) : Y_2 \rightarrow \nu F$  as follows:  $\overline{\mathfrak{c}}_1 := \overline{\mathfrak{b}}_1$ ,  $\overline{\mathfrak{c}}_2 := \overline{\mathfrak{a}}$ ,  $h_1(\mathfrak{c}_1, \mathfrak{c}_2) := u_1(\mathfrak{c}_1) \odot f_{11}^{(\mathfrak{c}_2)}$ ,  $h_2(\mathfrak{c}_1, \mathfrak{c}_2) := u_2(\overline{\mathfrak{c}}_1) \odot f_{12}^{(\mathfrak{c}_2)}$ , and  $h_3(\mathfrak{c}_1, \mathfrak{c}_2) := [u_1(\overline{\mathfrak{c}}_1), u_2(\overline{\mathfrak{c}}_1)] \odot \langle\langle f_{21}, f_{22} \rangle\rangle$  (see also Figure 4.1). Then  $p$  is a progress measure for the following HES.

$$\begin{aligned} h_1 &=_{\mu} (J\zeta)^{-1} \odot \overline{F} [[\text{id}_{\nu F}, \text{id}_{\nu F}] \odot \langle\langle h_1, h_2 \rangle\rangle, h_3] \odot d_{11} \in \mathcal{Kl}(T)(Y_1, \nu F) \\ h_2 &=_{\mu} (J\zeta)^{-1} \odot \overline{F} [[\text{id}_{\nu F}, \text{id}_{\nu F}] \odot \langle\langle h_1, h_2 \rangle\rangle, h_3] \odot d_{12} \in \mathcal{Kl}(T)(Y_1, \nu F) \\ h_3 &=_{\nu} (J\zeta)^{-1} \odot \overline{F} [[\text{id}_{\nu F}, \text{id}_{\nu F}] \odot \langle\langle h_1, h_2 \rangle\rangle, h_3] \odot d_2 \in \mathcal{Kl}(T)(Y_2, \nu F) \end{aligned} \quad (4.5)$$

**Proof.** We show that  $p$  satisfies the axioms of progress measure (Definition 4.3.10).

$$\begin{array}{ccc}
\begin{array}{c}
\overline{F}[\langle\langle f_{11}^{(c_2)}, f_{12}^{(c_2)} \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle] \overline{F}[u_1(c_1), u_2(c_1)] \\
\uparrow d_{11} \quad \sqsupseteq \quad \uparrow c_1 \quad \sqsupseteq \quad J\zeta \uparrow \cong \\
Y_1 \xrightarrow{f_{11}^{(c_2)}} X_1 \xrightarrow{u_1(c_1+1)} \nu F
\end{array} &
\begin{array}{c}
\overline{F}[\langle\langle f_{11}^{(c_2)}, f_{12}^{(c_2)} \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle] \overline{F}[u_1(\bar{c}_1), u_2(\bar{c}_1)] \\
\uparrow d_{12} \quad \sqsupseteq \quad \uparrow c_2 \quad \sqsupseteq \quad J\zeta \uparrow \cong \\
Y_1 \xrightarrow{f_{12}^{(c_2+1)}} X_2 \xrightarrow{u_2(\bar{c}_1)} \nu F
\end{array} \\
\underbrace{\hspace{15em}}_{h_1(c_1, c_2)} & \underbrace{\hspace{15em}}_{h_2(c_1, c_2)} \\
\begin{array}{c}
\overline{F}[\langle\langle f_{11}^{(c_2)}, f_{12}^{(c_2)} \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle] \overline{F}[u_1(\bar{c}_1), u_2(\bar{c}_1)] \\
\uparrow d_2 \quad \sqsupseteq \quad \uparrow c \quad \sqsupseteq \quad J\zeta \uparrow \cong \\
Y_2 \xrightarrow{\langle\langle f_{21}, f_{22} \rangle\rangle} X \xrightarrow{[u_1(\bar{c}_1), u_2(\bar{c}_1)]} \nu F
\end{array} \\
\underbrace{\hspace{15em}}_{h_3(c_1, c_2)}
\end{array}$$

Figure 4.1: The progress measure  $p$  in Lemma 4.3.15, pictorially.

1. **(Monotonicity)** Assume  $\mathbf{c}_1 \leq \mathbf{c}'_1$  and  $\mathbf{c}_2 \leq \mathbf{c}'_2$ . Then by Assumption 1 in Theorem 4.3.14, Condition 1 in Definition 4.3.10 and that  $(f_{11}^{(a)})_{a \leq \bar{a}}$  and  $(f_{12}^{(a)})_{a \leq \bar{a}}$  are increasing sequences, we have:

$$h_1(\mathbf{c}_1, \mathbf{c}_2) = u_1(\mathbf{c}_1) \odot f_{11}^{(c_2)} \sqsubseteq u_1(\mathbf{c}'_1) \odot f_{11}^{(c'_2)} = h_1(\mathbf{c}'_1, \mathbf{c}'_2).$$

Hence  $h_1$  is monotone. Monotonicity of  $h_2$  and  $h_3$  are similarly proved.

2. **( $\mu$ -variables, base case)** By Condition B(c) in Definition 4.3.13 and Condition 4 in Theorem 4.3.14, we have:

$$\begin{aligned}
h_1(0, \mathbf{c}_2) &= u_1(0) \odot f_{11}^{(c_2)} = \perp \odot f_{11}^{(c_2)} = \perp \quad \text{and} \\
h_2(\mathbf{c}_1, 0) &= u_2(\mathbf{c}_1) \odot f_{12}^{(0)} = u_2(\mathbf{c}_1) \odot \perp = \perp.
\end{aligned}$$

3. **( $\mu$ -variables, step case)** We have the following (see also Figure 4.1).

$$\begin{aligned}
&h_1(\mathbf{c}_1 + 1, \mathbf{c}_2) \\
&= u_1(\mathbf{c}_1 + 1) \odot f_{11}^{(c_2)} \quad (\text{by the definition of } h_1(\mathbf{c}_1, \mathbf{c}_2)) \\
&\sqsubseteq (J\zeta)^{-1} \odot \overline{F}[u_1(\mathbf{c}_1), u_2(\mathbf{c}_1)] \odot c_1 \odot f_{11}^{(c_2)} \quad (p_{\mathcal{X}} \text{ is a progress measure}) \\
&\sqsubseteq (J\zeta)^{-1} \odot \overline{F}[u_1(\mathbf{c}_1), u_2(\bar{c}_1)] \odot c_1 \odot f_{11}^{(c_2)} \quad (\text{by the monotonicity of } p_{\mathcal{X}}) \\
&\sqsubseteq (J\zeta)^{-1} \odot \overline{F}[u_1(\mathbf{c}_1), u_2(\bar{c}_1)] \odot \overline{F}[\langle\langle f_{11}^{(c_2)}, f_{12}^{(c_2)} \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle] \odot d_{11} \\
&\quad (\text{by Condition B(b) in Definition 4.3.13}) \\
&= (J\zeta)^{-1} \odot \overline{F}[\text{id}_{\nu F}, \text{id}_{\nu F}] \odot \langle\langle u_1(\mathbf{c}_1) \odot f_{11}^{(c_2)}, u_2(\bar{c}_1) \odot f_{12}^{(c_2)} \rangle\rangle, \\
&\quad [u_1(\mathbf{c}_1), u_2(\bar{c}_1)] \odot \langle\langle f_{21}, f_{22} \rangle\rangle] \odot d_{11} \\
&\sqsubseteq (J\zeta)^{-1} \odot \overline{F}[\text{id}_{\nu F}, \text{id}_{\nu F}] \odot \langle\langle u_1(\mathbf{c}_1) \odot f_{11}^{(c_2)}, u_2(\bar{c}_1) \odot f_{12}^{(c_2)} \rangle\rangle, \\
&\quad [u_1(\bar{c}_1), u_2(\bar{c}_1)] \odot \langle\langle f_{21}, f_{22} \rangle\rangle] \odot d_{11} \\
&\quad (\text{by the monotonicity of } p_{\mathcal{X}}) \\
&= (J\zeta)^{-1} \odot \overline{F}[\text{id}_{\nu F}, \text{id}_{\nu F}] \odot \langle\langle h_1(\mathbf{c}_1, \mathbf{c}_2), h_2(\mathbf{c}_1, \mathbf{c}_2) \rangle\rangle, h_3(\mathbf{c}_1, \mathbf{c}_2) \rangle \odot d_{11} \\
&\quad (\text{by the definitions of } h_1(\mathbf{c}_1, \mathbf{c}_2), h_2(\mathbf{c}_1, \mathbf{c}_2) \text{ and } h_3(\mathbf{c}_1, \mathbf{c}_2)).
\end{aligned}$$

We can similarly prove that there exists an ordinal  $\mathbf{c}'_1$  such that

$$h_2(\mathbf{c}_1, \mathbf{c}_2 + 1) \sqsubseteq \overline{F}[\langle\langle h_1(\mathbf{c}'_1, \mathbf{c}_2), h_2(\mathbf{c}_1, \mathbf{c}_2) \rangle\rangle, h_3(\mathbf{c}_1, \mathbf{c}_2) \rangle] \odot d_{12}.$$

4. ( $\mu$ -variables, limit case) Let  $\mathbf{c}_1$  be a limit ordinal. By  $\overline{\mathbf{c}}_1 = \overline{\mathbf{b}}_1 \leq \omega$  and Condition 5 in Theorem 4.3.14, Kleisli composition in  $\mathcal{Kl}(T)$  is  $\mathbf{c}_1$ -continuous. Hence for each ordinal  $\mathbf{c}_2$ , we have:

$$\begin{aligned} h_1(\mathbf{c}_1, \mathbf{c}_2) &= u_1(\mathbf{c}_1) \odot f_{11}^{(\mathbf{c}_2)} \sqsubseteq \left( \bigsqcup_{\mathbf{c}'_1 < \mathbf{c}_1} u_1(\mathbf{c}'_1) \right) \odot f_{11}^{(\mathbf{c}_2)} \\ &= \bigsqcup_{\mathbf{c}'_1 < \mathbf{c}_1} (u_1(\mathbf{c}'_1) \odot f_{11}^{(\mathbf{c}_2)}) = \bigsqcup_{\mathbf{c}'_1 < \mathbf{c}_1} h_1(\mathbf{c}'_1, \mathbf{c}_2). \end{aligned}$$

We can similarly prove that for an ordinal  $\mathbf{c}_1$  and a limit ordinal  $\mathbf{c}_2$ ,

$$h_2(\mathbf{c}_1, \mathbf{c}_2) = \bigsqcup_{\mathbf{c}'_2 < \mathbf{c}_2} h_2(\mathbf{c}_1, \mathbf{c}'_2).$$

5. ( $\nu$ -variables) In a similar manner to the step case of  $\mu$ -variables, we have:

$$h_3(\mathbf{c}_1, \mathbf{c}_2) \sqsubseteq (J\zeta)^{-1} \odot \overline{F}[\langle\langle h_1(\overline{\mathbf{c}}_1, \overline{\mathbf{c}}_2), h_2(\overline{\mathbf{c}}_1, \overline{\mathbf{c}}_2) \rangle\rangle, h_3(\overline{\mathbf{c}}_1, \overline{\mathbf{c}}_2)] \odot d_2.$$

Hence  $p$  is a progress measure for the HES (4.5).  $\square$

**Lemma 4.3.16.** *We assume the assumptions 1–5 in Theorem 4.3.14. Let  $f: Y \rightarrow X$  be an  $\bar{\mathbf{a}}$ -bounded fair simulation with dividing from  $\mathcal{X} = (X, c, (X_1, X_2))$  to  $\mathcal{Y} = (Y, c, (Y_1, Y_2))$ . Let  $(h_1^{\text{sol}}, h_2^{\text{sol}}, h_3^{\text{sol}})$  be the solution of the HES (4.5) in Lemma 4.3.15. Then we have:*

$$[\text{id}_{\nu F}, \text{id}_{\nu F}] \odot \langle\langle h_1^{\text{sol}}, h_2^{\text{sol}} \rangle\rangle \sqsubseteq \text{tr}_1^{\text{B}}(d) \quad \text{and} \quad h_3^{\text{sol}} \sqsubseteq \text{tr}_2^{\text{B}}(d). \quad (4.6)$$

**Proof.** In general, the least fixed point of a monotone function  $f: L_1 \times L_2 \rightarrow L_1 \times L_2$  with respect to the product order can be calculated in a “pointwise” manner. That is, the least fixed point of  $f$  is given by the solution of the following HES:  $\begin{cases} u_1 =_{\mu} f_1(u_1, u_2) \in (L_1, \sqsubseteq_1) \\ u_2 =_{\mu} f_2(u_1, u_2) \in (L_2, \sqsubseteq_2) \end{cases}$ . The proof is easy.

Because of this, the solution  $v_1^{\text{sol}}, v_2^{\text{sol}}, v_3^{\text{sol}}$  of the HES (4.5) in Lemma 4.3.15 is equivalent to the solution  $w_1^{\text{sol}}, w_2^{\text{sol}}$  of the following HES, in the sense that  $w_1^{\text{sol}} = (v_1^{\text{sol}}, v_2^{\text{sol}})$  and  $w_2^{\text{sol}} = v_3^{\text{sol}}$ .

$$\begin{aligned} w_1 =_{\mu} \left( \begin{array}{l} (J\zeta)^{-1} \odot \overline{F}[[\text{id}_{\nu F}, \text{id}_{\nu F}] \odot \langle\langle w_{11}, w_{12} \rangle\rangle, w_2] \odot d_{11}, \\ (J\zeta)^{-1} \odot \overline{F}[[\text{id}_{\nu F}, \text{id}_{\nu F}] \odot \langle\langle w_{11}, w_{12} \rangle\rangle, w_2] \odot d_{12} \end{array} \right) \\ \in \mathcal{Kl}(T)(Y_1, \nu F) \times \mathcal{Kl}(T)(Y_1, \nu F) \end{aligned} \quad (4.7)$$

$$w_2 =_{\nu} (J\zeta)^{-1} \odot \overline{F}[[\text{id}_{\nu F}, \text{id}_{\nu F}] \odot \langle\langle w_{11}, w_{12} \rangle\rangle, w_2] \odot d_2 \in \mathcal{Kl}(T)(Y_2, \nu F)$$

Here  $w_{11}$  and  $w_{12}$  denote the first and the second component of  $w_1 \in \mathcal{Kl}(T)(Y_1, \nu F) \times \mathcal{Kl}(T)(Y_1, \nu F)$  respectively.

By the completeness of progress measure (Theorem 4.3.11.2), there exists a progress measure  $q = ((\mathbf{a}), (w_1(\mathbf{a}), w_2(\mathbf{a}))_{\mathbf{a} \leq \bar{\mathbf{a}}})$  for (4.7) such that  $w_1(\mathbf{a}) = w_1^{\text{sol}} = (v_1^{\text{sol}}, v_2^{\text{sol}})$  and  $w_2(\bar{\mathbf{a}}) = w_2^{\text{sol}} = v_3^{\text{sol}}$ .

For each  $\mathbf{a} \leq \bar{\mathbf{a}}$ , let  $(w_{11}(\mathbf{a}), w_{12}(\mathbf{a})) := w_1(\mathbf{a})$  and define  $v'_1(\mathbf{a}): Y_1 \rightarrow \nu F$  and  $v'_2(\mathbf{a}): Y_2 \rightarrow \nu F$  by  $v'_1(\mathbf{a}) = [\text{id}_{\nu F}, \text{id}_{\nu F}] \odot \langle\langle w_{11}(\mathbf{a}), w_{12}(\mathbf{a}) \rangle\rangle$  and  $v'_2(\mathbf{a}) = w_2(\mathbf{a})$ .

We now show that  $p' := ((\bar{\mathbf{a}}), (v'_1(\mathbf{a}), v'_2(\mathbf{a}))_{\mathbf{a} \leq \bar{\mathbf{a}}})$  is a progress measure for the following HES, which defines  $\text{tr}_1^{\text{B}}(d): Y_1 \rightarrow \nu F$  and  $\text{tr}_2^{\text{B}}(d): Y_2 \rightarrow \nu F$  (see Definition 3.3.1).

$$\begin{cases} u'_1 =_{\mu} J(\zeta^F)^{-1} \odot \overline{F}[u'_1, u'_2] \odot d_1 \in (\mathcal{Kl}(T)(Y_1, \nu F), \sqsubseteq) \\ u'_2 =_{\nu} J(\zeta^F)^{-1} \odot \overline{F}[u'_1, u'_2] \odot d_2 \in (\mathcal{Kl}(T)(Y_2, \nu F), \sqsubseteq) \end{cases} \quad (4.8)$$

1. **(Monotonicity)** By the monotonicity of  $w_1(\mathbf{a})$  and  $w_2(\mathbf{a})$ ,  $v'_1(\mathbf{a})$  and  $v'_2(\mathbf{a})$  are also monotone.
2. **( $\mu$ -variables, base case)** We have  $(w_{11}(0), w_{12}(0)) = w_1(0) = (\perp, \perp)$  by the definition. Hence by Lemma 4.3.5.3 and Condition 4 of Theorem 4.3.14, we have:

$$v'_1(0) = [\text{id}_{\nu F}, \text{id}_{\nu F}] \odot \langle\langle w_{11}(0), w_{12}(0) \rangle\rangle = [\text{id}_{\nu F}, \text{id}_{\nu F}] \odot \langle\langle \perp, \perp \rangle\rangle = \perp.$$

3. **( $\mu$ -variables, step case)** For an ordinal  $\mathbf{a} < \bar{\mathbf{a}}$ , we have:

$$\begin{aligned} & v'_1(\mathbf{a} + 1) \\ &= [\text{id}_{\nu F}, \text{id}_{\nu F}] \odot \langle\langle w_{11}(\mathbf{a} + 1), w_{12}(\mathbf{a} + 1) \rangle\rangle \quad (\text{by the definition of } v'_1(\mathbf{a} + 1)) \\ &\sqsubseteq [\text{id}_{\nu F}, \text{id}_{\nu F}] \odot \langle\langle (J\zeta)^{-1} \odot \overline{F}[\text{id}_{\nu F}, \text{id}_{\nu F}] \odot \langle\langle w_{11}(\mathbf{a}), w_{12}(\mathbf{a}) \rangle\rangle, w_2 \rangle \odot d_{11}, \\ &\quad (J\zeta)^{-1} \odot \overline{F}[\text{id}_{\nu F}, \text{id}_{\nu F}] \odot \langle\langle w_{11}(\mathbf{a}), w_{12}(\mathbf{a}) \rangle\rangle, w_2 \rangle \odot d_{12} \rangle \\ &\quad \quad \quad (q \text{ is a progress measure}) \\ &= (J\zeta)^{-1} \odot \overline{F}[\text{id}_{\nu F}, \text{id}_{\nu F}] \odot \langle\langle w_{11}(\mathbf{a}), w_{12}(\mathbf{a}) \rangle\rangle, w_2 \rangle \\ &\quad \odot [\text{id}_{FY}, \text{id}_{FY}] \odot \langle\langle d_{11}, d_{12} \rangle\rangle \\ &\sqsubseteq (J\zeta)^{-1} \odot \overline{F}[\text{id}_{\nu F}, \text{id}_{\nu F}] \odot \langle\langle w_{11}(\mathbf{a}), w_{12}(\mathbf{a}) \rangle\rangle, w_2 \rangle \odot d_1 \\ &\quad \quad \quad (d_{11}, d_{12} \text{ are a dividing of } d_1) \\ &= (J\zeta)^{-1} \odot \overline{F}[v'_1(\mathbf{a}), v'_2] \odot d_1 \quad (\text{by definition}). \end{aligned}$$

4. **( $\mu$ -variables, limit case)** For a limit ordinal  $\mathbf{a} \leq \bar{\mathbf{a}}$ , we have:

$$\begin{aligned} & v'_1(\mathbf{a}) \\ &= [\text{id}_{\nu F}, \text{id}_{\nu F}] \odot \langle\langle w_{11}(\mathbf{a}), w_{12}(\mathbf{a}) \rangle\rangle \quad (\text{by the definition of } v'_1(\mathbf{a})) \\ &\sqsubseteq [\text{id}_{\nu F}, \text{id}_{\nu F}] \odot \langle\langle \bigsqcup_{\mathbf{b} < \mathbf{a}} w_{11}(\mathbf{b}), \bigsqcup_{\mathbf{b} < \mathbf{a}} v'_{12}(\mathbf{b}) \rangle\rangle \quad (q \text{ is a progress measure}) \\ &= \bigsqcup_{\mathbf{b} < \mathbf{a}} [\text{id}_{\nu F}, \text{id}_{\nu F}] \odot \langle\langle w_{11}(\mathbf{b}), w_{12}(\mathbf{b}) \rangle\rangle \\ &\quad \quad \quad (\text{by Conditions 1 and 2 of Theorem 4.3.14}) \\ &= \bigsqcup_{\mathbf{b} < \mathbf{a}} v'_1(\mathbf{b}) \quad (\text{by definition}). \end{aligned}$$

5. **( $\nu$ -variables)** For an ordinal  $\mathbf{a} \leq \bar{\mathbf{a}}$ , there exists an ordinal  $\mathbf{b} \leq \bar{\mathbf{a}}$  such that:

$$\begin{aligned} & v'_2(\mathbf{a}) = w_2(\mathbf{a}) \quad (\text{by the definition of } v'_2(\mathbf{a})) \\ &\sqsubseteq (J\zeta)^{-1} \odot \overline{F}[\text{id}_{\nu F}, \text{id}_{\nu F}] \odot \langle\langle w_{11}(\mathbf{b}), w_{12}(\mathbf{b}) \rangle\rangle, w_2(\mathbf{b}) \rangle \odot d_2 \\ &\quad \quad \quad (q \text{ is a progress measure}) \\ &= (J\zeta)^{-1} \odot \overline{F}[v'_1(\mathbf{b}), v'_2(\mathbf{b})] \odot d_2 \quad (\text{by definition}). \end{aligned}$$

Hence  $p' = ((\bar{\mathbf{a}}), (v'_1(\mathbf{a}), v'_2(\mathbf{a}))_{\mathbf{a}})$  is a progress measure for the HES (4.8). By soundness of progress measures (Theorem 4.3.11.1), we have (4.6).  $\square$

**Proof** (Theorem 4.3.14). By Lemma 4.3.15 and the soundness of progress measures (Theorem 4.3.11.1), for the progress measure  $p_{\mathcal{X}} = ((\bar{\mathbf{b}}_1), (u_1(\mathbf{b}_1), u_2(\mathbf{b}_1))_{\mathbf{b}_1 \leq \bar{\mathbf{b}}_1})$  in Lemma 4.3.15, we have:

$$\begin{aligned} u_1(\bar{\mathbf{b}}_1) \odot f_{11}^{(\bar{\mathbf{a}}_1)} \sqsubseteq v_1^{\text{sol}}, \quad u_2(\bar{\mathbf{b}}_1) \odot f_{12}^{(\bar{\mathbf{a}}_1)} \sqsubseteq v_2^{\text{sol}}, \\ \text{and} \quad [u_1(\bar{\mathbf{b}}_1), u_2(\bar{\mathbf{b}}_1)] \odot \langle\langle f_{21}, f_{22} \rangle\rangle \sqsubseteq v_3^{\text{sol}}. \end{aligned} \quad (4.9)$$

By Lemma 4.3.16, we have:

$$[\text{id}_{\nu F}, \text{id}_{\nu F}] \odot [v_1^{\text{sol}}, v_2^{\text{sol}}] \sqsubseteq \text{tr}_1^{\text{B}}(d) \quad \text{and} \quad v_3^{\text{sol}} \sqsubseteq \text{tr}_1^{\text{B}}(d). \quad (4.10)$$

Therefore we have:

$$\begin{aligned}
& [\mathrm{tr}_1^{\mathbf{B}}(c), \mathrm{tr}_2^{\mathbf{B}}(c)] \odot \langle\langle f_{11}, f_{12} \rangle\rangle \\
&= [\mathrm{tr}_1^{\mathbf{B}}(c), \mathrm{tr}_2^{\mathbf{B}}(c)] \odot \langle\langle f_{11}^{\langle \bar{\mathbf{a}}_1 \rangle}, f_{12}^{\langle \bar{\mathbf{a}}_1 \rangle} \rangle\rangle && \text{(by Definition 4.3.13)} \\
&= [u_1(\bar{\mathbf{b}}_1), u_2(\bar{\mathbf{b}}_1)] \odot \langle\langle f_{11}^{\langle \bar{\mathbf{a}}_1 \rangle}, f_{12}^{\langle \bar{\mathbf{a}}_1 \rangle} \rangle\rangle && \text{(by definition)} \\
&= [\mathrm{id}_{\nu F}, \mathrm{id}_{\nu F}] \odot \langle\langle u_1(\bar{\mathbf{b}}_1) \odot f_{11}^{\langle \bar{\mathbf{a}}_1 \rangle}, u_2 \odot f_{12}^{\langle \bar{\mathbf{a}}_1 \rangle} \rangle\rangle \\
&\sqsubseteq [\mathrm{id}_{\nu F}, \mathrm{id}_{\nu F}] \odot \langle\langle v_1^{\mathrm{sol}}, v_2^{\mathrm{sol}} \rangle\rangle && \text{(by (4.9))} \\
&\sqsubseteq \mathrm{tr}_1^{\mathbf{B}}(d) && \text{(by (4.10)).}
\end{aligned}$$

In a similar manner, we can prove  $[\mathrm{tr}_1^{\mathbf{B}}(c), \mathrm{tr}_2^{\mathbf{B}}(c)] \odot \langle\langle f_{21}, f_{22} \rangle\rangle \sqsubseteq \mathrm{tr}_2^{\mathbf{B}}(d)$ . Hence we have:

$$\begin{aligned}
\mathrm{tr}^{\mathbf{B}}(c) \odot f &= [\mathrm{tr}_1^{\mathbf{B}}(c), \mathrm{tr}_2^{\mathbf{B}}(c)] \odot [\langle\langle f_{11}, f_{12} \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle] && \text{(by definition)} \\
&= [[\mathrm{tr}_1^{\mathbf{B}}(c), \mathrm{tr}_2^{\mathbf{B}}(c)] \odot \langle\langle f_{11}, f_{12} \rangle\rangle, [\mathrm{tr}_1^{\mathbf{B}}(c), \mathrm{tr}_2^{\mathbf{B}}(c)] \odot \langle\langle f_{21}, f_{22} \rangle\rangle] \\
&\sqsubseteq [\mathrm{tr}_1^{\mathbf{B}}(d), \mathrm{tr}_2^{\mathbf{B}}(d)] && \text{(by the discussions above)} \\
&= \mathrm{tr}^{\mathbf{B}}(d) && \text{(by definition). } \square
\end{aligned}$$

#### 4.3.4 Kleisli Fair Simulation without Dividing

The coalgebraic simulation notion introduced in the previous section was the one *with dividing*. It required that a coalgebra  $d_1: Y_1 \rightarrow \overline{F}Y$  is equipped with a dividing  $d_{11}, d_{12}: Y_1 \rightarrow FY$ . However, this “dividing requirement” is problematic: it is often the case that the dividing requirement prevents us from giving a meaningful Kleisli fair simulation between Büchi  $(T, F)$ -systems (see Example 4.5.4). Hence the following coalgebraic simulation notion is more desirable for us.

**Definition 4.3.17** (fair simulation without dividing). We assume the same situation as Definition 4.3.13. A *(Kleisli  $\bar{\mathbf{a}}$ -bounded) fair simulation without dividing* from  $\mathcal{X} = (X, c, (X_1, X_2))$  to  $\mathcal{Y} = (Y, d, (Y_1, Y_2))$  is defined almost the same way as one with dividing in Definition 4.3.13, except that Condition B is replaced by the following condition.

B' There exists a pair of increasing transfinite sequences The components  $f_{11}: Y_1 \rightarrow X_1$  and  $f_{12}: Y_1 \rightarrow X_2$  come

$$f_{11}^{(0)} \sqsubseteq f_{11}^{(1)} \sqsubseteq \dots \sqsubseteq f_{11}^{(\bar{\mathbf{a}})}: Y_1 \rightarrow X_1 \text{ and } f_{12}^{(0)} \sqsubseteq f_{12}^{(1)} \sqsubseteq \dots \sqsubseteq f_{12}^{(\bar{\mathbf{a}})}: Y_1 \rightarrow X_2,$$

that satisfies Conditions B(a), B(c) and B(e) in Definition 4.3.13 and the following two conditions.

(b')  $(f_{11}^{(\mathbf{a})})$  For each  $\mathbf{a}$ ,  $c_1 \odot f_{11}^{(\mathbf{a})} \sqsubseteq \overline{F}[\langle\langle f_{11}^{(\mathbf{a})}, f_{12}^{(\mathbf{a})} \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle] \odot d_1$ .

(d')  $(f_{12}^{(\mathbf{a})}, \text{ the step case})$  If  $\mathbf{a}$  is a successor ordinal, then  $c_2 \odot f_{12}^{(\mathbf{a})} \sqsubseteq \overline{F}[\langle\langle f_{11}^{(\mathbf{a}-1)}, f_{12}^{(\mathbf{a}-1)} \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle] \odot d_{12}$ .

$$\begin{array}{ccc}
FY & \xrightarrow{\overline{F}[\langle\langle f_{11}^{(\mathbf{a})}, f_{12}^{(\mathbf{a})} \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle]} & FX \\
d_1 \uparrow & \sqsupseteq & \uparrow c_1 \\
Y_1 & \xrightarrow{f_{11}^{(\mathbf{a})}} & X_1
\end{array}
\qquad
\begin{array}{ccc}
FY & \xrightarrow{\overline{F}[\langle\langle f_{11}^{(\mathbf{a})}, f_{12}^{(\mathbf{a})} \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle]} & FX \\
d_1 \uparrow & \sqsupseteq & \uparrow c_2 \\
Y_1 & \xrightarrow{f_{12}^{(\mathbf{a}+1)}} & X_2
\end{array}$$

However, the assumptions of Theorem 4.3.14 are not sufficient to make the above simulation notion sound. In the rest of this section, we present two additional assumptions that respectively make it sound.

## Soundness when Kleisli Arrows are Idempotent

**Proposition 4.3.18.** *Assume the assumptions 1–5 in Theorem 4.3.14. We further assume that each arrow  $f: X \rightarrow Y$  in  $\mathcal{Kl}(T)$  is idempotent, i.e. the codomain join  $\langle\langle f, f \rangle\rangle: X \rightarrow Y + Y$  always exists and  $[\text{id}_Y, \text{id}_Y] \odot \langle\langle f, f \rangle\rangle = f$ . Then if  $f: Y \rightarrow X$  is a simulation without dividing from  $\mathcal{X} = (X, c, (X_1, X_2))$  to  $\mathcal{Y} = (Y, c, (Y_1, Y_2))$  then  $\text{tr}^B(c) \odot f \sqsubseteq \text{tr}^B(d)$ .  $\square$*

The proposition above is immediate by the soundness of forward fair simulation with dividing (Theorem 4.3.14) and the following lemma, which is easily proved by the definitions of simulation with dividing and one without dividing (Definition 4.3.13 and Definition 4.3.17).

**Lemma 4.3.19.** *If all the arrows in  $\mathcal{Kl}(T)$  are idempotent then a simulation without dividing from  $\mathcal{X}$  to  $\mathcal{Y}$  is a simulation with dividing  $\mathcal{X}$  to  $\mathcal{Y}$ . Here the dividing for the latter is given by  $d_{11} = d_{12} = d_1$ .  $\square$*

## Soundness when $Y_1$ is Trapping

**Proposition 4.3.20.** *Assume the assumptions 1–5 in Theorem 4.3.14. We further assume that  $d_1 = \overline{F}(\text{id}_{Y_1} + \perp_{Y_2, Y_2}) \odot d_1$ . Then existence of a simulation  $f: Y \rightarrow X$  without dividing from  $\mathcal{X} = (X, c, (X_1, X_2))$  to  $\mathcal{Y} = (Y, c, (Y_1, Y_2))$  implies  $\text{tr}^B(c) \odot f \sqsubseteq \text{tr}^B(d)$ .*

Intuitively, the condition  $d_1 = \overline{F}(\text{id}_{Y_1} + \perp_{Y_2, Y_2}) \odot d_1$  means that there exists no transition from  $Y_1$  to  $Y_2$ .

**Proof.** We define progress measures

$$\begin{aligned} p_{\mathcal{X}} &= \left( (\overline{\mathbf{b}}_1), (u_1(\mathbf{b}_1): X_1 \rightarrow \nu F, u_2(\mathbf{b}_1): X_2 \rightarrow \nu F)_{\mathbf{b}_1 \leq \overline{\mathbf{b}}_1} \right) \quad \text{and} \\ p &= \left( (\overline{\mathbf{c}}_1, \overline{\mathbf{c}}_2), (h_1(\mathbf{c}_1, \mathbf{c}_2), h_2(\mathbf{c}_1, \mathbf{c}_2), h_3(\mathbf{c}_1, \mathbf{c}_2))_{\mathbf{c}_1 \leq \overline{\mathbf{c}}_1, \mathbf{c}_2 \leq \overline{\mathbf{c}}_2} \right) \end{aligned}$$

as in Lemma 4.3.15.

We shall prove that the following statement holds under the assumption that  $d_1 = \overline{F}(\text{id}_{Y_1} + \perp_{Y_2, Y_2}) \odot d_1$ .

$$\forall \mathbf{c}_2 \leq \overline{\mathbf{c}}_2. h_1(\overline{\mathbf{c}}_1, \mathbf{c}_2) = h_2(\overline{\mathbf{c}}_1, \mathbf{c}_2) = \perp. \quad (4.11)$$

To this end, we first prove the following by the transfinite induction on  $\mathbf{c}_1$ :

$$\forall \mathbf{c}_2 \leq \overline{\mathbf{c}}_2. (h_2(\overline{\mathbf{c}}_1, \mathbf{c}_2) = \perp \Rightarrow \forall \mathbf{c}_1 \leq \overline{\mathbf{c}}_1. h_1(\mathbf{c}_1, \mathbf{c}_2) = \perp). \quad (4.12)$$

Let  $\mathbf{c}_2 \leq \overline{\mathbf{c}}_2$  and assume  $h_2(\overline{\mathbf{c}}_1, \mathbf{c}_2) = \perp$ .

**(base case)** If  $\mathbf{c}_1 = 0$ , we have:

$$\begin{aligned} h_1(\mathbf{c}_1, \mathbf{c}_2) &= u_1(0) \odot f_{11}^{(\mathbf{c}_2)} && \text{(by definition)} \\ &= \perp \odot f_{11}^{(\mathbf{c}_2)} && (p_{\mathcal{X}} \text{ is a progress measure)} \\ &= \perp && \text{(by Condition 4 in Theorem 4.3.14).} \end{aligned}$$

**(step case)** Assume  $h_1(\mathbf{c}_1, \mathbf{c}_2) = \perp$ . Then we have:

$$\begin{aligned} &h_1(\mathbf{c}_1 + 1, \mathbf{c}_2) \\ &= u_1(\mathbf{c}_1 + 1) \odot f_{11}^{(\mathbf{c}_2)} && \text{(by definition)} \\ &\sqsubseteq J\zeta^{-1} \odot \overline{F}[u_1(\mathbf{c}_1), u_2(\mathbf{c}_1)] \odot c_1 \odot f_{11}^{(\mathbf{c}_2)} && (p_{\mathcal{X}} \text{ is a progress measure)} \end{aligned}$$



$$\begin{aligned}
& \sqsubseteq J\zeta^{-1} \odot \overline{F}[u_1(\mathbf{c}_1), u_2(\mathbf{c}_1)] \odot \overline{F}[\langle\langle f_{11}^{(\mathbf{c}_2)}, f_{12}^{(\mathbf{c}_2)} \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle] \odot d_1 \\
& \hspace{15em} (f \text{ is a fair simulation without dividing}) \\
& = J\zeta^{-1} \odot \overline{F}[u_1(\mathbf{c}_1), u_2(\mathbf{c}_1)] \odot \overline{F}[\langle\langle f_{11}^{(\mathbf{c}_2)}, f_{12}^{(\mathbf{c}_2)} \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle] \odot \overline{F}(\text{id} + \perp) \odot d_1 \\
& \hspace{15em} (\text{by the assumption}) \\
& = J\zeta^{-1} \odot \overline{F}\left[ [u_1(\mathbf{c}_1), u_2(\mathbf{c}_1)] \odot \langle\langle f_{11}^{(\mathbf{c}_2)}, f_{12}^{(\mathbf{c}_2)} \rangle\rangle \odot \text{id}, \right. \\
& \quad \left. [u_1(\mathbf{c}_1), u_2(\mathbf{c}_1)] \odot \langle\langle f_{21}, f_{22} \rangle\rangle \odot \perp \right] \odot d_1 \\
& = J\zeta^{-1} \odot \overline{F}\left[ [\text{id}, \text{id}] \odot \langle\langle u_1(\mathbf{c}_1) \odot f_{11}^{(\mathbf{c}_2)}, u_2(\mathbf{c}_1) \odot f_{12}^{(\mathbf{c}_2)} \rangle\rangle, \perp \right] \odot d_1 \\
& \hspace{15em} (\text{by Assumption 4 of Theorem 4.3.14}) \\
& = J\zeta^{-1} \odot \overline{F}\left[ [\text{id}, \text{id}] \odot \langle\langle h_1(\mathbf{c}_1, \mathbf{c}_2), h_2(\overline{\mathbf{c}}_1, \mathbf{c}_2) \rangle\rangle, \perp \right] \odot d_1 \quad (\text{by definition}) \\
& = J\zeta^{-1} \odot \overline{F}\left[ [\text{id}, \text{id}] \odot \langle\langle \perp, \perp \rangle\rangle, \perp \right] \odot d_1 \\
& \hspace{15em} (\text{by the induction hypothesis and the assumption}) \\
& = \perp \quad (\text{by the assumptions 1 and 4 of Theorem 4.3.14}).
\end{aligned}$$

**(limit case)** Assume that  $\mathbf{c}_1$  is a limit ordinal and we have  $h_1(\mathbf{c}'_1, \mathbf{c}_2) = \perp$  for each  $\mathbf{c}'_1 < \mathbf{c}_1$ . Then we have:

$$\begin{aligned}
h_1(\mathbf{c}_1, \mathbf{c}_2) &= u_1(\mathbf{c}_1) \odot f_{11}^{(\mathbf{c}_2)} && (\text{by definition}) \\
&\sqsubseteq \left( \bigsqcup_{\mathbf{c}'_1 < \mathbf{c}_1} u_1(\mathbf{c}'_1) \right) \odot f_{11}^{(\mathbf{c}_2)} && (p_{\mathcal{X}} \text{ is a progress measure}) \\
&= \bigsqcup_{\mathbf{c}'_1 < \mathbf{c}_1} \left( u_1(\mathbf{c}'_1) \odot f_{11}^{(\mathbf{c}_2)} \right) && (\text{by Assumption 1 of Theorem 4.3.14}) \\
&= \bigsqcup_{\mathbf{c}'_1 < \mathbf{c}_1} h_1(\mathbf{c}'_1, \mathbf{c}_2) && (\text{by definition}) \\
&= \perp && (\text{by the induction hypothesis}).
\end{aligned}$$

Hence (4.12) holds. We next prove (4.11) by the transfinite induction on  $\mathbf{c}_2$ .

**(base case)** If  $\mathbf{c}_2 = 0$ , we have:

$$\begin{aligned}
h_2(\overline{\mathbf{c}}_1, \mathbf{c}_2) &= u_2(\overline{\mathbf{c}}_1) \odot f_{12}^{(\mathbf{c}_2)} && (\text{by definition}) \\
&= u_2(\overline{\mathbf{c}}_1) \odot \perp && (p_{\mathcal{X}} \text{ is a progress measure}) \\
&= \perp && (\text{by Condition 4 in Theorem 4.3.14}).
\end{aligned}$$

By (4.12), we also have  $h_1(\overline{\mathbf{c}}_1, \mathbf{c}_2) = \perp$ .

**(step case)** Assume  $h_1(\mathbf{c}_1, \mathbf{c}_2) = h_2(\overline{\mathbf{c}}_1, \mathbf{c}_2) = \perp$ . Then we have:

$$\begin{aligned}
& h_2(\mathbf{c}_1, \mathbf{c}_2 + 1) \\
&= u_2(\overline{\mathbf{c}}_1) \odot f_{12}^{(\mathbf{c}_2+1)} && (\text{by definition}) \\
&\sqsubseteq J\zeta^{-1} \odot \overline{F}[u_1(\overline{\mathbf{c}}_1), u_2(\overline{\mathbf{c}}_1)] \odot c_1 \odot f_{11}^{(\mathbf{c}_2+1)} && (p_{\mathcal{X}} \text{ is a progress measure}) \\
&\sqsubseteq J\zeta^{-1} \odot \overline{F}[u_1(\overline{\mathbf{c}}_1), u_2(\overline{\mathbf{c}}_1)] \odot \overline{F}[\langle\langle f_{11}^{(\mathbf{c}_2)}, f_{12}^{(\mathbf{c}_2)} \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle] \odot d_1 \\
& \hspace{15em} (f \text{ is a forward fair simulation without dividing}) \\
&= J\zeta^{-1} \odot \overline{F}[u_1(\overline{\mathbf{c}}_1), u_2(\overline{\mathbf{c}}_1)] \odot \overline{F}[\langle\langle f_{11}^{(\mathbf{c}_2)}, f_{12}^{(\mathbf{c}_2)} \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle] \odot \overline{F}(\text{id} + \perp) \odot d_1 \\
& \hspace{15em} (\text{by the assumption}) \\
&= J\zeta^{-1} \odot \overline{F}\left[ [u_1(\overline{\mathbf{c}}_1), u_2(\overline{\mathbf{c}}_1)] \odot \langle\langle f_{11}^{(\mathbf{c}_2)}, f_{12}^{(\mathbf{c}_2)} \rangle\rangle \odot \text{id}, \right. \\
& \quad \left. [u_1(\overline{\mathbf{c}}_1), u_2(\overline{\mathbf{c}}_1)] \odot \langle\langle f_{21}, f_{22} \rangle\rangle \odot \perp \right] \odot d_1
\end{aligned}$$

$$\begin{aligned}
&= J\zeta^{-1} \odot \overline{F} \left[ [\text{id}, \text{id}] \odot \langle \langle u_1(\overline{\mathbf{c}}_1) \odot f_{11}^{(\mathbf{c}_2)}, u_2(\overline{\mathbf{c}}_1) \odot f_{12}^{(\mathbf{c}_2)} \rangle \rangle, \perp \right] \odot d_1 \\
&\quad \text{(by Assumption 4 of Theorem 4.3.14)} \\
&= J\zeta^{-1} \odot \overline{F} \left[ [\text{id}, \text{id}] \odot \langle \langle h_1(\overline{\mathbf{c}}_1, \mathbf{c}_2), h_2(\overline{\mathbf{c}}_1, \mathbf{c}_2) \rangle \rangle, \perp \right] \odot d_1 \quad \text{(by definition)} \\
&= J\zeta^{-1} \odot \overline{F} \left[ [\text{id}, \text{id}] \odot \langle \langle \perp, \perp \rangle \rangle, \perp \right] \odot d_1 \quad \text{(by the induction hypothesis)} \\
&= \perp \quad \text{(by the assumptions 1 and 4 of Theorem 4.3.14).}
\end{aligned}$$

By (4.12), we also have  $h_1(\mathbf{c}_1, \mathbf{c}_2) \sqsubseteq h_1(\overline{\mathbf{c}}_1, \mathbf{c}_2 + 1) = \perp$ .

