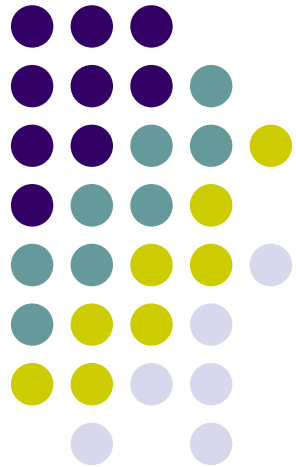
Probabilistic Anonymity via Coalgebraic Simulations

Ichiro Hasuo
Radboud Universiteit Nijmegen, the Netherlands
(From 2007.5, also at: RIMS, Kyoto Univ., Japan)

Japanese “France Telecom”

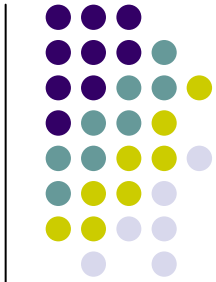
Joint work with Yoshinobu Kawabe
NTT Communication Science Laboratories, Japan



Radboud University Nijmegen

NTT

Online privacy
Online anonymity



is attracting ***growing***

- **Threats**
  - ISPs in EU are forced to keep logs of your web access
  - Your passport stores your fingerprint on a RFID chip
- **Public concerns**
- **Research interest**
  - See *Anonymity Bibliography* **`http://freehaven.net/anonbib/`**
  - The field is quite young
    - Compared to “traditional” security notions such as secrecy, authentication

Overview:

Probabilistic anonymity via coalgebraic simulations

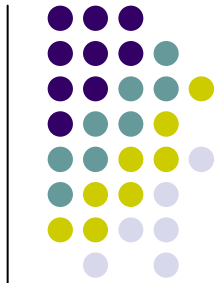
Simulation-based proof method for
non-deterministic = possibilistic
anonymity [KawabeMST06]

Generic, ***coalgebraic*** theory of traces and simulations
[IH,Jacobs,Sokolova]

- non-det. ➔ prob. is just a change of a parameter

Simulation-based proof method for
probabilistic
anonymity

For you to take home



- Probability in ***anonymity***
  - Subtlety in definition of “probabilistic anonymity”

- Power of ***categorical methods*** in computer science
  - Specifically, theory of ***coalgebras***
  - Abstraction and genericity
  - Category theory in action!

References

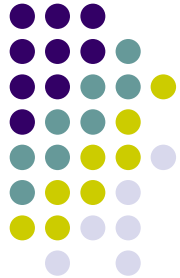
- Coalgebraic theory of traces and simulations
  - Ichiro Hasuo, Bart Jacobs and Ana Sokolova
    Generic Trace Theory
    CMCS'06, ENTCS 164
  - Ichiro Hasuo
    Generic Forward and Backward Simulations
    CONCUR 2006, LNCS 4137
- Simulation-based proof method of anonymity
  - Y. Kawabe, K. Mano, H. Sakurada and Y. Tsukada
    Theorem-proving anonymity of infinite state systems
    Information Processing Letters, to appear
  - Y. Kawabe, K. Mano, H. Sakurada and Y. Tsukada
    Backward simulations for anonymity
    WITS'06
  - Ichiro Hasuo and Yoshinobu Kawabe
    Probabilistic Anonymity via Coalgebraic Simulations
    ESOP'07, to appear

Y. Kawabe

B. Jacobs

A. Sokolova

Nina Hagen



Part I:

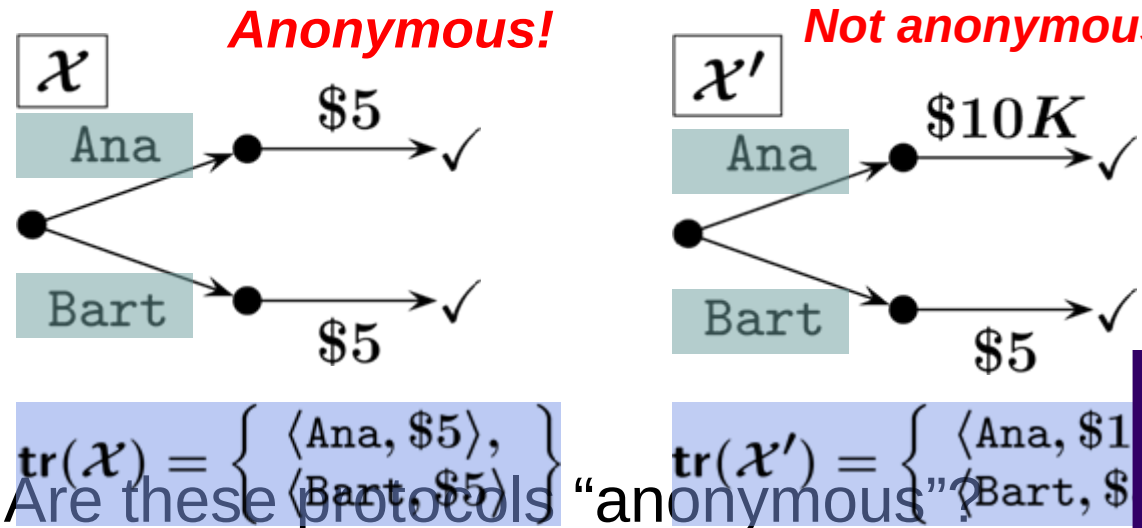
Non-deterministic version of

∃ simulation ➔ anonymity

Non-deterministic “trace” anonymity

[Schneider&Sidiropoulos,ESORICS’96]

Anonymous donation as an example



$$\mathbf{tr}(\mathcal{X}) = \{ \langle \text{Ana}, \$5 \rangle, \langle \text{Bart}, \$5 \rangle \}$$

$$\mathbf{tr}(\mathcal{X}') = \{ \langle \text{Ana}, \$1 \ldots \langle \text{Bart}, \$ \ldots \}$$

- Ana : ***actor*** action (invisible for adversary)
- $5: ***observable*** action

Observation $\vec{o} \cdot \vec{o'}$ can be attributed to anybody

Are these protocols “anonymous”?