**(limit case)** Assume that  $\mathbf{c}_2$  is a limit ordinal. Then we have:

$$\begin{aligned}
h_2(\overline{\mathbf{c}}_1, \mathbf{c}_2) &= u_2(\overline{\mathbf{c}}_1) \odot f_{12}^{(\mathbf{c}_2)} \quad \text{(by definition)} \\
&\sqsubseteq u_2(\overline{\mathbf{c}}_1) \odot \left( \bigsqcup_{\mathbf{c}'_2 < \mathbf{c}_2} f_{12}^{(\mathbf{c}'_2)} \right) \quad \text{(by definition)} \\
&= \bigsqcup_{\mathbf{c}'_2 < \mathbf{c}_2} \left( u_2(\overline{\mathbf{c}}_1) \odot f_{12}^{(\mathbf{c}'_2)} \right) \quad \text{(by Assumption 1 of Theorem 4.3.14)} \\
&= \bigsqcup_{\mathbf{c}'_2 < \mathbf{c}_2} h_2(\overline{\mathbf{c}}_1, \mathbf{c}_2) \quad \text{(by definition)} \\
&= \perp \quad \text{(by the induction hypothesis).}
\end{aligned}$$

Hence we have  $h_1(\overline{\mathbf{c}}_1, \mathbf{c}_2) = h_2(\overline{\mathbf{c}}_1, \mathbf{c}_2) = \perp$  for each  $\mathbf{c}_2 \leq \overline{\mathbf{c}}_2$ . Therefore we have:

$$\begin{aligned}
[\text{tr}_1^{\text{B}}(c), \text{tr}_2^{\text{B}}(c)] \odot \langle \langle f_{11}, f_{12} \rangle \rangle &= [u_1(\overline{\mathbf{b}}_1), u_2(\overline{\mathbf{b}}_1)] \odot \langle \langle f_{11}, f_{12} \rangle \rangle \\
&= [\text{id}_{\nu F}, \text{id}_{\nu F}] \odot \langle \langle h_1(\overline{\mathbf{c}}_1, \overline{\mathbf{c}}_2), h_2(\overline{\mathbf{c}}_1, \overline{\mathbf{c}}_2) \rangle \rangle = \perp \sqsubseteq \text{tr}_1^{\text{B}}(d). \quad (4.13)
\end{aligned}$$

Moreover, for  $h_3(\mathbf{c}_1, \mathbf{c}_2)$ , we have:

$$\begin{aligned}
h_3(\mathbf{c}_1, \mathbf{c}_2) &= [\text{tr}_1^{\text{B}}(c), \text{tr}_1^{\text{B}}(c)] \odot \langle \langle f_{21}, f_{22} \rangle \rangle \quad \text{(by definition)} \\
&\sqsubseteq J\zeta^{-1} \odot \overline{F} \left[ [\text{tr}_1^{\text{B}}(c), \text{tr}_1^{\text{B}}(c)] \odot [f_{11}, f_{12}], [\text{tr}_1^{\text{B}}(c), \text{tr}_1^{\text{B}}(c)] \odot [f_{21}, f_{22}] \right] \odot d_2 \\
&\quad \text{(similarly to the above)} \\
&= J\zeta^{-1} \odot \overline{F} [\perp, h_3(\mathbf{c}_1, \mathbf{c}_2)] \odot d_2 \quad \text{(by definition and the discussions above)} \\
&\sqsubseteq J\zeta^{-1} \odot \overline{F} [l_1^{(1)}(h_3(\mathbf{c}_1, \mathbf{c}_2)), h_3(\mathbf{c}_1, \mathbf{c}_2)] \odot d_2.
\end{aligned}$$

Here  $l_1^{(1)}: Y_1 \rightarrow \nu F$  denotes the interim solution (see Definition 2.3.6). By definition,  $\text{tr}_2^{\text{B}}(d): Y_2 \rightarrow \nu F$  is the greatest fixed point of the following function.

$$g \mapsto J\zeta^{-1} \odot \overline{F} [l_1^{(1)}(g), g] \odot d_2$$

Hence by the Knaster-Tarski theorem (Corollary 2.3.3), we have

$$[\text{tr}_1^{\text{B}}(c), \text{tr}_2^{\text{B}}(c)] \odot \langle \langle f_{21}, f_{22} \rangle \rangle = h_3(\mathbf{c}_1, \mathbf{c}_2) \sqsubseteq \text{tr}_2^{\text{B}}(d). \quad (4.14)$$

By (4.13) and (4.14), in a similar manner to the proof of Theorem 4.3.14, we have  $\text{tr}^{\text{B}}(c) \odot f \sqsubseteq \text{tr}^{\text{B}}(d)$ .  $\square$

#### 4.4 Kleisli Fair Simulation for NBTAs

We instantiate the framework developed in the previous section for nondeterministic Büchi tree automata (NBTAs). As expected, we obtain an almost the

same notion as the conventional fair simulation notion (Definition 4.1.3), except that while it is assumed that state spaces of NBTA's are finite in the original definitions in [50, 32, 117], we do not assume it.

Recall from Section 3.7 that we model an NBTA as a Büchi  $(\mathcal{P}, F_\Sigma)$ -system. The assumptions in Theorem 4.3.14 are satisfied.

**Proposition 4.4.1.** *If we let  $(T, F) = (\mathcal{P}, F_\Sigma)$ , the assumptions 1–5 in Theorem 4.3.14 are satisfied.*

**Proof.** The assumption 1 is already proved in [47].

For  $g : V \rightarrow X_1 + X_2$  and  $i \in \{1, 2\}$ ,  $g \upharpoonright^{X_i} : V \rightarrow X_i$  is given by  $g \upharpoonright^{X_i}(v) = g(v) \cap X_i$ . For  $g_1 : V \rightarrow X_1$  and  $g_2 : V \rightarrow X_2$ ,  $\langle\langle g_1, g_2 \rangle\rangle : V \rightarrow X_1 + X_2$  is given by  $\langle\langle g_1, g_2 \rangle\rangle(v) = g_1(v) \cup g_2(v)$ . For  $h_1 : X_1 \rightarrow W$  and  $h_2 : X_2 \rightarrow W$ ,  $[h_1, h_2] : X_1 + X_2 \rightarrow W$  is given by  $[h_1, h_2](x) = h_i(x)$  if  $x \in X_i$ . Using these characterizations, we can easily see that Assumption 2 is satisfied.

The above characterization also shows that codomain join is always defined. Hence Assumption 3 is satisfied.

The assumption 4 is immediate from that  $\perp : X \rightarrow Y$  is given by  $\perp(x) := \emptyset$  for each  $x \in X$ .

The assumption 5 is proved as follows: if  $f_i : X \rightarrow Y$  and  $g : Y \rightarrow Z$ , for  $x \in X$ , we have:

$$g \odot \left( \bigsqcup_{i < \mathfrak{a}} f_i \right)(x) = \bigcup_{y \in \bigcup_{i < \mathfrak{a}} f_i(x)} g(y) = \bigcup_{i < \mathfrak{a}} \bigcup_{y \in f_i(x)} g(y) = \bigsqcup_{i < \mathfrak{a}} (g \odot f_i(x)). \quad \square$$

Hence Kleisli fair simulation *with dividing* (Definition 4.3.13) is sound for Büchi  $(\mathcal{P}, F_\Sigma)$ -systems. The following lemma, together with Proposition 4.3.18, shows that Kleisli fair simulation *without dividing* (Definition 4.3.17) is also sound. The proof is easy.

**Lemma 4.4.2.** *Arrows in  $\mathcal{Kl}(\mathcal{P})$  are idempotent.* □

For Büchi  $(\mathcal{P}, F_\Sigma)$ -systems, we can further simplify the definition of simulation.

The definition of a Kleisli fair simulation  $f : X \rightarrow Y$  is very similar to that of a lattice-theoretic progress measure (Definition 4.3.10). However, in the definition of the former (Definition 4.3.17), the inequalities for “ $\mu$ -variables, step case” and “ $\nu$ -variables” were as follows.  $c_1 \odot f_{11}^{(\mathfrak{a})} \sqsubseteq \overline{F}[\langle\langle f_{11}^{(\mathfrak{a})}, f_{12}^{(\mathfrak{a})} \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle] \odot d_1$  and  $c_2 \odot f_{12}^{(\mathfrak{a})} \sqsubseteq \overline{F}[\langle\langle f_{11}^{(\mathfrak{a}-1)}, f_{12}^{(\mathfrak{a}-1)} \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle] \odot d_{12}$ . They are different from the corresponding conditions in Definition 4.3.10 where approximants appear on the left-hand side of inequalities alone. The former situation is somewhat problematic because it prevents us from calculating  $f_{11}^{(\mathfrak{a})}$  and  $f_{12}^{(\mathfrak{a})}$  in an inductive manner.

However, when  $T = \mathcal{P}$ , we can modify the inequalities and translate the definition of a Kleisli fair simulation in terms of progress measures so that  $f_{11}^{(\mathfrak{a})}$  and  $f_{12}^{(\mathfrak{a})}$  appear on the left-hand side alone. In the translation, the following property, which is specific to the powerset monad  $\mathcal{P}$ , is important. Intuitively, it claims that we can uniquely determine the *weakest precondition* of  $f : X \rightarrow Y$  in  $\mathcal{Kl}(\mathcal{P})$  with respect to the demonic nondeterminism. The proof is easy.

**Sublemma 4.4.3** (reversibility of Kleisli arrows). Let  $f : Y \rightarrow \nu F$  be an arrow in  $\mathcal{Kl}(\mathcal{P})$ . Define  $\square_f : \mathcal{Kl}(\mathcal{P})(X, \nu F) \rightarrow \mathcal{Kl}(\mathcal{P})(X, Y)$  by

$$\square_f(g)(x) := \{y \in Y \mid f(y) \subseteq g(x)\}.$$

$$\begin{array}{ccc} & & Y \\ & \nearrow \square_f(g) & \downarrow f \\ X & \xrightarrow{g} & \nu F \end{array}$$

Then we have the following:

1.  $f \odot \square_f(g) \sqsubseteq g$
2.  $\forall h : A \rightarrow B. f \odot h \sqsubseteq g \Rightarrow h \sqsubseteq \square_f(g)$ . □

Using  $\square_f$  defined above, we can rewrite Definition 4.3.17 as follows.

**Lemma 4.4.4.** *Let  $g_1^{\text{sol}} : Y_1 \rightarrow X_1$ ,  $g_2^{\text{sol}} : Y_1 \rightarrow X_2$ ,  $g_3^{\text{sol}} : Y_2 \rightarrow X_1$  and  $g_4^{\text{sol}} : Y_2 \rightarrow X_2$  be the solution of the following HES.*

$$\left\{ \begin{array}{l} g_1 =_{\nu} \square_{c_1}(\overline{F}[\langle\langle g_1, g_2 \rangle\rangle, \langle\langle g_3, g_4 \rangle\rangle] \odot d_1) \in \mathcal{Kl}(\mathcal{P})(Y_1, X_1) \\ g_2 =_{\mu} \square_{c_2}(\overline{F}[\langle\langle g_1, g_2 \rangle\rangle, \langle\langle g_3, g_4 \rangle\rangle] \odot d_1) \in \mathcal{Kl}(\mathcal{P})(Y_1, X_2) \\ g_3 =_{\nu} \square_{c_1}(\overline{F}[\langle\langle g_1, g_2 \rangle\rangle, \langle\langle g_3, g_4 \rangle\rangle] \odot d_2) \in \mathcal{Kl}(\mathcal{P})(Y_2, X_1) \\ g_4 =_{\nu} \square_{c_2}(\overline{F}[\langle\langle g_1, g_2 \rangle\rangle, \langle\langle g_3, g_4 \rangle\rangle] \odot d_2) \in \mathcal{Kl}(\mathcal{P})(Y_2, X_2) \end{array} \right. \quad (4.15)$$

Let  $g^{\text{sol}} = [\langle\langle g_1^{\text{sol}}, g_2^{\text{sol}} \rangle\rangle, \langle\langle g_3^{\text{sol}}, g_4^{\text{sol}} \rangle\rangle] : Y \rightarrow X$ . Then  $\text{tr}^{\text{B}}(c) \odot g^{\text{sol}} \sqsubseteq \text{tr}^{\text{B}}(d)$  if and only if there is a fair  $\bar{\alpha}$ -bounded simulation without dividing (Definition 4.3.17) from  $\mathcal{X}$  to  $\mathcal{Y}$  for some ordinal  $\bar{\alpha}$ .

**Proof.** We define  $f_{ji} : Y_j \rightarrow X_i$  as Definition 4.3.13.

( $\Rightarrow$ ) Assume  $\text{tr}^{\text{B}}(c) \odot g^{\text{sol}} \sqsubseteq \text{tr}^{\text{B}}(d)$ . By completeness of progress measures (Theorem 4.3.11.2), there exists a progress measure  $g = ((\bar{\alpha}), (g_i(\mathbf{a}))_{1 \leq i \leq 4, \mathbf{a} \leq \bar{\alpha}})$  for the HES (4.15) such that  $g_i^{\text{sol}} = g_i(\bar{\alpha})$  for each  $i$ . We define two sequences  $(f_{11}^{(\mathbf{a})} : Y_1 \rightarrow X_1)_{\mathbf{a} \leq \bar{\alpha}}$  and  $(f_{12}^{(\mathbf{a})} : Y_1 \rightarrow X_2)_{\mathbf{a} \leq \bar{\alpha}}$  of Kleisli arrows by  $f_{11}^{(\mathbf{a})} := g_1(\mathbf{a})$  and  $f_{12}^{(\mathbf{a})} := g_2(\mathbf{a})$ . We further define  $f : Y \rightarrow X$  by  $f = g^{\text{sol}}$ . We show that  $f$  is a fair simulation without dividing from  $\mathcal{X}$  to  $\mathcal{Y}$  whose approximation sequences are given by  $(f_{11}^{(\mathbf{a})})_{\mathbf{a} \leq \bar{\alpha}}$  and  $(f_{12}^{(\mathbf{a})})_{\mathbf{a} \leq \bar{\alpha}}$ .

We first show that  $f$  satisfies Condition A in Definition 4.3.13. We have:

$$\begin{aligned} c \odot f &= c \odot [\langle\langle f_{11}, f_{12} \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle] && \text{(by definition)} \\ &= c \odot [\langle\langle \square_{c_1}(\overline{F}f \odot d_1), \square_{c_2}(\overline{F}f \odot d_1) \rangle\rangle, \langle\langle \square_{c_1}(\overline{F}f \odot d_2), \square_{c_2}(\overline{F}f \odot d_2) \rangle\rangle] \\ &\quad (g_1^{\text{sol}}, g_2^{\text{sol}}, g_3^{\text{sol}} \text{ and } g_4^{\text{sol}} \text{ are the solution}) \\ &= [\langle\langle c_1 \odot \square_{c_1}(\overline{F}f \odot d_1), c_2 \odot \square_{c_2}(\overline{F}f \odot d_1) \rangle\rangle, \\ &\quad \langle\langle c_1 \odot \square_{c_1}(\overline{F}f \odot d_2), c_2 \odot \square_{c_2}(\overline{F}f \odot d_2) \rangle\rangle] \\ &\sqsubseteq [\langle\langle \overline{F}f \odot d_1, \overline{F}f \odot d_1 \rangle\rangle, \langle\langle \overline{F}f \odot d_2, \overline{F}f \odot d_2 \rangle\rangle] && \text{(by Sublemma 4.4.3.1)} \\ &= [\overline{F}f \odot d_1, \overline{F}f \odot d_2] && \text{(by Lemma 4.4.2)} \\ &= \overline{F}f \odot d. && (4.16) \end{aligned}$$

Next we show that  $f$  satisfies the Condition B' in Definition 4.3.17. Note that  $g = ((\bar{\alpha}), (g_i(\mathbf{a}))_{1 \leq i \leq 4, \mathbf{a} \leq \bar{\alpha}})$  satisfies the axioms of progress measure (Definition 4.3.10). It is immediate that this implies that conditions B(a), B(c) and B(e) in Definition 4.3.13 are satisfied. Moreover in a similar manner to (4.16) above, we can prove B'(b') and B'(d') in Definition 4.3.17. Therefore  $f$  is a fair  $\bar{\alpha}$ -bounded simulation without dividing.

( $\Leftarrow$ ) Conversely, let  $f : Y \rightarrow X$  be a fair simulation without dividing from  $\mathcal{X}$  to  $\mathcal{Y}$  whose approximation sequences are given by  $(f_{11}^{(\mathbf{a})})_{\mathbf{a} \leq \bar{\alpha}}$  and  $(f_{12}^{(\mathbf{a})})_{\mathbf{a} \leq \bar{\alpha}}$ . For each  $\mathbf{a} \leq \bar{\alpha}$ , we define arrows  $g_1(\mathbf{a}) : Y_1 \rightarrow X_1$ ,  $g_2(\mathbf{a}) : Y_1 \rightarrow X_2$ ,  $g_3(\mathbf{a}) : Y_2 \rightarrow X_1$  and  $g_4(\mathbf{a}) : Y_2 \rightarrow X_2$ , by  $g_1(\mathbf{a}) = f_{11}^{(\mathbf{a})}$ ,  $g_2(\mathbf{a}) = f_{12}^{(\mathbf{a})}$ ,  $g_3(\mathbf{a}) = f_{21}$  and  $g_4(\mathbf{a}) = f_{22}$ . Then by using Sublemma 4.4.3.2, we can show that Condition (A) in Definition 4.3.13, Condition (B') in Definition 4.3.17 and the monotonicity of  $f_{11}^{(\mathbf{a})}$  and  $f_{12}^{(\mathbf{a})}$  imply that  $g$  satisfies the axioms of a progress measure (Definition 4.3.10) with respect to the HES (4.15). □

By translating Lemma 4.4.4 with automata-theoretic terms, we can finally reach a fair simulation notion for NBTAs.

**Definition 4.4.5** (fair simulation for NBTAs). Let  $\mathcal{A} = (X, \tau, \text{Acc}_{\mathcal{A}})$  and  $\mathcal{B} = (Y, \sigma, \text{Acc}_{\mathcal{B}})$  be finite-state  $\Sigma$ -labeled NBTAs. Let  $X_1 := X \setminus \text{Acc}_{\mathcal{A}}$ ,  $X_2 := \text{Acc}_{\mathcal{A}}$ ,  $Y_1 := Y \setminus \text{Acc}_{\mathcal{B}}$  and  $Y_2 := \text{Acc}_{\mathcal{B}}$ . We define functions

$$\begin{aligned} \square_{\mathcal{X},i} &: \mathcal{P}(\left(\prod_{i \in \omega} \Sigma_i \times X^i\right) \times Y) \rightarrow \mathcal{P}(X_i \times Y), \\ \diamond_{\mathcal{Y},j} &: \mathcal{P}(\left(\prod_{i \in \omega} \Sigma_i \times X^i\right) \times \left(\prod_{i \in \omega} \Sigma_i \times Y^i\right)) \rightarrow \mathcal{P}(\prod_{i \in \omega} \Sigma_i \times X^i \times Y_j) \quad \text{and} \\ \bigwedge_{\Sigma} &: \mathcal{P}(X \times Y) \rightarrow \mathcal{P}(\left(\prod_{i \in \omega} \Sigma_i \times X^i\right) \times \left(\prod_{i \in \omega} \Sigma_i \times Y^i\right)) \end{aligned}$$

as follows.

$$\begin{aligned} \square_{\mathcal{X},i}(S) &:= \{(x, y) \in X_i \times Y \mid \forall (a, \mathbf{x}') \in \delta_{\mathcal{X}}(x). ((a, \mathbf{x}'), y) \in S\} \\ \diamond_{\mathcal{Y},j}(T) &:= \left\{ \begin{array}{l} ((a, \mathbf{x}'), y) \\ \in \left(\prod_{i \in \omega} \Sigma_i \times X^i\right) \times Y_j \end{array} \middle| \begin{array}{l} \exists (a', \mathbf{y}') \in \delta_{\mathcal{Y}}(y). \\ ((a, \mathbf{x}'), (a', \mathbf{y}')) \in T \end{array} \right\} \\ \bigwedge_{\Sigma}(U) &:= \left\{ \begin{array}{l} ((a, x_1, \dots, x_n), (a, y_1, \dots, y_n)) \\ \in \left(\prod_{i \in \omega} \Sigma_i \times X^i\right) \times \left(\prod_{i \in \omega} \Sigma_i \times Y^i\right) \end{array} \middle| \begin{array}{l} n \in \mathbb{N}, \\ \forall i. (x_i, y_i) \in U \end{array} \right\} \end{aligned}$$

A *fair simulation* from  $\mathcal{X}$  to  $\mathcal{Y}$  is a relation  $R \subseteq X \times Y$  such that  $R \subseteq u_1^{\text{sol}} \cup u_2^{\text{sol}} \cup u_3^{\text{sol}} \cup u_4^{\text{sol}}$  where  $u_1^{\text{sol}}, \dots, u_4^{\text{sol}}$  are the solution of the following HES.

$$\begin{cases} u_1 =_{\nu} \square_{\mathcal{X},1}(\diamond_{\mathcal{Y},1}(\bigwedge_{\Sigma}(u_1 \cup u_2 \cup u_3 \cup u_4))) \subseteq (X_1 \times Y_1, \subseteq) \\ u_2 =_{\mu} \square_{\mathcal{X},2}(\diamond_{\mathcal{Y},1}(\bigwedge_{\Sigma}(u_1 \cup u_2 \cup u_3 \cup u_4))) \subseteq (X_2 \times Y_1, \subseteq) \\ u_3 =_{\nu} \square_{\mathcal{X},1}(\diamond_{\mathcal{Y},2}(\bigwedge_{\Sigma}(u_1 \cup u_2 \cup u_3 \cup u_4))) \subseteq (X_1 \times Y_2, \subseteq) \\ u_4 =_{\nu} \square_{\mathcal{X},2}(\diamond_{\mathcal{Y},2}(\bigwedge_{\Sigma}(u_1 \cup u_2 \cup u_3 \cup u_4))) \subseteq (X_2 \times Y_2, \subseteq) \end{cases} \quad (4.17)$$

**Theorem 4.4.6** (soundness). *If  $R$  is a fair simulation from  $\mathcal{A}$  to  $\mathcal{B}$  in the sense of Definition 4.4.5,  $(x, y) \in R$  implies  $L_{\mathcal{A}}^{\text{B}}(x) \subseteq L_{\mathcal{B}}^{\text{B}}(y)$ .*

**Proof.** Let  $\mathcal{X} = (X, c, (X_1, X_2))$  and  $\mathcal{Y} = (Y, d, (Y_1, Y_2))$  be Büchi  $(\mathcal{P}, F_{\Sigma})$ -systems corresponding to  $\mathcal{A}$  and  $\mathcal{B}$  respectively (see Proposition 3.8.1).

Note that for each  $A, B \in \mathbf{Sets}$ , if we define a function  $\Delta_{A,B} : \mathcal{P}(A \times B) \rightarrow \mathcal{Kl}(\mathcal{P})(B, A)$  by  $\Delta_{A,B}(S)(b) := \{a \in A \mid (a, b) \in S\}$  then it is a bijection. Note also that  $F_{\Sigma}X = \prod_{i \in \omega} \Sigma_i \times X^i$  and  $F_{\Sigma}Y = \prod_{i \in \omega} \Sigma_i \times Y^i$ . It is easy to see that for the functions  $\square_f$  in Sublemma 4.4.3 and  $\square_{\mathcal{X},i}$ ,  $\diamond_{\mathcal{Y},j}$  and  $\bigwedge_{\Sigma}$  in Definition 4.4.5, the following properties hold.

$$\begin{aligned} \forall i \in 1, 2. \forall S \subseteq F_{\Sigma}X \times Y. \quad \Delta_{X_i, Y}(\square_{\mathcal{X},i}(S)) &= \square_{c_i}(\Delta_{F_{\Sigma}X, Y}(S)) \\ \forall j \in 1, 2. \forall T \subseteq F_{\Sigma}X \times F_{\Sigma}Y. \quad \Delta_{F_{\Sigma}X, Y_j}(\diamond_{\mathcal{Y},j}(T)) &= (\Delta_{F_{\Sigma}X, F_{\Sigma}Y}(T)) \odot d_j \\ \forall U \subseteq X \times Y. \quad \Delta_{F_{\Sigma}X, F_{\Sigma}Y}(\bigwedge_{\Sigma}(U)) &= \overline{F_{\Sigma}}(\Delta_{X, Y}(U)) \end{aligned}$$

Hence by Proposition 3.7.1, Proposition 4.3.18 and Lemma 4.4.4, we have:

$$\begin{aligned} R \subseteq X \times Y \text{ is a fair simulation from } \mathcal{X} \text{ to } \mathcal{Y} \text{ in the sense of Definition 4.4.5} \\ \Leftrightarrow R \subseteq u_1^{\text{sol}} \cup u_2^{\text{sol}} \cup u_3^{\text{sol}} \cup u_4^{\text{sol}} \text{ where } u_1^{\text{sol}}, \dots, u_4^{\text{sol}} \text{ are the solution of (4.17)} \\ \Leftrightarrow \Delta_{X, Y}(R) \subseteq [\langle\langle g_1^{\text{sol}}, g_2^{\text{sol}} \rangle\rangle, \langle\langle g_3^{\text{sol}}, g_4^{\text{sol}} \rangle\rangle] \\ \Leftrightarrow \Delta_{X, Y}(R) \text{ is an } \bar{\alpha}\text{-bounded fair simulation without dividing from } \mathcal{X} \text{ to } \mathcal{Y} \\ \text{for some } \bar{\alpha} \\ \Rightarrow \text{tr}^{\text{B}}(c) \odot \Delta_{X, Y}(R) \subseteq \text{tr}^{\text{B}}(d) \\ \Leftrightarrow \forall (x, y) \in R. L_{\mathcal{A}}^{\text{B}}(x) \subseteq L_{\mathcal{B}}^{\text{B}}(y). \quad \square \end{aligned}$$

By using the translation in Proposition 4.3.12, we can see that the simulation notion in Definition 4.4.5 is in fact essentially the same as the existing one (Definition 4.1.3), except that we do not assume that the state space is finite.

## 4.5 Kleisli Fair Simulation for PBTAs

We instantiate the simulation notion for PBTAs, which are modeled as Büchi  $(\mathcal{G}_s, F_\Sigma)$ -systems (see Proposition 3.8.1).

### 4.5.1 Kleisli Fair Simulation with Dividing for PBTAs

We first consider Kleisli fair simulation *with dividing*.

**Proposition 4.5.1.** *If we let  $(T, F) = (\mathcal{G}_s, F_\Sigma)$ , the assumptions 1–5 in Theorem 4.3.14 are satisfied.*

**Proof.** The assumption 1 is proved in a similar manner to the case when  $T = \mathcal{D}_s$ , which is proved in [47].

For  $g : V \rightarrow X_1 + X_2$  and  $i \in \{1, 2\}$ ,  $g \upharpoonright^{X_i} : V \rightarrow X_i$  is given by  $g \upharpoonright^{X_i}(v)(A) = g(v)(A)$ . For  $g_1 : V \rightarrow X_1$  and  $g_2 : V \rightarrow X_2$ ,  $\langle\langle g_1, g_2 \rangle\rangle : V \rightarrow X_1 + X_2$  is given by  $\langle\langle g_1, g_2 \rangle\rangle(v)(A) = g_1(v)(A \cap X_1) + g_2(v)(A \cap X_2)$  if  $g_1(v)(A \cap X_1) + g_2(v)(A \cap X_2) \leq 1$ , and it is undefined otherwise. For  $h_1 : X_1 \rightarrow W$  and  $h_2 : X_2 \rightarrow W$ ,  $[h_1, h_2] : X_1 + X_2 \rightarrow W$  is given by  $[h_1, h_2](x) = h_i(x)$  if  $x \in X_i$ . Using these characterizations, we can easily see that the assumptions 2–3 are satisfied.

The assumption 4 is immediate from that  $\perp : X \rightarrow Y$  is given by  $\perp(x)(A) := 0$  for each  $x \in X$  and  $A \in \mathfrak{F}_Y$ . The assumption 5 is proved using the dominated convergence theorem (see e.g. [8, Theorem 1.6.9]).  $\square$

By translating Definition 4.3.13 and Theorem 4.3.14 using automata-theoretic terms, we obtain the following simulation notion and soundness theorem.

**Definition 4.5.2** (fair simulation for PBTAs). Let  $\Sigma$  be a ranked alphabet,  $\mathcal{A} = ((X, \mathfrak{F}_X), \xi, \text{Acc}_{\mathcal{A}})$  and  $\mathcal{B} = ((Y, \mathfrak{F}_Y), \theta, \text{Acc}_{\mathcal{B}})$  be  $\Sigma$ -labeled PBTAs, and  $\bar{\alpha}$  be an ordinal. Let  $X_1 := X \setminus \text{Acc}_{\mathcal{A}}$ ,  $X_2 := \text{Acc}_{\mathcal{A}}$ ,  $Y_1 := Y \setminus \text{Acc}_{\mathcal{B}}$  and  $Y_2 := \text{Acc}_{\mathcal{B}}$ , and  $\mathfrak{F}_{X_1}$ ,  $\mathfrak{F}_{X_2}$ ,  $\mathfrak{F}_{Y_1}$  and  $\mathfrak{F}_{Y_2}$  be the canonical  $\sigma$ -algebras on them. An  $(\bar{\alpha}$ -bounded) *fair simulation with dividing* from  $\mathcal{A}$  to  $\mathcal{B}$  is a measurable function  $f : (Y, \mathfrak{F}_Y) \rightarrow \mathcal{G}_s(X, \mathfrak{F}_X)$  that satisfies the following conditions (for  $i, j \in \{1, 2\}$ , we define  $f_{ji} : (Y_j, \mathfrak{F}_{Y_j}) \rightarrow \mathcal{G}_s(X_i, \mathfrak{F}_{X_i})$  by  $f_{ji}(y)(A) := f(y)(A \cap X_i)$  for  $y \in Y_j$  and  $A \in \mathfrak{F}_{X_i}$ ).

A. For each  $y \in Y$ ,  $n \in \mathbb{N}$ ,  $a \in \Sigma_n$  and  $A_1, \dots, A_n \in \mathfrak{F}_X$ , we have:

$$\int_{x \in X} \tau(x)(\{a\} \times A_1 \times \dots \times A_n) f(y)(dx) \leq \int_{y_1, \dots, y_n \in Y} f(y_1)(A_1) \cdot \dots \cdot f(y_n)(A_n) \cdot \theta(y)(\{a\} \times dy_1 \times \dots \times dy_n).$$

B. There exists a pair  $\theta_{11}, \theta_{12} : Y_1 \rightarrow \mathcal{G}_s(\prod_{i \in \mathbb{N}} \Sigma_n \times Y^n)$  of measurable functions such that  $\theta_{11}(y)(A) + \theta_{12}(y)(A) = \theta(y)(A)$  for each  $y \in Y_1$  and  $A \in \mathfrak{F}_{\prod_{i \in \mathbb{N}} \Sigma_n \times Y^n}$ . There also exist increasing transfinite sequences

$$f_{11}^{(0)} \leq f_{11}^{(1)} \leq \dots \leq f_{11}^{(\bar{\alpha})} : Y_1 \rightarrow \mathcal{G}_s X_1 \text{ and } f_{12}^{(0)} \leq f_{12}^{(1)} \leq \dots \leq f_{12}^{(\bar{\alpha})} : Y_1 \rightarrow \mathcal{G}_s X_2,$$

of measurable functions with respect to the pointwise order such that the following conditions are satisfied:

(a) **(Approximate  $f_{11}$  and  $f_{12}$ )** We have  $f_{11}^{(\bar{\alpha})} = f_{11}$  and  $f_{12}^{(\bar{\alpha})} = f_{12}$ .

(b) ( $f_{11}^{(a)}$ ) For each  $\mathbf{a}$ ,  $y \in Y_1$  and  $A_1, \dots, A_n \in \mathfrak{F}_X$ ,

$$\int_{x \in X_1} \tau(x)(\{a\} \times A_1 \times \dots \times A_n) f_{11}^{(a)}(y)(dx) \leq \int_{y_1, \dots, y_n \in Y} f^{(a)}(y_1)(A_1) \dots f^{(a)}(y_n)(A_n) \cdot \theta_{11}(y)(\{a\} \times dy_1 \times \dots \times dy_n).$$

Here  $f^{(a)}: Y \rightarrow \mathcal{G}_s X$  is defined by

$$f^{(a)}(y)(A) := \begin{cases} f_{11}^{(a)}(y)(A) + f_{12}^{(a)}(y)(A) & (y \in Y_1) \\ f_{21}^{(a)}(y)(A) + f_{22}^{(a)}(y)(A) & (y \in Y_2). \end{cases}$$

(c) ( $f_{12}^{(a)}$ , **the base case**) If  $\mathbf{a} = 0$ , then  $f_{12}^{(a)}(y)(X_2) = 0$  for each  $y \in Y_1$ .

(d) ( $f_{12}^{(a)}$ , **the step case**) If  $\mathbf{a}$  is a successor ordinal, then for each  $y \in Y_1$  and  $A_1, \dots, A_n \in \mathfrak{F}_X$ ,

$$\int_{x \in X_2} \tau(x)(\{a\} \times A_1 \times \dots \times A_n) f_{12}^{(a)}(y)(dx) \leq \int_{y_1, \dots, y_n \in Y} f^{(a-1)}(y_1)(A_1) \dots f^{(a-1)}(y_n)(A_n) \cdot \theta_{12}(y)(\{a\} \times dy_1 \times \dots \times dy_n).$$

Here  $f^{(a)}$  is defined as above.

(e) ( $f_{12}^{(a)}$ , **the limit case**) If  $\mathbf{a}$  is a limit ordinal, then for each  $y \in Y_1$  and  $A \in \mathfrak{F}_{X_2}$ ,  $f_{12}^{(a)}(y)(A) \leq \bigvee_{\mathbf{a}' < \mathbf{a}} f_{12}^{(\mathbf{a}')} (y)(A)$ .

**Theorem 4.5.3.** *If  $f: Y \rightarrow \mathcal{G}_s X$  is a fair simulation with dividing from  $\mathcal{A} = ((X, \mathfrak{F}_X), \xi, \text{Acc}_{\mathcal{A}})$  to  $\mathcal{B} = ((Y, \mathfrak{F}_Y), \theta, \text{Acc}_{\mathcal{B}})$ , then for each  $y \in Y$  and  $A \in \mathfrak{F}_{\text{Tree}_{\Sigma}^{\infty}}$ , we have:*

$$\int_{x \in X} L_{\mathcal{A}}^{\mathbf{B}}(x)(A) df(y) \leq L_{\mathcal{B}}^{\mathbf{B}}(y)(A). \quad \square$$

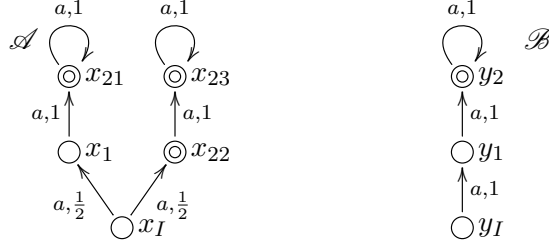
We can check a sort of “quantitative language inclusion” between PBTAs using the above simulation notion. For example, suppose that  $\mathcal{A}$  and  $\mathcal{B}$  are equipped with the initial states  $x_I \in X$  and  $y_I \in Y$ . If  $f(y_I)(\{x_I\}) = 1$ , then by the above theorem, we have  $L_{\mathcal{A}}^{\mathbf{B}}(x_I)(A) \leq L_{\mathcal{B}}^{\mathbf{B}}(y_I)(A)$  for each  $A \in \mathfrak{F}_{\text{Tree}_{\Sigma}^{\infty}}$ .

## 4.5.2 Kleisli Fair Simulation without Dividing for PBTAs

The simulation notion defined above is equipped with the “dividing requirement.” By instantiating Definition 4.3.17 for Büchi  $(\mathcal{G}_s, F_{\Sigma})$ -systems, we can also define a notion of fair simulation *without* dividing for PBTAs. The following example shows that the simulation notion with dividing is problematic compared to the one without dividing in the following sense: there exists a pair of PBTAs such that: (i) they exhibit quantitative language inclusion; (ii) a fair simulation *without* dividing exists between them; and (iii) a fair simulation *with* dividing does not exist between them.

**Example 4.5.4.** We define a ranked alphabet  $\mathbf{A}$  by  $\mathbf{A} := \{a\}$  and  $|a| := 1$ . Let  $\mathcal{A} = ((X, \mathfrak{F}_X), \xi, \text{Acc}_{\mathcal{A}})$  and  $\mathcal{B} = ((Y, \mathfrak{F}_Y), \theta, \text{Acc}_{\mathcal{B}})$  be  $\mathbf{A}$ -labeled PBTAs (i.e.

PBWAs) illustrated below ( $\odot$  denotes an accepting state).



Note that  $\text{Tree}_{\mathbf{A}}^{\infty} = \{a^{\omega}\}$ . It is easy to see that  $L_{\mathcal{A}}^{\mathbf{B}}(x_I)(\{a^{\omega}\}) = L_{\mathcal{B}}^{\mathbf{B}}(y_I)(\{a^{\omega}\}) = 1$  and hence they exhibit quantitative language inclusion. Suppose that we define  $f: Y \rightarrow \mathcal{G}_s X$  by  $f(y_I)(\{x_I\}) = 1$ ,

$$f(y_1)(\{x\}) := \begin{cases} \frac{1}{2} & (x \in \{x_1, x_{22}\}) \\ 0 & (\text{otherwise}) \end{cases} \quad \text{and} \quad f(y_2)(\{x\}) := \begin{cases} \frac{1}{2} & (x \in \{x_{21}, x_{23}\}) \\ 0 & (\text{otherwise}), \end{cases}$$

and  $f(y)(\{x\}) = 0$  for the other combinations. Then  $f$  is a fair simulation *without* dividing from  $\mathcal{A}$  to  $\mathcal{B}$  (see Proposition 3.8.1). In contrast,  $f$  is not a fair simulation *with* dividing. In fact, there exists no fair simulation with dividing from  $\mathcal{A}$  to  $\mathcal{B}$ .

Hence a simulation notion without dividing is desirable. We found that a fair simulation without dividing is sound when we focus on probabilistic *word* automata with a *finite* state space.

**Theorem 4.5.5.** *Let  $\mathcal{A} = ((X, \mathfrak{F}_X), \xi, \text{Acc}_{\mathcal{A}})$  and  $\mathcal{B} = ((Y, \mathfrak{F}_Y), \theta, \text{Acc}_{\mathcal{B}})$  be  $\Sigma$ -labeled PBTA's. We assume the following conditions.*

1.  $\mathcal{A}$  and  $\mathcal{B}$  are PBWA's, i.e.  $|a| = 1$  for each  $a \in \Sigma$ .
2.  $(Y, \mathfrak{F}_Y)$  is a finite set equipped with the discrete  $\sigma$ -algebras.

We write  $\mathbf{A}$  for the underlying set of  $\Sigma$ . Then if  $f: Y \rightarrow \mathcal{G}_s X$  is a fair simulation without dividing from  $\mathcal{A}$  to  $\mathcal{B}$ , for each  $y \in Y$  and  $A \in \mathfrak{F}_{\text{Tree}_{\Sigma}^{\infty}} = \mathfrak{F}_{\mathbf{A}^{\omega}}$ , we have:

$$\int_{x \in X} L_{\mathcal{A}}^{\mathbf{B}}(x)(A) \, df(y) \leq L_{\mathcal{B}}^{\mathbf{B}}(y)(A). \quad (4.18)$$

The key lemma to prove the above theorem is given below. It tells us that under the assumptions in the theorem, we can modify the PBWA  $\mathcal{B}$  to a PBWA  $\mathcal{B}'$  with the same state space such that: (i) the languages  $L_{\mathcal{B}}^{\mathbf{B}}$  and  $L_{\mathcal{B}'}^{\mathbf{B}}$  coincide; (ii) a function  $f: Y \rightarrow \mathcal{G}_s X$  is a fair simulation without dividing from  $\mathcal{A}$  to  $\mathcal{B}$  if and only if it is one from  $\mathcal{A}$  to  $\mathcal{B}'$ ; and (iii) the assumption of Proposition 4.3.20 is satisfied by  $\mathcal{B}'$ .

**Lemma 4.5.6.** *We assume the assumptions in Theorem 4.5.5. Let  $y_{>0} \in Y_1$  and assume  $\theta(y_{>0})(\mathbf{A} \times Y_2) > 0$ . We define an  $\mathbf{A}$ -labeled PBWA  $\mathcal{B}' = ((Y', \mathfrak{F}_{Y'}), \theta', \text{Acc}_{\mathcal{B}'})$  by  $Y' = Y$ ,  $\theta' = \theta$  and  $\text{Acc}_{\mathcal{B}'} = \text{Acc}_{\mathcal{B}} \cup \{y_{>0}\}$ . Then we have:*

- I. for each  $y \in Y$  and  $A \in \mathfrak{F}_{\mathbf{A}^{\omega}}$ ,  $L_{\mathcal{B}}^{\mathbf{B}}(y)(A) = L_{\mathcal{B}'}^{\mathbf{B}}(y)(A)$ ; and
- II. if  $f: Y \rightarrow \mathcal{G}_s X$  is a fair simulation without dividing from  $\mathcal{A}$  to  $\mathcal{B}$ , then it is so from  $\mathcal{A}$  to  $\mathcal{B}'$ .



**Proof.** By “forgetting” the labels, we can induce a Markov chain from the PBWA  $\mathcal{B}$ . More concretely, we define a Markov chain  $\mathcal{M}_{\mathcal{B}}$  whose state space is given by  $Y_{\perp} = Y + \{\perp\}$  and transition function  $\bar{\theta}: Y_{\perp} \times Y_{\perp} \rightarrow [0, 1]$  is given by

$$\bar{\theta}(y, y') = \begin{cases} \sum_{a \in A} \theta(y)(\{(a, y')\}) & (y, y' \in Y) \\ 1 - \sum_{y' \in Y} \sum_{a \in A} \theta(y)(\{(a, y')\}) & (y \in Y, y' = \perp) \\ 1 & (y = y' = \perp) \\ 0 & (y = \perp, y' \in Y). \end{cases}$$

We define  $\mathcal{M}_{\mathcal{B}'}$  similarly.

A subset  $B \subseteq Y$  is called a *strongly connected component* (SCC for short) if for all  $y, y' \in S$ , there exist  $y_0, y_1, \dots, y_n$  such that  $y_0 = y$ ,  $y_n = y'$  and  $\bar{\theta}(y_i, y_{i+1}) > 0$  for each  $i$ . An SCC  $B$  is called a *bottom strongly connected component* (BSCC for short) if  $\bar{\theta}(y, y') = 0$  for each  $y \in B$  and  $y' \notin B$ . See e.g. [12] for more details.

For  $y \in Y$  and  $Y' \subseteq Y$ , we write  $\Pr(y \models \text{GF}Y')$  for the probability where a state in  $Y'$  is visited infinitely often on  $\mathcal{M}_{\mathcal{B}}$  from  $y$ . By definition, we have

$$\begin{aligned} L_{\mathcal{B}}^{\text{B}}(y)(A^{\omega}) &= \Pr(y \models \text{GF} \text{Acc}_{\mathcal{B}}) \quad \text{and} \\ L_{\mathcal{B}'}^{\text{B}}(y)(A^{\omega}) &= \Pr(y \models \text{GF}(\text{Acc}_{\mathcal{B}} + \{y_{>0}\})). \end{aligned}$$

We define  $U, U' \subseteq Y$  by

$$\begin{aligned} U &:= \bigcup \{B \subseteq Y \mid B \text{ is a BSCC and } B \cap \text{Acc}_{\mathcal{B}} \neq \emptyset\} \quad \text{and} \\ U' &:= \bigcup \{B \subseteq Y \mid B \text{ is a BSCC and } B \cap (\text{Acc}_{\mathcal{B}} + \{y_{>0}\}) \neq \emptyset\}. \end{aligned}$$

We write  $\Pr(y \models \text{FU})$  for the probability where a state in  $U$  is reached from  $y$ . It is known that  $\Pr(y \models \text{GF} \text{Acc}_{\mathcal{B}}) = \Pr(y \models \text{FU})$  (see e.g. [12, Corollary 10.34]). Similarly, we have  $\Pr(y \models \text{GF}(\text{Acc}_{\mathcal{B}} + \{y_{>0}\})) = \Pr(y \models \text{FU}')$ .

Assume that  $y_{>0} \in B$  for some BSCC  $B$  in  $\mathcal{M}_{\mathcal{B}}$ . As  $B$  is a BSCC, it has no outgoing transition. Moreover, by  $\theta(y_{>0})(A \times Y_2) > 0$ ,  $y_{>0}$  has an accepting successor state. Hence we have  $B \cap \text{Acc}_{\mathcal{B}} \neq \emptyset$ , and by the definition of  $U$  this implies  $U = U'$ .

If  $y_{>0} \notin B$  for any BSCC  $B$ , then by the definitions of  $U$  and  $U'$  we have  $U = U'$ .

Therefore in both cases, for each  $y \in Y$ , we have:

$$\begin{aligned} L_{\mathcal{B}}^{\text{B}}(y)(A^{\omega}) &= \Pr(y \models \text{GF} \text{Acc}_{\mathcal{B}}) = \Pr(y \models \text{FU}) \\ &= \Pr(y \models \text{FU}') = \Pr(y \models \text{GF}(\text{Acc}_{\mathcal{B}} + \{y_{>0}\})) = L_{\mathcal{B}'}^{\text{B}}(y)(A^{\omega}). \end{aligned}$$

It remains to prove  $L_{\mathcal{B}}^{\text{B}}(y)(A) = L_{\mathcal{B}'}^{\text{B}}(y)(A)$  for each measurable set  $A \subseteq A^{\omega}$ . To this end, by the Kolmogorov extension theorem (see [103] for example), it suffices to prove  $L_{\mathcal{B}}^{\text{B}}(y)(wA^{\omega}) = L_{\mathcal{B}'}^{\text{B}}(y)(wA^{\omega})$  for each  $w \in A^*$ .

We inductively define a function  $\chi_{\mathcal{B}}: Y \times A^* \rightarrow \mathcal{G}_s Y$  by

$$\begin{aligned} \chi_{\mathcal{B}}(y, \langle \rangle)(\{y'\}) &= \begin{cases} 1 & (y = y') \\ 0 & (\text{otherwise}), \end{cases} \quad \text{and} \\ \chi_{\mathcal{B}}(y, aw)(\{y'\}) &= \sum_{y'' \in Y} d(y)(\{(a, y'')\}) \cdot \chi_{\mathcal{B}}(y'', w)(\{y'\}) \end{aligned}$$

where  $a \in A$  and  $w \in A^*$ .

Then for each  $y \in Y$  and  $w \in \mathbf{A}^*$ , we have:

$$\begin{aligned} L_{\mathcal{B}}^{\mathbf{B}}(y)(w\mathbf{A}^\omega) &= \sum_{y' \in Y} \chi_{\mathcal{B}}(y, w)(y') \cdot L_{\mathcal{B}}^{\mathbf{B}}(y)(y')(\mathbf{A}^\omega) \\ &= \sum_{y' \in Y} \chi_{\mathcal{B}'}(y, w)(y') \cdot L_{\mathcal{B}'}^{\mathbf{B}}(y)(y')(\mathbf{A}^\omega) \\ &= L_{\mathcal{B}'}^{\mathbf{B}}(y)(w\mathbf{A}^\omega). \end{aligned}$$

By the Kolmogorov extension theorem, this implies  $L_{\mathcal{B}}^{\mathbf{B}}(y)(A) = L_{\mathcal{B}'}^{\mathbf{B}}(y)(A)$  for each measurable set  $A$ . Hence we have  $L_{\mathcal{B}}^{\mathbf{B}} = L_{\mathcal{B}'}^{\mathbf{B}}$ .

It is immediate by definition that  $f$  is also a fair simulation without dividing from  $\mathcal{A}$  to  $\mathcal{B}'$ .  $\square$

**Proof** (Theorem 4.5.5). We define  $Y_{12} \subseteq Y \setminus \text{Acc}_{\mathcal{B}}$  by

$$Y_{12} := \left\{ y \notin \text{Acc}_{\mathcal{B}} \mid \exists y_0, \dots, y_n \in Y. \left( y = y_0, y_n \in \text{Acc}_{\mathcal{B}}, \text{ and } \forall i \in \{0, \dots, n-1\}. \theta(y_i)(\mathbf{A} \times \{y_{i+1}\}) > 0 \right) \right\}.$$

As  $Y_1$  is finite,  $Y_{12}$  is also finite. We define an  $\mathbf{A}$ -labeled PBWA  $\mathcal{B}' = ((Y', \mathfrak{F}_{Y'}), \theta', \text{Acc}_{\mathcal{B}'})$  by  $Y' = Y$ ,  $\mathfrak{F}_{Y'} = \mathfrak{F}_Y$ ,  $\theta' = \theta$  and  $\text{Acc}_{\mathcal{B}'} = \text{Acc}_{\mathcal{B}} \cup Y_{12}$ . As  $Y_{12}$  is finite, by repeatedly applying Lemma 4.5.6, we can prove  $L_{\mathcal{B}}^{\mathbf{B}} = L_{\mathcal{B}'}^{\mathbf{B}}$  and that  $f$  is a fair simulation without dividing from  $\mathcal{A}$  to  $\mathcal{B}'$ .

Let  $\mathcal{X} = (X, c, (X_1, X_2))$  and  $\mathcal{Y}' = (Y', d', (Y'_1, Y'_2))$  be the corresponding Büchi  $(\mathcal{G}_s, F_{\mathbf{A}})$ -systems to  $\mathcal{A}$  and  $\mathcal{B}'$ . Then  $f: Y \rightarrow X$  is a fair simulation without dividing from  $\mathcal{A}$  to  $\mathcal{B}'$ . By definition, we have  $\text{tr}^{\mathbf{B}}(y)(Y_2) = 0$  for each  $y \in Y_1'$ , and therefore by Proposition 4.3.20, soundness of fair simulations without dividing holds. Hence we have  $\text{tr}^{\mathbf{B}}(c) \odot f \sqsubseteq \text{tr}^{\mathbf{B}}(d') = \text{tr}^{\mathbf{B}}(d)$ . This implies the inequality (4.18).  $\square$

In [111], Kleisli simulation was instantiated for quantitative systems called *weighted automata*, and the resulting simulation notion was stated in terms of matrices. We conclude this section by doing the same thing for fair simulation.

**Definition 4.5.7** (fair simulation for PBWAs). Let  $\mathcal{A} = (X, \xi, \text{Acc}_{\mathcal{A}})$  and  $\mathcal{B} = (Y, \theta, \text{Acc}_{\mathcal{B}})$  be  $\mathbf{A}$ -labeled PBWAs whose state spaces are equipped with discrete  $\sigma$ -algebras. Let  $X_1 := X \setminus \text{Acc}_{\mathcal{A}}$ ,  $X_2 := \text{Acc}_{\mathcal{A}}$ ,  $Y_1 := Y \setminus \text{Acc}_{\mathcal{B}}$  and  $Y_2 := \text{Acc}_{\mathcal{B}}$ . For each  $a \in \mathbf{A}$ , let  $M_{\mathcal{A}}(a) \in [0, 1]^{X \times X}$  and  $M_{\mathcal{B}}(a) \in [0, 1]^{Y \times Y}$  be the *transition matrices* of  $\mathcal{A}$  and  $\mathcal{B}$ , i.e.  $(M_{\mathcal{A}}(a))_{x, x'} = \xi(x)(\{(a, x')\})$  and  $(M_{\mathcal{B}}(a))_{y, y'} = \theta(y)(\{(a, y')\})$ . A *fair matrix simulation from  $\mathcal{A}$  to  $\mathcal{B}$*  is a matrix  $A \in [0, 1]^{Y \times X}$  satisfying the following conditions. (Here  $M_{\mathcal{A}, i}(a) \in [0, 1]^{X_i \times X}$ ,  $M_{\mathcal{B}, j}(a) \in [0, 1]^{Y_j \times Y}$  and  $A_{ji} \in [0, 1]^{Y_j \times X_i}$  are the obvious partial matrices of  $M_{\mathcal{A}}(a) \in [0, 1]^{X \times X}$ ,  $M_{\mathcal{B}}(a) \in [0, 1]^{Y \times Y}$  and  $A \in [0, 1]^{Y \times X}$ , respectively. Moreover,  $\leq$  denotes the elementwise order between matrices.)

- O. The matrix  $A$  is a substochastic matrix, i.e.  $\forall y \in Y. \sum_{x \in X} A_{y, x} \leq 1$ .
- A. The matrix  $A$  is a *forward matrix simulation from  $\mathcal{A}$  to  $\mathcal{B}$*  [111], i.e.  $\forall a \in \mathbf{A}. A \cdot M_{\mathcal{X}}(a) \leq M_{\mathcal{Y}}(a) \cdot A$ .
- B. There exists a pair of increasing sequences of matrices of length  $\bar{\alpha} \leq \omega$

$$\begin{aligned} A_{11}^{(0)} &\leq A_{11}^{(1)} \leq \dots \leq A_{11}^{(\bar{\alpha})} \in [0, 1]^{Y_1 \times X_1} \quad \text{and} \\ A_{12}^{(0)} &\leq A_{12}^{(1)} \leq \dots \leq A_{12}^{(\bar{\alpha})} \in [0, 1]^{Y_1 \times X_2} \end{aligned}$$

such that:

- (a) (**Approximate  $A_{11}$  and  $A_{12}$** ) We have  $A_{11}^{(\bar{a})} = A_{11}$  and  $A_{12}^{(\bar{a})} = A_{12}$ .  
(b) ( $A_{11}^a$ ) For each  $a \leq \bar{a}$  and  $a \in \mathbf{A}$  we have:

$$A_{11}^{(a)} \cdot M_{\mathcal{X},1}(a) \leq M_{\mathcal{Y},1}(a) \cdot \begin{pmatrix} A_{11}^{(a)} & A_{12}^{(a)} \\ A_{21} & A_{22} \end{pmatrix}.$$

- (c) ( $A_{12}^a$ , **the base case**) The 0-th approximant  $A_{12}^{(0)}$  is the zero matrix.  
(d) ( $A_{12}^a$ , **the step case**) For each  $a < \bar{a}$  and  $a \in \mathbf{A}$ :

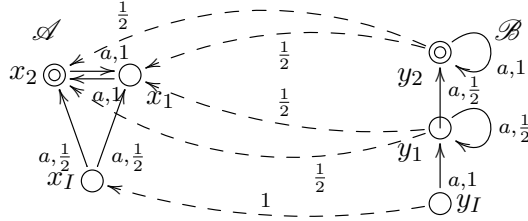
$$A_{12}^{(a+1)} \cdot M_{\mathcal{X},2}(a) \leq M_{\mathcal{Y},1}(a) \cdot \begin{pmatrix} A_{11}^{(a)} & A_{12}^{(a)} \\ A_{21} & A_{22} \end{pmatrix}.$$

- (e) ( $A_{12}^a$ , **the limit case**) When  $\bar{a} = \omega$ ,  $(A_{12}^{(\omega)})_{y,x} = \sup_{a' < \omega} (A_{12}^{(a')})_{y,x}$  for each  $y \in Y_1$  and  $x \in X_2$ .

**Theorem 4.5.8** (soundness). *Let  $\mathcal{A} = (X, \xi, \text{Acc}_{\mathcal{A}})$  and  $\mathcal{B} = (Y, \theta, \text{Acc}_{\mathcal{B}})$  be  $\mathbf{A}$ -labeled PBWAs whose state spaces are equipped with discrete  $\sigma$ -algebras. We further assume that  $\mathcal{B}$  has a finite state space. If  $A \in [0, 1]^{Y \times X}$  is a fair matrix simulation from  $\mathcal{A}$  to  $\mathcal{B}$ , then for each  $y \in Y$  and  $P \in \mathfrak{F}_{\mathbf{A}^\omega}$ , we have:*

$$\sum_{x \in X} A_{y,x} \cdot L_{\mathcal{A}}^{\mathbf{B}}(x)(P) \leq L_{\mathcal{B}}^{\mathbf{B}}(y)(P). \quad \square$$

**Example 4.5.9.** Let  $\mathbf{A} := \{a\}$ . Let  $\mathcal{A}$  and  $\mathcal{B}$  be  $\mathbf{A}$ -labeled PBWAs illustrated below (ignore the dashed lines for the moment).



We define  $A \in [0, 1]^{\{y_I, y_1, y_2\} \times \{x_I, x_1, x_2\}}$  by  $A_{y_i, x_j} = \frac{1}{2}$  for each  $i, j \in \{1, 2\}$ ,  $A_{y_I, x_I} = 1$  and  $A_{y, x} = 0$  for the other combinations (see also the dashed lines above). Then  $A$  is a fair matrix simulation from  $\mathcal{X}$  to  $\mathcal{Y}$ . Here the approximation sequences  $A_{11}^{(0)} \sqsubseteq A_{11}^{(1)} \sqsubseteq \dots \sqsubseteq A_{11}^{(\omega)} \in [0, 1]^{\{y_I, y_1\} \times \{x_I, x_1\}}$  and  $A_{12}^{(0)} \sqsubseteq A_{12}^{(1)} \sqsubseteq \dots \sqsubseteq A_{12}^{(\omega)} \in [0, 1]^{\{y_I, y_1\} \times \{x_2\}}$  are given by:  $(A_{11}^{(i)})_{y_I, x_I} = 1 - (\frac{1}{2})^i$ ,  $(A_{11}^{(i)})_{y_1, x_1} = \frac{1}{2}(1 - (\frac{1}{2})^i)$  and  $(A_{11}^{(i)})_{y, x} = 0$  for the other combinations; and  $(A_{12}^{(i)})_{y_1, x_2} = \frac{1}{2}(1 - (\frac{1}{2})^i)$ , for each  $i \leq \omega$  (here we let  $(\frac{1}{2})^\omega = 0$ ).

## 4.6 Conclusion and Related Work

Using the logical fixed point-based characterization of languages of Büchi automata, we have categorically generalized the notion of fair simulation. We then concretized it for a probabilistic variant of Büchi automata, and obtained a fair simulation notion for them. We have two types of categorical generalizations of fair simulation—one *with dividing* (Definition 4.3.13) and one *without dividing* (Definition 4.3.17). The former requires fewer axioms to prove its soundness than the latter does. However, its complicated definition limits its applicability. For nondeterministic Büchi tree automata, both simulation notions are always sound. For probabilistic Büchi tree automata, a simulation notion induced by the categorical simulation notion with dividing is always sound. In contrast, the one induced by the notion without dividing is proved to be sound when we focus on *finite-state word* automata.

**Related Work** We have introduced the first simulation notion for probabilistic Büchi word and tree automata to the best of our knowledge. However, there exist several simulation notions for other probabilistic systems. In [66], a simulation notion was introduced for PTSs whose states are labeled by a set of atomic propositions. In [55] a simulation notion between probabilistic systems is studied from a coalgebraic perspective. In [48, 111], a simulation notion for weighted automata, which encompass probabilistic automata, is introduced by concretizing Kleisli simulation (Definition 4.2.1). Comparison between these three simulation notions is found in [44].

Using our simulation notion, we can prove (quantitative) language inclusion between *generative* probabilistic Büchi automata (see also Chapter 7 for a discussion about generative and reactive systems). Its applicability is yet to be studied, but one candidate is security verification. For example, quantitative language inclusion-checking between (ordinary) probabilistic automata is known to be useful for proving *probable innocence*, a kind of anonymity [48].

Compared with generative ones, *reactive* probabilistic Büchi automata are more extensively studied because it is useful as a qualitative language acceptor [11]. More concretely, we can define a language of a reactive probabilistic Büchi automaton as the set of infinite words that are accepted by the automaton with positive probabilities. Interestingly, it is known that probabilistic Büchi automata are more expressible than nondeterministic Büchi automata as qualitative language acceptors [11]. Moreover, inclusion between such qualitative languages of probabilistic Büchi automata is known to be undecidable [12]. Hence a study of simulation notion for reactive probabilistic Büchi automata would be also interesting as future work.

## Chapter 5

# Categorical Ranking Function

In this chapter, we categorically generalize *ranking functions* [35].

Ranking functions are commonly used to prove termination of transition systems. They are especially useful for proving termination of infinite-state systems like while programs. Termination of such systems are often undecidable (e.g. [108]) and therefore a sound and complete method for proving termination does not exist. A ranking function provides us with one of such methods.

Sections 5.1–5.2 are devoted to preliminaries. We first review the definition of ranking function for reachability games, and then review categorical characterization of reachability to accepting states. A categorical generalization of ranking function is discussed in Section 5.3. Similarly to the previous chapter, we first concretize it for reachability games (Section 5.4). In Section 5.5 we concretize the categorical generalization for PTSs to obtain new ranking function-like notions for them.

The contents of this chapter are based on [109].

### 5.1 Ranking Function

We first review the notion of ranking function. Its categorical generalization is the main goal of this chapter.

**Definition 5.1.1** (ranking function [35]). Let  $\mathcal{T} = (X^{\text{Max}}, X^{\text{Min}}, E^{\text{Max}}, E^{\text{Min}}, \text{Acc})$  be a reachability game. Fix an ordinal  $\mathfrak{z}$  and let  $[\mathfrak{z}] := \{\mathfrak{n} \mid \mathfrak{n} \leq \mathfrak{z}\}$ . A *ranking function* for  $\mathcal{T}$  is a function  $b: X_{\text{max}} \rightarrow [\mathfrak{z}]$  such that

$$\forall x \in X^{\text{Max}} \setminus \text{Acc}. \quad \min_{y:(x,y) \in E^{\text{Max}}} \sup_{x':(y,x') \in E^{\text{Min}}} b(x') \hat{+} 1 \leq b(x). \quad (5.1)$$

Here  $b(x') \hat{+} 1 := \min\{b(x') + 1, \mathfrak{z}\}$ .

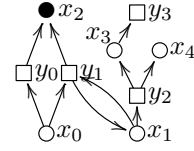
The existence of a ranking function that assigns a state  $x$  an ordinal strictly less than  $\mathfrak{z}$  implies that the game is winning for Player Max from the state.

**Theorem 5.1.2** (soundness [35]). *Let  $\mathfrak{z}$  be an ordinal and  $b: X \rightarrow [\mathfrak{z}]$  be a ranking function for a reachability game  $\mathcal{T} = (X^{\text{Max}}, X^{\text{Min}}, E^{\text{Max}}, E^{\text{Min}}, \text{Acc})$ . For each  $x \in X^{\text{Max}}$ ,  $b(x) < \mathfrak{z}$  implies  $x \in \text{Win}_{\mathcal{T}}$ .  $\square$*

**Remark 5.1.3.** A ranking function such that  $b(x) < \mathfrak{z}$  induces a positional winning strategy from  $x$ . More concretely, if  $b(x) < \mathfrak{z}$  then a strategy  $\mathfrak{s}^{\text{Max}} \in \mathfrak{G}_{\mathfrak{G}}^{\text{Max}}$  for Player Max defined as follows is winning from  $x$ .

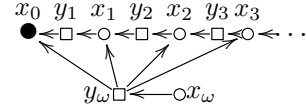
$$\mathfrak{s}^{\text{Max}}(x_0 y_0 \dots x_{i-1} y_{i-1} x_i) := \arg \min_{y:(x_i,y) \in E^{\text{Max}}} \sup_{x':(y,x') \in E^{\text{Min}}} b(x')$$

**Example 5.1.4.** We define a reachability game  $\mathcal{T} = (X^{\text{Max}}, X^{\text{Min}}, E^{\text{Max}}, E^{\text{Min}}, \text{Acc})$  by  $X^{\text{Max}} = \{x_0, x_1, x_2, x_3, x_4\}$ ,  $X^{\text{Min}} = \{y_0, y_1, y_2, y_3\}$ ,  $E^{\text{Max}} = \{(x_0, y_0), (x_0, y_1), (x_1, y_1), (x_1, y_2), (x_3, y_3)\}$ ,  $E^{\text{Min}} = \{(y_0, x_2), (y_1, x_1), (y_1, x_2), (y_2, x_3), (y_2, x_4)\}$ , and  $\text{Acc} = \{x_2\}$ . The game is pictured on the right. If we define a function  $b: X \rightarrow [\omega]$  by  $b(x_0) = 1$ ,  $b(x_2) = 0$ , and  $b(x_1) = b(x_3) = b(x_4) = \omega$ , then  $b$  is a ranking function. Hence by Theorem 5.1.2 we have  $\{x_0, x_2, x_3\} \subseteq \text{Win}_{\mathcal{T}}$ .



We note that *completeness* does not necessarily hold in the following sense: for every ordinal  $\mathfrak{z}$ , there exist a reachability game  $\mathcal{T}$  and a state  $x$  of  $\mathcal{T}$  such that  $x \in \text{Win}_{\mathcal{T}}$  but there exists no ranking function  $b: X^{\text{Max}} \rightarrow [\mathfrak{z}]$  such that  $b(x) < \mathfrak{z}$ . A counterexample is as follows.

**Example 5.1.5.** Let  $\mathfrak{z} = \omega$ . We define a reachability game  $\mathcal{T} = (X^{\text{Max}}, X^{\text{Min}}, E^{\text{Max}}, E^{\text{Min}}, \text{Acc})$  by  $X^{\text{Max}} := \{x_{\mathfrak{a}} \mid \mathfrak{a} \leq \omega\}$ ,  $X^{\text{Min}} := \{y_{\mathfrak{a}} \mid 1 \leq \mathfrak{a} \leq \omega\}$ ,  $E^{\text{Max}} := \{(x_{\mathfrak{a}}, y_{\mathfrak{a}}) \mid \mathfrak{a} \leq \omega\}$  and  $E^{\text{Min}} = \{(y_{\mathfrak{a}}, y_{\mathfrak{a}'}) \mid \mathfrak{a}' < \mathfrak{a}\}$  and  $\text{Acc} := \{x_0\}$ . Note that all the choices are made by Player Max. By the well-foundedness of  $\omega$ , we have  $x_{\omega} \in \text{Win}_{\mathcal{T}}$ . However, we can prove by contradiction that if  $b: X^{\text{Max}} \rightarrow [\omega]$  is a ranking function for  $\mathcal{T}$  then  $b(x_{\omega}) = \omega$ . Hence there exists no ranking function  $b: X^{\text{Max}} \rightarrow [\omega]$  that proves  $x_{\omega} \in \text{Win}_{\mathcal{T}}$ .



**Remark 5.1.6.** Completeness in the following sense *does* hold: for every reachability game  $\mathcal{T}$  and state  $x$  of  $\mathcal{T}$  such that  $x \in \text{Win}_{\mathcal{T}}$ , there exists an ordinal  $\mathfrak{z}$  and a ranking function  $b: X \rightarrow [\mathfrak{z}]$  such that  $b(x) < \mathfrak{z}$ .

## 5.2 Modalities and Fixed-Point Properties

We have seen that ranking function is used to prove reachability to accepting states. This means that categorically generalizing ranking function requires us to categorically characterizing reachability first. In this thesis, we follow an existing framework that is standard (see e.g. [45, 53]). Differently from the categorical framework reviewed in Section 2.4.2, the framework we review here uses *algebras* instead of final coalgebras, and characterize behaviors of systems as coalgebra-algebra homomorphisms (Definition 2.4.21).

In general, there can exist multiple homomorphisms from a coalgebra to an algebra. We choose the *least* one by assuming that each homset to the carrier of the algebra carries a partial order.

**Definition 5.2.1.** A *truth-value domain* is a pair  $(\Omega, \sqsubseteq_{\Omega})$  of an object  $\Omega \in \mathbb{C}$  and a family  $\sqsubseteq_{\Omega} = (\sqsubseteq_{X, \Omega})_{X \in \mathbb{C}}$  of partial orders, where  $\sqsubseteq_{X, \Omega}$  is defined over a homset  $\mathbb{C}(X, \Omega)$ . If no confusion is likely, we write  $\sqsubseteq$  for  $\sqsubseteq_{\Omega}$  and  $\sqsubseteq_{X, \Omega}$ . For  $F: \mathbb{C} \rightarrow \mathbb{C}$ , an *F-modality over*  $(\Omega, \sqsubseteq)$  is an *F-algebra*  $\sigma: F\Omega \rightarrow \Omega$ .

**Definition 5.2.2** ( $\llbracket \mu \sigma \rrbracket_c$ ). Let  $(\Omega, \sqsubseteq_{\Omega})$  be a truth-value domain and  $\sigma: F\Omega \rightarrow \Omega$  be an *F-modality over*  $(\Omega, \sqsubseteq)$ . We say that  $\sigma$  has *least fixed points* if for each  $c: X \rightarrow FX$ ,  $\Phi_{c, \sigma}: \mathbb{C}(X, \Omega) \rightarrow \mathbb{C}(X, \Omega)$  is a monotone function and has the least fixed point with respect to  $\sqsubseteq_{X, \Omega}$ . The least fixed point is called the (*coalgebraic*) *least fixed-point property* in  $c$  specified by  $\sigma$ , and denoted by  $\llbracket \mu \sigma \rrbracket_c: X \rightarrow \Omega$ .

$$\begin{array}{ccc} FX & \xrightarrow{F\llbracket \mu \sigma \rrbracket_c} & F\Omega \\ \uparrow c & =_{\mu} & \sigma \downarrow \\ X & \xrightarrow{\llbracket \mu \sigma \rrbracket_c} & \Omega \end{array}$$

Note that  $\text{tr}(c)$  and  $\text{tr}^*(c)$  in Section 3.1 are describable in the above framework: the former is  $\llbracket \mu(J(\iota^F)^{-1}) \rrbracket_c$ , and the latter is  $\llbracket \mu(J\zeta^F) \rrbracket_c$ .

**Example 5.2.3.** Let  $F = \{0, 1\} \times \mathcal{P}X : \mathbf{Sets} \rightarrow \mathbf{Sets}$ . Then an  $F$ -coalgebra  $c: X \rightarrow \{0, 1\} \times X$  is understood as a nondeterministic transition system with accepting states, by regarding  $x$  as accepting if  $\pi_1(c(x)) = 1$ . Note that  $F$  does not have a final coalgebra (cf. Section 1.3.1).

However, the framework above allows us to categorically capture behaviors of  $F$ -coalgebras. We regard  $\{0, 1\}$  as a truth-value domain by defining an order  $\sqsubseteq$  over  $\mathbf{Sets}(X, \{0, 1\})$  by  $f \sqsubseteq g \stackrel{\text{def.}}{\iff} \forall x. f(x) \leq g(x)$  for each  $X \in \mathbf{Sets}$ . We define an  $F$ -modality  $\sigma : \{0, 1\} \times \mathcal{P}\{0, 1\} \rightarrow \{0, 1\}$  over  $(\{0, 1\}, \sqsubseteq)$  by

$$\sigma(a, A) := \begin{cases} 1 & (a = 1) \\ \max A & (a = 0) \end{cases}$$

Then the function  $\Phi_{c, \sigma} : \mathbf{Sets}(X, \{0, 1\}) \rightarrow \mathbf{Sets}(X, \{0, 1\})$  (Definition 2.4.22) is given as follows:

$$\Phi_{c, \sigma}(f)(x) = \begin{cases} 1 & (\pi_1(c(x)) = 1) \\ \max\{f(a) \mid a \in \pi_2(c(x))\} & (\pi_1(c(x)) = 0). \end{cases}$$

In terms of *predicate transformer semantics* (see e.g. [77]), this indicates that  $\Phi_{c, \sigma}$  calculates the *weakest precondition* of a predicate  $f$  with respect to the angelic nondeterminism. The least fixed point property  $\llbracket \mu\sigma \rrbracket_c : X \rightarrow \{0, 1\}$  assigns 1 to  $x$  if and only if an accepting state is reachable from  $x$  when we choose successor states appropriately.

In contrast, if we define an  $F$ -modality  $\sigma' : \{0, 1\} \times \{0, 1\} \rightarrow \{0, 1\}$  over the same truth-value domain by

$$\sigma'(a, A) := \begin{cases} 1 & (a = 1) \\ \min A & (a = 0) \end{cases}$$

then

$$\Phi_{c, \sigma'}(f)(x) = \begin{cases} 1 & (\pi_1(c(x)) = 1) \\ \min\{f(a) \mid a \in \pi_2(c(x))\} & (\pi_1(c(x)) = 0). \end{cases}$$

This is understood as calculating the weakest precondition with respect to the demonic nondeterminism, and  $\llbracket \mu\sigma' \rrbracket_c : X \rightarrow \{0, 1\}$  assigns 1 to  $x$  if and only if an accepting state is reached from  $x$  regardless of the nondeterministic choice.

We later see that a winning region  $\text{Win}_G$  of a reachability game and a reachability probability function  $\text{Reach}_{\mathcal{T}}$  of a PTS are characterized by this framework.

### 5.3 Categorical Generalization of Ranking Function

In this section, we categorically generalize ranking functions.

We first explain the intuition using a reachability game. Let  $\mathcal{T} = (X^{\text{Max}}, X^{\text{Min}}, E^{\text{Max}}, E^{\text{Min}}, \text{Acc})$  be a reachability game and  $\mathfrak{z}$  be an ordinal. Define  $q_{\mathfrak{g}, \mathfrak{z}} : [\mathfrak{z}] \rightarrow \{0, 1\}$  by  $q_{\mathfrak{g}, \mathfrak{z}}(\mathfrak{z}) := 0$  and  $q_{\mathfrak{g}, \mathfrak{z}}(\mathfrak{a}) := 1$  for  $\mathfrak{a} < \mathfrak{z}$ . Then we can rewrite the soundness theorem of ranking functions (Theorem 5.1.2) as follows: if  $b : X^{\text{Max}} \rightarrow [\mathfrak{z}]$  is a ranking function, then  $q_{\mathfrak{g}, \mathfrak{z}} \circ b : X^{\text{Max}} \rightarrow \{0, 1\}$  underapproximates the characteristic function  $X^{\text{Max}} \rightarrow \{0, 1\}$  of  $\text{Win}_{\mathcal{T}} \subseteq X^{\text{Max}}$ .

Our categorical generalization of ranking functions is based on this observation. As we will see in the next section, we can characterize the characteristic function of  $\text{Win}_{\mathcal{T}}$  as a coalgebraic least fixed-point property  $\llbracket \mu\sigma \rrbracket_c$  (Definition 5.2.2). Hence our goal is to underapproximate  $\llbracket \mu\sigma \rrbracket_c$  using a categorical ranking function.

Towards this goal, we use the notion of *corecursive algebra*. In the proof of the soundness theorem of ranking functions, *well-foundedness* of a poset  $([\mathfrak{J}], \leq)$  plays a very important role. The notion of corecursive algebra is a categorical counterpart of well-foundedness.

**Definition 5.3.1** (corecursive algebra, [16]). An  $F$ -algebra  $r : FR \rightarrow R$  is *corecursive* if for an arbitrary coalgebra  $c : X \rightarrow FX$  there exists a unique arrow  $f : X \rightarrow R$  such that  $f = r \circ Ff \circ c$ . We write  $(c)_r$  for the unique arrow.

$$\begin{array}{ccc} FX & \xrightarrow{Ff} & FR \\ \uparrow c & \stackrel{f}{=} & r \downarrow \\ X & \xrightarrow{\quad} & R \end{array}$$

In other words,  $r$  is corecursive if  $\Phi_{c,r}$  has the unique fixed point for each  $c$ . Using the notion of corecursive algebra, we introduce a notion of *ranking domain*. It is an  $F$ -algebra  $r : FR \rightarrow R$  equipped with several data. Intuitively, the carrier  $R$  corresponds to  $[\mathfrak{J}]$  in Definition 5.1.1 and  $\Phi_{c,r} : \mathbb{C}(X, R) \rightarrow \mathbb{C}(X, R)$  corresponds to the left-hand side of (5.1) in the definition.

**Definition 5.3.2** (ranking domain). Let  $F : \mathbb{C} \rightarrow \mathbb{C}$ . Let  $\sigma : F\Omega \rightarrow \Omega$  be an  $F$ -modality over a truth-value domain  $(\Omega, \sqsubseteq_\Omega)$  that has least fixed points. Let  $r : FR \rightarrow R$  be an  $F$ -algebra,  $q : R \rightarrow \Omega$  be an arrow, and  $\sqsubseteq_R = (\sqsubseteq_{X,R})_{X \in \mathbb{C}}$  be a family of partial orders where  $\sqsubseteq_{X,R}$  is define over  $\mathbb{C}(X, R)$ . A triple  $(r, q, \sqsubseteq_R)$  is called a *ranking domain* for  $\sigma$  if the following conditions are satisfied.

1. We have  $q \circ r \sqsubseteq_{FR, \Omega} \sigma \circ Fq$  (see the diagram on the right).
2. For each  $c : X \rightarrow FX$ ,  $\Phi_{c,r}$  is monotone and  $(\mathbb{C}(X, R), \sqsubseteq_{X,R})$  has the least element  $\perp_{X,R}$ . Moreover, either of the following conditions is satisfied:
  - (a)  $(\mathbb{C}(X, R), \sqsubseteq_{X,R})$  is  $\omega$ -complete and  $\Phi_{c,r}$  is  $\omega$ -continuous; or
  - (b)  $(\mathbb{C}(X, R), \sqsubseteq_{X,R})$  is directed complete.
3. For each  $X \in \mathbb{C}$ , a function  $q \circ (\_) : \mathbb{C}(X, R) \rightarrow \mathbb{C}(X, \Omega)$  is monotone (i.e.  $f \sqsubseteq_{X,R} g$  implies  $q \circ f \sqsubseteq_{X, \Omega} q \circ g$ ) and strict (i.e.  $q \circ \perp_{X,R} = \perp_{X, \Omega}$ ). Moreover,
  - if 2(a) is satisfied, it is  $\omega$ -continuous (i.e.  $q \circ \bigsqcup_{i \in \omega} f_i = \bigsqcup_{i \in \omega} q \circ f_i$ ); and
  - if 2(b) is satisfied, it is directed continuous (i.e.  $q \circ \bigsqcup_{f \in A} f = \bigsqcup_{f \in A} q \circ f$  for each directed subset  $A$ ).
4. The algebra  $r : FR \rightarrow R$  is corecursive.

$$\begin{array}{ccc} FR & \xrightarrow{Fq} & F\Omega \\ \downarrow r & \sqsubseteq & \sigma \downarrow \\ R & \xrightarrow{q} & \Omega \end{array}$$

The intuition is as follows: Condition 2 ensures that we can obtain the least fixed point of  $\Phi_{c,r}$  using Theorem 2.3.2.2–3. By Theorem 2.3.4, Conditions 1 and 3 imply that the least fixed point is preserved by  $q$ . Finally, Condition 4 implies that the least fixed point of  $\Phi_{c,r}$  is in fact the *unique* fixed point.

A *ranking arrow*, a categorical generalization of ranking functions, is defined with respect to a ranking domain.

**Definition 5.3.3** (ranking arrow). Let  $(r, q, \sqsubseteq_R)$  be a ranking domain for an  $F$ -modality  $\sigma : F\Omega \rightarrow \Omega$ , and  $c : X \rightarrow FX$  be an  $F$ -coalgebra. An arrow  $b : X \rightarrow R$  is called a *ranking arrow* for  $c$  with respect to  $(r, q, \sqsubseteq_R)$  if it satisfies  $b \sqsubseteq_R \Phi_{c,r}(b)$ .

$$\begin{array}{ccccc} FX & \xrightarrow{Fb} & FR & \xrightarrow{Fq} & F\Omega \\ \uparrow c & \sqsubseteq & r \downarrow & \sqsubseteq & \sigma \downarrow \\ X & \xrightarrow{b} & R & \xrightarrow{q} & \Omega \end{array}$$

Recall that the soundness theorem of ranking functions claims that  $q_{\mathfrak{g}, \mathfrak{J}} \circ b$  underapproximates the characteristic function  $X^{\text{Max}} \rightarrow \{0, 1\}$  of  $\text{Win}_{\mathcal{T}} \subseteq X^{\text{Max}}$ .



The following theorem is its categorical counterpart: it claims that  $q \circ b$  underapproximates the coalgebraic least fixed-point property (Definition 5.2.2).

**Theorem 5.3.4** (soundness). *Let  $F : \mathbb{C} \rightarrow \mathbb{C}$  and  $\sigma : F\Omega \rightarrow \Omega$  be an  $F$ -modality over a truth-value domain  $(\Omega, \sqsubseteq_\Omega)$  that has least fixed points. Let  $(r, q, \sqsubseteq_R)$  be a ranking domain for  $\sigma$ . Let  $c : X \rightarrow FX$  be an  $F$ -coalgebra and  $b : X \rightarrow R$  be a ranking arrow for  $c$ . Then we have:*

$$q \circ b \sqsubseteq_\Omega \llbracket \mu\sigma \rrbracket_c.$$

**Proof.** We are assuming that  $r$  is a corecursive algebra (Condition 4 in Definition 5.3.2). This means that  $\Phi_{c,r}$  has a unique fixed point  $(\lfloor c \rfloor)_r : X \rightarrow R$ . Obviously, it is also the least fixed point and the greatest fixed point of  $\Phi_{c,r}$ , i.e.  $(\lfloor c \rfloor)_r = \mu\Phi_{c,r} = \nu\Phi_{c,r}$ . Moreover, by Condition 2 in Definition 5.3.2 and the Knaster-Tarski theorem (the dual of Corollary 2.3.3.1), we have  $b \sqsubseteq_R \nu\Phi_{c,r}$ . By Condition 3,  $q \circ (\_)$  is a monotone function. Hence we have:

$$q \circ b \sqsubseteq_{X,\Omega} q \circ \nu\Phi_{c,r} = q \circ \mu\Phi_{c,r}. \quad (5.2)$$

We next prove  $q \circ \mu\Phi_{c,r} \sqsubseteq_{X,\Omega} \llbracket \mu\sigma \rrbracket_c$ . Note that the right-hand side is a fixed point of  $\Phi_{c,\sigma}$  by definition. We shall prove the inequality using Theorem 2.3.4.

We first prove that Condition 2 of Theorem 2.3.4 is satisfied. For  $l \in \mathbb{C}(X, R)$ , we have:

$$\begin{aligned} q \circ \Phi_{c,r}(l) &= q \circ r \circ Fl \circ c && \text{(by definition)} \\ &\sqsubseteq_{X,\Omega} \sigma \circ Fq \circ Fl \circ c && \text{(by Condition 1 in Definition 5.3.2)} \\ &= \Phi_{c,\sigma}(q \circ l) && \text{(by definition)}. \end{aligned}$$

Conditions 1 and 3 of Theorem 2.3.4 are immediate by Condition 3 in Definition 5.3.2.

Hence by Theorem 2.3.4, we have  $q \circ \mu\Phi_{c,r} \sqsubseteq_{X,\Omega} \llbracket \mu\sigma \rrbracket_c$ . Together with (5.2), we have  $q \circ b \sqsubseteq_\Omega \llbracket \mu\sigma \rrbracket_c$ .  $\square$

**Remark 5.3.5.** The unique fixed point  $(\lfloor c \rfloor)_r : X \rightarrow R$  of  $\Phi_{c,r}$  is the ‘‘optimal’’ ranking arrow. That is, by the Knaster-Tarski theorem (Corollary 2.3.3.1), if  $b : X \rightarrow R$  is a ranking arrow, then we have  $b \sqsubseteq_R (\lfloor c \rfloor)_r$ . This implies  $q \circ b \sqsubseteq q \circ (\lfloor c \rfloor)_r$ , which means that  $q \circ (\lfloor c \rfloor)_r$  gives better bound for  $\llbracket \mu\sigma \rrbracket_c$  than  $q \circ b$ .

The converse of Theorem 5.3.4, i.e. *completeness*, does not necessarily hold. This means that it is possible that there exists no  $b : X \rightarrow R$  such that  $q \circ b = \llbracket \mu\sigma \rrbracket_c$ . By the above remark, it is equivalent to that  $q \circ (\lfloor c \rfloor)_r \neq \llbracket \mu\sigma \rrbracket_c$ . The following proposition shows a sufficient condition for the completeness.

**Proposition 5.3.6.** *Assume that the following equality holds in Theorem 5.3.4:*

$$q \circ r = \sigma \circ Fq. \quad (5.3)$$

*Then we have  $q \circ (\lfloor c \rfloor)_r = \llbracket \mu\sigma \rrbracket_c$ .*

$$\begin{array}{ccccc} FX & \xrightarrow{F\llbracket \mu\sigma \rrbracket_c} & FR & \xrightarrow{Fq} & F\Omega \\ \uparrow c & \begin{array}{c} F(\lfloor c \rfloor)_r \\ = \\ (\lfloor c \rfloor)_r \end{array} & \downarrow r & \begin{array}{c} Fq \\ = \\ q \end{array} & \downarrow \sigma \\ X & \xrightarrow{(\lfloor c \rfloor)_r} & R & \xrightarrow{q} & \Omega \\ & \xrightarrow{\llbracket \mu\sigma \rrbracket_c} & & & \end{array}$$

**Proof.** By  $(\lfloor c \rfloor)_r = r \circ F(\lfloor c \rfloor)_r \circ c$  and  $q \circ r = \sigma \circ Fq$ , we have:

$$q \circ (\lfloor c \rfloor)_r = \Phi_{c,\sigma}(q \circ (\lfloor c \rfloor)_r).$$

This means that  $q \circ \langle c \rangle_r$  is a fixed point of  $\Phi_{c,\sigma}$ . As  $\llbracket \mu\sigma \rrbracket_c$  is the least fixed point of  $\Phi_{c,\sigma}$ , we have

$$q \circ \langle c \rangle_r \sqsupseteq_{\Omega} \llbracket \mu\sigma \rrbracket_c.$$

Together with Theorem 5.3.4, we have  $q \circ \langle c \rangle_r = \llbracket \mu\sigma \rrbracket_c$ .  $\square$

## 5.4 Concretization to Reachability Games

This section is devoted to a “sanity-check”: we will see that there exists a ranking domain for  $F_{\mathbf{g}}$ -coalgebras (see Example 2.4.14) such that the resulting definition of ranking arrow coincides with the conventional definition of ranking function (Definition 5.1.1).

Recall that we can represent a reachability game  $\mathcal{T} = (X^{\text{Max}}, X^{\text{Min}}, E^{\text{Max}}, E^{\text{Min}}, \text{Acc})$  as an  $F_{\mathbf{g}}$ -coalgebra  $c_{\mathcal{T}} : X^{\text{Max}} \rightarrow \mathcal{P}^2 X^{\text{Max}} \times \{0, 1\}$  (Example 2.4.14). Then the winning region  $\text{Win}_{\mathcal{T}} \subseteq X^{\text{Max}}$  (Definition 2.2.22) is characterized as a coalgebraic least fixed-point property as follows.

**Proposition 5.4.1.** *For each  $X \in \mathbf{Sets}$ , we define a partial order  $\leq_X$  over  $\mathbf{Sets}(X, \{0, 1\})$  by  $f \leq_X g \stackrel{\text{def.}}{\iff} \forall x \in X. f(x) \leq g(x)$ , and let  $\leq := (\leq_X)_{X \in \mathbf{Sets}}$ . We define an  $F_{\mathbf{g}}$ -modality  $\sigma_{\mathbf{g}} : F_{\mathbf{g}}\{0, 1\} \rightarrow \{0, 1\}$  over a truth-value domain  $(\{0, 1\}, \leq)$  by*

$$\sigma_{\mathbf{g}}(\Gamma, t) = \begin{cases} 1 & (t = 1) \\ \max_{A \in \Gamma} \min_{a \in A} a & (t = 0). \end{cases}$$

*Then  $\sigma_{\mathbf{g}}$  has least fixed points. Moreover, for a reachability game  $\mathcal{T} = (X^{\text{Max}}, X^{\text{Min}}, E^{\text{Max}}, E^{\text{Min}}, \text{Acc})$ , if we construct an  $F_{\mathbf{g}}$ -coalgebra  $c_{\mathcal{T}}$  as in Example 2.4.14, then the least fixed-point property  $\llbracket \mu\sigma_{\mathbf{g}} \rrbracket_{c_{\mathcal{T}}} : X^{\text{Max}} \rightarrow \{0, 1\}$  in  $c_{\mathcal{T}}$  specified by  $\sigma_{\mathbf{g}}$  is given by the characteristic function of  $\text{Win}_{\mathcal{T}}$ , i.e.*

$$\llbracket \mu\sigma_{\mathbf{g}} \rrbracket_{c_{\mathcal{T}}}(x) = \begin{cases} 1 & (x \in \text{Win}_{\mathcal{T}}) \\ 0 & (x \notin \text{Win}_{\mathcal{T}}). \end{cases}$$

**Proof.** It is not hard to see that  $\Phi_{c,\sigma}$  is monotone.

It is also easy to see that for each  $c : X \rightarrow F_{\mathbf{g}}X$  there exists a reachability game  $\mathcal{T}$  such that  $c = c_{\mathcal{T}}$ . Hence it suffices to show that for each reachability game  $\mathcal{T} = (X^{\text{Max}}, X^{\text{Min}}, E^{\text{Max}}, E^{\text{Min}}, \text{Acc})$  if we define  $f : X^{\text{Max}} \rightarrow \{0, 1\}$  by

$$f(x) := \begin{cases} 1 & (x \in \text{Win}_{\mathcal{T}}) \\ 0 & (x \notin \text{Win}_{\mathcal{T}}), \end{cases}$$

then  $f$  is the least fixed point of  $\Phi_{c_{\mathcal{T}}, \sigma_{\mathbf{g}}} : \mathbf{Sets}(X^{\text{Max}}, \{0, 1\}) \rightarrow \mathbf{Sets}(X^{\text{Max}}, \{0, 1\})$ .

We first show that  $f$  is a fixed point of  $\Phi_{c,\sigma_{\mathbf{g}}}$ . For each  $x \in X^{\text{Max}}$ , we have:

$$\begin{aligned} \Phi_{c_{\mathcal{T}}, \sigma_{\mathbf{g}}}(f)(x) &= 1 \\ \Leftrightarrow (\sigma_{\mathbf{g}} \circ F_{\mathbf{g}}f \circ c_{\mathcal{T}})(x) &= 1 && \text{(by the definition of } \Phi_{c_{\mathcal{T}}, \sigma_{\mathbf{g}}}\text{)} \\ \Leftrightarrow \pi_2(c_{\mathcal{T}}(x)) = 1, \text{ or } \exists A \in \pi_1(c_1(x)). \forall x' \in A. f(x') = 1 &&& \text{(by definition)} \\ \Leftrightarrow x \in \text{Acc} \text{ or } \exists y \in X^{\text{Min}} \text{ s.t. } (x, y) \in E^{\text{Max}}. \forall x' \in X^{\text{Max}} \text{ s.t. } (y, x') \in E^{\text{Min}}. &&& \\ & f(x') = 1 && \text{(by the definition of } c_{\mathcal{T}}\text{)} \\ \Leftrightarrow x \in \text{Acc} \text{ or } \exists y \in X^{\text{Min}} \text{ s.t. } (x, y) \in E^{\text{Max}}. \forall x' \in X^{\text{Max}} \text{ s.t. } (y, x') \in E^{\text{Min}}. &&& \\ & \exists \mathfrak{s}^{\text{Max}} \in \mathfrak{S}_{\mathcal{T}}^{\text{Max}}. \forall \mathfrak{s}^{\text{Min}} \in \mathfrak{S}_{\mathcal{T}}^{\text{Min}}. \rho_{x', \mathfrak{s}^{\text{Max}}, \mathfrak{s}^{\text{Min}}} \text{ is winning} &&& \\ & && \text{(by the definition of } f\text{)} \end{aligned}$$

$\Leftrightarrow x \in \text{Acc}$  or

$$\begin{aligned} & \exists y \in X^{\text{Min}} \text{ s.t. } (x, y) \in E^{\text{Max}}. \exists (\mathfrak{s}_{x'}^{\text{Max}} \in \mathfrak{S}_{\mathcal{T}}^{\text{Max}})_{x' \in \{x' \in X^{\text{Max}} \mid (y, x') \in E^{\text{Min}}\}}. \\ & \quad \forall x' \in X^{\text{Max}} \text{ s.t. } (y, x') \in E^{\text{Min}}. \forall \mathfrak{s}^{\text{Min}} \in \mathfrak{S}_{\mathcal{T}}^{\text{Min}}. \rho_{x', \mathfrak{s}^{\text{Max}}, \mathfrak{s}^{\text{Min}}} \text{ is winning} \end{aligned}$$

$\Leftrightarrow x \in \text{Acc}$  or  $\exists \mathfrak{s}^{\text{Max}} \in \mathfrak{S}_{\mathcal{T}}^{\text{Max}}. \forall \mathfrak{s}^{\text{Min}} \in \mathfrak{S}_{\mathcal{T}}^{\text{Min}}. \rho_{x, \mathfrak{s}^{\text{Max}}, \mathfrak{s}^{\text{Min}}}$  is winning

$\Leftrightarrow f(x) = 1$  (by the definition of  $f$ ).

Hence  $f$  is a fixed point of  $\Phi_{c_{\mathcal{T}}, \sigma_{\mathfrak{g}}}$ .

It remains to show that  $f : X^{\text{Max}} \rightarrow \{0, 1\}$  is the least fixed point. Let  $f' : X^{\text{Max}} \rightarrow \{0, 1\}$  be a fixed point of  $\Phi_{c_{\mathcal{T}}, \sigma_{\mathfrak{g}}}$ . To prove  $f \leq_{X^{\text{Max}}} f'$ , it suffices to prove  $f'(x) = 0$  implies  $f(x) = 0$  for each  $x \in X$ .

For each  $x' \in X^{\text{Max}}$ , we have:

$$\begin{aligned} f'(x') &= 0 \\ \Leftrightarrow \Phi_{c_{\mathcal{T}}, \sigma_{\mathfrak{g}}}(f')(x') &= 0 && (f' \text{ is a fixed point of } \Phi_{c_{\mathcal{T}}, \sigma_{\mathfrak{g}}}) \\ \Leftrightarrow (\sigma_{\mathfrak{g}} \circ F_{\mathfrak{g}} f' \circ c_{\mathcal{T}})(x') &= 0 && (\text{by the definition } \Phi_{c_{\mathcal{T}}, \sigma_{\mathfrak{g}}}) \\ \Leftrightarrow \pi_2(c_{\mathcal{T}}(x')) = 0 \text{ and } \forall A \in \pi_1(c_{\mathcal{T}}(x')). \exists x'' \in A. f'(x'') = 0 && (\text{by definition}) \\ \Leftrightarrow x' \notin \text{Acc} \text{ and } \forall y \in X^{\text{Min}} \text{ s.t. } (x', y) \in E^{\text{Max}}. \exists x'' \in X^{\text{Max}} \text{ s.t. } (y, x'') \in E^{\text{Min}}. \\ & \quad f'(x'') = 0 && (\text{by definition}). \end{aligned}$$

This means that if  $f'(x') = 0$ , then  $x' \notin \text{Acc}$  and for each  $y \in X^{\text{Min}}$  there exists  $x'' \in X^{\text{Max}}$  such that  $f'(x'') = 0$ . Hence for each  $x \in X^{\text{Max}}$  such that  $f'(x) = 0$ , we can inductively define a strategy  $\mathfrak{s}^{\text{Min}} \in \mathfrak{S}_{\mathcal{T}}^{\text{Min}}$  so that for each strategy  $\mathfrak{s}^{\text{Max}} \in \mathfrak{S}_{\mathcal{T}}^{\text{Max}}$  the resulting run  $\rho_{x, \mathfrak{s}^{\text{Max}}, \mathfrak{s}^{\text{Min}}}$  from  $x$  is not winning.

Therefore by definition, we have  $f(x) = 0$ . This concludes the proof.  $\square$

A ranking domain for  $\sigma_{\mathfrak{g}}$  inducing the conventional definition of ranking function is given as follows.

**Proposition 5.4.2.** *Fix an ordinal  $\mathfrak{z}$ . We define an  $F_{\mathfrak{g}}$ -algebra  $r_{\mathfrak{g}, \mathfrak{z}} : F_{\mathfrak{g}}[\mathfrak{z}] \rightarrow [\mathfrak{z}]$ , a function  $q_{\mathfrak{g}, \mathfrak{z}} : F_{\mathfrak{g}}[\mathfrak{z}] \rightarrow \{0, 1\}$  and a family  $\sqsubseteq_{[\mathfrak{z}]} = (\sqsubseteq_{X, [\mathfrak{z}]})_{X \in \mathbf{Sets}}$  of partial orders as follows.*

- $r_{\mathfrak{g}, \mathfrak{z}} : \mathcal{P}^2[\mathfrak{z}] \times \{0, 1\} \rightarrow [\mathfrak{z}]$  is defined as follows:

$$r_{\mathfrak{g}, \mathfrak{z}}(\Gamma, t) := \begin{cases} 0 & (t = 1) \\ \min_{A \in \Gamma} \sup_{a \in A} (\mathfrak{a} \hat{+} 1) & (\text{otherwise}). \end{cases}$$

Recall that  $\mathfrak{a} \hat{+} 1$  denotes  $\min\{\mathfrak{a} + 1, \mathfrak{z}\}$ .

- $q_{\mathfrak{g}, \mathfrak{z}} : [\mathfrak{z}] \rightarrow \{0, 1\}$  is defined by

$$q_{\mathfrak{g}, \mathfrak{z}}(\mathfrak{a}) := \begin{cases} 0 & (\mathfrak{a} = \mathfrak{z}) \\ 1 & (\text{otherwise}). \end{cases}$$

- For  $f, g \in \mathbf{Sets}(X, [\mathfrak{z}])$ ,  $f \sqsubseteq_{X, [\mathfrak{z}]} g \stackrel{\text{def.}}{\Leftrightarrow} \forall x \in X. f(x) \geq g(x)$  (be aware of the direction).

Then the triple  $(r_{\mathfrak{g}, \mathfrak{z}}, q_{\mathfrak{g}, \mathfrak{z}}, \sqsubseteq_{[\mathfrak{z}]})$  is a ranking domain.

**Proof.** We prove that Conditions 1–4 of Definition 5.3.2 are satisfied.

**Condition 1** It suffices to prove that for each  $(\Gamma, t) \in F_{\mathbf{g}}[\mathfrak{z}] = \mathcal{P}^2 r_{\mathbf{g}, \mathfrak{z}} \times \{0, 1\}$ , if  $\sigma_{\mathbf{g}} \circ F_{\mathbf{g}} q_{\mathbf{g}, \mathfrak{z}}(\Gamma, t) = 0$  then  $q_{\mathbf{g}, \mathfrak{z}} \circ r_{\mathbf{g}, \mathfrak{z}}(\Gamma, t) = 0$ .

$$\begin{aligned}
& \sigma_{\mathbf{g}} \circ F_{\mathbf{g}} q_{\mathbf{g}, \mathfrak{z}}(\Gamma, t) = 0 \\
& \Leftrightarrow t = 0 \text{ and } \forall A \in \Gamma. \mathfrak{z} \in A \quad (\text{by the definitions of } \sigma_{\mathbf{g}}, q_{\mathbf{g}, \mathfrak{z}} \text{ and } F_{\mathbf{g}}) \\
& \Rightarrow t = 0 \text{ and } \min_{A \in \Gamma} \sup_{\mathfrak{a} \in A} (\mathfrak{a} \hat{+} 1) = \mathfrak{z} \quad (5.4) \\
& \Leftrightarrow r_{\mathbf{g}, \mathfrak{z}}(\Gamma, t) = \mathfrak{z} \quad (\text{by the definition of } r_{\mathbf{g}, \mathfrak{z}}) \\
& \Leftrightarrow q_{\mathbf{g}, \mathfrak{z}} \circ r_{\mathbf{g}, \mathfrak{z}}(\Gamma, t) = 0 \quad (\text{by the definition of } q_{\mathbf{g}, \mathfrak{z}}).
\end{aligned}$$

**Condition 2** Assume that  $f \sqsubseteq_{X, [\mathfrak{z}]} g$ . Then for each  $x \in X$ , we have:

$$\begin{aligned}
\Phi_{c,r}(f)(x) &= \sigma \circ F_{\mathbf{g}} f \circ c(x) \\
&= \begin{cases} 0 & (\pi_2(c(x)) = 1) \\ \min_{A \in \pi_1(c(x))} \sup_{x \in A} (f(x) \hat{+} 1) & (\pi_2(c(x)) = 0) \end{cases} \\
&\geq \begin{cases} 0 & (\pi_2(c(x)) = 1) \\ \min_{A \in \pi_1(c(x))} \sup_{x \in A} (g(x) \hat{+} 1) & (\pi_2(c(x)) = 0) \end{cases} \\
&= \sigma \circ F_{\mathbf{g}} g \circ c(x) = \Phi_{c,r}(g)(x).
\end{aligned}$$

Hence we have  $\Phi_{c,r}(f) \sqsubseteq_{X, [\mathfrak{z}]} \Phi_{c,r}(g)$ , and therefore  $\Phi_{c,r}$  is monotone.

It is easy to see that a function that maps each  $x \in X$  to  $\mathfrak{z}$  is the least element in  $(\mathbf{Sets}(X, [\mathfrak{z}]), \sqsubseteq_{X,R})$ .

Finally, it is also easy to see that for  $A \subseteq \mathbf{Sets}(X, [\mathfrak{z}])$ , its supremum  $\bigsqcup_{f \in A} f: X \rightarrow R$  with respect to  $\sqsubseteq_{X,R}$  is given by  $(\bigsqcup_{f \in A} f)(x) = \bigwedge_{f \in A} (f(x))$ . Hence Condition 2(b) is satisfied.

**Condition 3** By the definition of  $q_{\mathbf{g}, \mathfrak{z}}: [\mathfrak{z}] \rightarrow \{0, 1\}$ , for  $f_1, f_2 \in \mathbf{Sets}(X, [\mathfrak{z}])$  and  $x \in X$ ,  $f_1(x) \leq f_2(x)$  (with respect to the ordinary order) implies  $q_{\mathbf{g}, \mathfrak{z}} \circ f_1(x) \geq q_{\mathbf{g}, \mathfrak{z}} \circ f_2(x)$ . Therefore  $q_{\mathbf{g}, \mathfrak{z}} \circ (\_)$  is monotone.

By definition, if  $f(x) = \mathfrak{z}$  then  $q_{\mathbf{g}, \mathfrak{z}} \circ f(x) = 0$ . Hence  $q_{\mathbf{g}, \mathfrak{z}} \circ (\_)$  is strict.

Let  $K \subseteq \mathbf{Sets}(X, [\mathfrak{z}])$ . We have:

$$\begin{aligned}
q_{\mathbf{g}, \mathfrak{z}} \left( \bigsqcup_{f \in K} f(x) \right) = 1 &\Leftrightarrow \exists \mathfrak{n} < \mathfrak{z}. \left( \bigwedge_{f \in K} f(x) \right) = \mathfrak{n} \Leftrightarrow \exists f \in K. \exists \mathfrak{n} < \mathfrak{z}. f(x) = \mathfrak{n} \\
&\Leftrightarrow \exists f \in K. q_{\mathbf{g}, \mathfrak{z}} \circ f(x) = 1 \Leftrightarrow \bigvee_{f \in K} q_{\mathbf{g}, \mathfrak{z}} \circ f(x) = 1.
\end{aligned}$$

Hence we have  $\bigsqcup_{f \in K} (q_{\mathbf{g}, \mathfrak{z}} \circ f) = q_{\mathbf{g}, \mathfrak{z}} \circ (\bigsqcup_{f \in K} f)$ , and  $q_{\mathbf{g}, \mathfrak{z}} \circ (\_)$  is continuous.

**Condition 4** Let  $c: X \rightarrow F_{\mathbf{g}} X$  be an  $F_{\mathbf{g}}$ -coalgebra. We prove that  $\Phi_{c, r_{\mathbf{g}, \mathfrak{z}}}$  has a unique fixed point.

By Condition 2 and Theorem 2.3.2.3,  $\Phi_{c, r_{\mathbf{g}, \mathfrak{z}}}$  has the least fixed point.

We will show that  $\Phi_{c, r_{\mathbf{g}, \mathfrak{z}}}$  has a unique fixed point. Let  $f_1, f_2: X \rightarrow [\mathfrak{z}]$  be fixed points of  $\Phi_{c, r_{\mathbf{g}, \mathfrak{z}}}$ . We prove  $f_1(x) = \mathfrak{a} \Leftrightarrow f_2(x) = \mathfrak{a}$  for each  $x \in X$  and  $\mathfrak{a} < \mathfrak{z}$  by the transfinite induction on  $\mathfrak{a}$ .

When  $\mathfrak{a} = 0$  we have:

$$\begin{aligned}
f_1(x) = 0 &\Leftrightarrow r_{\mathbf{g}, \mathfrak{z}} \circ F_{\mathbf{g}} f_1 \circ c(x) = 0 && (f_1 \text{ is a fixed point of } \Phi_{c, r_{\mathbf{g}, \mathfrak{z}}}) \\
&\Leftrightarrow \pi_2(c_2(x)) = 1 && (\text{by definition}) \\
&\Leftrightarrow r_{\mathbf{g}, \mathfrak{z}} \circ F_{\mathbf{g}} f_2 \circ c(x) = 0 && (\text{by definition}) \\
&\Leftrightarrow f_2(x) = 0 && (f_2 \text{ is a fixed point of } \Phi_{c, r_{\mathbf{g}, \mathfrak{z}}}).
\end{aligned}$$

Let  $\mathfrak{a}$  be a successor ordinal such that  $\mathfrak{a} < \mathfrak{z}$ , and assume  $f(x) = \mathfrak{a}' \Leftrightarrow g(x) = \mathfrak{a}'$  for each  $x \in X$  and  $\mathfrak{a}' < \mathfrak{a}$ . Note that by  $\mathfrak{a} < \mathfrak{z}$ , we have  $\mathfrak{a} \hat{+} 1 = \mathfrak{a} + 1$ .

$$\begin{aligned}
f_1(x) = \mathfrak{a} &\Leftrightarrow r_{\mathfrak{g},\mathfrak{z}} \circ F_{\mathfrak{g}} f_1 \circ c(x) = \mathfrak{a} && (f_1 \text{ is a fixed point of } \Phi_{c,r_{\mathfrak{g},\mathfrak{z}}}) \\
&\Leftrightarrow \pi_2(c_2(x)) = 0 \text{ and } \min_{A \in \pi_1(c(x))} \sup_{x' \in A} f_1(x') \hat{+} 1 = \mathfrak{a} && (\text{by definition}) \\
&\Leftrightarrow \pi_2(c_2(x)) = 0, \forall A \in \pi_1(c(x)). \exists x' \in A. \neg(f_1(x') < \mathfrak{a} - 1) \\
&\quad \text{and } \exists A \in \pi_1(c(x)). \forall x' \in A. f_1(x') \leq \mathfrak{a} - 1 \\
&\Leftrightarrow \pi_2(c_2(x)) = 0, \forall A \in \pi_1(c(x)). \exists x' \in A. \neg(f_2(x') < \mathfrak{a} - 1) \\
&\quad \text{and } \exists A \in \pi_1(c(x)). \forall x' \in A. f_2(x') \leq \mathfrak{a} - 1 \\
&&& (\text{by the induction hypothesis}) \\
&\Leftrightarrow \pi_2(c_2(x)) = 0 \text{ and } \min_{A \in \pi_1(c(x))} \sup_{x' \in A} f_2(x') \hat{+} 1 = \mathfrak{a} \\
&\Leftrightarrow r_{\mathfrak{g},\mathfrak{z}} \circ F_{\mathfrak{g}} f_2 \circ c(x) = \mathfrak{a} && (\text{by definition}) \\
&\Leftrightarrow f_2(x) = \mathfrak{a} && (f_2 \text{ is a fixed point of } \Phi_{c,r_{\mathfrak{g},\mathfrak{z}}}).
\end{aligned}$$

Let  $\mathfrak{a}$  be a limit ordinal such that  $\mathfrak{a} < \mathfrak{z}$ , and assume  $f(x) = \mathfrak{a}' \Leftrightarrow g(x) = \mathfrak{a}'$  for each  $x \in X$  and  $\mathfrak{a}' < \mathfrak{a}$ . We have:

$$\begin{aligned}
f_1(x) = \mathfrak{a} &\Leftrightarrow r_{\mathfrak{g},\mathfrak{z}} \circ F_{\mathfrak{g}} f_1 \circ c(x) = \mathfrak{a} && (f_1 \text{ is a fixed point of } \Phi_{c,r_{\mathfrak{g},\mathfrak{z}}}) \\
&\Leftrightarrow \pi_2(c_2(x)) = 0 \text{ and } \min_{A \in \pi_1(c(x))} \sup_{x' \in A} f_1(x') \hat{+} 1 = \mathfrak{a} && (\text{by definition}) \\
&\Leftrightarrow \pi_2(c_2(x)) = 0, \forall A \in \pi_1(c(x)). \forall \mathfrak{a}' < \mathfrak{a}. \exists x' \in A. \neg(f_1(x') < \mathfrak{a}') \\
&\quad \text{and } \exists A \in \pi_1(c(x)). \forall x' \in A. f_1(x') < \mathfrak{a} \\
&\Leftrightarrow \pi_2(c_2(x)) = 0, \forall A \in \pi_1(c(x)). \forall \mathfrak{a}' < \mathfrak{a}. \exists x' \in A. \neg(f_2(x') < \mathfrak{a}') \\
&\quad \text{and } \exists A \in \pi_1(c(x)). \forall x' \in A. f_2(x') < \mathfrak{a} \\
&&& (\text{by the induction hypothesis}) \\
&\Leftrightarrow \pi_2(c_2(x)) = 0 \text{ and } \min_{A \in \pi_1(c(x))} \sup_{x' \in A} f_2(x') \hat{+} 1 = \mathfrak{a} \\
&\Leftrightarrow r_{\mathfrak{g},\mathfrak{z}} \circ F_{\mathfrak{g}} f_2 \circ c(x) = \mathfrak{a} && (\text{by definition}) \\
&\Leftrightarrow f_2(x) = \mathfrak{a} && (f_2 \text{ is a fixed point of } \Phi_{c,r_{\mathfrak{g},\mathfrak{z}}}).
\end{aligned}$$

Hence we have  $f_1(x) = \mathfrak{a} \Leftrightarrow f_2(x) = \mathfrak{a}$  for each  $x \in X$  and  $\mathfrak{a} < \mathfrak{z}$ . This immediately implies that  $f_1(x) = \mathfrak{z} \Leftrightarrow f_2(x) = \mathfrak{z}$  for each  $x \in X$ .  $\square$

We finally show that ranking arrows with respect to the ranking domain coincide with ranking functions in Definition 5.1.1.

**Proposition 5.4.3.** *Let  $\mathcal{T} = (X^{\text{Max}}, X^{\text{Min}}, E^{\text{Max}}, E^{\text{Min}}, \text{Acc})$  be a reachability game and  $\mathfrak{z}$  be an ordinal. A function  $b : X^{\text{Max}} \rightarrow [\mathfrak{z}]$  is a ranking arrow for  $c_{\mathcal{T}}$  (see Example 2.4.14) with respect to  $(r_{\mathfrak{g},\mathfrak{z}}, q_{\mathfrak{g},\mathfrak{z}}, \sqsubseteq_{[\mathfrak{z}]})$  if and only if  $b$  is a ranking function for  $\mathcal{T}$ . Moreover, for  $x \in X^{\text{Max}}$ ,  $b(x) < \mathfrak{z}$  if and only if  $q_{\mathfrak{g},\mathfrak{z}} \circ b(x) = 1$ .*

**Proof.** We have:

$$\begin{aligned}
b : X^{\text{Max}} \rightarrow [\mathfrak{z}] &\text{ is a ranking arrow for } c_{\mathcal{T}} \\
&\Leftrightarrow \forall x \in X^{\text{Max}}. b(x) \sqsubseteq r_{\mathfrak{g},\mathfrak{z}} \circ F_{\mathfrak{g}} b \circ c(x) && (\text{by Definition 5.3.3}) \\
&\Leftrightarrow \forall x \in X^{\text{Max}}. r_{\mathfrak{g},\mathfrak{z}} \circ F_{\mathfrak{g}} b \circ c(x) \leq b(x) && (\text{by the definition of } \sqsubseteq) \\
&\Leftrightarrow \forall x \in X^{\text{Max}}. (\pi_2(c(x)) = 0 \Rightarrow \min_{A \in \pi_1(c(x))} \sup_{x' \in A} b(x') \hat{+} 1 \leq b(x)) \\
&&& (\text{by the definitions of } r_{\mathfrak{g},\mathfrak{z}} \text{ and } F_{\mathfrak{g}})
\end{aligned}$$

$$\begin{aligned}
&\Leftrightarrow \forall x \in X^{\text{Max}}. (x \notin \text{Acc} \Rightarrow \min_{y:(x,y) \in E^{\text{Max}}} \sup_{x':(y,x') \in E^{\text{Min}}} b(x') \hat{+} 1 \leq b(x)) \\
&\hspace{25em} \text{(by the definition of } c_{\mathcal{T}}) \\
&\Leftrightarrow b: X \rightarrow [\mathfrak{z}] \text{ is a ranking function for } \mathcal{T} \hspace{10em} \text{(by Definition 5.1.1).}
\end{aligned}$$

It is immediate by definition that  $b(x) < \mathfrak{z} \Leftrightarrow q_{\mathfrak{g},\mathfrak{z}} \circ b(x) = 1$ .  $\square$

Note that the inequality  $q_{\mathfrak{g},\mathfrak{z}} \circ r_{\mathfrak{g},\mathfrak{z}} \sqsubseteq \sigma_{\mathfrak{g}} \circ F_{\mathfrak{g}} q_{\mathfrak{g},\mathfrak{z}}$  is strict. (see (5.4) in the proof of Proposition 5.4.2). Hence we cannot prove completeness using Proposition 5.3.6. Indeed, Example 5.1.5 shows that it does not hold.

## 5.5 Concretization to PTSs

In this section, we apply the framework in Section 5.3 to PTSs (Definition 2.2.26). We introduce two ranking domains for PTSs. They induce different notions of “ranking function” for PTSs respectively.

We model a PTS  $\mathcal{T} = ((X, \mathfrak{F}_X), \xi, \text{Acc})$  as an  $F_{\mathfrak{p}}$ -coalgebra  $c_{\mathcal{T}}: X \rightarrow \mathcal{G}X \times \{0, 1\}$  as in Example 2.4.14. We first characterize reachability probability function  $\text{Reach}_{\mathcal{T}}: X \rightarrow [0, 1]$  (Definition 2.2.27) as the least fixed point property.

**Proposition 5.5.1.** *For each  $X \in \mathbf{SB}$ , we define a partial order  $\leq_X$  over  $\mathbf{SB}(X, [0, 1])$  (here  $[0, 1]$  is equipped with the standard  $\sigma$ -algebra) by  $f \leq_X g \stackrel{\text{def.}}{\Leftrightarrow} \forall x \in X. f(x) \leq g(x)$  where the last  $\leq$  denotes the ordinary order. Let  $\leq := (\leq_X)_{X \in \mathbf{SB}}$ . We define an  $F_{\mathfrak{p}}$ -modality  $\sigma_{\mathfrak{p}}: \mathcal{G}[0, 1] \times \{0, 1\} \rightarrow [0, 1]$  over a truth-value domain  $([0, 1], \leq)$  as follows:*

$$\sigma_{\mathfrak{p}}(\varphi, t) := \begin{cases} 1 & (t = 1) \\ \int_{a \in [0, 1]} a \, \text{d}\varphi & (t = 0). \end{cases}$$

Then  $\sigma_{\mathfrak{p}}$  has least fixed points. Moreover, for a PTS  $\mathcal{T} = ((X, \mathfrak{F}_X), \xi, \text{Acc})$ , if we define an  $F_{\mathfrak{p}}$ -coalgebra  $c_{\mathcal{T}}: X \rightarrow \mathcal{G}X \times \{0, 1\}$  as in Example 2.4.14, then the least fixed point property  $\llbracket \mu \sigma_{\mathfrak{p}} \rrbracket_{c_{\mathcal{T}}}: X \rightarrow [0, 1]$  coincides with the reachability probability function  $\text{Reach}_{\mathcal{T}}$ .

**Proof.** Let  $f, g \in \mathbf{SB}(X, [0, 1])$  and assume  $f \leq g$ . Then for each  $x \in X$ ,

$$\begin{aligned}
&\Phi_{c, \sigma}(f)(x) \\
&= \sigma \circ F_{\mathfrak{p}} f \circ c(x) \hspace{15em} \text{(by definition)} \\
&= \begin{cases} 1 & (\pi_2(c(x)) = 1) \\ \int_{x' \in X} f(x') \, \text{d}(\pi_1(c(x))) & (\pi_2(c(x)) = 0) \end{cases} \hspace{2em} \text{(by the definitions of } \sigma_{\mathfrak{p}} \text{ and } F_{\mathfrak{p}}) \\
&\leq \begin{cases} 1 & (\pi_2(c(x)) = 1) \\ \int_{x' \in X} g(x') \, \text{d}(\pi_1(c(x))) & (\pi_2(c(x)) = 0) \end{cases} \hspace{2em} \text{(by } f \leq g) \\
&= \sigma \circ F_{\mathfrak{p}} g \circ c(x) \hspace{15em} \text{(by the definitions of } \sigma_{\mathfrak{p}} \text{ and } F_{\mathfrak{p}}) \\
&= \Phi_{c, \sigma}(g)(x) \hspace{15em} \text{(by definition).}
\end{aligned}$$

Hence  $\Phi_{c, \sigma_{\mathfrak{p}}}$  is monotone.

We prove  $\llbracket \mu \sigma_{\mathfrak{p}} \rrbracket_{c_{\mathcal{T}}} = \text{Reach}_{\mathcal{T}}$ . It is easy to see that for each  $F_{\mathfrak{g}}$ -coalgebra  $c: X \rightarrow F_{\mathfrak{p}}X$  there exists a PTS  $\mathcal{T}$  such that  $c = c_{\mathcal{T}}$ . Hence it suffices to show that for each PTS  $\mathcal{T} = ((X, \mathfrak{F}_X), \xi, \text{Acc})$ , the reachability probability function  $\text{Reach}_{\mathcal{T}}: X \rightarrow [0, 1]$  is the least fixed point of  $\Phi_{c_{\mathcal{T}}, \sigma_{\mathfrak{p}}}: \mathbf{SB}(X, [0, 1]) \rightarrow \mathbf{Sets}(X, [0, 1])$ .

We first show that  $\text{Reach}_{\mathcal{T}}$  is a fixed point of  $\Phi_{c_{\mathcal{T}}, \sigma_{\mathfrak{p}}}$ . For  $x \in X$ , we have:

$$\begin{aligned}
& \Phi_{c_{\mathcal{T}}, \sigma_{\mathfrak{p}}}(\text{Reach}_{\mathcal{T}})(x) \\
&= (\sigma_{\mathfrak{p}} \circ F_{\mathfrak{p}} f \circ c_{\mathcal{T}})(x) && \text{(by definition)} \\
&= \begin{cases} 1 & (x \in \text{Acc}) \\ \int_{x' \in X} \text{Reach}_{\mathcal{T}}(x') d(\xi(x)) & (x \notin \text{Acc}) \end{cases} \\
& && \text{(by the definitions of } c_{\mathcal{T}}, \sigma_{\mathfrak{p}} \text{ and } F_{\mathfrak{p}}) \\
&= \begin{cases} 1 & (x \in \text{Acc}) \\ \int_{x' \in X} \lim_{k \rightarrow \infty} \text{Reach}_{\mathcal{T}}^k(x') d(\xi(x)) & (x \notin \text{Acc}) \end{cases} \\
& && \text{(by the definition of } \text{Reach}_{\mathcal{T}}) \\
&= \begin{cases} 1 & (x \in \text{Acc}) \\ \lim_{k \rightarrow \infty} \int_{x' \in X} \text{Reach}_{\mathcal{T}}^k(x') d(\xi(x)) & (x \notin \text{Acc}) \end{cases} \\
& && \text{(by the dominated convergence theorem, e.g. [8, Theorem 1.6.9])} \\
&= \begin{cases} 1 & (x \in \text{Acc}) \\ \lim_{k \rightarrow \infty} \text{Reach}_{\mathcal{T}}^{k+1}(x) & (x \notin \text{Acc}) \end{cases} && \text{(by the definition of } \text{Reach}_{\mathcal{T}}^k(x)) \\
&= \lim_{k \rightarrow \infty} \text{Reach}_{\mathcal{T}}^k(x) && \text{(by the definition of } \text{Reach}_{\mathcal{T}}^k(x)) \\
&= \text{Reach}_{\mathcal{T}}(x) && \text{(by the definition of } \text{Reach}_{\mathcal{T}}).
\end{aligned}$$

Hence  $\text{Reach}_{\mathcal{T}}$  is a fixed point of  $\Phi_{c_{\mathcal{T}}, \sigma_{\mathfrak{p}}}$ .

It remains to show that it is the least fixed point. Let  $f: X \rightarrow [0, 1]$  be a fixed point of  $\Phi_{c_{\mathcal{T}}, \sigma_{\mathfrak{p}}}$ . We will prove  $\text{Reach}_{\mathcal{T}} \leq_X f$ . To this end, by the definition of  $\text{Reach}_{\mathcal{T}}$ , it suffices to prove  $\text{Reach}_{\mathcal{T}}^k(x) \leq f(x)$  for each  $x \in X$  and  $k \in \mathbb{N}$ . We prove this by the induction on  $n$ .

For  $k = 0$ , it is immediate from that  $\text{Reach}_{\mathcal{T}}^0(x) = 0$ .

For  $k > 0$ , we have:

$$\begin{aligned}
& \text{Reach}_{\mathcal{T}}^k(x) \\
&= \begin{cases} 1 & (x \in \text{Acc}) \\ \int_{x' \in X} \text{Reach}_{\mathcal{T}}^{k-1}(x') d\xi(x) & (x \notin \text{Acc}) \end{cases} && \text{(by the definition of } \text{Reach}_{\mathcal{T}}^k(x)) \\
&\leq \begin{cases} 1 & (x \in \text{Acc}) \\ \int_{x' \in X} f(x') d\xi(x) & (x \notin \text{Acc}) \end{cases} && \text{(by the induction hypothesis)} \\
&= (\sigma_{\mathfrak{p}} \circ F_{\mathfrak{p}} f \circ c)(x) && \text{(by the definitions of } c_{\mathcal{T}}, F_{\mathfrak{p}} \text{ and } \sigma_{\mathfrak{p}}) \\
&= f(x) && (f \text{ is a fixed point of } \Phi_{c_{\mathcal{T}}, \sigma_{\mathfrak{p}}}).
\end{aligned}$$

Hence we have  $\text{Reach}_{\mathcal{T}}^k(x) \leq f(x)$  for each  $x \in X$  and  $k \in \mathbb{N}$ , and therefore  $\text{Reach}_{\mathcal{T}}$  is the least fixed point of  $\Phi_{c_{\mathcal{T}}, \sigma_{\mathfrak{p}}}$ .  $\square$

### 5.5.1 Distribution-valued Ranking Supermartingale

In this and the next sections, we introduce ranking domains for  $\sigma_{\mathfrak{p}}$  respectively. In this section, we introduce a ranking domain such that  $R$  is the set of distributions over  $\mathbb{N} \cup \{\infty\}$ .

We first introduce some notations.

**Notation 5.5.2.** We write  $\mathbb{N}_{\infty}$  for  $\mathbb{N} \cup \{\infty\}$ . We define a  $\sigma$ -algebra  $\mathfrak{F}_{\mathcal{DN}_{\infty}}$  over  $\mathcal{DN}_{\infty} = \{f : \mathbb{N}_{\infty} \rightarrow [0, 1] \mid \sum_{n \in \mathbb{N}_{\infty}} f(n) = 1\}$  as the smallest  $\sigma$ -algebra that makes a function  $\text{ev}_a : \mathcal{DN}_{\infty} \rightarrow [0, 1]$  defined by  $\text{ev}_a(\varphi) := \varphi(a)$  measurable for

each  $a \in \mathbb{N}_\infty$  (cf. the definition of  $\mathcal{G}$ ). As mentioned in Section 2.1, we might write  $\mathcal{DN}_\infty$  for  $(\mathcal{DN}_\infty, \mathfrak{F}_{\mathcal{DN}_\infty})$ . For  $\varphi \in \mathcal{DN}_\infty$  and  $A \subseteq \mathbb{R} \cup \{\infty\}$ ,  $\varphi(A)$  denotes  $\sum_{a \in A \cap \mathbb{N}_\infty} \varphi(a)$ . We let  $[a, b] = \emptyset$  when  $b < a$ .

In this section, we consider the following ranking domain.

**Proposition 5.5.3.** *We define an  $F_p$ -algebra  $r_p : F_p \mathcal{DN}_\infty \rightarrow \mathcal{DN}_\infty$ , a function  $q_p : \mathcal{DN}_\infty \rightarrow [0, 1]$  and a family  $\sqsubseteq_{\mathcal{DN}_\infty}$  of partial orders as follows.*

- $r_p : \mathcal{GDN}_\infty \times \{0, 1\} \rightarrow \mathcal{DN}_\infty$  is defined by:

$$r_p(\Gamma, t)(a) := \begin{cases} 1 & (t = 1) \\ 0 & (t = 0, a = 0) \\ \int_{\varphi \in \mathcal{DN}_\infty} \varphi(a - 1) d\Gamma & (t = 0, a > 0). \end{cases}$$

- $q_p(\varphi) := \varphi([0, \infty))$ .
- We define a partial order  $\sqsubseteq_{\mathcal{DN}_\infty}$  over  $\mathcal{DN}_\infty$  by

$$\varphi \sqsubseteq_{\mathcal{DN}_\infty} \varphi' \stackrel{\text{def.}}{\iff} \forall a \in \mathbb{N}. \varphi([0, a]) \leq \varphi'([0, a]),$$

Moreover, for each  $X \in \mathbf{SB}$ , we define a partial order  $\sqsubseteq_{X, \mathcal{DN}_\infty}$  by  $f \sqsubseteq_{X, \mathcal{DN}_\infty} g \stackrel{\text{def.}}{\iff} \forall x \in X. f(x) \sqsubseteq_{\mathcal{DN}_\infty} g(x)$ , and define a family  $\sqsubseteq_{\mathcal{DN}_\infty}$  of partial orders by  $\sqsubseteq_{\mathcal{DN}_\infty} := (\sqsubseteq_{X, \mathcal{DN}_\infty})_{X \in \mathbf{SB}}$ .

Then a triple  $(r_p, q_p, \sqsubseteq_{\mathcal{DN}_\infty})$  is a ranking domain for  $\sigma_p$ . Moreover we have  $q_p \circ r_p = \sigma_p \circ F_p q_p$  (cf. Proposition 5.3.6).

We use the following lemma in the proof of the above proposition.

**Lemma 5.5.4.** *For every nondecreasing function  $G : \mathbb{N} \rightarrow [0, 1]$ , there exists a unique distribution  $\varphi$  over  $\mathbb{N}_\infty$  such that  $\varphi([0, a]) = G(a)$  for each  $a \in \mathbb{N}$ .*

**Proof.** We define  $\varphi : \mathbb{N}_\infty \rightarrow \mathbb{R}$  by

$$\varphi(a) = \begin{cases} G(a) & (a = 0) \\ G(a) - G(a - 1) & (0 < a < \infty) \\ 1 - \lim_{a' \rightarrow \infty} G(a') & (a = \infty). \end{cases}$$

By definition and that  $G$  is nondecreasing,  $0 \leq \varphi(a) \leq 1$  for each  $a$ . By its definition, we have  $\sum_{a \in \mathbb{N}_\infty} \varphi(a) = 1$ . Hence  $\varphi \in \mathcal{DN}_\infty$ .

Let  $\varphi' \in \mathcal{DN}_\infty$  and assume  $\varphi'([0, a]) = G(a)$  for each  $a \in \mathbb{N}$ . Then we have  $\varphi(0) = G(0) = \varphi'([0, 0]) = \varphi'(0)$ . Moreover for each  $a \in \mathbb{N} \setminus \{0\}$ , we have:

$$\varphi(a) = G(a) - G(a - 1) = \varphi'([0, a]) - \varphi'([0, a - 1]) = \varphi'(a).$$

Therefore we have  $\varphi(a) = \varphi'(a)$  for each  $a \in \mathbb{N}$ , and this implies  $\varphi(\infty) = \varphi'(\infty)$ . Hence the uniqueness is proved, and this concludes the proof.  $\square$

We next prove that  $\sqsubseteq_{X, \mathcal{DN}_\infty}$  is indeed a partial order.

**Lemma 5.5.5.** *The order  $\sqsubseteq_{X, \mathcal{DN}_\infty}$  is a partial order.*

**Proof.** We first prove that  $\sqsubseteq_{X, \mathcal{DN}_\infty}$  is a partial order. Reflexivity and transitivity are immediate from those of the standard order  $\leq$  over  $[0, 1]$ . Assume that  $f \sqsubseteq_{X, \mathcal{DN}_\infty} g$  and  $g \sqsubseteq_{X, \mathcal{DN}_\infty} f$ . By the definition of  $\sqsubseteq_{X, \mathcal{DN}_\infty}$ , we have  $f(x)([0, a]) = g(x)([0, a])$  for each  $x \in X$  and  $a \in [0, \infty]$ . Then by Lemma 5.5.4, we have  $f = g$ . Hence antisymmetry is also satisfied.  $\square$

**Proof** (Proposition 5.5.3). We prove that Conditions 1–4 are satisfied.



**Condition 1** Let  $(\Gamma, t) \in F_{\mathbf{p}}\mathcal{DN}_{\infty} = \mathcal{G}\mathcal{DN}_{\infty} \times \{0, 1\}$ . If  $t = 1$  then by the definitions of  $r_{\mathbf{p}}$  and  $\sigma_{\mathbf{p}}$ , we have the following (recall that  $\delta_0$  denotes the Dirac distribution):

$$q_{\mathbf{p}} \circ r_{\mathbf{p}}(\Gamma, t) = q_{\mathbf{p}}(\delta_0) = 1 = \sigma_{\mathbf{p}}(\mathcal{G}q_{\mathbf{p}}(\Gamma), t) = \sigma_{\mathbf{p}} \circ F_{\mathbf{p}}q_{\mathbf{p}}(\Gamma, t).$$

Assume  $t = 0$ . Then we have:

$$\begin{aligned} q_{\mathbf{p}} \circ r_{\mathbf{p}}(\Gamma, t) &= r_{\mathbf{p}}(\Gamma, 1)([0, \infty)) && \text{(by the definition of } q_{\mathbf{p}}) \\ &= \sum_{a=0}^{\infty} r_{\mathbf{p}}(\Gamma, 1)(a) \\ &= \sum_{a=1}^{\infty} \int_{\varphi \in \mathcal{DN}_{\infty}} \varphi(a-1) \, d\Gamma && \text{(by the definition of } r_{\mathbf{p}}) \\ &= \int_{\varphi \in \mathcal{DN}_{\infty}} \sum_{a=1}^{\infty} \varphi(a-1) \, d\Gamma \\ &= \int_{\varphi \in \mathcal{DN}_{\infty}} q_{\mathbf{p}}(\varphi) \, d\Gamma && \text{(by the definition of } q_{\mathbf{p}}) \\ &= \sigma_{\mathbf{p}} \circ F_{\mathbf{p}}q_{\mathbf{p}}(\Gamma, t) && \text{(by the definition of } \sigma_{\mathbf{p}}). \end{aligned}$$

Hence Condition 1 is satisfied. We can also see that  $q_{\mathbf{p}} \circ r_{\mathbf{p}} = \sigma_{\mathbf{p}} \circ F_{\mathbf{p}}q_{\mathbf{p}}$ .

**Condition 2** Let  $f_1, f_2: X \rightarrow \mathcal{DN}_{\infty}$  and assume  $f_1 \sqsubseteq_{X, \mathcal{DN}_{\infty}} f_2$ . Let  $x \in X$  and assume that  $c(x) = (\psi, t) \in F_{\mathbf{p}}X = \mathcal{G}X \times \{0, 1\}$ .

If  $t = 1$  then by definition we have:

$$r_{\mathbf{p}} \circ F_{\mathbf{p}}f_1(\psi, t) = r_{\mathbf{p}} \circ F_{\mathbf{p}}f_2(\psi, t) = \delta_0.$$

Hence we have  $\Phi_{c, \sigma_{\mathbf{p}}}(f_1)(x) = \Phi_{c, \sigma_{\mathbf{p}}}(f_2)(x)$ .

Assume  $t = 0$ . Let  $a \in \mathbb{N}$ . If  $a = 0$  then by the definition of  $r_{\mathbf{p}}$ , we have:

$$r_{\mathbf{p}} \circ F_{\mathbf{p}}f_1(\psi, t)([0, a]) = r_{\mathbf{p}} \circ F_{\mathbf{p}}f_2(\psi, t)([0, a]) = 0.$$

If  $a > 0$ , we have:

$$\begin{aligned} &r_{\mathbf{p}} \circ F_{\mathbf{p}}f_1(\psi, t)([0, a]) \\ &= \int_{x' \in X} f_1(x')([0, a-1]) \, d\psi && \text{(by the definition of } r_{\mathbf{p}}) \\ &\leq \int_{x' \in X} f_2(x')([0, a-1]) \, d\psi && \text{(by } f_1 \sqsubseteq_{X, \mathcal{DN}_{\infty}} f_2) \\ &= r_{\mathbf{p}} \circ F_{\mathbf{p}}f_2(\psi, t)([0, a]) && \text{(by the definition of } r_{\mathbf{p}}). \end{aligned}$$

Therefore by the definition of  $\sqsubseteq_{\mathcal{DN}_{\infty}}$ , we have  $r_{\mathbf{p}} \circ F_{\mathbf{p}}f_1(\psi, t) \sqsubseteq_{\mathcal{DN}_{\infty}} r_{\mathbf{p}} \circ F_{\mathbf{p}}f_2(\psi, t)$ , and  $\Phi_{c, r_{\mathbf{p}}}$  is monotone.

It is easy to see that a function  $f: X \rightarrow \mathcal{DN}_{\infty}$  defined by  $f(x) := \delta_{\infty}$  is the least element of  $(\mathbf{SB}(X, \mathcal{DN}_{\infty}), \sqsubseteq_{X, \mathcal{DN}_{\infty}})$ .

We prove that  $(\mathbf{SB}(X, \mathcal{DN}_{\infty}), \sqsubseteq_{X, \mathcal{DN}_{\infty}})$  is  $\omega$ -complete and  $\Phi_{c, r_{\mathbf{p}}}$  is  $\omega$ -continuous. Let  $f_0, f_1, \dots \in \mathbf{SB}(X, \mathcal{DN}_{\infty})$  and assume that they constitute an increasing sequence, i.e.  $f_0 \sqsubseteq_{X, \mathcal{DN}_{\infty}} f_1 \sqsubseteq_{\mathcal{DN}_{\infty}} \dots$ .

We define  $G: X \rightarrow [0, 1]^{\mathbb{N}}$  by  $G(x)(a) = \lim_{n \rightarrow \infty} f_n(x)([0, a])$ . Note that for each  $x \in X$  and  $f \in K$ ,  $a \leq b$  implies  $f(x)([0, a]) \leq f(x)([0, b])$ . Hence by the monotonicity of supremums,  $G$  is nondecreasing. Therefore by Lemma 5.5.4, for

each  $x \in X$  there exists unique  $\varphi_x \in \mathcal{DN}_\infty$  such that  $\varphi_x([0, a]) = G(x)(a)$  for each  $a \in [0, \infty]$ . We define  $f: X \rightarrow \mathcal{DN}_\infty$  by  $f(x) := \varphi_x$ .

By its definition,  $f$  is the supremum of  $f_0, f_1, \dots$  if we ignore measurability. We prove measurability of  $f$ . We have:

$$\begin{aligned}
& f: X \rightarrow \mathcal{DN}_\infty \text{ is measurable} \\
& \Leftrightarrow \forall A \in \mathfrak{F}_{\mathcal{DN}_\infty}. f^{-1}(A) \in \mathfrak{F}_X \quad (\text{by definition}) \\
& \Leftrightarrow \forall a \in \mathbb{N}_\infty. \forall B \in \mathfrak{F}_{[0,1]}. \{x \in X \mid f(x)(a) \in B\} \in \mathfrak{F}_X \quad (\text{by definition}) \\
& \Leftrightarrow \forall a \in \mathbb{N}_\infty. f(\_)(a): X \rightarrow [0, 1] \text{ is measurable} \\
& \Leftrightarrow \forall a \in \mathbb{N}_\infty. f(\_)([0, a]): X \rightarrow [0, 1] \text{ is measurable.}
\end{aligned}$$

As the limit of Borel-measurable functions is Borel measurable (see e.g. [8, Theorem 1.5.4]), the last statement holds. Hence  $f$  is measurable.

We next prove that  $\Phi_{c,r_p}$  is  $\omega$ -continuous. For each  $x \in X$  and  $a \in \mathbb{N}$ ,

$$\begin{aligned}
& \Phi_{c,r_p}\left(\bigsqcup_{i \in \omega} f_i\right)(x)([0, a]) \\
& = r_p \circ F_p\left(\bigsqcup_{i \in \omega} f_i\right) \circ c(x) \quad (\text{by definition}) \\
& = \begin{cases} 1 & (\pi_2(c(x)) = 1) \\ \int_{x' \in X} \lim_{i \rightarrow \infty} f_i(x')([0, a-1]) d(\pi_1(c(x))) & (\pi_2(c(x)) = 0) \end{cases} \\
& \quad (\text{by the definitions of } \sigma_p \text{ and } F_p) \\
& = \begin{cases} 1 & (\pi_2(c(x)) = 1) \\ \lim_{i \rightarrow \infty} \int_{x' \in X} f_i(x')([0, a-1]) d(\pi_1(c(x))) & (\pi_2(c(x)) = 0) \end{cases} \\
& \quad (\text{by the dominated convergence theorem}) \\
& = \bigsqcup_{i \in \omega} r_p \circ F_p f_i \circ c(x)([0, a]) \quad (\text{by the definitions of } \sigma_p \text{ and } F_p) \\
& = \bigsqcup_{i \in \omega} \Phi_{c,\sigma}(f_i)(x)([0, a]) \quad (\text{by definition}).
\end{aligned}$$

Hence by Lemma 5.5.4, we have  $\Phi_{c,r_p}(\bigsqcup_{i \in \omega} f) = \bigsqcup_{i \in \omega} \Phi_{c,r_p}(f_i)$ . Therefore  $\Phi_{c,r_p}$  is  $\omega$ -continuous.

**Condition 3** We first prove that  $q_p$  is monotone. Let  $f_1, f_2: X \rightarrow \mathcal{DN}_\infty$  and assume that  $f_1 \sqsubseteq_{\mathcal{DN}_\infty} f_2$ . Then we have:

$$\begin{aligned}
q_p \circ f_1(x) & = f_1(x)([0, \infty)) \quad (\text{by the definition of } q_p) \\
& = \lim_{a \rightarrow \infty} f_1(x)([0, a]) \\
& \leq \lim_{a \rightarrow \infty} f_2(x)([0, a]) \quad (\text{by } f_1 \sqsubseteq_{\mathcal{DN}_\infty} f_2) \\
& = f_2(x)([0, \infty)) \\
& = q_p \circ f_2(x). \quad (\text{by the definition of } q_p)
\end{aligned}$$

Hence we have  $q_p \circ f_1 \leq q_p \circ f_2$ , and therefore  $q_p \circ (\_)$  is monotone.

Recall that the least element  $\perp_{X, \mathcal{DN}_\infty}$  of  $\mathcal{Kl}(\mathcal{G}_s)(X, \mathcal{DN}_\infty)$  is given by  $\perp_{X, \mathcal{DN}_\infty}(x) = \delta_\infty$  (i.e. the Dirac distribution). By the definition of  $q_p$ , we have  $q_p(\delta_\infty) = 0$ . Hence  $q_p \circ (\_)$  is strict.

We prove that  $q_p$  is  $\omega$ -continuous. We define  $f_0, f_1, \dots$  and  $f$  in  $\mathbf{SB}(X, \mathcal{DN}_\infty)$  as in the proof of Condition 2. As the limit of Borel-measurable functions is

Borel measurable (see e.g. [8, Theorem 1.5.4]), we can calculate the supremum of  $q_p \circ f_0 \leq_X q_p \circ f_1 \leq_X \dots$  in the pointwise manner. For each  $x \in X$ , we have:

$$\begin{aligned}
q_p \left( \bigsqcup_{i \in \omega} f_i \right) (x) &= \lim_{a \rightarrow \infty} \left( \bigsqcup_{i \in \omega} f_i \right) (x) ([0, a]) && \text{(by the definition of } q_p) \\
&= \lim_{a \rightarrow \infty} \lim_{i \rightarrow \infty} f_i(x) ([0, a]) && \text{(by the proof of Condition 2)} \\
&= \lim_{i \rightarrow \infty} \lim_{a \rightarrow \infty} f_i(x) ([0, a]) \\
&= \lim_{i \rightarrow \infty} q_p \circ f_i(x) && \text{(by the definition of } q_p) \\
&= \left( \bigvee_{i \in \omega} q_p \circ f_i \right) (x).
\end{aligned}$$

Therefore  $q_p \circ (\_)$  is  $\omega$ -continuous.

**Condition 4** Let  $c: X \rightarrow F_p X$  be an  $F_p$ -coalgebra. It suffices to show that the function  $\Phi_{c, r_p}: \mathbf{SB}(X, \mathcal{DN}_\infty) \rightarrow \mathbf{SB}(X, \mathcal{DN}_\infty)$  has a unique fixed point.

By Condition 2 proved above and Theorem 2.3.2.2,  $\Phi_{c, r_{\mathbf{g}, \delta}}$  has the least fixed point. We prove that this is the unique fixed point. Let  $f_1, f_2: X \rightarrow \mathcal{DN}_\infty$  be fixed points of  $\Phi_{c, r_p}$ . We prove

$$f_1(x)(a) = f_2(x)(a) \quad (5.5)$$

for each  $x \in X$  and  $a \in \mathbb{N}$ .

Let  $x \in X$  and assume  $\pi_2(c(x)) = 1$ . Then we have:

$$\begin{aligned}
f_1(x) &= r_p \circ F_p f_1 \circ c(x) && (f_1 \text{ is a fixed point of } \Phi_{c, r_p}) \\
&= \delta_0 && \text{(by definition)} \\
&= r_p \circ F_p f_2 \circ c(x) && \text{(by definition)} \\
&= f_2(x) && (f_2 \text{ is a fixed point of } \Phi_{c, r_p}).
\end{aligned}$$

Hence we have  $f_1(x)(a) = f_2(x)(a)$  for each  $a$ .

Let  $x \in X$  and assume  $\pi_2(c(x)) = 0$ . We prove (5.5) for each  $a \in \mathbb{N}$  by the induction on  $a$ .

If  $a = 0$  then by the definition of  $r_p$ , we have

$$f_1(x)(a) = r_p \circ F_p f_1 \circ c(x)(a) = 0 = r_p \circ F_p f_2 \circ c(x)(a) = f_2(x)(a).$$

Let  $a > 0$  and assume  $f_1(x')(a') = f_2(x')(a')$  for each  $x' \in X$  and  $a' < a$ .

$$\begin{aligned}
f_1(x)(a) &= r_p \circ F_p f_1 \circ c(x)(a) && (f_1 \text{ is a fixed point of } \Phi_{c, r_p}) \\
&= \int_{x' \in X} f_1(x')(a-1) d(\pi_1(c(x))) && \text{(by the definition of } r_p) \\
&= \int_{x' \in X} f_2(x')(a-1) d(\pi_1(c(x))) && \text{(by the induction hypothesis)} \\
&= r_p \circ F_p f_2 \circ c(x)(a) && \text{(by the definition of } r_p) \\
&= f_2(x)(a) && (f_2 \text{ is a fixed point of } \Phi_{c, r_p}).
\end{aligned}$$

Therefore we have  $f_1(x)(a) = f_2(x)(a)$  for each  $x \in X$  and  $a \in \mathbb{N}$ , and this implies that  $f_x(x)(\infty) = f_2(x)(\infty)$  for each  $x \in X$ . This concludes the proof.  $\square$

In automata-theoretic terms, a ranking arrow with respect to  $(r_p, q_p, \sqsubseteq_{\mathcal{DN}_\infty})$  and its soundness theorem are given as follows.

**Definition 5.5.6** (distribution-valued ranking function). Let  $\mathcal{T} = ((X, \mathfrak{F}_X), \xi, \text{Acc})$  be a PTS. A *distribution-valued ranking function* for  $\mathcal{T}$  is a measurable function  $b: X \rightarrow \mathcal{DN}_\infty$  that satisfies the following:

$$\forall x \in X \setminus \text{Acc}. \forall a \in \mathbb{N}. b(x)([0, a]) \leq \int_{x' \in X} b(x')([0, a-1]) d\xi(x). \quad (5.6)$$

**Theorem 5.5.7.** *Let  $b: X \rightarrow \mathcal{DN}_\infty$  be a distribution-valued ranking function for a PTS  $\mathcal{T} = ((X, \mathfrak{F}_X), \xi, \text{Acc})$ . Then*

$$\forall x \in X. b(x)([0, \infty)) \leq \text{Reach}_{\mathcal{T}}(x).$$

**Proof.** We define an  $F_{\mathfrak{p}}$ -coalgebra  $c_{\mathcal{T}}: X \rightarrow F_{\mathfrak{p}}X$  as in Example 2.4.14. By the definitions of  $r_{\mathfrak{p}}$  and  $F_{\mathfrak{p}}$ , we can easily see that  $b: X \rightarrow \mathcal{DN}_\infty$  is a ranking arrow for  $c_{\mathcal{T}}$  with respect to  $(r_{\mathfrak{p}}, q_{\mathfrak{p}}, \sqsubseteq_{\mathcal{DN}_\infty})$  (i.e.  $b(x)([0, a]) \leq r_{\mathfrak{p}} \circ F_{\mathfrak{p}}b \circ c_{\mathcal{T}}(x)([0, a])$  for each  $x \in X \setminus \text{Acc}$  and  $a \in \mathbb{N}$ ) if and only if  $b$  is a distribution-valued ranking function for  $\mathcal{T}$ . Moreover, by the definition of  $q_{\mathfrak{p}}$ , we have  $b(x)([0, \infty)) = q_{\mathfrak{p}} \circ b(x)$ .

Hence the theorem is immediate by Theorem 5.3.4.  $\square$

By the above soundness theorem, we can use distribution-valued ranking function for underapproximating the reachability probability of PTSs. This means that we can do *quantitative reasoning* using distribution-valued ranking function. This is in contrast to a well-known ranking function-like notion for PTSs called *ranking supermartingales* [19, 34] which can verify a *qualitative* property called *almost-sure termination*, i.e. that the reachability probability is 1 (see also Definition 6.0.1 and Theorem 6.0.2).

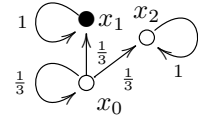
As we have seen in Proposition 5.5.3, we have  $q_{\mathfrak{p}} \circ r_{\mathfrak{p}} = \sigma_{\mathfrak{p}} \circ F_{\mathfrak{p}}q_{\mathfrak{p}}$ . Hence by Proposition 5.3.6, we have the following completeness theorem.

**Theorem 5.5.8.** *For a PTS  $\mathcal{T} = ((X, \mathfrak{F}_X), \xi, \text{Acc})$ , there exists a distribution-valued ranking function  $b: X \rightarrow \mathcal{DN}_\infty$  such that*

$$\forall x \in X. b(x)([0, \infty)) = \text{Reach}_{\mathcal{T}}(x). \quad \square$$

**Example 5.5.9.** We define a PTS  $\mathcal{T} = ((X, \mathfrak{F}_X), \xi, \text{Acc})$

by  $X = \{x_0, x_1, x_2\}$ ,  $\mathfrak{F}_X = \mathcal{P}X$ ,  $\xi(x_0) = [x_0 \mapsto \frac{1}{3}, x_1 \mapsto \frac{1}{3}, x_2 \mapsto \frac{1}{3}]$ ,  $\xi(x_1) = [x_1 \mapsto 1]$  and  $\xi(x_2) = [x_2 \mapsto 1]$  and  $\text{Acc} = \{x_1\}$ . The function  $b: X \rightarrow \mathcal{DN}_\infty$  defined by  $b(x_0) = [i \mapsto 1/3^{i+1}, \infty \mapsto 1/2]$ ,  $b(x_1) = [0 \mapsto 1]$  and  $b(x_2) = [\infty \mapsto 1]$  is a distribution-valued ranking function. Hence we can conclude  $\text{Reach}_{\mathcal{T}}(x_0) \geq (1 - \frac{1}{2}) = \frac{1}{2}$ . (As mentioned in Section 2.1, we are identifying  $\mathcal{G}X$  with  $\mathcal{D}X$ .)



## 5.5.2 $\gamma$ -Scaled Submartingale

Fix a real number  $\gamma \in [0, 1]$ . We next consider the following ranking domain.

**Proposition 5.5.10.** *We assume that  $[0, 1]$  is equipped with the standard  $\sigma$ -algebra. We define an  $F_{\mathfrak{p}}$ -algebra  $r'_{\mathfrak{p}, \gamma}: F_{\mathfrak{p}}[0, 1] \rightarrow [0, 1]$ , a function  $q_{\mathfrak{p}}: [0, 1] \rightarrow [0, 1]$  and a family  $\leq$  of partial orders as follows.*

- $r'_{\mathfrak{p}, \gamma}: \mathcal{G}[0, 1] \times \{0, 1\} \rightarrow [0, 1]$  is define as follows:

$$r'_{\mathfrak{p}, \gamma}(\varphi, t) = \begin{cases} 1 & (t = 1) \\ \gamma \cdot \int_{a \in [0, 1]} a d\varphi & (\text{otherwise}). \end{cases}$$

- $q'_p : [0, 1] \rightarrow [0, 1]$  is the identity function  $\text{id}_{[0,1]}$ .
- $\leq := (\leq_X)_{X \in \mathbf{SB}}$  is defined as in Proposition 5.5.1.

Then  $(r'_{p,\gamma}, q'_p, \leq)$  is a ranking domain for  $\sigma_p$ .

**Proof.** We prove that Conditions 1–4 are satisfied.

**Condition 1** Let  $(\varphi, t) \in F_p[0, 1] = \mathcal{G}[0, 1] \times \{0, 1\}$ . Then we have:

$$\begin{aligned}
q'_p \circ r'_{p,\gamma}(\varphi, t) &= \begin{cases} q'_p(1) & (t = 1) \\ q'_p(\gamma \cdot \int_{a \in [0,1]} a \, d\varphi) & (t = 0) \end{cases} && \text{(by the definition of } r'_{p,\gamma}) \\
&= \begin{cases} 1 & (t = 1) \\ \gamma \cdot \int_{a \in [0,1]} a \, d\varphi & (t = 0) \end{cases} && \text{(by the definition of } q'_p) \\
&\leq \begin{cases} 1 & (t = 1) \\ \int_{a \in [0,1]} a \, d\varphi & (t = 0) \end{cases} && \text{(by } \gamma < 1) \\
&= \sigma_p(\varphi, t) && \text{(by the definition of } \sigma_p) \\
&= \sigma_p \circ F_p q'_p(\varphi, t) && \text{(by the definition of } q'_p).
\end{aligned}$$

Hence we have  $q'_p \circ r'_{p,\gamma} \leq \sigma_p \circ F_p q'_p$ .

**Condition 2** We first prove that  $\Phi_{c,r'_{p,\gamma}}$  is monotone. Let  $f_1, f_2: X \rightarrow [0, 1]$  and assume that  $f_1 \leq_X f_2$ . Let  $x \in X$  and assume that  $c(x) = (\psi, t) \in F_p X = \mathcal{G}X \times \{0, 1\}$ . Then we have:

$$\begin{aligned}
r'_{p,\gamma} \circ F_p f_1(\psi, t) &= \begin{cases} 1 & (t = 1) \\ \int_{x \in X} f_1(x) \, d\psi & (t = 0) \end{cases} && \text{(by the definitions of } F_p \text{ and } r'_{p,\gamma}) \\
&\leq \begin{cases} 1 & (t = 1) \\ \int_{x \in X} f_2(x) \, d\psi & (t = 0) \end{cases} && \text{(by } f_1 \leq f_2) \\
&= r'_{p,\gamma} \circ F_p f_2(\psi, t) && \text{(by the definitions of } r'_{p,\gamma} \text{ and } F_p).
\end{aligned}$$

Hence  $\Phi_{c,r'_{p,\gamma}}$  is monotone.

It is easy to see that the least element of  $(\mathbf{SB}(X, [0, 1]), \leq_X)$  is given by the function that maps each  $x \in X$  to 0.

We show that  $(\mathbf{SB}(X, [0, 1]), \leq_X)$  is  $\omega$ -complete and  $\Phi_{c,r'_{p,\gamma}}$  is  $\omega$ -continuous. Let  $f_0, f_1, \dots \in \mathbf{SB}(X, [0, 1])$  and assume  $f_0 \leq_X f_1 \leq_X \dots$ . We define  $f: X \rightarrow [0, 1]$  by  $f(x) := \bigvee_{i \in \omega} f_i(x)$ . As the limit of Borel-measurable functions is Borel measurable (see e.g. [8, Theorem 1.5.4]),  $f$  is measurable. It is easy to see that  $f$  is the supremum of a chain  $f_0, f_1, \dots$ .

Moreover, we have:

$$\begin{aligned}
\Phi_{c,r'_{p,\gamma}}\left(\bigvee_{i \in \omega} f_i\right)(x) &= r'_{p,\gamma} \circ F_p\left(\bigvee_{i \in \omega} f_i\right) \circ c(x) && \text{(by definition)} \\
&= \begin{cases} 1 & (\pi_2(c(x)) = 1) \\ \gamma \cdot \int_{x' \in X} \lim_{i \rightarrow \infty} f_i(x') \, d(\pi_1(c(x))) & (\pi_2(c(x)) = 0) \end{cases} && \text{(by the definitions of } r'_{p,\gamma} \text{ and } F_p) \\
&= \begin{cases} 1 & (\pi_2(c(x)) = 1) \\ \lim_{i \rightarrow \infty} \gamma \cdot \int_{x' \in X} f_i(x') \, d(\pi_1(c(x))) & (\pi_2(c(x)) = 0) \end{cases} && \text{(by the dominated convergence theorem)}
\end{aligned}$$

$$\begin{aligned}
&= \bigvee_{i \in \omega} r'_{\mathfrak{p}, \gamma} \circ F_{\mathfrak{p}} f_i \circ c(x) && \text{(by the definitions of } r'_{\mathfrak{p}, \gamma} \text{ and } F_{\mathfrak{p}}) \\
&= \bigvee_{i \in \omega} \Phi_{c, r'_{\mathfrak{p}, \gamma}}(f_i)(x) && \text{(by definition)}.
\end{aligned}$$

Hence  $\Phi_{c, r'_{\mathfrak{p}, \gamma}}$  is  $\omega$ -continuous.

**Condition 3** Immediate from that  $q'_{\mathfrak{p}} = \text{id}_{[0,1]}$ .

**Condition 4** Let  $c: X \rightarrow F_{\mathfrak{p}}X$  be an  $F_{\mathfrak{p}}$ -coalgebra. We prove that  $\Phi_{c, r'_{\mathfrak{p}, \gamma}}$  has the unique fixed point.

By Condition 2 proved above and Theorem 2.3.2.2,  $\Phi_{c, r'_{\mathfrak{p}, \gamma}}$  has the least fixed point  $f: X \rightarrow [0, 1]$ . We prove that this is the unique fixed point. Let  $g: X \rightarrow [0, 1]$  be a fixed point of  $\Phi_{c, r'_{\mathfrak{p}, \gamma}}$ . As  $f$  is the least fixed point, we have  $f(x) \leq g(x)$  for each  $x \in X$ . Define  $h: X \rightarrow [0, 1]$  by  $h(x) = g(x) - f(x)$ . Then,

$$\begin{aligned}
\sup_{x \in X} h(x) &= \sup_{x \in X} (g(x) - f(x)) && \text{(by the definition of } h) \\
&= \sup_{x \in X} (\Phi_{c, r'_{\mathfrak{p}, \gamma}}(g)(x) - \Phi_{c, r'_{\mathfrak{p}, \gamma}}(f)(x)) && (f \text{ and } g \text{ are fixed points)} \\
&\leq \sup_{x \in X} \left( \gamma \cdot \int_{x' \in X} g(x') \, d\pi_1(c(x)) - \gamma \cdot \int_{x' \in X} f(x') \, d\pi_1(c(x)) \right) \\
&\text{(by the definition of } r'_{\mathfrak{p}, \gamma} \text{ and that } \pi_2(c(x)) = 1 \Rightarrow \Phi_{c, r}(f)(x) = \Phi_{c, r}(g)(x) = 1) \\
&= \gamma \cdot \sup_{x \in X} \int_{x' \in X} (g(x') - f(x')) \, d\pi_1(c(x)) \\
&= \gamma \cdot \sup_{x \in X} \int_{x' \in X} h(x') \, d\pi_1(c(x)) && \text{(by the definition of } h) \\
&\leq \gamma \cdot \sup_{x \in X} \sup_{x' \in X} h(x') && \text{(by } \int_{x' \in X} d\pi_1(c(x)) = 1) \\
&= \gamma \cdot \sup_{x \in X} h(x).
\end{aligned}$$

As  $0 \leq \gamma < 1$ , we have  $\sup_{x \in X} h(x) = 0$ . Hence we have  $f = g$ .  $\square$

A ranking arrow with respect to  $(r'_{\mathfrak{p}, \gamma}, q'_{\mathfrak{p}}, \leq)$  and its soundness theorem are as follows.

**Definition 5.5.11** ( $\gamma$ -scaled submartingale). Let  $\mathcal{T} = ((X, \mathfrak{F}_X), \xi, \text{Acc})$  be a PTS. Let  $\gamma \in (0, 1)$ . A  $\gamma$ -scaled submartingale for  $\mathcal{T}$  is a measurable function  $b: X \rightarrow [0, 1]$  that satisfies the following:

$$\forall x \in X \setminus \text{Acc}. b(x) \leq \gamma \cdot \int_{x' \in X} b(x') \, d\xi(x). \quad (5.7)$$

**Theorem 5.5.12.** Let  $b: X \rightarrow [0, 1]$  be a  $\gamma$ -scaled submartingale for a PTS  $\mathcal{T} = ((X, \mathfrak{F}_X), \xi, \text{Acc})$ . Then for each  $x \in X$ ,

$$b(x) \leq \text{Reach}_{\mathcal{T}}(x).$$

**Proof.** We define an  $F_{\mathfrak{p}}$ -coalgebra  $c_{\mathcal{T}}: X \rightarrow F_{\mathfrak{p}}X$  as in Example 2.4.14. By the definitions of  $r'_{\mathfrak{p}, \gamma}$  and  $F_{\mathfrak{p}}$ , we can easily see that  $b: X \rightarrow [0, 1]$  is a ranking arrow for  $c_{\mathcal{T}}$  with respect to  $(r'_{\mathfrak{p}, \gamma}, q'_{\mathfrak{p}}, \leq)$ , (i.e.  $b(x) \leq r'_{\mathfrak{p}, \gamma} \circ F_{\mathfrak{p}} b \circ c_{\mathcal{T}}(x)$  for each  $x \in X$ ) if and only if  $b$  is a  $\gamma$ -scaled submartingale for  $\mathcal{T}$ . Moreover, by the definition of  $q'_{\mathfrak{p}}$ , we have  $b(x) = q'_{\mathfrak{p}} \circ b(x)$ .

Hence the theorem is immediate by Theorem 5.3.4.  $\square$

**Example 5.5.13.** Consider the PTS  $\mathcal{T}$  in Example 5.5.9. For each  $\gamma \in [0, 1)$ , we define  $b_\gamma: X \rightarrow [0, 1]$  by  $b_\gamma(x_0) = \frac{\gamma}{3-\gamma}$ ,  $b_\gamma(x_1) = 1$  and  $b_\gamma(x_2) = 0$ . Then  $b_\gamma$  is a  $\gamma$ -scaled submartingale. Hence by Theorem 5.5.12 we have  $\text{Reach}_{\mathcal{T}}(x_0) \geq \frac{\gamma}{3-\gamma}$ .

### Properties of $\gamma$ -scaled Submartingales

In the rest of this section, we present three properties of  $\gamma$ -scaled submartingales. The following proposition shows that the bigger  $\gamma$  we take, the better bound we can obtain.

**Proposition 5.5.14.** *Let  $\mathcal{T} = ((X, \mathfrak{F}_X), \xi, \text{Acc})$  be a PTS. Let  $\gamma_1, \gamma_2 \in [0, 1)$  and assume  $\gamma_1 \leq \gamma_2$ . If  $b_1: X \rightarrow [0, 1]$  is a  $\gamma_1$ -scaled submartingale for  $\mathcal{T}$  then there exists a  $\gamma_2$ -scaled submartingale  $b_2: X \rightarrow [0, 1]$  for  $\mathcal{T}$  such that  $b_1(x) \leq b_2(x)$  for each  $x \in X$ .*

**Proof.** Immediate by the Knaster-Tarski theorem (Corollary 2.3.3.1) and that  $b_1$  is a post fixed point of  $\Phi_{c, r_p, \gamma_2}$ .  $\square$

We move onto the second property of  $\gamma$ -scaled submartingales. We can see in the proof of Proposition 5.5.10 that the inequality  $q'_p \circ r'_{p, \gamma} \leq \sigma_p \circ F_p q'_p$  is strict. Hence we cannot imply completeness using Proposition 5.3.6. Indeed, in Example 5.5.13, the probability bound  $\frac{\gamma}{3-\gamma}$  of  $\text{Reach}_{\mathcal{T}}(x_0)$  given by  $b_\gamma$  is strictly smaller than the true reachability probability  $\frac{1}{2}$ .

However, if we let  $\gamma \rightarrow 1$  then we have  $\frac{\gamma}{3-\gamma} \rightarrow \frac{1}{2}$ . As  $\frac{\gamma}{3-\gamma} \leq \text{Reach}_{\mathcal{T}}(x_0)$  holds for each  $\gamma < 1$ , this implies  $\frac{1}{2} \leq \text{Reach}_{\mathcal{T}}(x_0)$ . The following proposition shows that the completeness in such an asymptotic sense does hold.

**Proposition 5.5.15.** *Let  $\mathcal{T} = ((X, \mathfrak{F}_X), \xi, \text{Acc})$  be a PTS. Let  $(\gamma_i \in [0, 1))_{i \in \omega}$  be an increasing sequence of real numbers that converges to 1, i.e.  $0 \leq \gamma_i < 1$  for each  $i$ ,  $\gamma_0 \leq \gamma_1 \leq \gamma_2 \leq \dots$  and  $\lim_{i \rightarrow \infty} \gamma_i = 1$ . Then there exists a sequence  $(b_i: X \rightarrow [0, 1])_{i \in \omega}$  of measurable functions such that:*

1. for each  $i \in \omega$ ,  $b_i: X \rightarrow [0, 1]$  is a  $\gamma_i$ -scaled submartingale; and
2.  $(b_i)_{i \in \omega}$  is an increasing sequence and converges to  $\text{Reach}_{\mathcal{T}}$ , i.e. for each  $x \in X$ ,  $b_0(x) \leq b_1(x) \leq \dots$  and  $\lim_{i \rightarrow \infty} b_i(x) = \text{Reach}_{\mathcal{T}}(x)$ .

We prove the above proposition as a corollary of the following categorically general result. The theorem shows a setting where such the completeness in an asymptotic sense holds.

**Theorem 5.5.16.** *Let  $F: \mathbb{C} \rightarrow \mathbb{C}$ ,  $(\Omega, \sqsubseteq_\Omega)$  be a truth-value domain and  $\sigma: F\Omega \rightarrow \Omega$  be an  $F$ -modality over  $(\Omega, \sqsubseteq_\Omega)$  that has least fixed points. Let  $r': FR \rightarrow R$  be an  $F$ -algebra,  $q: R \rightarrow \Omega$  be an arrow, and  $\sqsubseteq_R = (\sqsubseteq_{X, R})_{X \in \mathbb{C}}$  be a family of partial orders where  $\sqsubseteq_{X, R}$  is define over  $\mathbb{C}(X, R)$ . We assume the following conditions.*

1. There exists an increasing sequence  $r_0 \sqsubseteq_{FR, R} r_1 \sqsubseteq_{FR, R} \dots \in \mathbb{C}(FR, R)$  of arrows such that  $r' = \bigsqcup_{i \in \omega} r_i$ .
2. For each  $i \in \omega$ , a triple  $(r_i, q, \sqsubseteq_R)$  is a ranking domain.
3. For each  $f: X \rightarrow Y$ , a function  $(\_) \circ f: \mathbb{C}(Y, R) \rightarrow \mathbb{C}(X, R)$  is monotone and  $\omega$ -continuous with respect to  $\sqsubseteq_{Y, R}$  and  $\sqsubseteq_{X, R}$ .
4. For an  $F$ -coalgebra  $c: X \rightarrow FX$ , the function  $\Phi_{c, r'}: \mathbb{C}(X, R) \rightarrow \mathbb{C}(X, R)$  is monotone and  $\omega$ -continuous with respect to  $\sqsubseteq_{X, R}$ .

$$5. q \circ r' = \sigma \circ Fq.$$

Let  $c: X \rightarrow FX$  be an  $F$ -coalgebra. By the definition of a ranking domain, for each  $i \in \omega$ , there exists a unique arrow  $\langle c \rangle_{r_i}: X \rightarrow R$  such that  $\Phi_{c,r_i}(\langle c \rangle_{r_i}) = \langle c \rangle_{r_i}$ . Then we have:

I.  $(\langle c \rangle_{r_i}: X \rightarrow R)_{i \in \omega}$  is an increasing sequence with respect to  $\sqsubseteq_{X,R}$ .

II.  $\bigsqcup_{i \in \omega} (q \circ \langle c \rangle_{r_i}) = \llbracket \mu\sigma \rrbracket_c$ .

**Proof.** We first prove  $\langle c \rangle_{r_i} \sqsubseteq \langle c \rangle_{r_{i+1}}$  for each  $i \in \omega$ . Note that  $\langle c \rangle_{r_i}$  is the least fixed point of  $\Phi_{c,r_i}$ . Hence by the Knaster-Tarski theorem (Corollary 2.3.3.1), it suffices to show that  $\langle c \rangle_{r_{i+1}}$  is a pre-fixed point of  $\Phi_{c,r_i}$ . We have:

$$\begin{aligned} \Phi_{c,r_i}(\langle c \rangle_{r_{i+1}}) &= r_i \circ F\langle c \rangle_{r_{i+1}} \circ c && \text{(by definition)} \\ &\sqsubseteq_{X,R} r_{i+1} \circ F\langle c \rangle_{r_{i+1}} \circ c && \text{(by the assumption)} \\ &= \Phi_{c,r_{i+1}}(\langle c \rangle_{r_{i+1}}) && \text{(by definition)} \\ &= \langle c \rangle_{r_{i+1}} && (r_{i+1} \text{ is a fixed point}). \end{aligned}$$

Hence  $(\langle c \rangle_{r_i}: X \rightarrow R)_{i \in \omega}$  is an increasing sequence.

We next prove  $\bigsqcup_{i \in \omega} (q \circ \langle c \rangle_{r_i}) = \llbracket \mu\sigma \rrbracket_c$ . Note that for each  $i$ ,  $\langle c \rangle_{r_i}: X \rightarrow R$  is a ranking arrow for  $c$  with respect to  $(r_i, q, \sqsubseteq_R)$ . Hence by the soundness of ranking arrows (Theorem 5.3.4), we have  $q \circ \langle c \rangle_{r_i} \sqsubseteq_{X,\Omega} \llbracket \mu\sigma \rrbracket_c$  for each  $i$ . Hence we have  $\bigsqcup_{i \in \omega} (q \circ \langle c \rangle_{r_i}) \sqsubseteq_{\Omega} \llbracket \mu\sigma \rrbracket_c$ .

We shall prove the opposite direction. To this end, as  $\llbracket \mu\sigma \rrbracket_c$  is defined as the least fixed point of  $\Phi_{c,\sigma}$ , it suffices to show that  $\bigsqcup_{i \in \omega} (q \circ \langle c \rangle_{r_i})$  is a fixed point of  $\Phi_{c,\sigma}$ . We have:

$$\begin{aligned} \Phi_{c,\sigma}\left(\bigsqcup_{i \in \omega} (q \circ \langle c \rangle_{r_i})\right) &= \sigma \circ F\left(\bigsqcup_{i \in \omega} (q \circ \langle c \rangle_{r_i})\right) \circ c && \text{(by the definition of } \Phi_{c,\sigma}\text{)} \\ &= \sigma \circ Fq \circ F\left(\bigsqcup_{i \in \omega} \langle c \rangle_{r_i}\right) \circ c && \\ & && \text{(by Condition 3 of Definition 5.3.2)} \\ &= q \circ r' \circ F\left(\bigsqcup_{i \in \omega} \langle c \rangle_{r_i}\right) \circ c && \text{(by Assumption 5)} \\ &= q \circ \bigsqcup_{i \in \omega} \left(r' \circ F\langle c \rangle_{r_i} \circ c\right) && \text{(by Assumption 4)} \\ &= q \circ \bigsqcup_{i \in \omega} \left(\left(\bigsqcup_{j \in \omega} r_j\right) \circ F\langle c \rangle_{r_i} \circ c\right) && \text{(by Assumption 1)} \\ &= q \circ \bigsqcup_{i \in \omega} \bigsqcup_{j \in \omega} \left(r_j \circ F\langle c \rangle_{r_i} \circ c\right) && \text{(by Assumption 3)} \\ &= q \circ \bigsqcup_{i \in \omega} \left(r_i \circ F\langle c \rangle_{r_i} \circ c\right) \\ &= q \circ \bigsqcup_{i \in \omega} \langle c \rangle_{r_i} && (\langle c \rangle_{r_i} \text{ is a fixed point of } \Phi_{c,r_i}) \\ &= \bigsqcup_{i \in \omega} (q \circ \langle c \rangle_{r_i}) && \text{(by Condition 3 of Definition 5.3.2)}. \end{aligned}$$

This concludes the proof.  $\square$



**Proof** (Proposition 5.5.15). We define  $r'_p : \mathcal{G}[0, 1] \times \{0, 1\} \rightarrow [0, 1]$  as follows:

$$r'_p(\varphi, t) = \begin{cases} 1 & (t = 1) \\ \int_{a \in [0, 1]} a \, d\varphi & (\text{otherwise}). \end{cases}$$

Moreover, we define a function  $q'_p : [0, 1] \rightarrow [0, 1]$  and a family  $\leq = (\leq_X)_{X \in \mathbb{C}}$  of partial orders as in Proposition 5.5.10. We prove that Conditions 1–5 of Theorem 5.5.16 are satisfied.

For each  $i \in \omega$  we define  $r_i : F_p[0, 1] \rightarrow [0, 1]$  by  $r_i := r'_{p, 1-1/2^i}$  where the right-hand side is defined as in Proposition 5.5.10. It is easy to see that  $(r_i)_{i \in \omega}$  is an increasing sequence and  $\bigvee_{i \in \omega} r_i = r'_p$  (Condition 1).

By Proposition 5.5.10, for each  $i \in \omega$ ,  $(r_i, q'_p, \leq)$  is a ranking domain (Condition 2).

Let  $f : X \rightarrow Y$  and  $x \in X$ . For  $g, g' : Y \rightarrow [0, 1]$  such that  $g \leq_Y g'$ , we have

$$g \circ f(x) = g(f(x)) \leq g'(f(x)) = g' \circ f(x).$$

For  $g_1, g_2, \dots : Y \rightarrow [0, 1]$  such that  $g_1 \leq_Y g_2 \leq_Y \dots$ , we have

$$\left( \bigvee_{i \in \omega} g_i \right) \circ f(x) = \bigvee_{i \in \omega} (g_i(f(x))) = \bigvee_{i \in \omega} (g_i \circ f(x)) = \bigvee_{i \in \omega} (g_i \circ f)(x).$$

Hence Condition 3 is satisfied.

In a similar manner to the proof of Proposition 5.5.10 (Condition 2 of Definition 5.3.2), we can show that Condition 4 is satisfied.

In a similar manner to the proof of Proposition 5.5.10 (Condition 1 of Definition 5.3.2), we can prove  $q'_p \circ r'_p = \sigma_p \circ Fq'_p$ . Hence Condition 5 holds.

For each  $i \in \omega$ , we let  $b_i := (c_{\mathcal{G}})_{r_i}$ . By definition, each  $b_i$  is a  $(1 - 1/2^i)$ -scaled submartingale. Moreover, by Theorem 5.5.16,  $(b_i : X \rightarrow [0, 1])_{i \in \omega}$  is an increasing sequence, and for each  $x \in X$  we have:

$$\begin{aligned} \text{Reach}_{\mathcal{G}}(x) &= \llbracket \mu \sigma_p \rrbracket_{c_{\mathcal{G}}}(x) && \text{(by Proposition 5.5.1)} \\ &= \bigvee_{i \in \omega} q'_p \circ b_i(x) && \text{(by Theorem 5.5.16)} \\ &= b_i(x) && \text{(by definition).} \end{aligned}$$

This concludes the proof.  $\square$

We conclude this section by showing that we can relax the definition of  $\gamma$ -scaled submartingale. Definition 5.5.11 fixes the range of a  $\gamma$ -scaled submartingale  $b$  to  $[0, 1]$ . The following proposition shows that we can extend it to  $[-\infty, 1]$ . This relaxation is convenient when we synthesize  $\gamma$ -scaled submartingales.

**Proposition 5.5.17.** *Theorem 5.5.12 still holds if we extend the range of  $b$  in Definition 5.5.11 to  $b : X \rightarrow (-\infty, 1]$ .*

**Proof.** Immediate from that if  $b : X \rightarrow (-\infty, 1]$  satisfies the inequality (5.7) in Definition 5.5.11 then  $b' : X \rightarrow [0, 1]$  defined by  $b'(x) = \max\{0, b(x)\}$  also satisfies the inequality and hence is a  $\gamma$ -scaled submartingale.  $\square$

## 5.6 Conclusion and Related Work

We have categorically generalized the notion of ranking function and proved its soundness in the categorical level. In the generalization, a categorical notion of *corecursive algebra* played the central role. We then instantiated the generalization for PTSs, and obtained two ranking function-like notions for them. The induced notions were named *distribution-valued ranking function* (Definition 5.5.6) and  *$\gamma$ -scaled submartingale* (Definition 5.5.11).

**Related Work** For the soundness of ranking function (Theorem 5.1.2), well-foundedness of [3] played an important role. We have used *corecursive algebras* to characterize well-foundedness categorically. In fact, in category theory, a notion called *well-founded coalgebra* exists [106]. It is known that under some weak assumptions, the notion of well-founded coalgebra coincides with that of *recursive coalgebra* [106]. As its name suggests, the notion of recursive coalgebra is dual to that of corecursive algebra. A relationship between *anti-founded algebra* (the dual notion of well-founded coalgebra) and corecursive algebra is studied in [16]. It is remarkable that well-foundedness of coalgebras is commonly used to prove well-foundedness (i.e. termination) of the coalgebra itself. In contrast, in this thesis, we have used corecursive algebras to prove termination of another coalgebra.

We have induced kinds of martingales from our categorical framework. A categorical study of martingales is also found in [72]. There, a relationship between two classical results in the measure theory called Kolmogorov extension theorem and Doob's martingale convergence theorem is investigated.

In Chapter 5 we have categorically modeled a modality as an  $F$ -algebra  $\sigma : F\Omega \rightarrow \Omega$ . Another standard modeling is one by a *predicate lifting*, a natural transformation  $\sigma_X : \Omega^X \Rightarrow \Omega^{FX}$  (see e.g. [95]). These two modelings are related by the Yoneda lemma (see e.g. [45]).

## Chapter 6

# $\gamma$ -Scaled Submartingale for Probabilistic Programs and its Synthesis

In this chapter, we discuss an algorithm for synthesizing  $\gamma$ -scaled submartingales for *probabilistic programs*. A probabilistic program is a variant of a while program augmented with *probabilistic assignments* and *probabilistic branchings*. Probabilistic programs can model not only randomized algorithms but also systems including *physical phenomena* [92].

Our algorithm is adapted from existing ones. It is obtained by modifying template-based synthesis algorithms for *ranking supermartingales* [19]. Ranking supermartingale is a well-known ranking function-like notion for probabilistic systems. We can use it for proving almost-sure termination of a probabilistic system (i.e. that the system terminates in probability 1).

**Definition 6.0.1** (ranking supermartingale [19, 34]). Let  $\mathcal{T} = ((X, \mathfrak{F}_X), \xi, \text{Acc})$  be a PTS. An (*additive*) *ranking supermartingale* for  $\mathcal{T}$  is a measurable function  $b: X \rightarrow [0, \infty]$  that satisfies the following condition:

$$\forall x \in X \setminus \text{Acc}. b(x) \geq \int_{x' \in X} b(x') d\xi(x) + 1. \quad (6.1)$$

**Theorem 6.0.2** (soundness, [19, 34]). *Let  $b: X \rightarrow [0, \infty]$  be a ranking supermartingale for a PTS  $\mathcal{T} = ((X, \mathfrak{F}_X), \xi, \text{Acc})$ . Then for each  $x \in X$ ,*

$$b(x) < \infty \Rightarrow \text{Reach}_{\mathcal{T}}(x) = 1. \quad \square$$

The definitions of ranking supermartingale and  $\gamma$ -scaled submartingale are very similar. We found that an existing synthesis algorithm for ranking supermartingales can be easily adapted for  $\gamma$ -scaled submartingales.

In this chapter, we first give a *linear template-based algorithm* based on an algorithm in [19, 23]. It fixes a linear template for a  $\gamma$ -scaled submartingale and searches for a valuation of parameters (unknown coefficients) that makes the template a ranking supermartingale using a linear programming (LP) solver. We have implemented the algorithm and tested it for several probabilistic programs. We have also compared it with another algorithm in [23] that is for the same purpose, i.e. underapproximating termination probability. The algorithm in [23] is similar to ours: it synthesizes a *repulsing supermartingale*, yet another ranking function-like notion for probabilistic systems, using a linear template. We will compare the lower bounds of the termination probability calculated by our and their algorithms.

For ranking supermartingales over probabilistic programs, a *polynomial template-based synthesis algorithm* is also known [21]. It fixes a polynomial template for a ranking supermartingale and searches for a valuation of parameters

using a semidefinite programming (SDP) solver. Similarly to the linear template-based algorithm, it can be easily adapted for  $\gamma$ -scaled submartingales.

We also implemented the polynomial template-based algorithm and tested it for probabilistic programs. However, it turned out that our implementation does not work well because of numerical errors that seem to be caused by the SDP solver. We nevertheless present our polynomial template-based algorithm, (failed) efforts to remedy the problems caused by numerical errors, and the experimental results for the record.

This chapter consists of two sections. In Section 6.1, we present our linear template-based algorithm, explain its implementation, and give the experimental results. In Section 6.2 we do the same for a polynomial template-based algorithm.

This chapter extends a part of [102].

## 6.1 Linear Template-Based Algorithm

In this section, we consider synthesizing a  $\gamma$ -scaled submartingale for probabilistic programs using a linear template.

### 6.1.1 Syntax of Probabilistic Programs

The syntax of probabilistic programs we use in this thesis mainly follows the one in [23] except that we do not include nondeterministic assignments and branchings so that the theory developed in the previous chapter is applicable.<sup>1</sup>

**Definition 6.1.1** (linear probabilistic program). Let  $\mathcal{V}$  be a countably infinite set of variables. A *linear probabilistic program* (LPP) is a program  $\langle prog \rangle$  defined by the following BNF notation:

$$\begin{aligned}
\langle prog \rangle &::= \langle stmt \rangle \\
\langle stmt \rangle &::= \langle stmt \rangle; \langle stmt \rangle \mid \langle assgn \rangle \mid \mathbf{skip} \mid \mathbf{while} \langle pbepr \rangle \mathbf{do} \langle stmt \rangle \mathbf{od} \\
&\quad \mid \mathbf{if} \langle pbepr \rangle \mathbf{then} \langle stmt \rangle \mathbf{else} \langle stmt \rangle \mathbf{fi} \tag{6.2} \\
\langle assgn \rangle &::= \langle pvar \rangle := \langle expr \rangle \mid \langle pvar \rangle := \mathbf{sample}(\langle dist \rangle) \\
\langle expr \rangle &::= \langle const \rangle \mid \langle pvar \rangle \mid \langle const \rangle \cdot \langle pvar \rangle \\
&\quad \mid \langle expr \rangle + \langle expr \rangle \mid \langle expr \rangle - \langle expr \rangle \tag{6.3} \\
\langle pbepr \rangle &::= \mathbf{prob}(p) \mid \langle bepr \rangle \quad (\text{where } p \in [0, 1]) \\
\langle bepr \rangle &::= \langle conjexpr \rangle \mid \langle conjexpr \rangle \mathbf{or} \langle bepr \rangle \\
\langle conjexpr \rangle &::= \langle literal \rangle \mid \langle literal \rangle \mathbf{and} \langle conjexpr \rangle \\
\langle literal \rangle &::= \langle expr \rangle \leq \langle expr \rangle \mid \langle expr \rangle < \langle expr \rangle \\
\langle pvar \rangle &::= v \in \mathcal{V} \quad \langle dist \rangle ::= d \in \mathcal{GR} \quad \langle const \rangle ::= c \in \mathbb{R}.
\end{aligned}$$

We further assume the following:

- (†) for each probability measure  $d$  appearing in  $\langle prog \rangle$ , an algorithm that calculates the expectation of  $d$  is given.

We write  $\{\{ stmt \}\}_{\text{lin}}$ ,  $\{\{ assgn \}\}_{\text{lin}}$ ,  $\{\{ expr \}\}_{\text{lin}}$ , etc... for the sets of formulas defined by the BNF notation above (i.e.  $\{\{ pvar \}\}_{\text{lin}} = \{v \in \mathcal{V}\}$  for example). We call an element in  $\{\{ expr \}\}_{\text{lin}}$  a *linear expression*.

<sup>1</sup>In fact, it was proved later by another author that a  $\gamma$ -scaled submartingale can be defined for probabilistic systems with nondeterminism and its soundness theorem is provable [102]. Moreover, the original algorithms in [19, 23, 21] from which our algorithm is adapted can deal with probabilistic programs with demonic nondeterminism. Hence all the discussions in this chapter are also applicable for probabilistic programs with nondeterminism.

For  $\langle expr \rangle \in \{\{expr\}\}_{\text{lin}}$ , its semantics  $\llbracket \langle expr \rangle \rrbracket : \mathbb{R}^{\mathcal{V}} \rightarrow \mathbb{R}$  is defined in the standard manner (we omit the definition). Similarly, the semantics  $\llbracket b \rrbracket \subseteq \mathbb{R}^{\mathcal{V}}$  of  $b \in \{\{bexpr\}\}_{\text{lin}} \cup \{\{conjexpr\}\}_{\text{lin}} \cup \{\{literal\}\}_{\text{lin}}$  is ordinarily defined.

In the rest of this chapter, we may identify semantically equivalent expressions like  $x_1 \cdot x_2$  and  $x_2 \cdot x_1$ , or  $x_1 + x_2$  and  $x_2 + x_1$ . No confusion is likely.

### 6.1.2 Problem

We formalize the problem. The problem is stated in terms of *probabilistic control flow graph*, a syntactic object that is induced from an LPP.

**Definition 6.1.2** ([5]). A *linear probabilistic control flow graph* (linear pCFG) is a tuple  $\Gamma = (L_A, L_B, l_{\text{init}}, \tau_A, \tau_B)$  consisting of the following components:

- finite sets  $L_A$  and  $L_B$  of *assignment locations* and *branching locations*;
- an *initial location*  $l_{\text{init}} \in L_A \cup L_B$ ; and
- *transition functions*  $\tau_A : L_A \rightarrow \{\{assgn\}\}_{\text{lin}} \times L$  and  $\tau_B : L_B \rightarrow \{\{pbexpr\}\}_{\text{lin}} \times L \times L$ .

We write  $V_\Gamma$  for the finite set of variables in  $\mathcal{V}$  appearing in  $\Gamma$ . We call  $l \in L_A \cup L_B$  a *location*,  $\lambda \in \mathbb{R}^{V_\Gamma}$  a *valuation*, and a pair  $(l, \lambda) \in (L_A \cup L_B) \times \mathbb{R}^{V_\Gamma}$  a *configuration*.

An LPP induces a pCFG as follows.

**Definition 6.1.3.** For an LPP  $\langle prog \rangle$  a countably infinite set  $\mathcal{L}$  and a symbol  $l_{\text{term}} \notin \mathcal{L}$ , we inductively define a linear pCFG  $\Gamma_{\langle prog \rangle}^{\mathcal{L}, l_{\text{term}}} = (L_A, L_B, l_{\text{init}}, \tau_A, \tau_B)$  as follows.

- Assume  $\langle prog \rangle = \langle stmt \rangle^1; \langle stmt \rangle^2$ . Fix countably infinite sets  $\mathcal{L}_1, \mathcal{L}_2 \subseteq \mathcal{L}$  so that  $\mathcal{L}_1 \cap \mathcal{L}_2 = \emptyset$ . Assume  $\Gamma_{\langle stmt \rangle^2}^{\mathcal{L}_2, l_{\text{term}}} = (L_A^2, L_B^2, l_{\text{init}}^2, \tau_A^2, \tau_B^2)$  and  $\Gamma_{\langle stmt \rangle^1}^{\mathcal{L}_1, l_{\text{init}}^2} = (L_A^1, L_B^1, l_{\text{init}}^1, \tau_A^1, \tau_B^1)$ . Then we let  $L_A := L_A^1 \cup L_A^2$ ,  $L_B := L_B^1 \cup L_B^2$ ,  $l_{\text{init}} := l_{\text{init}}^1$ ,  $\tau_A(l) := \tau_A^1(l)$  if  $l \in L_A^1$  and  $\tau_A^2(l)$  if  $l \in L_A^2$ , and  $\tau_B(l) := \tau_B^1(l)$  if  $l \in L_B^1$  and  $\tau_B^2(l)$  if  $l \in L_B^2$ .
- Assume  $\langle prog \rangle = \langle assgn \rangle$ . Then we choose  $l_{\text{new}} \in \mathcal{L}$  and let  $L_A := \{l_{\text{new}}, l_{\text{term}}\}$ ,  $L_B := \emptyset$ ,  $l_{\text{init}} := l_{\text{new}}$ ,  $\tau_A(l_{\text{new}}) = (\langle assgn \rangle, l_{\text{term}})$  and  $\tau_B$  be the empty function.
- If  $\langle prog \rangle = \mathbf{skip}$  then we choose  $v \in \mathcal{V}$  and let  $\Gamma_{\langle stmt \rangle} := \Gamma_{v:=v}$ .
- Assume  $\langle prog \rangle = \mathbf{while} \langle pbexpr \rangle \mathbf{do} \langle stmt \rangle' \mathbf{od}$ . We choose  $l_{\text{new}} \in \mathcal{L}$  and assume  $\Gamma_{\langle stmt \rangle'}^{\mathcal{L} \setminus \{l_{\text{new}}\}, l_{\text{new}}} = (L'_A, L'_B, l'_{\text{init}}, \tau'_A, \tau'_B)$ . Then let  $L_A := L'_A$ ,  $L_B := L'_B$ ,  $l_{\text{init}} := l_{\text{new}}$ ,  $\tau_A(l) := \tau'_A(l)$  and  $\tau_B(l) := (\langle pbexpr \rangle, l'_{\text{init}}, l_{\text{term}})$  if  $l = l_{\text{new}}$  and  $\tau_B(l) := \tau'_B(l)$  otherwise.
- Assume  $\langle prog \rangle = \mathbf{if} \langle pbexpr \rangle \mathbf{then} \langle stmt \rangle^1 \mathbf{else} \langle stmt \rangle^2 \mathbf{fi}$ . Fix countably infinite sets  $\mathcal{L}_1, \mathcal{L}_2 \subseteq \mathcal{L}$  and  $l_{\text{new}} \in \mathcal{L}$  so that  $\mathcal{L}_1 \cap \mathcal{L}_2 = \emptyset$  and  $l_{\text{new}} \notin \mathcal{L}_1 \cup \mathcal{L}_2$ . Assume  $\Gamma_{\langle stmt \rangle^i}^{\mathcal{L}_i, l_{\text{term}}} = (L_A^i, L_B^i, l_{\text{init}}^i, \tau_A^i, \tau_B^i)$  for  $i \in \{1, 2\}$ . We then let  $L_A := L_A^1 \cup L_A^2$ ,  $L_B := \{l_{\text{new}}\} \cup L_B^1 \cup L_B^2$ ,  $l_{\text{init}} := l_{\text{new}}$ ,  $\tau_A(l) := \tau_A^i(l)$  when  $l \in L_A^i$  and  $\tau_B(l) := (\langle pbexpr \rangle, l_{\text{init}}^1, l_{\text{init}}^2)$  if  $l = l_{\text{new}}$  and  $\tau_B(l) := \tau_B^i(l)$  if  $l \in L_B^i$  (here  $i \in \{1, 2\}$ ).

We next introduce notions of *invariant* and *terminal configuration*. An invariant specifies ranges of variables while a terminal configuration specifies accepting states. Both of them are defined as *predicate maps*, functions that assign Boolean expressions to each location of a pCFG.

**Definition 6.1.4** (linear predicate map). Let  $\Gamma = (L_A, L_B, l_{\text{init}}, \tau_A, \tau_B)$  be a pCFG. A *linear predicate map* is a function  $\mathfrak{p} : L_A + L_B \rightarrow \{\{bexpr\}\}_{\text{lin}}$ .

We first explain invariants. For example, in the LPP on the right, we always have  $x > 0$  at line 3. Hence a linear predicate map assigning “ $x > 0$ ” to the corresponding location is an invariant.

An invariant allows us to find a better  $\gamma$ -scaled submartingale that gives tighter bound for the reachability probability. In this thesis, we assume that a correct invariant is provided by the user and do not discuss how to obtain a correct invariant. A synthesis algorithm for invariants is found in e.g. [69].

```

1  x := 5;
2  while x > 0 do
3    if prob(0.5) then
4      x := x + 1
5    else
6      x := x - 1
7    fi
8  od

```

**Definition 6.1.5** (invariant). Let  $\Gamma = (L_A, L_B, l_{\text{init}}, \tau_A, \tau_B)$  be a linear pCFG. A linear predicate map  $\mathfrak{J} : L_A + L_B \rightarrow \{\{expr\}\}_{\text{lin}}$  is called a *linear invariant* if for each  $l \in L_A \cup L_B$  and  $\lambda \in \llbracket \mathfrak{J}(l) \rrbracket$ , the following conditions are satisfied.

- If  $l \in L_A$  and  $\tau_A(l) = (\langle assign \rangle, l')$  then
  - if  $\langle assign \rangle = (v := \langle expr \rangle)$  then  $\lambda[v \mapsto \llbracket \langle expr \rangle \rrbracket(\lambda)] \in \llbracket \mathfrak{J}(l') \rrbracket$ , and
  - if  $\langle assign \rangle = (v := \mathbf{sample}(\langle dist \rangle))$  where  $\langle \mathcal{D} \rangle = d$  then  $\lambda[v \mapsto c] \in \llbracket \mathfrak{J}(l') \rrbracket$  for each  $c \in \text{supp}(d) := \{r \in \mathbb{R} \mid \forall O: \text{open} \subseteq \mathbb{R}. r \in O \Rightarrow d(O) > 0\}$ .
- If  $l \in L_B$  and  $\tau_B(l) = (\langle pbexpr \rangle, l_1, l_2)$  then
  - if  $\langle pbexpr \rangle = \mathbf{prob}(p)$  then  $\lambda \in \llbracket \mathfrak{J}(l_i) \rrbracket$  for each  $i \in \{1, 2\}$ ,
  - if  $\langle pbexpr \rangle = \langle bexpr \rangle$  and  $\lambda \in \llbracket \langle bexpr \rangle \rrbracket$  then  $\lambda \in \llbracket \mathfrak{J}(l_1) \rrbracket$ , and
  - if  $\langle pbexpr \rangle = \langle bexpr \rangle$  and  $\lambda \notin \llbracket \langle bexpr \rangle \rrbracket$  then  $\lambda \in \llbracket \mathfrak{J}(l_2) \rrbracket$ .

Finally, a terminal configuration specifies accepting states. It is simply defined as a predicate map. A pCFG, an invariant and a terminal configuration together induce a PTS as follows.

**Definition 6.1.6.** Let  $\Gamma = (L_A, L_B, l_{\text{init}}, \tau_A, \tau_B)$  be a linear pCFG,  $\mathfrak{J} : L_A + L_B \rightarrow \{\{expr\}\}_{\text{lin}}$  be a linear invariant, and  $\mathfrak{T} : L_A + L_B \rightarrow \{\{expr\}\}_{\text{lin}}$  be a linear predicate map. Let  $L = L_A + L_B$ . We define a PTS  $\mathcal{T}_{\Gamma, \mathfrak{J}, \mathfrak{T}} = ((X, \mathfrak{F}_X), \xi, \text{Acc})$  as follows (recall that  $\delta_x$  denotes the Dirac measure):

- $(X, \mathfrak{F}_X) := \coprod_{l \in L} (\llbracket \mathfrak{J}(l) \rrbracket, \mathfrak{F}_l)$  where  $\mathfrak{F}_l = \{A \cap \llbracket \mathfrak{J}(l) \rrbracket \mid A \subseteq \mathbb{R}^{V_\Gamma}, A \text{ is measurable}\}$ .
- Let  $l \in L_A + L_B$  and  $\lambda \in \llbracket \mathfrak{J}(l) \rrbracket$ .
  - If  $l \in L_A$  and  $\tau_A(l) = (\langle assign \rangle, l')$  then
    - \* if  $\langle assign \rangle = (v := \langle expr \rangle)$  then  $\xi(l, \lambda) := \delta_{(v, \lambda[v \mapsto \llbracket \langle expr \rangle \rrbracket(\lambda)])}$ , and
    - \* if  $\langle assign \rangle = (v := \mathbf{sample}(\langle dist \rangle))$  then  $\xi(l, \lambda)(\{l''\} \times A) := \langle dist \rangle(\{\lambda'(v) \mid \lambda' \in A, \forall v' \in V_\Gamma \setminus \{v\}. \lambda'(v') = \lambda(v')\})$  if  $l'' = l'$  and 0 otherwise.

- If  $l \in L_B$  and  $\tau_B(l) = (\langle p\text{bexpr} \rangle, l_1, l_2)$  then
  - \* if  $\langle p\text{bexpr} \rangle = \mathbf{prob}(p)$  then  $\xi(l, \lambda)(\{l''\} \times A) := p$  if  $l'' = l_1$  and  $\lambda \in A$ ;  $1 - p$  if  $l'' = l_2$  and  $\lambda \in A$ ; and 0 otherwise.
  - \* if  $\langle p\text{bexpr} \rangle = \langle \text{bexpr} \rangle$  and  $\lambda \in \llbracket \langle \text{bexpr} \rangle \rrbracket$  then  $\xi(l, \lambda) := \delta_{(l_1, \lambda)}$ , and
  - \* if  $\langle p\text{bexpr} \rangle = \langle \text{bexpr} \rangle$  and  $\lambda \notin \llbracket \langle \text{bexpr} \rangle \rrbracket$  then  $\xi(l, \lambda) := \delta_{(l_2, \lambda)}$ .

The probability measure  $\xi(l, \lambda)$  is well-defined because  $\mathfrak{I}$  is an invariant.

- $\text{Acc} := \{(l, \lambda) \mid \lambda \in \llbracket \mathfrak{I}(l) \rrbracket\}$ .

We can now formalize the problem that we tackle in this section.

**Problem 6.1.7.**     input:    a linear pCFG  $\Gamma = (L_A, L_B, l_{\text{init}}, \tau_A, \tau_B)$   
                           an initial valuation  $\lambda_{\text{init}} : V_\Gamma \rightarrow \mathbb{R}$   
                           a linear invariant  $\mathfrak{I}$  for a pCFG  $\Gamma$   
                           a linear predicate map  $\mathfrak{F}$  for a pCFG  $\Gamma$   
                           output:     $p \in \mathbb{R}$  such that  $p \leq \text{Reach}_{\mathcal{F}_\Gamma, \mathfrak{I}, \mathfrak{F}}(l_{\text{init}}, \lambda_{\text{init}})$

### 6.1.3 Algorithm

Our algorithm is almost the same as the one for ranking supermartingales in [19], although Lemma 6.1.12 given later seems to be new.

Throughout this section, let  $\Gamma = (L_A, L_B, l_{\text{init}}, \tau_A, \tau_B)$  be a linear pCFG,  $L = L_A + L_B$  and  $\{x_1, \dots, x_n\} = V_\Gamma$ . Let  $\mathfrak{I} : L \rightarrow \{\{\text{expr}\}\}_{\text{lin}}$  be a linear invariant and  $\mathfrak{F} : L \rightarrow \{\{\text{expr}\}\}_{\text{lin}}$  be a linear predicate map, and assume that an initial valuation  $\lambda_{\text{init}} \in \llbracket \mathfrak{I}(l_{\text{init}}) \rrbracket (\subseteq \mathbb{R}^{V_\Gamma})$  is given. We fix  $\gamma \in [0, 1)$ . The algorithm consists of four steps.

#### Fix a Template

The algorithm first fixes a linear template for a  $\gamma$ -scaled submartingale.

**Definition 6.1.8** (linear expression map and linear template).

- A *linear expression map* for  $\Gamma$  is a function  $\mathfrak{f} : L \rightarrow \{\{\text{expr}\}\}_{\text{lin}}$ . We define  $\llbracket \mathfrak{f} \rrbracket : L \times \mathbb{R}^{V_\Gamma} \rightarrow \mathbb{R}$  by  $\llbracket \mathfrak{f} \rrbracket(l, \lambda) := \llbracket \mathfrak{f}(l) \rrbracket(\lambda)$ .
- A *linear template* over  $\Gamma$  is a family  $\mathfrak{t} = (\mathfrak{t}(l))_{l \in L}$  of formulas of a form

$$\mathfrak{t}(l) = a_1^l x_1 + \dots + a_n^l x_n + b^l.$$

For each  $l \in L$ ,  $a_1^l, \dots, a_n^l, b^l$  are new variables. We call each of them a *parameter* and write  $P_\Gamma$  for the set of parameters. For a valuation  $\chi : P_\Gamma \rightarrow \mathbb{R}$  of parameters, we write  $\mathfrak{t}_\chi$  for a linear expression map  $l \mapsto \chi(a_1^l)x_1 + \dots + \chi(a_n^l)x_n + \chi(b^l)$ .

Suppose that we fix a linear template for  $\Gamma$ . Recall that the state space of a PTS  $\mathcal{F}_\Gamma, \mathfrak{I}, \mathfrak{F}$  is given by  $\coprod_{l \in L} \llbracket \mathfrak{I}(l) \rrbracket \subseteq L \times \mathbb{R}^{V_\Gamma}$ . Our goal is to synthesize a valuation  $\chi : P_\Gamma \rightarrow \mathbb{R}$  of parameters so that:

- $\llbracket \mathfrak{t}_\chi \rrbracket|_{\coprod_{l \in L} \llbracket \mathfrak{I}(l) \rrbracket} : \coprod_{l \in L} \llbracket \mathfrak{I}(l) \rrbracket \rightarrow \mathbb{R}$  is a  $\gamma$ -scaled submartingale over the PTS  $\mathcal{F}_\Gamma, \mathfrak{I}, \mathfrak{F}$  (in this case we say that  $\mathfrak{t}_\chi$  is a  $\gamma$ -scaled submartingale); and
- maximize the probability bound  $\llbracket \mathfrak{t}_\chi \rrbracket(l_{\text{init}}, \lambda_{\text{init}}) \in \mathbb{R}$  given by the  $\gamma$ -scaled submartingale.

## Collect Conditions for Parameters

We next turn the axioms of  $\gamma$ -scaled submartingales to conditions on the parameters. For example, suppose that there exists a state  $l \in L_B$  such that

- $\mathfrak{I}(l) = (\langle \text{conjexpr} \rangle_1 \text{ or } \cdots \text{ or } \langle \text{conjexpr} \rangle_k)$ ;
- $\overline{\mathfrak{I}(l)} = (\langle \text{conjexpr} \rangle_{k+1} \text{ or } \cdots \text{ or } \langle \text{conjexpr} \rangle_{k+k'})$ ; and
- $\tau_B(l) = (\text{prob}(p), l_1, l_2)$ .

Here  $\overline{\mathfrak{I}(l)} \in \{\{\langle \text{bexpr} \rangle\}_{\text{lin}}\}$  is a “negation” of  $\mathfrak{I}(l)$  in the following sense:

**Lemma 6.1.9.** *For  $\langle \text{bexpr} \rangle \in \{\{\langle \text{bexpr} \rangle\}_{\text{lin}}\}$ , there exists  $\overline{\langle \text{bexpr} \rangle} \in \{\{\langle \text{bexpr} \rangle\}_{\text{lin}}\}$  such that  $\llbracket \langle \text{bexpr} \rangle \rrbracket = \mathbb{R}^{V_\Gamma} \setminus \llbracket \overline{\langle \text{bexpr} \rangle} \rrbracket$ .  $\square$*

In general, the formula  $\overline{\langle \text{bexpr} \rangle}$  can be exponentially larger than  $\langle \text{bexpr} \rangle$ .

Then the inequality (5.7) in Definition 5.5.11 for the location  $l$  boils down to the following formula.

$$\forall \lambda \in \mathbb{R}^{V_\Gamma}. \forall i \in \{1, \dots, k, k+1, \dots, k+k'\}.$$

$$\lambda \in \llbracket \langle \text{conjexpr} \rangle_i(l) \rrbracket \Rightarrow \llbracket \mathfrak{t}_\chi(l) \rrbracket(\lambda) \leq \gamma \cdot (p \cdot \llbracket \mathfrak{t}_\chi(l_1) \rrbracket(\lambda) + (1-p) \cdot \llbracket \mathfrak{t}_\chi(l_2) \rrbracket(\lambda))$$

Note that the premise is representable as a conjunction of linear inequalities over  $V_\Gamma$  without parameters, and the consequence is representable as a linear inequality over  $V_\Gamma$  whose coefficients are linear expressions over  $P_\Gamma$ . We will later observe that this is the case for all the cases. Therefore if we collect all the conditions for the parameters so that  $\mathfrak{t}_\chi$  is a  $\gamma$ -scaled submartingale, then we obtain a formula that is a conjunction of formulas of the following form:

$$(\mathfrak{d}_1 \triangleright_1 0 \wedge \cdots \wedge \mathfrak{d}_k \triangleright_k 0) \Rightarrow \mathfrak{e} \geq 0 \quad (6.4)$$

Here  $\triangleright_i \in \{\geq, >\}$ ,  $\mathfrak{d}_1, \dots, \mathfrak{d}_k$  are linear expressions over  $V_\Gamma$  without parameters, and  $\mathfrak{e}$  is a formula of the following form:

$$\mathfrak{e} = p_1 x_1 + \cdots + p_n x_n + q \quad (6.5)$$

where each coefficient  $p_i$  is a linear expression over  $P_\Gamma$ . For a valuation  $\chi : P_\Gamma \rightarrow \mathbb{R}$  of the parameters, we write  $\chi(\mathfrak{e})$  for a linear expression  $\llbracket p_1 \rrbracket(\chi)x_1 + \cdots + \llbracket p_n \rrbracket(\chi)x_n + \llbracket q \rrbracket(\chi)$ .

We write down all the concrete constructions of such formulas for a record.

**Definition 6.1.10.** Let  $\Gamma = (L_A, L_B, l_{\text{init}}, \tau_A, \tau_B)$  be a linear pCFG,  $\mathfrak{I}$  and  $\mathfrak{E}$  be linear predicate maps, and  $\gamma \in [0, 1)$ . Let  $L = L_A + L_B$  and  $\{x_1, \dots, x_n\} = V_\Gamma$ . We write  $\text{Fml}^{\Rightarrow}$  for the set of formulas of a form as in (6.4). Without loss of generality, for each  $l \in L$ , let

$$\mathfrak{I}(l) = \left( (\alpha_{1,1}^l \triangleright_{1,1}^l 0 \text{ and } \cdots \text{ and } \alpha_{1,M_1^l}^l \triangleright_{1,M_1^l}^l 0) \text{ or } \cdots \right.$$

$$\left. \text{or } (\alpha_{N^l,1}^l \triangleright_{N^l,1}^l 0 \text{ and } \cdots \text{ and } \alpha_{N^l,M_{N^l}^l}^l \triangleright_{N^l,M_{N^l}^l}^l 0) \right); \text{ and}$$

$$\overline{\mathfrak{I}(l)} = \left( (\beta_{1,1}^l \triangleright_{1,1}^l 0 \text{ and } \cdots \text{ and } \beta_{1,M_1^l}^l \triangleright_{1,M_1^l}^l 0) \text{ or } \cdots \right.$$

$$\left. \text{or } (\beta_{N^l,1}^l \triangleright_{N^l,1}^l 0 \text{ and } \cdots \text{ and } \beta_{N^l,M_{N^l}^l}^l \triangleright_{N^l,M_{N^l}^l}^l 0) \right).$$

where  $\triangleright_{i,j}^l, \triangleright_{i,j}^l \in \{\geq, >\}$  and  $\alpha_{j,k}^l$  and  $\beta_i^l$  are linear expressions for each  $l$ .



For each  $l \in L$ , we define  $A_1^l, A_2^l \subseteq \text{Fml}^{\Rightarrow}$  as follows. Firstly,  $A_1^l$  is defined by:

$$A_1^l := \left\{ \bigwedge_{j=1}^{M_i^l} (\alpha_{i,j}^l \triangleright_{i,j}^l 0) \Rightarrow ((-a_1^l)x_1 + \cdots + (-a_n^l)x_n + (1 - b^l) \geq 0) \mid 1 \leq i \leq N^l \right\}.$$

Moreover,  $A_2^l$  is defined as follows.

- If  $l \in L_A$ ,  $\tau_A(l) = (\langle \text{assgn} \rangle, l')$ ,  $\langle \text{assgn} \rangle = (x_k := \langle \text{expr} \rangle)$  and  $\langle \text{expr} \rangle = r_1x_1 + \cdots + r_nx_n + r$  where  $r_1, \dots, r_n, r \in \mathbb{R}$ , then

$$A_2^l := \left\{ \begin{array}{l} \bigwedge_{j=1}^{M_i^l} (\alpha_{i,j}^l \triangleright_{i,j}^l 0) \wedge \bigwedge_{j=1}^{M_{i'}^l} (\beta_{i',j}^l \triangleright_{i',j}^l 0) \Rightarrow \\ \left( \begin{array}{l} (\gamma(a_1^{l'} + r_1a_k^{l'}) - a_1^l)x_1 + \cdots \\ + (\gamma(a_{k-1}^{l'} + r_{k-1}a_k^{l'}) - a_{k-1}^l)x_{k-1} \\ + (\gamma r_k a_k^{l'} - a_k^l)x_k + (\gamma(a_{k+1}^{l'} + r_{k+1}a_k^{l'}) - a_{k+1}^l)x_{k+1} \\ + \cdots + (\gamma(a_n^{l'} + r_n a_k^{l'}) - a_n^l)x_n + (\gamma(b^{l'} + r a_k^{l'}) - b^l) \end{array} \right) \geq 0 \end{array} \right\} \begin{array}{l} 1 \leq i \leq N^l, \\ 1 \leq i' \leq N^{l'} \end{array}.$$

- If  $l \in L_A$ ,  $\tau_A(l) = (\langle \text{assgn} \rangle, l')$ ,  $\langle \text{assgn} \rangle = (x_k := \mathbf{sample}(\langle \text{dist} \rangle))$  and the expectation of  $\langle \text{dist} \rangle$  is  $r \in \mathbb{R}$ , then

$$A_2^l := \left\{ \begin{array}{l} \bigwedge_{j=1}^{M_i^l} (\alpha_{i,j}^l \triangleright_{i,j}^l 0) \wedge \bigwedge_{j=1}^{M_{i'}^l} (\beta_{i',j}^l \triangleright_{i',j}^l 0) \Rightarrow \\ \left( \begin{array}{l} (\gamma a_1^{l'} - a_1^l)x_1 + \cdots + (\gamma a_{k-1}^{l'} - a_{k-1}^l)x_{k-1} - a_k^l x_k \\ + (\gamma a_{k+1}^{l'} - a_{k+1}^l)x_{k+1} + \cdots + (\gamma a_n^{l'} - a_n^l)x_n \\ + (\gamma(b^{l'} + r a_k^{l'}) + (\gamma r - b^l)) \end{array} \right) \geq 0 \end{array} \right\} \begin{array}{l} 1 \leq i \leq N^l, \\ 1 \leq i' \leq N^{l'} \end{array}.$$

- If  $l \in L_B$ ,  $\tau_B(l) = (\langle \text{pbexpr} \rangle, l_1, l_2)$  and  $\langle \text{pbexpr} \rangle = \mathbf{prob}(p)$ , then

$$A_2^l := \left\{ \begin{array}{l} \bigwedge_{j=1}^{M_i^l} (\alpha_{i,j}^l \triangleright_{i,j}^l 0) \wedge \bigwedge_{j=1}^{M_{i'}^l} (\beta_{i',j}^l \triangleright_{i',j}^l 0) \Rightarrow \\ \left( \begin{array}{l} (\gamma(p a_1^{l_1} + (1-p)a_1^{l_2}) - a_1^l)x_1 + \cdots \\ + (\gamma(p a_n^{l_1} + (1-p)a_n^{l_2}) - a_n^l)x_n \\ + (\gamma(p b^{l_1} + (1-p)b^{l_2}) - b^l) \end{array} \right) \geq 0 \end{array} \right\} \begin{array}{l} 1 \leq i \leq N^l, \\ 1 \leq i' \leq N^{l'} \end{array}.$$

- If  $l \in L_B$ ,  $\tau_B(l) = (\langle \text{pbexpr} \rangle, l_1, l_2)$ ,  $\langle \text{pbexpr} \rangle = \langle \text{bexpr} \rangle$ ,  $\llbracket \langle \text{bexpr} \rangle \rrbracket = \bigvee_{s=1}^S \bigwedge_{t=1}^{T_s} (\varepsilon_{s,t} \triangleright_{s,t} 0)$  and  $\llbracket \overline{\langle \text{bexpr} \rangle} \rrbracket = \bigvee_{s=1}^{S'} \bigwedge_{t=1}^{T'_s} (\varepsilon'_{s,t} \triangleright'_{s,t} 0)$  where  $\varepsilon_{s,t}, \varepsilon'_{s,t}$  are linear expressions over  $V_{\Gamma}$ , then

$$A_2^l := \left\{ \begin{array}{l} \bigwedge_{j=1}^{M_i^l} (\alpha_{i,j}^l \triangleright_{i,j}^l 0) \wedge \bigwedge_{j=1}^{M_{i'}^l} (\beta_{i',j}^l \triangleright_{i',j}^l 0) \wedge \bigwedge_{t=1}^{T_s} (\varepsilon_{s,t} \triangleright_{s,t} 0) \Rightarrow \\ \left( \begin{array}{l} ((\gamma a_1^{l_1} - a_1^l)x_1 + \cdots + (\gamma a_n^{l_1} - a_n^l)x_n + (\gamma b^{l_1} - b^l)) \geq 0 \end{array} \right) \end{array} \right\} \begin{array}{l} 1 \leq i \leq N^l, \\ 1 \leq i' \leq N^{l'}, \\ 1 \leq s \leq S \end{array} \\ \cup \left\{ \begin{array}{l} \bigwedge_{j=1}^{M_i^l} (\alpha_{i,j}^l \triangleright_{i,j}^l 0) \wedge \bigwedge_{j=1}^{M_{i'}^l} (\beta_{i',j}^l \triangleright_{i',j}^l 0) \wedge \bigwedge_{t=1}^{T'_s} (\varepsilon'_{s,t} \triangleright'_{s,t} 0) \Rightarrow \\ \left( \begin{array}{l} ((\gamma a_1^{l_2} - a_1^l)x_1 + \cdots + (\gamma a_n^{l_2} - a_n^l)x_n + (\gamma b^{l_2} - b^l)) \geq 0 \end{array} \right) \end{array} \right\} \begin{array}{l} 1 \leq i \leq N^l, \\ 1 \leq i' \leq N^{l'}, \\ 1 \leq s \leq S' \end{array}.$$

**Proposition 6.1.11.** Let  $\mathbf{t} = (a_1^l x_1 + \cdots + a_n^l x_n + b^l)_{l \in L}$  and  $\chi : P_\Gamma \rightarrow \mathbb{R}$ . If

$$\forall \lambda : V_\Gamma \rightarrow \mathbb{R}. \forall \left( (\mathfrak{d}_1 \triangleright_1 0 \wedge \cdots \wedge \mathfrak{d}_k \triangleright_k 0) \Rightarrow \mathbf{e}_i \geq 0 \right) \in \bigcup_{l \in L} (A_1^l \cup A_2^l).$$

$$(\llbracket \mathfrak{d}_1 \rrbracket(\lambda) \triangleright_1 0 \wedge \cdots \wedge \llbracket \mathfrak{d}_k \rrbracket(\lambda) \triangleright_k 0) \implies \llbracket \chi(\mathbf{e}) \rrbracket(\lambda) \geq 0,$$

then  $\llbracket \mathbf{t}_\chi \rrbracket$  is a  $\gamma$ -scaled submartingale for  $\mathcal{T}_{\Gamma, \mathfrak{J}, \mathfrak{S}}$ .  $\square$

### Relax Strict Inequalities

We next relax each formula of a form (6.4) to the following formula.

$$(\mathfrak{d}_1 \geq 0 \wedge \cdots \wedge \mathfrak{d}_k \geq 0) \Rightarrow \mathbf{e} \geq 0 \quad (6.6)$$

The same relaxation is done in the algorithm for additive ranking supermartingales in [19, 23].

Obviously, if a valuation  $\chi : P_\Gamma \rightarrow \mathbb{R}$  makes (6.6) hold then it also makes (6.4) hold. Its converse does not necessarily hold, and hence completeness is lost in general. The following lemma, which is easy to prove but seems new to the best of our knowledge, presents a sufficient condition for the converse to hold.

**Lemma 6.1.12.** Assume that  $\llbracket \mathfrak{d}_1 \geq 0 \wedge \cdots \wedge \mathfrak{d}_k \geq 0 \rrbracket \subseteq \mathbb{R}^{V_\Gamma}$  is an  $n$ -manifold with a boundary, i.e. each  $\lambda \in \llbracket \mathfrak{d}_1 \geq 0 \wedge \cdots \wedge \mathfrak{d}_k \geq 0 \rrbracket$  has a neighborhood that is homeomorphic to an open subset in either  $\mathbb{R}^n$  or  $\mathbb{R}^{n-1} \times [0, \infty)$ . Then for each  $\chi : P_\Gamma \rightarrow \mathbb{R}$  and  $\lambda : V_\Gamma \rightarrow \mathbb{R}$ ,

$$\begin{aligned} & \left( \lambda \in \llbracket \mathfrak{d}_1 \triangleright_1 0 \wedge \cdots \wedge \mathfrak{d}_k \triangleright_k 0 \rrbracket \Rightarrow \llbracket \chi(\mathbf{e}) \rrbracket(\lambda) \geq 0 \right) \\ & \implies \left( \lambda \in \llbracket \mathfrak{d}_1 \geq 0 \wedge \cdots \wedge \mathfrak{d}_k \geq 0 \rrbracket \Rightarrow \llbracket \chi(\mathbf{e}) \rrbracket(\lambda) \geq 0 \right). \end{aligned}$$

**Proof.** Assume  $\lambda \in \llbracket \mathfrak{d}_1 \triangleright_1 0 \wedge \cdots \wedge \mathfrak{d}_k \triangleright_k 0 \rrbracket \Rightarrow \mathbf{e}(\chi)(\lambda) \geq 0$ . It suffices to prove that for each  $\lambda \in \llbracket \mathfrak{d}_1 \geq 0 \wedge \cdots \wedge \mathfrak{d}_k \geq 0 \rrbracket \setminus \llbracket \mathfrak{d}_1 \triangleright_1 0 \wedge \cdots \wedge \mathfrak{d}_k \triangleright_k 0 \rrbracket$  and  $\varepsilon > 0$  we have  $\llbracket \chi(\mathbf{e}) \rrbracket(\lambda) \geq -\varepsilon$ .

By the assumption that  $\llbracket \mathfrak{d}_1 \geq 0 \wedge \cdots \wedge \mathfrak{d}_k \geq 0 \rrbracket$  is an  $n$ -manifold with a boundary, for each  $\delta > 0$  there exists  $\lambda' \in \llbracket \mathfrak{d}_1 \triangleright_1 0 \wedge \cdots \wedge \mathfrak{d}_k \triangleright_k 0 \rrbracket$  such that  $\|\lambda - \lambda'\| \leq \delta$ . Moreover, by the continuity of  $\llbracket \chi(\mathbf{e}) \rrbracket : \mathbb{R}^{V_\Gamma} \rightarrow \mathbb{R}$ , there exists  $\delta > 0$  such that for each  $\lambda' \in \llbracket \mathfrak{d}_1 \triangleright_1 0 \wedge \cdots \wedge \mathfrak{d}_k \triangleright_k 0 \rrbracket$ , if  $\|\lambda - \lambda'\| \leq \delta$  then  $\|\llbracket \chi(\mathbf{e}) \rrbracket(\lambda) - \llbracket \chi(\mathbf{e}) \rrbracket(\lambda')\| \leq \varepsilon$ . Hence we have:

$$\mathbf{e}(\chi)(\lambda) \geq \mathbf{e}(\chi)(\lambda') - \varepsilon \geq -\varepsilon. \quad \square$$

### Reduce to LP problem

Using matrices, we can express a formula of a form (6.6) as follows:

$$\forall \mathbf{x} \in \mathbb{R}^n. \mathbf{A}\mathbf{x} \leq \mathbf{b} \Rightarrow \mathbf{c}^T \mathbf{x} \leq d. \quad (6.7)$$

Here  $\mathbf{A} \in \mathbb{R}^{m \times n}$  is a matrix and  $\mathbf{b} \in \mathbb{R}^m$  is a column vector whose elements are real numbers, and  $\mathbf{c}$  is a column vector and  $d$  is a scalar whose elements are linear expressions over  $P_\Gamma$ . Recall that our goal is to find a valuation  $\chi : P_\Gamma \rightarrow \mathbb{R}$  that makes (6.7) hold.

In [19], it was translated to an LP problem. We explain the translation. We first observe that the following is obviously a sufficient condition for (6.7):

$$\forall \mathbf{x} \in \mathbb{R}^n. \exists \mathbf{y} \in \mathbb{R}^m. \exists z \in \mathbb{R}. (d - \mathbf{c}^T \mathbf{x}) = z + \mathbf{y}^T (\mathbf{b} - \mathbf{A}\mathbf{x}) \wedge \mathbf{y} \geq \mathbf{0} \wedge z \geq 0. \quad (6.8)$$

A natural question would be about the completeness of the above reduction. The following theorem partially answers the question.

**Theorem 6.1.13** (affine form of Farkas lemma, see e.g. [94, Corollary 7.1h]). *If  $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$  is nonempty, then (6.7) implies (6.8).  $\square$*

By comparing the coefficients on both sides of (6.8), we can see that it is equivalent to the following:

$$\exists \mathbf{y} \in \mathbb{R}^m. \mathbf{A}^T \mathbf{y} = \mathbf{c} \wedge \mathbf{b}^T \mathbf{y} \geq d \wedge \mathbf{y} \geq \mathbf{0}. \quad (6.9)$$

Hence if we find a valuation  $P_\Gamma \rightarrow \mathbb{R}$  and a vector  $\mathbf{y} \in \mathbb{R}^m$  satisfying (6.9) then (6.7) is satisfied. As parameters in  $P_\Gamma$  do not appear in  $\mathbf{A}$  or  $\mathbf{b}$ , (in)equalities in (6.9) are linear with respect to  $\mathbf{y}$  and parameters in  $P_\Gamma$ . Hence the satisfiability problem is efficiently solvable using an LP solver.

Recall that we wish to maximize the probability bound  $\llbracket \mathbf{t}_\chi \rrbracket(l_{\text{init}}, \lambda_{\text{init}})$  given by the synthesized  $\gamma$ -scaled submartingale. We can achieve this task by setting  $\llbracket \mathbf{t}_\chi \rrbracket(l_{\text{init}}, \lambda_{\text{init}})$ , which is a linear expression over the parameters, to the objective function of the LP problem.

To summarize, the linear template-based algorithm is as follows:

**Algorithm 0** Turn the axioms of  $\gamma$ -scaled submartingale for the pCFG into a conjunction of formulas of a form (6.7). For each of the formulas, fix a vector  $\mathbf{y}$  consisting of new parameters, and collect linear equalities as in (6.9) for all of such formulas. We then ask the LP solver to maximize  $\llbracket \mathbf{t}_\chi \rrbracket(l_{\text{init}}, \lambda_{\text{init}})$  under the collected conditions.

#### 6.1.4 Implementation

We have implemented the algorithm described in the previous section.

So that we can designate an invariant and a terminal configuration, we have augmented the syntax of LPPs in Definition 6.1.1 with two components  $\{\langle \text{expr} \rangle\}$  and  $\llbracket \langle \text{expr} \rangle \rrbracket$ . Concretely, we replaced (6.2) in Definition 6.1.1 with the following.

$$\begin{aligned} \langle \text{stmt} \rangle ::= & \langle \text{stmt} \rangle ; \langle \text{stmt} \rangle \mid \langle \text{assgn} \rangle \mid \mathbf{skip} \mid \mathbf{while} \langle \text{pbexpr} \rangle \mathbf{do} \langle \text{stmt} \rangle \mathbf{od} \\ & \mid \mathbf{if} \langle \text{pbexpr} \rangle \mathbf{then} \langle \text{stmt} \rangle \mathbf{else} \langle \text{stmt} \rangle \mathbf{fi} \mid \{\langle \text{expr} \rangle\} \mid \llbracket \langle \text{expr} \rangle \rrbracket \end{aligned}$$

A statement  $\{\langle \text{expr} \rangle\}$  assigns the expression  $\langle \text{expr} \rangle$  to the location just after it as an invariant. A statement  $\llbracket \langle \text{expr} \rangle \rrbracket$  assigns  $\langle \text{expr} \rangle$  to the location just after it as a terminal configuration. An example of input is shown on the right. It implements a simple unbounded random walk.

We have implemented the algorithm by modifying an existing implementation. *Nonnegative repulsing supermartingale* is a ranking function-like notion for PTSs introduced in [102]. We can use it for overapproximating reachability probabilities of PTSs.

```

1  x := 1;
2  {0 <= x}
3  while x > 0 do
4    {1 <= x}
5    if prob(0.75) then
6      {1 <= x}
7      x := x + 1
8    else
9      {1 <= x}
10     x := x - 1
11   fi;
12   {0 <= x}
13 od;
14 {x <= 0} [true]

```

**Definition 6.1.14** (nonnegative repulsing supermartingale). Let  $\mathcal{T} = ((X, \mathfrak{F}_X), \xi, \text{Acc})$  be a PTS. A *nonnegative repulsing supermartingale* for  $\mathcal{T}$  is a measurable function  $b: X \rightarrow [0, \infty]$  that satisfies the following condition:

$$\forall x \in X \setminus \text{Acc}. b(x) \geq \int_{x' \in X} b(x') d\xi(x) \quad \text{and} \quad \forall x \in \text{Acc}. b(x) \geq 1.$$

**Theorem 6.1.15** ([102]). *Let  $b: X \rightarrow [0, \infty]$  be a nonnegative repulsing supermartingale for a PTS  $\mathcal{T} = ((X, \mathfrak{F}_X), \xi, \text{Acc})$ . Then for each  $x \in X$ ,  $\text{Reach}_{\mathcal{T}}(x) \geq b(x)$ .  $\square$*

As we can see, the definition of nonnegative repulsing supermartingale is very similar to those of  $\gamma$ -scaled submartingale (Definition 5.5.11) and ranking supermartingale (Definition 6.0.1). As we have done in the previous section, in [102], a linear (and polynomial) template-based synthesis algorithm for finding a nonnegative repulsing supermartingale is given by modifying the algorithm for ranking supermartingales in [19, 21]. In [102], an implementation of the algorithm, which is written in OCaml, was also given.

By modifying the implementation in [102], we have implemented a program that takes an LPP and  $\gamma \in [0, 1)$  as inputs and outputs an LP problem as is described in the previous section. If we feed the output to an LP solver, then its optimal solution  $p \in [0, 1]$  satisfies  $p \leq \text{Reach}_{\mathcal{T}}(l_{\text{init}}, \vec{0})$ .

Note that our program assumes that the initial valuation is  $\vec{0}$ . This is not a restriction because we can fix the initial valuation by inserting an assignment instruction to the beginning of the LPP.

### 6.1.5 Experiments I: Probabilistic Programs in the Literature

As described in the previous section, our implementation outputs input to an LP solver. We have used glpk [40] (version 4.63) as an LP solver. The experiments were conducted on a MacBook Pro laptop with a Core i5 processor (2.6 GHz, 2 cores) and 16 GB RAM.

We first tested our implementation for the following probabilistic programs.

1. **(bounded random walk)** A simple bounded random walk (it is sometimes called *gambler's ruin problem* [7]). A pebble is at a position  $x = n$  first. In each turn, it moves to the right with a probability  $p$  and moves to the left with the remaining probability. If the pebble reaches  $x = 0$  then the program terminates, and if it reaches  $x = 100$  the program diverges. The code is shown in Figure 6.1. The true termination probability is  $(1 - (\frac{p}{1-p})^{100-n}) / (1 - (\frac{p}{1-p})^{100})$  if  $p \neq 0.5$  and  $1 - n/100$  if  $p = 0.5$  (see e.g. [7]).
2. **(room temperature)** A model of an air-conditioning control system for two adjacent rooms. A similar model is used in [20]. Each room exchange heat with the other room and the open air. An air conditioner is installed to each room, and it tries to keep the temperature of the room to  $19.5^\circ\text{C}$ . The room temperatures perturb following a uniform distribution. We are interested in the probability where the temperatures of the rooms go out of specified ranges until 100 seconds. The code is in Figure 6.2. It is parameterized by a constant  $C$  that determines the magnitude of the perturbations.
3. **(simple pendulum)** An approximated model of a pendulum. A similar model is used in [100]. It perturbs following a uniform distribution. We wish to know the probability where the position of the weight goes out of a specified range in 100 seconds. The code is in Figure 6.3. It is parameterized by constants  $C$  and  $D$  that determine the magnitude of the perturbations.

The results are in Table 6.1. We let  $\gamma = 0.999$ . The first column shows the input probabilistic program (the numbers correspond to the above enumeration) and the values of constants. The next two columns show the results: the total

```

1 x := n;
2 { 1 <= x and x <= 99 }
3 while 0 <= x do
4   { 1 <= x and x <= 99 }
5   if prob(p) then
6     { 1 <= x and x <= 99 }
7     x := x + 1
8   else
9     { 1 <= x and x <= 99 }
10    x := x - 1
11  fi;
12  { 0 <= x and x <= 100 }
13  while x >= 100 do
14    { 100 <= x }
15    skip
16  od
17 od;
18 { x < 0 } [ true ]

```

Figure 6.1: code for 1

	param.	time (s)	bound	true prob.
1	n = 10 p = 0.1	0.023638	$\geq 0.90437$	$1 - 1.3127 \times 10^{-86}$
	n = 90 p = 0.1	0.021892	$\geq 0.10757$	$1 - 2.8680 \times 10^{-10}$
	n = 10 p = 0.9	0.018067	$\geq 0$	$2.8680 \times 10^{-10}$
	n = 50 p = 0.5	0.018341	$\geq 0$	0.5
2	C = 1	0.047402	$\geq 0$	—
	C = 10	0.049987	$\geq 0.75037$	—
	C = 20	0.053965	$\geq 0.93285$	—
	C = 100	0.071837	$\geq 0.95676$	—
3	C = -0.01 D = 0.01	0.028786	$\geq 0$	—
	C = -1 D = 1	0.027086	$\geq 0$	—
	C = -1 D = 9	0.025237	$\geq 0$	—
	C = -1 D = 99	0.025537	$\geq 0$	—

Table 6.1: experimental results for the linear template-based algorithm

time consumed to translate an LPP to an LP problem and calculate a probability bound using glpk, and the calculated probability bounds. As we can easily calculate the true termination probabilities for 1, they are also shown.

For 1 (bounded random walk), a nontrivial probability bound was given when the true reachability probability is close to 1. For 2 (room temperature), a nontrivial bound was given when the perturbation is large (i.e. when the reachability probability is large). These two examples show that the probability bound given by a linear  $\gamma$ -scaled submartingale increases discontinuously as the true reachability increases.

For 3 (simple pendulum), our algorithm could not give any nontrivial probability bound. This would be because of the shape of the function of the true probabilities. In the code in Figure 6.3, the probability to reach the terminal configuration within 100 seconds is 1 at positions `theta1` = 0.01 and -0.01, and it takes the minimum value at `theta1` = 0. This means that the function giving the true probabilities is U-shaped, and this might make it hard to approximate it with a linear function.

### 6.1.6 Experiments II: Comparison with Existing Work

Compared to those for proving almost-sure termination, the number of algorithms for underapproximating the termination probability is limited. One of them is in [23]. In this section, we compare it with our algorithm.

We first briefly explain the theory and the algorithm in [23]. The key notions are *repulsing supermartingale* and *stochastic invariant*. The former is a new ranking function-like notion that can *over*-approximate the reachability probability.

**Definition 6.1.16** ([23], repulsing supermartingale). Let  $\mathcal{T} = ((X, \mathfrak{F}_X), \xi, \text{Acc})$  be a PTS such that  $\mathfrak{F}_X$  is a Borel  $\sigma$ -algebra with respect to a topology  $\mathcal{O}_X$  over  $X$ . A *repulsing supermartingale* for  $\mathcal{T}$  is a measurable function  $b: X \rightarrow \mathbb{R} \cup \{\infty\}$  that satisfies the following condition:

$$\forall x \in X \setminus \text{Acc}. b(x) \geq \int_{x' \in X} b(x') d\xi(x) + 1 \quad \text{and} \quad \forall x \in \text{Acc}. b(x) \geq 0.$$

For  $\kappa > 0$ , we say that  $b$  has  $\kappa$ -bounded differences (with respect to  $\mathcal{O}_X$ ) if

$$\forall x \in X. \forall x' \in \text{supp}(\xi(x)). |b(x) - b(x')| \leq \kappa.$$

```

1 x0 := 6; x1 := 18; x2 := 19; t := 0;
2 { x0 = 6 and 17 <= x1 and x1 <= 22 and 16 <= x2 and x2 <= 23 and 0 <= t
3   and t <= 101 }
4 while t <= 100 do
5   { x0 = 6 and 17 <= x1 and x1 <= 22 and 16 <= x2 and x2 <= 23 and
6     0 <= t and t <= 100 }
7   controller1 := 19.5 - x1;
8   { x0 = 6 and 17 <= x1 and x1 <= 22 and 16 <= x2 and x2 <= 23 and
9     -2.5 <= controller1 and controller1 <= 2.5 and 0 <= t and t <= 100 }
10  controller2 := 19.5 - x2;
11  { x0 = 6 and 17 <= x1 and x1 <= 22 and 16 <= x2 and x2 <= 23 and
12    -2.5 <= controller1 and controller1 <= 2.5 and -3.5 <= controller2
13    and controller2 <= 3.5 and 0 <= t and t <= 100 }
14  noise1 := Unif(-1,C);
15  { x0 = 6 and 17 <= x1 and x1 <= 22 and 16 <= x2 and x2 <= 23 and
16    -2.5 <= controller1 and controller1 <= 2.5 and -3.5 <= controller2
17    and controller2 <= 3.5 and -1 <= noise1 and noise1 <= C and 0 <= t
18    and t <= 100 }
19  noise2 := Unif(-C,1);
20  { x0 = 6 and 17 <= x1 and x1 <= 22 and 16 <= x2 and x2 <= 23 and
21    -2.5 <= controller1 and controller1 <= 2.5 and -3.5 <= controller2
22    and controller2 <= 3.5 and -1 <= noise1 and noise1 <= C and
23    -C <= noise2 and noise2 <= 1 and 0 <= t and t <= 100 }
24  x1 := x1 + 0.0375 * x0 - 0.0375 * x1 + 0.0625 * x2 - 0.0625 * x1
25    + 0.5 * controller1 + noise1;
26  { x0 = 6 and 16 <= x1 and x1 <= 22+C and 16 <= x2 and x2 <= 23 and
27    -3.5 <= controller2 and controller2 <= 3.5 and -C <= noise2 and
28    noise2 <= 1 and 0 <= t and t <= 100 }
29  x2 := x2 + 0.025 * x0 - 0.025 * x2 + 0.0625 * x1 - 0.0625 * x2
30    + 0.5 * controller2 + noise2;
31  { x0 = 6 and 16 <= x1 and x1 <= 22+C and 16-C <= x2 and
32    x2 <= 23+1.0625*C and 0 <= t and t <= 100 }
33  t := t + 1;
34  { x0 = 6 and 16 <= x1 and x1 <= 22+C and 16-C <= x2 and
35    x2 <= 23+1.0625*C and 0 <= t and t <= 101 }
36  [ x1 < 17 or x1 > 22 or x2 < 16 or x2 > 23 ]
37  skip
38 od

```

Figure 6.2: code for 2

```

1 theta1 := 0; dt_theta := 0; t := 0;
2 { -0.01 + 0.01 * dt_theta <= theta1 and theta1 <= 0.01 + 0.01 * dt_theta
3   and -2+C <= dt_theta and dt_theta <= 2+D and 0 <= t and t <= 100.01 }
4 [ theta1 > 0.01 or theta1 < -0.01 ]
5 while t <= 100 do
6   { -0.01 <= theta1 and theta1 <= 0.01 and -2 <= dt_theta and
7     dt_theta <= 2 and 0 <= t and t <= 100 }
8   w1 := Unif (C,D);
9   { -0.01 <= theta1 and theta1 <= 0.01 and -2 <= dt_theta and
10    dt_theta <= 2 and C <= w1 and w1 <= D and 0 <= t and t <= 100 }
11  dt_theta := dt_theta - 0.1 * theta1 + w1;
12  { -0.01 <= theta1 and theta1 <= 0.01 and -2+C <= dt_theta and
13    dt_theta <= 2+D and 0 <= t and t <= 100 }
14  theta1 := theta1 + 0.01 * dt_theta;
15  { -0.01 + 0.01 * dt_theta <= theta1 and
16    theta1 <= 0.01 + 0.01 * dt_theta and -2+C <= dt_theta and
17    dt_theta <= 2+D and 0 <= t and t <= 100 }
18  t := t + 0.01
19 od

```

Figure 6.3: code for 3

Here  $\text{supp}(\xi(x)) \subseteq X$  denotes the support of  $\xi(x)$ , i.e.  $\text{supp}(\xi(x)) = \{x' \in X \mid \forall O \in \mathcal{O}_X. x' \in O \Rightarrow \xi(x)(O) > 0\}$ .

**Theorem 6.1.17** ([23]). *Let  $b: X \rightarrow \mathbb{R} \cup \{\infty\}$  be a repulsing supermartingale having  $\kappa$ -bounded differences for a PTS  $\mathcal{T} = ((X, \mathfrak{F}_X), \xi, \text{Acc})$  as in Definition 6.1.16.*

1. For each  $x \in X$ ,

$$\text{Reach}_{\mathcal{T}}(x) \leq e^{\frac{b(x)}{(\kappa+1)^2}} \cdot \frac{(e^{-\frac{1}{2(\kappa+1)^2}})^{\lceil \frac{|b(x)|}{\kappa} \rceil}}{1 - (e^{-\frac{1}{2(\kappa+1)^2}})}. \quad (6.10)$$

2. For each  $x \in X$ ,  $b(x) < \infty \Rightarrow \text{Reach}_{\mathcal{T}}(x) < 1$ .  $\square$

In [23], it is shown that we can *under*-approximate a reachability probability by *over*-approximating a reachability probability to some set of states and then synthesizing a ranking supermartingale (Definition 6.0.1). The main idea is to find a *stochastic invariant*.

**Definition 6.1.18** (stochastic invariant). Let  $\mathcal{T} = ((X, \mathfrak{F}_X), \xi, \text{Acc})$  be a PTS. A *stochastic invariant* is a pair  $(PI, p)$  of  $PI \subseteq X$  and  $p \in [0, 1]$  such that  $\text{Reach}_{\mathcal{T}}(x, X \setminus PI) \leq p$ . We call a *pure invariant* for an invariant in Definition 6.1.5 for distinction.

Note that we can prove that a given pair  $(PI, p)$  is a stochastic invariant by overapproximating the reachability probability to  $X \setminus PI$ .

Using stochastic invariants, we can *under*-approximate reachability probabilities as follows.

**Theorem 6.1.19** ([23]). *Let  $\mathcal{T} = ((X, \mathfrak{F}_X), \xi, \text{Acc})$  be a PTS, and  $(PI_1, p_1), \dots, (PI_n, p_n)$  be stochastic invariants. Let  $b: X \rightarrow [0, \infty]$  be a ranking supermartingale supported by  $PI_1 \cap \dots \cap PI_n$ , i.e. it satisfies  $b(x) \geq \int_{x' \in X} b(x') d\xi(x) + 1$  for each  $x \in (PI_1 \cap \dots \cap PI_n) \setminus \text{Acc}$ . Then for each  $x \in X$ ,  $\text{Reach}_{\mathcal{T}}(x) \geq 1 - \sum_{i=1}^n p_i$ .  $\square$*

Using the above results, the following algorithms are given in [23].

**Algorithm A** For a given linear pCFG  $\Gamma$ , a linear pure invariant  $\mathfrak{J}$  and linear predicate maps  $\mathfrak{T}$  and  $\mathfrak{J}'$ , it computes  $p \in [0, 1]$  that makes  $(\llbracket \mathfrak{J}' \rrbracket, p)$  a stochastic invariant over  $\mathcal{T}_{\Gamma, \mathfrak{J}, \mathfrak{T}}$ .

**Algorithm B** For a given linear pCFG  $\Gamma$ , a linear pure invariant  $\mathfrak{J}$  and a linear predicate map  $\mathfrak{T}$ , it computes a linear predicate map  $\mathfrak{J}'$  and  $p \in [0, 1]$  such that  $(\llbracket \mathfrak{J}' \rrbracket, p)$  is a stochastic invariant over  $\mathcal{T}_{\Gamma, \mathfrak{J}, \mathfrak{T}}$ .

Algorithm A computes  $p$  by synthesizing a repulsing supermartingale for the inputs. The synthesis can be done in a very similar manner to the standard synthesis algorithm for ranking supermartingales, from which our algorithms are also adapted. While our algorithm tries to maximize the probability bound, Algorithm A tries to minimize the probability bound. However, as a repulsing supermartingale is parameterized by  $\kappa$  (see Definition 6.1.16), unlike our setting, the probability bound given by a repulsing supermartingale (6.10) is not linear nor polynomial with respect to the parameters in the template and  $\kappa$ . Hence we cannot naively set it to the objective function of an LP solver. In [23] the

following heuristics is used: we first synthesize a repulsing supermartingale so that  $\kappa$  is minimized (we let  $\kappa_{\min}$  be the calculated  $\kappa$ ). Then for each  $\kappa \in \{\kappa_{\min}, \kappa_{\min} + 1, \dots, \kappa_{\min} + N\}$ , we synthesize a repulsing supermartingale so that  $\frac{|b(l_{\text{init}}, x_I)|}{\kappa}$  is minimized, and then take the minimum probability bound among them. They let  $N = 1000$  in [23].

Algorithm B is more complicated. It fixes a template for a ranking supermartingale, a repulsing supermartingale and a stochastic invariant. Then it reduces the axioms of ranking supermartingale and repulsing supermartingale to inequalities over the parameters in the templates. Unlike Algorithm A, the resulting inequalities are not necessarily linear but quadratic in general. However, they are solvable as the first order theories of real numbers are decidable [104].

Unfortunately, no implementation of Algorithm B was accessible for us and it seems hard to implement, while an implementation of Algorithm A can be easily obtained by modifying our implementation for  $\gamma$ -scaled submartingale. Moreover, in [23] experiments were conducted only for Algorithm A, and only the resulting probability bounds are presented as experimental results. Hence we decided to compare our and their algorithms as follows: we first implemented the algorithm in [23] by modifying our implementation. In [23], their implementation of Algorithm A is tested for three probabilistic programs equipped with pure and stochastic invariants such that a ranking supermartingale supported by the stochastic invariant exists. For each combination of a probabilistic program (it induces a pCFG  $\Gamma$ , an invariant  $\mathfrak{I}$  and a terminal configuration  $\mathfrak{T}$ ) and a stochastic invariant  $\mathfrak{J}'$ , we have compared: i)  $1 - p$ , where  $p \in [0, 1]$  is calculated by our implementation of Algorithm A so that it makes  $(\llbracket \mathfrak{J}' \rrbracket, p)$  a stochastic invariant over  $\mathcal{T}_{\Gamma, \mathfrak{J}, \mathfrak{T}}$ ; and ii) an underapproximation of the termination probability calculated by a 0.999-scaled submartingale. We have not compared time consumption as the efficiency of Algorithm B would depend on the way to solve the quadratic inequalities, that is not explicitly specified in [23].

The probabilistic programs and their stochastic invariants used in [23] are shown in Figures 6.4–6.6. The programs model variants of 1D, 2D, and 3D random walks respectively. Stochastic invariants are represented as  $/ \dots /$ . We have modified some pure invariants in the probabilistic programs in [23] because our implementation of Algorithm A could not find a repulsing supermartingale for the original probabilistic programs. The probabilistic programs are parameterized, and in [23] experiments were conducted for three combinations of parameters for each.

We have also tested the algorithm in [23] and our algorithm for the bounded random walk in Figure 6.1 by fixing a seemingly reasonable stochastic invariant  $\mathfrak{J}'$ . Concretely, the stochastic invariant  $\mathfrak{J}'$  assigns  $x \geq 99.5$  to line 13, and true to the other locations.

The experimental results are shown in Table 6.2. A term “> 0” means that a repulsing supermartingale failed to give a probability bound but it proved that it terminates in a positive probability using Theorem 6.1.17.2. A term “infeasible” in the last row means that no repulsing supermartingale was found.

For all the probabilistic programs used in [23], the algorithm in [23] gave tighter probability bounds than ours. In contrast, for a bounded random walk, our algorithm gave better bounds for some parameters.



```

1 x := 100;
2 { -2 <= x }
3 while x >= 0 do
4   { 0 <= x }
5   if x <= 1000.5 then
6     { 0 <= x and x <= 1000 }
7     if prob(0.5) then
8       { 0 <= x and x <= 1000 }
9       x := x - 2
10    else
11      { 0 <= x and x <= 1000 }
12      x := x + 1
13    fi
14  else
15    { 1001 <= x }
16    if prob(0.5) then
17      { 1001 <= x }
18      x := x - 1
19    else
20      { 1001 <= x }
21      x := x + 2
22    fi
23  fi
24 od;
25 { x <= -1 } [ true ]

```

Figure 6.4: 1D random walk

```

1 x := 400; y := 50;
2 { 0 <= y }
3 while 1 <= y do
4   { 1 <= y }
5   if prob(0.5) then
6     { 1 <= y }
7     if prob(0.75) then
8       { 1 <= y }
9       x := x + 1
10    else
11      { 1 <= y }
12      x := x - 1
13    fi
14  else
15    { 1 <= y }
16    if prob(0.75) then
17      { 1 <= y }
18      y := y - 1
19    else
20      { 1 <= y }
21      y := y + 1
22    fi
23  fi;
24 { 0 <= x and 0 <= y } / x < 1 /
25 while x <= 0 do
26   { x <= 0 and 0 <= y }
27   x := 0
28 od
29 od;
30 { 1 <= x and y <= 0 } [ true ]

```

Figure 6.5: 2D random walk

```

1 x := 300; y := 100; z := 150;
2 { -1 <= x and -1 <= y and -1 <= z } / x + y + z > 1000 /
3 while 0 <= x and 0 <= y and
4   0 <= z do
5   { 0 <= x and 0 <= y and 0 <= z }
6   { 0 <= x and 0 <= y and 0 <= z }
7   { 0 <= x and 0 <= y and 0 <= z }
8   { -1 <= x and 0 <= y and 0 <= z }
9
10  { 0 <= x and 0 <= y and 0 <= z }
11
12
13  { 0 <= x and 0 <= y and 0 <= z }
14  { 0 <= x and 0 <= y and 0 <= z }
15  { 0.1 <= x and 0 <= y and 0 <= z }
16
17  { 0 <= x and 0 <= y and 0 <= z }
18
19
20
21 { x <= -1 and y <= -1 and z <= -1 } [ true ]

```

Figure 6.6: 3D random walk

## 6.2 Polynomial Template-Based Algorithm

The experimental result and the discussion for a probabilistic program modeling a pendulum in Section 6.1.5 encourages us to fix a more complicated template for a  $\gamma$ -scaled submartingale. In this section, we focus on a polynomial template-based synthesis of  $\gamma$ -scaled submartingales.

A polynomial template-based synthesis algorithm for ranking supermartingales is found in [21]. We have adapted the algorithm for our setting again, and implemented it. However, to get straight to the result, our implementation did not work well, possibly because of numerical errors. We made an attempt to remedy the situation, but it also failed. We still present them for record.

	param.	algorithm in [23]	our algorithm	true prob.
Fig. 6.4	$x = 10$	$\geq 1 - 5.2959 \times 10^{-15}$	$\geq 0.90347$	—
	$x = 50$	$\geq 1 - 1.25427 \times 10^{-14}$	$\geq 0.58836$	—
	$x = 100$	$\geq 1 - 1.8083 \times 10^{-13}$	$\geq 0.19448$	—
Fig. 6.5	$x, y = 1000, 10$	$\geq 1 - 1.7674 \times 10^{-16}$	$\geq 0$	—
	$x, y = 500, 40$	$\geq 1 - 1.2930 \times 10^{-6}$	$\geq 5.9952 \times 10^{-15}$	—
	$x, y = 400, 50$	$\geq 1 - 1.4439 \times 10^{-4}$	$\geq 0$	—
Fig. 6.6	$x, y, z = 100, 100, 100$	$\geq 1 - 1.91158 \times 10^{-70}$	$\geq 6.5725 \times 10^{-14}$	—
	$x, y, z = 100, 150, 200$	$\geq 1 - 1.5420 \times 10^{-54}$	$\geq 3.2085 \times 10^{-14}$	—
	$x, y, z = 300, 100, 150$	$\geq 1 - 2.1891 \times 10^{-44}$	$\geq 0$	—
Fig. 6.1	$n, p = 10, 0.1$	$\geq 0.010200$	$\geq 0.90437$	$1 - 1.3127 \times 10^{-86}$
	$n, p = 90, 0.1$	$> 0$	$\geq 0.10757$	$1 - 2.8680 \times 10^{-10}$
	$n, p = 10, 0.9$	$\geq 0$	$\geq 0$	$2.8680 \times 10^{-10}$
	$n, p = 50, 0.5$	infeasible	$\geq 0$	0.5

Table 6.2: comparison with the algorithm in [23]

### 6.2.1 Syntax of Probabilistic Programs

If we use a polynomial template, we can deal with programs including polynomial expressions.

**Definition 6.2.1** (polynomial probabilistic program). A *polynomial probabilistic program* (PPP) is defined in almost the same way as Definition 6.1.1 except that: i) line (6.3) in Definition 6.1.1 is replaced by:

$$\langle expr \rangle ::= \langle const \rangle \mid \langle pvar \rangle \mid \langle expr \rangle \cdot \langle expr \rangle \mid \langle expr \rangle + \langle expr \rangle \mid \langle expr \rangle - \langle expr \rangle$$

and the (†) in Definition 6.1.1 is replaced by the following:

(†') for each probability measure  $d$  appearing in  $\langle prog \rangle$  and each  $n \in \mathbb{N}$ , an algorithm that calculates the  $n$ -th moment of  $d$  is given.

We write  $\{\{stmt\}\}_{poly}$ ,  $\{\{assgn\}\}_{poly}$ ,  $\{\{expr\}\}_{poly}$ , etc... for the sets of formulas defined by this new BNF notation. We call an element in  $\{\{expr\}\}_{poly}$  a *polynomial expression*.

We define  $V_\Gamma$  and notions of location, valuation and configuration as in Definition 6.1.1. We define  $\llbracket \langle expr \rangle \rrbracket : \mathbb{R}^\nu \rightarrow \mathbb{R}$  and  $\llbracket b \rrbracket \subseteq \mathbb{R}^\nu$  in the standard manners.

### 6.2.2 Problem

A *polynomial probabilistic control flow graph* is defined in the same manner as Definition 6.1.2, except that  $\{\{assgn\}\}_{lin}$  and  $\{\{pbexpr\}\}_{lin}$  in the third bullet of Definition 6.1.2 are replaced by  $\{\{assgn\}\}_{poly}$  and  $\{\{pbexpr\}\}_{poly}$  respectively. Moreover, a polynomial pCFG induced by a PPP, a *polynomial predicate map*  $\mathbf{p} : L_A + L_B \rightarrow \{\{expr\}\}_{poly}$ , a *polynomial invariant*  $\mathfrak{I} : L_A + L_B \rightarrow \{\{expr\}\}_{poly}$  for a polynomial pCFG, and a PTS  $\mathcal{T}_{\Gamma, \mathfrak{I}, \mathfrak{P}}$  induced by a polynomial pCFG, a polynomial invariant and a polynomial predicate map are all defined in the same way as Definitions 6.1.3–6.1.6.

We can now state the problem as follows.

**Problem 6.2.2.**   input:   a polynomial pCFG  $\Gamma = (L_A, L_B, l_{init}, \tau_A, \tau_B)$   
                                   an initial valuation  $\lambda_{init} : V_\Gamma \rightarrow \mathbb{R}$   
                                   a polynomial invariant  $\mathfrak{I}$  for a pCFG  $\Gamma$   
                                   a polynomial predicate map  $\mathfrak{P}$  for a pCFG  $\Gamma$   
                                   output:     $p \in \mathbb{R}$  such that  $p \leq \text{Reach}_{\mathcal{T}_{\Gamma, \mathfrak{I}, \mathfrak{P}}}(l_{init}, \lambda_{init})$

### 6.2.3 Algorithm

The algorithm is adapted from the existing one for ranking supermartingales in [21]. Throughout this section, let  $\Gamma = (L_A, L_B, l_{\text{init}}, \tau_A, \tau_B)$  be a polynomial pCFG,  $L = L_A + L_B$  and  $\{x_1, \dots, x_n\} = V_\Gamma$ . Let  $\mathfrak{J} : L \rightarrow \{\{expr\}\}_{\text{poly}}$  be a polynomial invariant and  $\mathfrak{T} : L \rightarrow \{\{expr\}\}_{\text{poly}}$  be a polynomial predicate map, and assume that an initial valuation  $\lambda_{\text{init}} \in \llbracket \mathfrak{J}(l_{\text{init}}) \rrbracket (\subseteq \mathbb{R}^{V_\Gamma})$  is given.

The algorithm is similar to the linear template-based one.

#### Fix a Template

In this case, a template has a polynomial shape.

**Definition 6.2.3** (polynomial expression map and polynomial template).

- A *polynomial expression map* for  $\Gamma$  is a function  $\mathfrak{f} : L \rightarrow \{\{expr\}\}_{\text{poly}}$ . We define  $\llbracket \mathfrak{f} \rrbracket : L \times \mathbb{R}^{V_\Gamma} \rightarrow \mathbb{R}$  as in Definition 6.1.8.
- Let  $d \in \mathbb{N}$ . We write  $\mathcal{M}_{\leq d}$  for the set of monomials over  $V_\Gamma$  whose degrees are no greater than  $d$ , i.e.

$$\mathcal{M}_{\leq d} := \left\{ x_1^{d_1} \cdots x_n^{d_n} \mid d_1, \dots, d_n \geq 0, d_1 + \dots + d_n \leq d \right\}.$$

A *polynomial template* over  $\Gamma$  is a family  $\mathfrak{t} = (\mathfrak{t}(l))_{l \in L}$  of formulas of a form  $\mathfrak{t}(l) = \prod_{h \in \mathcal{M}_{\leq d}} a_h^l h$ . Here each  $a_h^l$  is a new variable called a *parameter*. We write  $P_\Gamma$  for the set of all parameters. For a valuation  $\chi : P_\Gamma \rightarrow \mathbb{R}$  of parameters, we write  $\mathfrak{t}_\chi$  for a polynomial expression map  $l \mapsto \prod_{h \in \mathcal{M}_{\leq d}} \chi(a_h^l) h$ .

Similarly to the linear case, we wish to synthesize  $\chi : P_\Gamma \rightarrow \mathbb{R}$  that makes  $\mathfrak{t}_\chi$  a “ $\gamma$ -scaled submartingale” in the sense that  $\llbracket \mathfrak{t}_\chi \rrbracket |_{\prod_{l \in L} \llbracket \mathfrak{J}(l) \rrbracket}$  is a  $\gamma$ -scaled submartingale, and maximizes  $\llbracket \mathfrak{t}_\chi \rrbracket (l_{\text{init}}, \lambda_{\text{init}}) \in \mathbb{R}$ .

#### Collect Conditions for Parameters

We reduce the axioms of  $\gamma$ -scaled submartingales to conditions on the parameters. In a similar manner to the linear case, we can reduce the axioms to a conjunction of formulas of a form as in (6.4), but in the current setting,  $\mathfrak{d}_1, \dots, \mathfrak{d}_k$  are *polynomial* expressions over  $V_\Gamma$ . Moreover,  $\mathfrak{e}$  has a form

$$\sum_{h \in \mathcal{M}_{\leq d}} p_h \cdot h \tag{6.11}$$

where each coefficient  $p_h$  is a *linear* expression over  $P_\Gamma$ . We will write  $\chi(\mathfrak{e})$  for a polynomial expression  $\sum_{h \in \mathcal{M}_{\leq d}} \llbracket p_h \rrbracket \cdot h$ .

The reduction is similar to Definition 6.1.10, so we do not write down all the concrete constructions. The most non-trivial case would be the one of probabilistic assignment. If  $l \in L_A$ ,  $\tau_A(l) = (\langle \text{assgn} \rangle, l')$ ,  $\langle \text{assgn} \rangle = (x_k := \mathbf{sample}(\langle \text{dist} \rangle))$  and the  $t$ -th moment of  $\langle \text{dist} \rangle$  is  $r_t \in \mathbb{R}$  for each  $t \in \mathbb{N}$ , a set of formulas that is analogous to  $A_2^l$  in Definition 6.1.10 is given as follows:

$$\left\{ \begin{array}{l} \bigwedge_{j=1}^{M_i^l} (\alpha_{i,j}^l \triangleright_{i,j}^l 0) \wedge \bigwedge_{j=1}^{M_{i'}^{l'}} (\beta_{i',j}^{l'} \triangleright_{i',j}^{l'} 0) \Rightarrow \\ \left( \sum_{\substack{h=x_1^{d_1} \cdots x_n^{d_n} \\ \in \mathcal{M}_{\leq d, d_k=0}}} \left( \gamma \sum_{d'_k=0}^{d-(d_1+\dots+d_n)} a_{x_1^{d_1} \cdots x_{k-1}^{d_{k-1}} x_k^{d'_k} x_{k+1}^{d_{k+1}} \cdots x_n^{d_n}} r_{d'_k} - a_h^l \right) h \right) \\ - \sum_{h=x_1^{d_1} \cdots x_n^{d_n} \in \mathcal{M}_{\leq d, d_k \neq 0}} a_h^l h \end{array} \right\} \geq 0 \quad \left| \begin{array}{l} 1 \leq i \leq N^l, \\ 1 \leq i' \leq N^{l'} \end{array} \right.$$

## Relax Strict Inequalities

As in the linear case, we then relax the strict inequalities and obtain a conjunction of formulas of a form (6.6), where  $\mathfrak{d}_1, \dots, \mathfrak{d}_k, \mathfrak{e}$  are polynomial. An analogous statement to Lemma 6.1.12 holds.

## Reduce to SDP problem

In the linear case, we reduced a conjunction of formulas of a form (6.6) to an LP problem using the Farkas lemma. In the polynomial case, we reduce it to a *semidefinite programming* (SDP) problem using a theorem called *Positivstellensatz*.

There exist several variants for Positivstellensatz. In [21] three of them, *Schmüdgen's Positivstellensatz*, *Putinar's Positivstellensatz* and *Handelman's Positivstellensatz*, were used for the synthesis of ranking supermartingales. They induce different algorithms from each other. In this thesis, we use the former two. Use of Handelman's Positivstellensatz is left as future work.

We now sketch the reduction to an SDP problem following [21]. Here the notion of *sum of square* is important.

**Definition 6.2.4** (SOS). A polynomial expression  $\mathfrak{h}$  over  $V_\Gamma$  is *sum of square* (SOS) if it has a form  $\mathfrak{h} = \sum_{i=1}^s \mathfrak{i}_i^2$  where each  $\mathfrak{i}_i$  is a polynomial expression.

We write  $\text{SOS}(V_\Gamma)$  for the set of SOS polynomial expressions over  $V_\Gamma$ .

We can easily see that if  $\mathfrak{h}$  is SOS then  $\llbracket \mathfrak{h} \rrbracket(\lambda) \geq 0$  for each  $\lambda : V_\Gamma \rightarrow \mathbb{R}$ . Hence for each  $\chi : P_\Gamma \rightarrow \mathbb{R}$  and  $\lambda : V_\Gamma \rightarrow \mathbb{R}$ , we have

$$\lambda \in \llbracket \mathfrak{d}_1 \geq 0 \wedge \dots \wedge \mathfrak{d}_k \geq 0 \rrbracket \Rightarrow \llbracket \chi(\mathfrak{e}) \rrbracket(\lambda) \geq 0 \quad (6.12)$$

if either of the following conditions is satisfied:

$$\exists (\mathfrak{h}_i \in \text{SOS}(V_\Gamma))_{i \in \{1, \dots, k\}} \cdot \mathfrak{e} = \mathfrak{h}_0 + \sum_{i \in \{1, \dots, k\}} \mathfrak{h}_i \mathfrak{d}_i, \quad \text{or} \quad (6.13)$$

$$\exists (\mathfrak{h}_{w_1, \dots, w_k} \in \text{SOS}(V_\Gamma))_{w_1, \dots, w_k \in \{0, 1\}} \cdot \mathfrak{e} = \sum_{w_1, \dots, w_k \in \{0, 1\}} \mathfrak{h}_{w_1, \dots, w_k} \prod_{i=1}^k \mathfrak{d}_i^{w_i}. \quad (6.14)$$

Recall that in the linear case, completeness was partially ensured by the Farkas lemma. In the polynomial case, the role is played by Positivstellensatz's.

**Theorem 6.2.5** (Positivstellensatz's).

1. (Putinar, [87]) If  $\llbracket \mathfrak{d}_i \geq 0 \rrbracket$  is compact for some  $i$ , then  $\lambda \in \llbracket \mathfrak{d}_1 \geq 0 \wedge \dots \wedge \mathfrak{d}_k \geq 0 \rrbracket \Rightarrow \llbracket \chi(\mathfrak{e}) \rrbracket(\lambda) > 0$  implies (6.13).
2. (Schmüdgen, [93]) If  $\llbracket \mathfrak{d}_1 \geq 0 \wedge \dots \wedge \mathfrak{d}_k \geq 0 \rrbracket$  is compact, then  $\lambda \in \llbracket \mathfrak{d}_1 \geq 0 \wedge \dots \wedge \mathfrak{d}_k \geq 0 \rrbracket \Rightarrow \llbracket \chi(\mathfrak{e}) \rrbracket(\lambda) > 0$  implies (6.14).  $\square$

Reducing positivity to sum-of-square is useful in implementation as the condition that a polynomial is SOS is representable using positive semidefiniteness.

**Proposition 6.2.6** (see e.g. [54]). Let  $\mathfrak{h}$  be a polynomial expression whose degree is no greater than  $2m$ . The following are equivalent.

- $\mathfrak{h}$  is SOS.
- There exists a positive semidefinite matrix  $\mathbf{A} \in \mathbb{R}^{\mathcal{M}_{\leq m} \times \mathcal{M}_{\leq m}}$  such that  $\mathfrak{h} = \mathbf{y}_{\leq m}^T \mathbf{A} \mathbf{y}_{\leq m}$ , where  $\mathbf{y}_{\leq m}$  is a vector whose components consist of all the elements of  $\mathcal{M}_{\leq m}$ .  $\square$

Hence we can obtain the following two algorithms that synthesize  $\chi : P_\Gamma \rightarrow \mathbb{R}$  so that  $\mathfrak{t}_\chi$  is a  $\gamma$ -scaled submartingale.

**Algorithm 1** Turn the pCFG to a conjunction of formulas of a form (6.12). Fix a maximum degree  $m \in \mathbb{N}$  for SOS polynomial expressions. For each a formula of a form (6.12), fix a set  $U := \{\alpha_{t,i,j} \mid t \in \{0, 1, \dots, k\}, 1 \leq i \leq |\mathcal{M}_{\leq k}|, 1 \leq j \leq |\mathcal{M}_{\leq k}|\}$  of new parameters, and collect linear equalities over  $P_\Gamma \cup U$  by comparing coefficients of an equality  $\mathbf{e} = \mathbf{y}_m^T \mathbf{A}_0 \mathbf{y}_m + \sum_{t \in \{1, \dots, k\}} \mathbf{y}_m^T \mathbf{A}_t \mathbf{y}_m \cdot \mathbf{d}_t$ , where  $\mathbf{A}_t$  is a  $|\mathcal{M}_{\leq m}| \times |\mathcal{M}_{\leq m}|$ -matrix whose  $(i, j)$ -element is  $\alpha_{t,i,j}$ . We then ask the SDP solver to maximize  $\llbracket \mathbf{t}_\chi \rrbracket(l_{\text{init}}, \lambda_{\text{init}})$  under the collected conditions, assuming that each  $\mathbf{A}_t^l$  is positive semidefinite.

**Algorithm 2** Turn the pCFG to a conjunction of formulas of a form (6.12). Fix a maximum degree  $m \in \mathbb{N}$  for SOS polynomial expressions. For each a formula of a form (6.12), fix a set  $U := \{\alpha_{w,i,j} \mid w \in \{0, 1\}^k, 1 \leq i \leq |\mathcal{M}_{\leq k}|, 1 \leq j \leq |\mathcal{M}_{\leq k}|\}$  of new parameters, and collect linear equalities over  $P_\Gamma \cup U$  by comparing coefficients of an equality  $\mathbf{e} = \sum_{w=w_1 \dots w_k \in \{0,1\}^k} \mathbf{y}_m^T \mathbf{A}_w \mathbf{y}_m \cdot \prod_{j=1}^k \mathfrak{d}_j^{w_j}$ , where  $\mathbf{A}_w$  is a  $|\mathcal{M}_{\leq m}| \times |\mathcal{M}_{\leq m}|$ -matrix whose  $(i, j)$ -element is  $\alpha_{w,i,j}$ . We then ask the SDP solver to maximize  $\llbracket \mathbf{t}_\chi \rrbracket(l_{\text{init}}, \lambda_{\text{init}})$  under the collected conditions, assuming that each  $\mathbf{A}_w^l$  is positive semidefinite.

These two algorithms have weak points and strong points.

A weak point of Algorithm 1 would be that the condition where the corresponding Positivstellensatz holds is weaker (see Theorem 6.2.5). Note that (6.13) is a special case of (6.14). Hence Algorithm 2 always gives tighter bound for the reachability probability than Algorithm 1.

A weak point of Algorithm 2 would be that the size of a family  $(\mathfrak{h}_{w_1, \dots, w_k})_{w_1, \dots, w_k \in \{0,1\}^k}$  in (6.14) can be easily large. Note that  $k$  is the number of literals in the premise, and it is mainly determined by an invariant  $\mathfrak{J}$  and a terminal configuration  $\mathfrak{T}$ . For example, suppose that the program has  $n$  variables  $x_1, \dots, x_n$  all of which are assumed to be in  $[0, 1]$ . If we naively specify the assumption as an invariant, then it would result in a conjunction of  $2n$  inequalities  $0 \leq x_1, x_1 \leq 1, \dots, 0 \leq x_n, x_n \leq 1$ . This means that the family  $(\mathfrak{h}_{w_1, \dots, w_k})_{w_1, \dots, w_k \in \{0,1\}^k}$  has a size  $2^{2n}$  at least.

To summarize, there is a trade-off between speed and quality.

#### 6.2.4 Improvement of Polynomial Template-based Algorithm

The last section was almost a review of existing algorithms for ranking supermartingale. In this section, we give new results: we will present an improvement of the algorithms that is peculiar to  $\gamma$ -scaled submartingale.

In the algorithms given in the previous sections,  $\gamma$  was regarded as a constant and the user was required to provide it to the algorithm. As we have discussed in Section 5.5.2, the bigger value we fix  $\gamma$  to, the better bound we can obtain. We cannot fix  $\gamma$  to 1, but it is allowed to let  $\gamma \rightarrow 1$  after symbolically synthesizing a  $\gamma$ -scaled submartingale. This observation encourages us to modify the polynomial template-based algorithm so that it regards  $\gamma$  as a *variable* (such as  $x \in V_\Gamma$ ) that ranges over  $[\gamma_{\min}, 1)$  for some  $\gamma_{\min} < 1$ . Note that this is difficult in the linear case because the resulting inequalities are no longer linear.

The synthesized  $\gamma$ -scaled submartingale is a polynomial function and hence continuous. Therefore letting  $\gamma \rightarrow 1$  is equivalent to letting  $\gamma = 1$ . We can achieve this task by defining the initial valuation  $\lambda_{\text{init}} : V_\Gamma \cup \{\gamma\} \rightarrow \mathbb{R}$  so that  $\lambda_{\text{init}}(\gamma) = 1$ .

To summarize, we can modify the polynomial template-based algorithm in the previous section as follows:

- We add  $\gamma$  to  $V_\Gamma$ .

- When collecting conditions on parameters, we regard  $\gamma$  as a variable instead of a constant. Moreover, we add  $\gamma_{\min} \leq \gamma$  and  $\gamma \leq 1$  to the premises of collected formulas of the form (6.6).
- update the initial valuation  $\lambda_{\text{init}}$  so that  $\lambda_{\text{init}}(\gamma) = 1$ .

We can easily see that the probability bound calculated by the above new algorithm is no worse than the bound calculated by the algorithm in the previous section letting  $\gamma = \gamma_{\min}$ . However, we should also note that this modification increases the number of variables and hence can slow down the algorithm: we can again see a speed-quality trade-off.

### 6.2.5 Implementation

Similarly to the linear case, we have augmented the syntax of PPPs so that we can designate an invariant and a terminal configuration. In the polynomial case, they are allowed to be polynomial expressions. By modifying an implementation of a polynomial template-based algorithm for nonnegative repulsing supermartingales in [102], we have obtained a program that executes one of the following tasks depending on the provided option:

- take a PPP, the maximum degrees of a template and SOS polynomials and  $\gamma \in [0, 1)$  as inputs, and by constructing a  $\gamma$ -scaled submartingale from a polynomial template, outputs input to MATLAB [79]. If we feed the output to MATLAB, then it calculates  $p \in [0, 1]$  such that  $p \leq \text{Reach}_{\mathcal{T}}(l_{\text{init}}, \vec{0})$  using an external SDP solver.
- take a PPP, the maximum degrees of a template and SOS polynomials and  $\gamma_{\min} \in [0, 1)$  as inputs, and by constructing a  $\gamma$ -scaled submartingale from a polynomial template regarding  $\gamma$  as a variable ranging over  $[\gamma_{\min}, 1]$ , outputs an input to MATLAB. If we feed the output to MATLAB, it calculates  $p \in [0, 1]$  such that  $p \leq \text{Reach}_{\mathcal{T}}(l_{\text{init}}, \vec{0})$  using an SDP solver.

We have also made use of a MATLAB toolbox called *SOSTOOLS* [83] (version 3.03) in the implementation. It helps us to turn SOS equalities ((6.13) or (6.14)) to an SDP problem.

We have modified some codes of *SOSTOOLS* for the sake of speedup. We found that *SOSTOOLS* is very slow in adding a new constraint to a problem, and one of its reason was a procedure to substitute all the variables in a term with 0. In the original code, the procedure tries to substitute all the variables that have been declared so far, but we modified it so that it only substitutes variables appearing in the input term.

### 6.2.6 Experiments

We have used SeDuMi [98] (version 1.32) as an SDP solver. The version and the release number of MATLAB were 9.4.0.813654 (R2018a). The experiments were conducted on a MacBook Pro laptop with a Core i5 processor (2.6 GHz, 2 cores) and 16 GB RAM.

We tested our implementation for bounded random walks that were used in the linear case (Figure 6.1). Recall that we can choose: i) a Positivstellensatz between Putinar's one or Schmüdgen's one, and ii) how to deal with  $\gamma$  between regarding it as a constant or regarding it as a variable. Hence we have four choices in total. We have fixed  $\gamma$  and  $\gamma_{\min}$  to 0.99, and we fixed the degrees of a template

		Putinar				Schmüdgen				true prob.
		$\gamma$ :constant		$\gamma$ :variable		$\gamma$ :constant		$\gamma$ :variable		
param.		time (s)	bound	time (s)	bound	time (s)	bound	time (s)	bound	
1	$n = 10$ $p = 0.1$	29.650	$\geq 0.4616$	133.254	$\geq 0.9998$	92.633	$\geq 0.9866$	711.143	$\geq \mathbf{1.0}$	$1 - 1.3127 \times 10^{-86}$
	$n = 90$ $p = 0.1$	31.267	$\geq \mathbf{1.0}$	116.936	$\geq \mathbf{1.0}$	94.035	$\geq \mathbf{1.0}$	693.341	$\geq \mathbf{1.0}$	$1 - 2.8680 \times 10^{-10}$
	$n = 10$ $p = 0.9$	29.745	$\geq \mathbf{0.7389}$	105.798	$\geq \mathbf{0.9999}$	92.046	$\geq \mathbf{0.9998}$	730.187	$\geq \mathbf{1.0}$	$2.8680 \times 10^{-10}$
	$n = 50$ $p = 0.5$	28.736	$\geq \mathbf{1.0}$	108.437	$\geq \mathbf{1.0}$	96.824	$\geq \mathbf{1.0}$	799.545	$\geq \mathbf{1.0}$	0.5

Table 6.3: experimental results for the polynomial template-based algorithm

and SOS polynomials to 2. The results are shown in Table 6.3. As in the linear case, the true termination probabilities are also shown.

We can see that our implementation returns unsound probability bounds for many cases (they are written with bold letters in the table). It seems that these incorrect bounds are due to numerical errors caused by the SDP solver. Hence we cannot use the implementation as it is.

### 6.2.7 (Failed) Attempt to Remedy the Situation

Although they are very fast, SDP solvers are not suitable for verifications compared to SMT solvers because of numerical errors. Nevertheless, there exist several studies for using SDP solvers in verifications. One of them is found in a series of papers by Jansson and his coauthors [62, 64]. The techniques introduced in those papers resulted in a tool named VSDP [63, 118]. It is a wrapper of existing SDP solvers, which is written in MATLAB using INTLAB [91]. It processes an output of an SDP solver and calculates rigorous error bounds of the optimal value of the objective function.

We have modified SOSTOOLS so that it calls SDP solvers via VSDP, and tested it for the bounded random walks used in the previous section. However, it did not work well again: the lower bounds of the termination probabilities calculated by considering error bounds using VSDP were all  $-\infty$ , which is trivial.

## 6.3 Conclusion and Related Work

We have developed algorithms for synthesizing  $\gamma$ -scaled submartingales for probabilistic programs following existing synthesis algorithms for ranking supermartingales. We considered two algorithms: linear template-based one and polynomial template-based one. We have implemented the algorithms and conducted several experiments. While the linear template-based algorithm exhibited its applicability, the polynomial template-based algorithm did not work well because of numerical errors.

**Related Work** We have reviewed a notion of *ranking supermartingale* in Definition 6.0.1. The ranking supermartingale-based reachability analysis is studied in [5, 19, 21, 22, 9]. We have shown that distribution-valued ranking functions and  $\gamma$ -scaled submartingales satisfy completeness (see Theorem 5.5.8 and Proposition 5.5.15). It is known that completeness also holds for ranking supermartingales in the following sense: if a PTS (without nondeterminism) exhibits strongly almost-sure termination (i.e. the expected number of steps to an accepting state is finite), then there exists an additive ranking supermartingale that proves it [86, 100].

In the context of probabilistic systems, the Büchi condition is sometimes called *recurrence property*. Similarly, the co-Büchi condition (i.e. accepting states are visited only finitely many times) is called *persistence property*. It is known that ranking supermartingales can be also used for proving almost-sure recurrence or persistence of probabilistic systems [20].



# Chapter 7

## Related Work

We have already reviewed much related work in the chapters so far. Here we discuss related work that were omitted in individual chapters.

**Applications of “Generalize-and-concretize” Framework** Throughout this thesis, we have followed the categorical “generalize-and-concretize” framework. This framework is not new: as we have mentioned in Section 1.2, a conventional notion of forward and backward simulation [74] was categorically generalized in [43] and concretized for weighted automata in [111]. We shall give two more examples where we can find this framework.

One of the most popular examples would be *bisimulation* (see e.g. [12]). As its name suggests, a bisimulation is understood as a “bidirectional simulation.” While the existence of a simulation implies behavioral inclusion between transition systems, that of a bisimulation implies *language equivalence*. A bisimulation notion was originally defined for *CCS* (*calculus of communicating systems*, a kind of process algebra) [84, 81], and its variant was defined for various systems including nondeterministic automata [71] and (discrete-state) probabilistic transition systems [73]. A categorical generalization of bisimulation is found in [1], and it was used in [29] to define a bisimulation notion for *continuous-state* probabilistic transition systems. Another categorical generalization of bisimulation is in [51]. A relationship between multiple categorical bisimulation notions is studied in [99].

Another example is *Scott domain* [97]. A Scott domain is a special poset satisfying certain properties. It is one of the most common *domains* in *domain theory*, where the basis for defining denotational semantics for programs are studied. Studies of categorical generalizations of Scott domain are found in [2, 107, 68, 18]. In [18], the developed framework is said to constitute a basis for extending domain theory for concurrency. In [56], in addition to the possibility of obtaining a new variant by concretization, several other reasons for generalizing domain theory are discussed.

**Predicate Transformer Semantics** We have used different frameworks for categorically characterizing behaviors of systems in Chapters 3–4 (Definition 3.3.1) and Chapters 5–6 (Definition 5.2.2). However, if we focus on the fact that the coalgebra  $J\zeta^F : \nu F \rightarrow F\nu F$  in Definition 3.6.3 is an isomorphism and hence invertible, we can see not only  $[\mu\sigma]_c : X \rightarrow \Omega$  in Definition 5.2.2 but also  $[\text{tr}_1^B(c), \text{tr}_2^B(c)] : X \rightarrow \nu F$  in Definition 3.3.1 can be characterized as a fixed point of a function  $\Phi_{c,a}$  (Definition 2.4.22) for a suitable algebra  $a$ .

If we regard an arrow of a type  $X \rightarrow \nu F$  or  $X \rightarrow \Omega$  as a (multi-valued) predicate over  $X$ , then  $\Phi_{c,a}$  can be thought of as a *transformer* that transforms

predicates in a backward manner. A technique to define semantics of a program as such a predicate transformer is known as *predicate transformer semantics* [77]. Categorical studies of predicate transformer semantics are found in [37, 82, 78, 57]. We can also define semantics of a program as a *state transformer*, which transforms (a superposition of) states in a forward manner. A categorical relationship between state- and predicate-transformer semantics is studied in [45, 61].

**Various Probabilistic Transition Systems** In Chapters 3–4, we have considered PBTAs whose transition function has a type  $X \rightarrow \mathcal{G} \prod_{n \in \omega} \Sigma_n \times X^n$  (see Definition 2.2.12). In Chapters 5–6, we considered PTSs whose transition function has a type  $X \rightarrow \mathcal{G}X \times \{0, 1\}$  (see Definition 2.2.26 and Example 2.4.14). Both of them can be regarded as *generative pure* probabilistic systems: in each state, the system randomly chooses a next transition without taking any input.

Of course, there exist many other types of probabilistic transition systems. For example, a probabilistic system with a transition type  $X \rightarrow (\mathcal{G}X)^A$  can be thought of as a *reactive* probabilistic automaton: in each state, the system takes an input from the alphabet  $A$ , and randomly choose a successor state according to the input. Another example is a system with a transition function  $X \rightarrow \mathcal{P}\mathcal{G}X$ . It can be regarded as a system that exhibits both probabilistic and nondeterministic branching (such a system is often called *Markov decision process*, see e.g. [12]).

While the notions that we have induced aim at generative probabilistic systems, there exist many verification methods aiming at other types of probabilistic systems. In [66] simulation notions are introduced for probabilistic systems. It is introduced for systems with a transition type  $X \rightarrow \mathcal{D}X \times 2^A$ . A probabilistic bisimulation notion in [73] aims at *reactive* probabilistic systems. The notions of ranking supermartingale (Definition 6.0.1) and nonnegative repulsing supermartingale (Definition 6.1.14) were originally defined for Markov decision processes. As we have remarked in the footnote of page 116, it is known that  $\gamma$ -scaled submartingale can be also generalized for Markov decision processes. A comprehensive study of various types of probabilistic transition systems and hierarchy between them is found in [13].

## Chapter 8

### Conclusion and Future Work

We have categorically generalized notions of fair simulation and ranking function.

For the former, we had to categorically characterize behaviors of Büchi automata first (Chapter 3). We gave two categorical characterizations—the *logical fixed point-based* one and the *categorical fixed point-based* one. They differ in how to categorically involve the notion of *alternating fixed point*, which is known to be strongly related to Büchi and parity automata. The logical fixed point-based characterization regards a homomorphism as a fixed point and considers an alternating fixed point in a homset by assuming an order on the homset. In contrast, the categorical fixed point-based characterization refers to a well-known analogy between categories and posets and considers a sort of an alternating fixed point of a functor. We have proved that the latter characterization induces the former and hence they can be thought of as essentially the same characterization.

In Chapter 5, using the developed categorical characterization of Büchi automata, we have categorically generalized fair simulation, a simulation notion for Büchi automata. We have introduced two categorical fair simulation notions: one *with* dividing and one *without* dividing. The latter is more practical but we need more axioms than the former does to prove its soundness categorically. We then concretized them for probabilistic systems. Categorical fair simulation with dividing resulted in a simulation notion that is sound for general probabilistic Büchi tree automata while one without dividing produced a notion that is sound only for *finite-state* probabilistic Büchi *word* automata.

For categorically generalizing ranking function, we have used existing characterization for capturing behaviors of systems (Chapter 5). The key in the generalization was the categorical notion of *corecursive algebra*. Intuitively, its role was to merge the least and the greatest fixed points into one unique fixed point so that a categorical ranking function, which is defined as a post-fixed point of a certain function, can underapproximate the reachability, which is characterized as the least fixed point. We then concretized them for probabilistic systems and induced several new notions for probabilistic transition systems: *distribution-valued ranking function* and  *$\gamma$ -scaled submartingale*.

For  $\gamma$ -scaled submartingales, we gave a synthesis algorithm for probabilistic programs (Chapter 6). We found that existing linear and polynomial template-based synthesis algorithms in [19, 23, 21] for ranking supermartingales, yet another ranking function-like notion for probabilistic systems, can be easily adapted for our setting. We have implemented them and tested them for several probabilistic programs. While the linear template-based algorithm achieved a certain result, the polynomial template-based one did not work seemingly because of numerical errors.

**Future Work** Applying the “generalize-and-concretize” approach used throughout this thesis for other verification notions is one main possible direction of future work. One common scenario of the “generalize-and-concretize” approach is, as we have done in this thesis, to transfer a qualitative existing method to a quantitative method. Another possible scenario would be to first unify “seemingly similar” notions in different fields of computer science using category theory and extend the notions using the categorical characterization.

Here is an example. We have seen that there exists a notion of ranking function that can be used to prove termination of nondeterministic transition systems. A seemingly similar notion called *Lyapunov function* is known for ordinary differential equations (ODEs). We can use a Lyapunov function for proving the *stability* of ODEs. If we succeeded in categorically unifying definitions of ranking function and Lyapunov function, then it might be possible to obtain counterparts of progress measures (Definition 4.3.7), ranking supermartingales (Definition 6.0.1) or  $\gamma$ -scaled submartingales (Definition 5.5.11) for ODEs.

When we are generalizing or unifying existing notions, the theoretical basis developed in this thesis in the course of generalization might be helpful. One of the candidates is the categorical fixed point-based characterization of the parity condition, which was used only for proving the correctness of the logical fixed point-based characterization. We can aim at investigating another usage of it. For example, as we have done with the logical fixed point-based characterization, we can use the categorical fixed point-based characterization for categorically generalizing existing verification techniques. A candidate is *delayed simulation* [32, 36], a yet another simulation notion for Büchi automata that is known to be useful for state-space reduction of Büchi automata.

We gave no implementation for probabilistic fair simulation or distribution-valued ranking function. Their algorithms and implementations are future work.

In Chapters 3–4, we have focused on systems with simple branching types. Extending this for systems with more complicated branching types like two-player games, systems including both probabilistic branching and demonic nondeterminism (*Markov decision process*), or ones including probabilistic branching and both angelic and demonic nondeterminism (sometimes called  $2\frac{1}{2}$ -*player game*) would be interesting. Similarly, we can consider extending Chapters 5–6 to Markov decision processes or  $2\frac{1}{2}$ -player games.

We can possibly combine the frameworks in Chapter 3 and Chapters 5–6. As we have noted in the previous chapter, a ranking supermartingale (Definition 6.0.1) is known to be also useful for proving almost-sure recurrence (i.e. the Büchi condition) and persistence (i.e. the coBüchi condition) [20]. Extending our categorical framework so that it induces a technique for underapproximating the recurrence or persistence probability would be interesting.

In Chapter 4, we have focused on simulation between Büchi automata. Extending it for parity automata would be challenging, partly because a fair simulation between parity automata is not representable as a parity game.

We are also interested in the decidability and the complexity of probabilistic fair simulations. As the problem of determining the winner of a finite-state parity game is decidable and in  $\text{NP} \cap \text{co-NP}$  (see e.g. [67]), a fair simulation between NBTAs is also decidable and in  $\text{NP} \cap \text{co-NP}$ . We wish to study the same thing for the probabilistic case.

We mentioned in Chapter 7 that there exist three related notions to corecursive algebra: well-founded coalgebra, recursive coalgebra, and anti-founded algebra. Studying their relationship would be useful for extending Chapter 5.

There would be room for improving our algorithm and implementation in

Chapter 6. Our implementation of the polynomial template-based algorithm failed because of numerical errors of an SDP solver. We have tried to remedy the situation using a tool called VSDP, but it also failed. However, there exist other works that consider using SDP solvers for verifications, e.g. [89, 90]. Especially in [89], it is proposed to integrate an SDP solver with an SMT solver. It might be also possible to use other templates than linear and polynomial ones, e.g. exponential one. Another possible way to remedy the situation would be to rely on *Handelman's Positivstellensatz* instead of Schmüdgen's or Putinar's Positivstellensatz. Handelman's Positivstellensatz was used for the synthesis of ranking supermartingales in [21].

Another direction of future work would be to extend the framework in this thesis for higher-order programs. An obstacle towards this direction would be that the category **Meas** of measurable spaces and functions does not have an exponential object  $X^Y$ . A categorical framework for dealing with probabilistic higher-order functions is introduced in [52]. There, a measurable set  $X$  is replaced by a subset  $M_X \subseteq [\mathbb{R} \rightarrow X]$  of the set of functions from  $\mathbb{R}$  to  $X$ . It allows us to make use of the measurable structure over  $\mathbb{R}$ , which is well-behaved.

## References

- [1] Peter Aczel and Nax Mendler. A final coalgebra theorem. In David H. Pitt, David E. Rydeheard, Peter Dybjer, Andrew M. Pitts, and Axel Poigné, editors, *Category Theory and Computer Science*, pages 357–365, Berlin, Heidelberg, 1989. Springer Berlin Heidelberg.
- [2] Jirí Adámek. A categorical generalization of scott domains. *Mathematical Structures in Computer Science*, 7(5):419–443, 1997.
- [3] Jirí Adámek and Václav Koubek. Least fixed point of a functor. *J. Comput. Syst. Sci.*, 19(2):163–178, 1979.
- [4] Jirí Adámek, Stefan Milius, and Lawrence S. Moss. Fixed points of functors. *Journal of Logical and Algebraic Methods in Programming*, 95:41 – 81, 2018.
- [5] Sheshansh Agrawal, Krishnendu Chatterjee, and Petr Novotný. Lexicographic ranking supermartingales: an efficient approach to termination of probabilistic programs. *PACMPL*, 2(POPL):34:1–34:32, 2018.
- [6] André Arnold and Damian Niwiński. *Rudiments of  $\mu$ -Calculus*. Studies in Logic and the Foundations of Mathematics. Elsevier, Amsterdam, 2001.
- [7] Robert B. Ash. *Basic Probability Theory*. Wiley, 1970.
- [8] Robert B. Ash and Catherine A. Doléans-Dade. *Probability and Measure Theory*. Academic Press, 2 edition, 1999.
- [9] Martin Avanzini, Ugo Dal Lago, and Akihisa Yamada. On probabilistic term rewriting. In *FLOPS*, volume 10818 of *Lecture Notes in Computer Science*, pages 132–148. Springer, 2018.
- [10] Steve Awodey. *Category Theory*. Oxford Logic Guides. Oxford Univ. Press, 2006.
- [11] Christel Baier and Marcus Größer. Recognizing omega-regular languages with probabilistic automata. In *LICS*, pages 137–146. IEEE Computer Society, 2005.
- [12] Christel Baier and Joost-Pieter Katoen. *Principles of model checking*. MIT Press, 2008.
- [13] Falk Bartels, Ana Sokolova, and Erik P. de Vink. A hierarchy of probabilistic system types. *Theor. Comput. Sci.*, 327(1-2):3–22, 2004.
- [14] Francis Borceux. *Handbook of Categorical Algebra*., volume 2. Cambridge University Press, Cambridge, 11 1994.

- [15] Julius R. Büchi. On a Decision Method in Restricted Second-Order Arithmetic. In *International Congress on Logic, Methodology, and Philosophy of Science*, pages 1–11. Stanford University Press, 1962.
- [16] Venanzio Capretta, Tarmo Uustalu, and Varmo Vene. Corecursive algebras: A study of general structured corecursion. In *SBMF*, volume 5902 of *Lecture Notes in Computer Science*, pages 84–100. Springer, 2009.
- [17] Arnaud Carayol, Axel Haddad, and Olivier Serre. Randomization in automata on infinite trees. *ACM Trans. Comput. Log.*, 15(3):24:1–24:33, 2014.
- [18] Gian Luca Cattani and Glynn Winskel. Profunctors, open maps and bisimulation. *Mathematical Structures in Computer Science*, 15(3):553–614, 2005.
- [19] Aleksandar Chakarov and Sriram Sankaranarayanan. Probabilistic program analysis with martingales. In *CAV*, volume 8044 of *Lecture Notes in Computer Science*, pages 511–526. Springer, 2013.
- [20] Aleksandar Chakarov, Yuen-Lam Voronin, and Sriram Sankaranarayanan. Deductive proofs of almost sure persistence and recurrence properties. In Marsha Chechik and Jean-François Raskin, editors, *Tools and Algorithms for the Construction and Analysis of Systems - 22nd International Conference, TACAS 2016, Proceedings*, volume 9636 of *LNCS*, pages 260–279. Springer, 2016.
- [21] Krishnendu Chatterjee, Hongfei Fu, and Amir Kafshdar Goharshady. Termination analysis of probabilistic programs through positivstellensatz’s. In *CAV (1)*, volume 9779 of *Lecture Notes in Computer Science*, pages 3–22. Springer, 2016.
- [22] Krishnendu Chatterjee, Hongfei Fu, Petr Novotný, and Rouzbeh Hasheminezhad. Algorithmic analysis of qualitative and quantitative termination problems for affine probabilistic programs. In *POPL*, pages 327–342. ACM, 2016.
- [23] Krishnendu Chatterjee, Petr Novotný, and Dorde Zikelic. Stochastic invariants for probabilistic termination. In Giuseppe Castagna and Andrew D. Gordon, editors, *Proc. of POPL 2017*, pages 145–160. ACM, 2017.
- [24] Vincenzo Ciancia and Yde Venema. Stream automata are coalgebras. In *CMCS*, volume 7399 of *Lecture Notes in Computer Science*, pages 90–108. Springer, 2012.
- [25] Corina Cîrstea. From branching to linear time, coalgebraically. In *FICS*, volume 126 of *EPTCS*, pages 11–27, 2013.
- [26] Rance Cleaveland, Marion Klein, and Bernhard Steffen. Faster model checking for the modal mu-calculus. In *CAV*, volume 663 of *Lecture Notes in Computer Science*, pages 410–422. Springer, 1992.
- [27] Patrick Cousot and Radhia Cousot. Constructive versions of Tarski’s fixed point theorems. *Pacific Journal of Mathematics*, 82(1):43–57, 1979.
- [28] Fredrik Dahlqvist, Vincent Danos, Ilias Garnier, and Alexandra Silva. Borel kernels and their approximation, categorically. *CoRR*, abs/1803.02651, 2018.

- [29] E.P. de Vink and J.J.M.M. Rutten. Bisimulation for probabilistic transition systems: a coalgebraic approach. *Theoretical Computer Science*, 221(1):271 – 293, 1999.
- [30] J.L. Doob. *Measure Theory*. Graduate Texts in Mathematics. Springer New York, 1994.
- [31] E. A. Emerson and C. S. Jutla. Tree automata, mu-calculus and determinacy. In *Proceedings of the 32Nd Annual Symposium on Foundations of Computer Science*, SFCS '91, pages 368–377, Washington, DC, USA, 1991. IEEE Computer Society.
- [32] Kousha Etessami, Thomas Wilke, and Rebecca A. Schuller. Fair simulation relations, parity games, and state space reduction for Büchi automata. *SICOMP*, 34(5):1159–1175, 2005.
- [33] Berndt Farwer.  $\omega$ -Automata, pages 3–21. Springer Berlin Heidelberg, Berlin, Heidelberg, 2002.
- [34] Luis María Ferrer Fioriti and Holger Hermanns. Probabilistic termination: Soundness, completeness, and compositionality. In *POPL*, pages 489–501. ACM, 2015.
- [35] Robert W. Floyd. Assigning meanings to programs. In J.T. Schwartz, editor, *Mathematical Aspects of Computer Science*, volume 19 of *Proceedings of Symposium on Applied Mathematics*, pages 19–32, 1967.
- [36] Carsten Fritz and Thomas Wilke. Simulation relations for alternating büchi automata. *Theor. Comput. Sci.*, 338(1-3):275–314, 2005.
- [37] Paul H.B. Gardiner, Clare E. Martin, and Oege de Moor. An algebraic construction of predicate transformers. *Science of Computer Programming*, 22(1):21 – 44, 1994.
- [38] Neil Ghani, Peter Hancock, and Dirk Pattinson. Representations of stream processors using nested fixed points. *Logical Methods in Computer Science*, 5(3), 2009.
- [39] Michele Giry. A categorical approach to probability theory. In *Proc. Categorical Aspects of Topology and Analysis*, volume 915 of *Lect. Notes Math.*, pages 68–85, 1982.
- [40] The GNU linear programming kit. <http://www.gnu.org/software/glpk>.
- [41] Erich Grädel and Igor Walukiewicz. Positional determinacy of games with infinitely many priorities. *Logical Methods in Computer Science*, 2(4), 2006.
- [42] Martin Grohe, Eric Koskinen, and Natarajan Shankar, editors. *Proceedings of the 31st Annual ACM/IEEE Symposium on Logic in Computer Science, LICS '16, New York, NY, USA, July 5-8, 2016*. ACM, 2016.
- [43] Ichiro Hasuo. Generic forward and backward simulations. In Christel Baier and Holger Hermanns, editors, *CONCUR*, volume 4137 of *Lecture Notes in Computer Science*, pages 406–420. Springer, 2006.
- [44] Ichiro Hasuo. Generic forward and backward simulations II: probabilistic simulation. In *CONCUR*, volume 6269 of *Lecture Notes in Computer Science*, pages 447–461. Springer, 2010.



- [45] Ichiro Hasuo. Generic weakest precondition semantics from monads enriched with order. *Theor. Comput. Sci.*, 604:2–29, 2015.
- [46] Ichiro Hasuo and Bart Jacobs. Context-free languages via coalgebraic trace semantics. In *Proc. CALCO 2005*, volume 3629 of *LNCS*, pages 213–231. Springer, 2005.
- [47] Ichiro Hasuo, Bart Jacobs, and Ana Sokolova. Generic trace semantics via coinduction. *Logical Methods in Computer Science*, 3(4:11), 2007.
- [48] Ichiro Hasuo, Yoshinobu Kawabe, and Hideki Sakurada. Probabilistic anonymity via coalgebraic simulations. *Theor. Comput. Sci.*, 411(22–24):2239–2259, 2010.
- [49] Ichiro Hasuo, Shunsuke Shimizu, and Corina Cîrstea. Lattice-theoretic progress measures and coalgebraic model checking. In *POPL*, pages 718–732. ACM, 2016.
- [50] Thomas A. Henzinger, Orna Kupferman, and Sriram K. Rajamani. Fair simulation. *Inf. Comput.*, 173(1):64–81, 2002.
- [51] Claudio Hermida and Bart Jacobs. Structural induction and coinduction in a fibrational setting. *Inf. Comput.*, 145(2):107–152, September 1998.
- [52] Chris Heunen, Ohad Kammar, Sam Staton, and Hongseok Yang. A convenient category for higher-order probability theory. In *LICS*, pages 1–12. IEEE Computer Society, 2017.
- [53] Wataru Hino, Hiroki Kobayashi, Ichiro Hasuo, and Bart Jacobs. Healthiness from duality. In Grohe et al. [42], pages 682–691.
- [54] Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, New York, NY, USA, 2nd edition, 2012.
- [55] Jesse Hughes and Bart Jacobs. Simulations in coalgebra. *Theor. Comput. Sci.*, 327(1-2):71–108, 2004.
- [56] Martin Hyland. Some reasons for generalising domain theory. *Mathematical Structures in Computer Science*, 20(2):239–265, 2010.
- [57] Pierre Hyvernat. Predicate transformers and linear logic, yet another denotational model. *CoRR*, abs/0905.3998, 2009.
- [58] Bart Jacobs. Trace semantics for coalgebras. *Electr. Notes Theor. Comput. Sci.*, 106:167–184, 2004.
- [59] Bart Jacobs. From coalgebraic to monoidal traces. In *Coalgebraic Methods in Computer Science (CMCS 2010)*, volume 264 of *Elect. Notes in Theor. Comp. Sci.*, pages 125–140. Elsevier, Amsterdam, 2010.
- [60] Bart Jacobs. *Introduction to Coalgebra: Towards Mathematics of States and Observation*, volume 59 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, 2016.
- [61] Bart Jacobs. A recipe for state-and-effect triangles. *Logical Methods in Computer Science*, 13(2), 2017.

- [62] Christian Jansson. Termination and verification for ill-posed semidefinite programming problems, 2005.
- [63] Christian Jansson. VSDP: A MATLAB software package for verified semidefinite programming. *NOLTA*, pages 327–330, 2006.
- [64] Christian Jansson, Denis Chaykin, and Christian Keil. Rigorous error bounds for the optimal value in semidefinite programming. *SIAM J. Numerical Analysis*, 46(1):180–200, 2007.
- [65] Thomas Jech. *Set Theory*. Springer, the third millennium edition, 2003.
- [66] Bengt Jonsson and Kim Guldstrand Larsen. Specification and refinement of probabilistic processes. In *LICS*, pages 266–277. IEEE Computer Society, 1991.
- [67] Marcin Jurdzinski. Small progress measures for solving parity games. In Horst Reichel and Sophie Tison, editors, *STACS*, volume 1770 of *Lecture Notes in Computer Science*, pages 290–301. Springer, 2000.
- [68] Panagis Karazeris. Categorical domain theory: Scott topology, powercategories, coherent categories. *Theory and Applications of Categories*, 9, 01 2002.
- [69] Joost-Pieter Katoen, Annabelle McIver, Larissa Meinicke, and Carroll C. Morgan. Linear-invariant generation for probabilistic programs: - automated support for proof-based methods. In *SAS*, volume 6337 of *Lecture Notes in Computer Science*, pages 390–406. Springer, 2010.
- [70] Stefan Kiefer, Andrzej S. Murawski, Joël Ouaknine, Björn Wachter, and James Worrell. Language equivalence for probabilistic automata. In *CAV*, volume 6806 of *Lecture Notes in Computer Science*, pages 526–540. Springer, 2011.
- [71] Dexter Kozen. *Automata and Computability*. Springer-Verlag, Berlin, Heidelberg, 1st edition, 1997.
- [72] Dexter Kozen. Kolmogorov extension, martingale convergence, and compositionality of processes. In Grohe et al. [42], pages 692–699.
- [73] Kim G. Larsen and Arne Skou. Bisimulation through probabilistic testing. *Information and Computation*, 94(1):1–28, 1991.
- [74] Nancy A. Lynch and Frits W. Vaandrager. Forward and backward simulations. I. Untimed systems. *Inf. Comput.*, 121(2):214–233, 1995.
- [75] Nancy A. Lynch and Frits W. Vaandrager. Forward and backward simulations. II. Timing based systems. *Inf. Comput.*, 128(1):1–25, 1996.
- [76] Saunders Mac Lane. *Categories for the Working Mathematician*. Springer, Berlin, 2nd edition, 1998.
- [77] Ernest. G. Manes. *Predicate Transformer Semantics*. Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, 1992.
- [78] Clare Martin. Towards a calculus of predicate transformers. In Jiří Wiedermann and Petr Hájek, editors, *Mathematical Foundations of Computer Science 1995*, pages 489–498, Berlin, Heidelberg, 1995. Springer Berlin Heidelberg.

- [79] MATLAB. <https://www.mathworks.com/products/matlab.html>.
- [80] Stefan Milius. Completely iterative algebras and completely iterative monads. *Information and Computation*, 196(1):1 – 41, 2005.
- [81] R. Milner. *A Calculus of Communicating Systems*. Springer-Verlag, Berlin, Heidelberg, 1982.
- [82] David A. Naumann. A recursion theorem for predicate transformers on inductive data types. *Inf. Process. Lett.*, 50(6):329–336, 1994.
- [83] A. Papachristodoulou, J. Anderson, G. Valmorbida, S. Prajna, P. Seiler, and P. A. Parrilo. *SOSTOOLS: Sum of squares optimization toolbox for MATLAB*. <http://arxiv.org/abs/1310.4716>, 2013. Available from <http://www.eng.ox.ac.uk/control/sostools>, <http://www.cds.caltech.edu/sostools> and <http://www.mit.edu/~parrilo/sostools>.
- [84] David Park. Concurrency and automata on infinite sequences. In Peter Deussen, editor, *Theoretical Computer Science*, pages 167–183, Berlin, Heidelberg, 1981. Springer Berlin Heidelberg.
- [85] John Power and Daniele Turi. A coalgebraic foundation for linear time semantics. *Electr. Notes Theor. Comput. Sci.*, 29:259–274, 1999.
- [86] Stephen Prajna, Ali Jadbabaie, and George J. Pappas. Stochastic safety verification using barrier certificates. In *2004 43rd IEEE Conference on Decision and Control. IEEE , Piscataway*, pages 929–934, 2004.
- [87] Mihai Putinar. Positive polynomials on compact semi-algebraic sets. *Indiana University Mathematics Journal*, 42(3):969–984, 1993.
- [88] Michael K. Reiter and Aviel D. Rubin. Crowds: Anonymity for web transactions. *ACM Trans. Inf. Syst. Secur.*, 1(1):66–92, 1998.
- [89] Pierre Roux, Mohamed Iguernlala, and Sylvain Conchon. A non-linear arithmetic procedure for control-command software verification. In *TACAS (2)*, volume 10806 of *Lecture Notes in Computer Science*, pages 132–151. Springer, 2018.
- [90] Pierre Roux, Yuen-Lam Voronin, and Sriram Sankaranarayanan. Validating numerical semidefinite programming solvers for polynomial invariants. *Formal Methods in System Design*, 53(2):286–312, 2018.
- [91] S.M. Rump. INTLAB - INTerval LABoratory. In Tibor Csendes, editor, *Developments in Reliable Computing*, pages 77–104. Kluwer Academic Publishers, Dordrecht, 1999. <http://www.ti3.tuhh.de/rump/>.
- [92] Sriram Sankaranarayanan, Aleksandar Chakarov, and Sumit Gulwani. Static analysis for probabilistic programs: inferring whole program properties from finitely many paths. In *PLDI*, pages 447–458. ACM, 2013.
- [93] Konrad Schmüdgen. Thek-moment problem for compact semi-algebraic sets. *Mathematische Annalen*, 289(1):203–206, Mar 1991.
- [94] Alexander Schrijver. *Theory of Linear and Integer Programming*. John Wiley & Sons, Inc., New York, NY, USA, 1986.

- [95] Lutz Schröder and Dirk Pattinson. PSPACE bounds for rank-1 modal logics. *ACM Trans. Comput. Log.*, 10(2), 2009.
- [96] Christoph Schubert. Terminal coalgebras for measure-polynomial functors. In Jianer Chen and S. Barry Cooper, editors, *Proc. TAMC 2009*, volume 5532 of *Lect. Not. in Comp. Sci.*, pages 325–334. Springer, 2009.
- [97] Dana S. Scott. Domains for denotational semantics. In Mogens Nielsen and Erik Meineche Schmidt, editors, *Automata, Languages and Programming*, pages 577–610, Berlin, Heidelberg, 1982. Springer Berlin Heidelberg.
- [98] SeDuMi. <http://sedumi.ie.lehigh.edu/>.
- [99] Sam Staton. Relating coalgebraic notions of bisimulation. *Logical Methods in Computer Science*, 7(1), 2011.
- [100] Jacob Steinhardt and Russ Tedrake. Finite-time regional verification of stochastic non-linear systems. *I. J. Robotics Res.*, 31(7):901–923, 2012.
- [101] Viggo Stoltenberg-Hansen, Ingrid Lindström, and Edward R. Griffor. *Mathematical Theory of Domains*. Cambridge University Press, New York, NY, USA, 1994.
- [102] Toru Takisaka, Yuichiro Oyabu, Natsuki Urabe, and Ichiro Hasuo. Ranking and repulsing supermartingales for reachability in probabilistic programs. In *ATVA*, volume 11138 of *Lecture Notes in Computer Science*, pages 476–493. Springer, 2018.
- [103] Terence Tao. *An Introduction to Measure Theory*. Graduate studies in mathematics. American Mathematical Society, 2011.
- [104] Alfred Tarski. *A Decision Method for Elementary Algebra and Geometry*. University of California Press, 1951.
- [105] Alfred Tarski. A lattice-theoretical fixpoint theorem and its applications. *Pacific Journal of Mathematics*, 5, 06 1955.
- [106] Paul Taylor. *Practical foundations of mathematics*. Cambridge studies in advanced mathematics. Cambridge University Press, Cambridge, New York (N. Y.), Melbourne, 1999.
- [107] Vera Trnková and Jiri Velebil. On categories generalizing universal domains. *Mathematical Structures in Computer Science*, 9(2):159–175, 1999.
- [108] A. M. Turing. On computable numbers, with an application to the entscheidungsproblem. *Proceedings of the London Mathematical Society*, s2-42(1):230–265, 1937.
- [109] Natsuki Urabe, Masaki Hara, and Ichiro Hasuo. Categorical liveness checking by corecursive algebras. In *LICS*, pages 1–12. IEEE Computer Society, 2017.
- [110] Natsuki Urabe and Ichiro Hasuo. Fair simulation for nondeterministic and probabilistic buchi automata: a coalgebraic perspective. *Logical Methods in Computer Science*, 13(3), 2017.
- [111] Natsuki Urabe and Ichiro Hasuo. Quantitative simulations by matrices. *Inf. Comput.*, 252:110–137, 2017.

- [112] Natsuki Urabe and Ichiro Hasuo. Categorical büchi and parity conditions via alternating fixed points of functors. In *CMCS*, volume 11202 of *Lecture Notes in Computer Science*, pages 214–234. Springer, 2018.
- [113] Natsuki Urabe and Ichiro Hasuo. Coalgebraic Infinite Traces and Kleisli Simulations. *Logical Methods in Computer Science*, Volume 14, Issue 3, September 2018.
- [114] Natsuki Urabe, Shunsuke Shimizu, and Ichiro Hasuo. Coalgebraic trace semantics for büchi and parity automata. In *CONCUR*, volume 59 of *LIPICs*, pages 24:1–24:15. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2016.
- [115] Moshe Y. Vardi and Pierre Wolper. Automata-theoretic techniques for modal logics of programs. *J. Comput. Syst. Sci.*, 32(2):183–221, 1986.
- [116] Yde Venema. Automata and fixed point logic: A coalgebraic perspective. *Inf. Comput.*, 204(4):637–678, 2006.
- [117] Thomas von Bomhard. Minimization of tree automata. BSc thesis, Universität des Saarlandes, September 2008.
- [118] VSDP. <http://www.ti3.tuhh.de/jansson/vsdp/>.