Definition (*Trace anonymity*)

$\mathcal{X}$ is *anonymous* $\Leftrightarrow$

$$\forall \vec{o}, \vec{o'}.\ \langle \vec{o}, \text{Ana}, \vec{o'} \rangle \in \mathbf{tr}(\mathcal{X}) \implies \langle \vec{o}, \text{Bart}, \vec{o'} \rangle \in \mathbf{tr}(\mathcal{X}), \langle \vec{o}, \text{Chris}, \vec{o'} \rangle \in \mathbf{tr}(\mathcal{X}), \ldots$$

Non-deterministic (∃ simulation ➔ anonymity)

[Kawabe,Mano,Sakurada,Tsukada 2006]

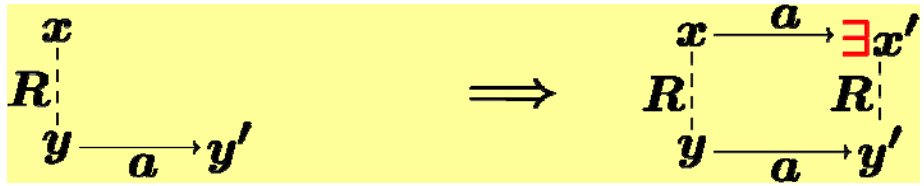
An automaton which models an anonymizing protocol

Theorem

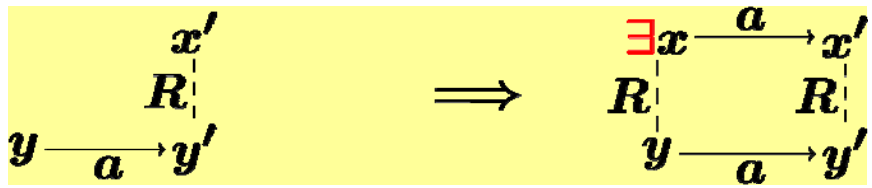
∃ forward/backward simulation from an($\mathcal{X}$) to $\mathcal{X}$

➔ $\mathcal{X}$ is anonymous

- ***Theory of traces and simulations***
  - Forward simulation ***R*** is such that:



  - Backward simulation ***R*** is such that:



  - ***Soundness theorem:***

    ∃ fwd/bwd simulation ➔ trace inclusion
  - [Lynch&Vaandrager, Inf.&Comp. 1995]

Non-deterministic
(∃ simulation ➔ anonymity)
[Kawabe,Mano,Sakurada,Tsukada 2006]

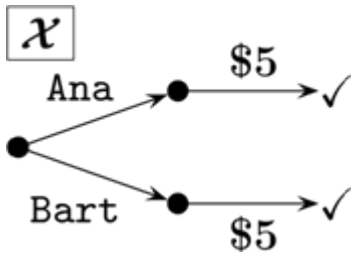
Theorem

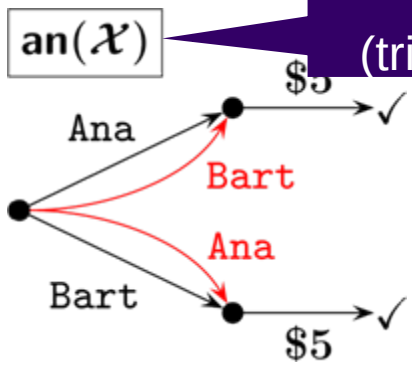
∃ forward/backward simulation from an($\mathcal{X}$) to $\mathcal{X}$

➔ $\mathcal{X}$ is anonymous

"**anonymized**" version of ***X***

(trivially anonymous)





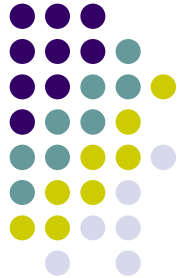
Proof.

$\mathcal{X}$ is anonymous

🡐 tr($\mathcal{X}$) = tr(an($\mathcal{X}$)) (anonymity is trace-based)

🡐 tr($\mathcal{X}$) ⊇ tr(an($\mathcal{X}$)) (⊆ is trivial)

🡐 ∃ simulation from an($\mathcal{X}$) to $\mathcal{X}$ (**soundness theorem**!)

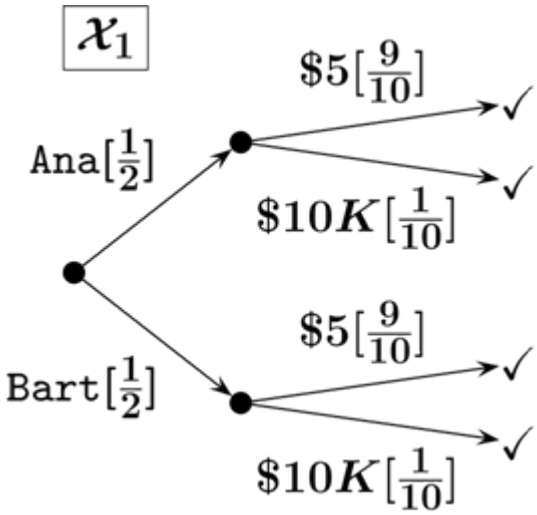


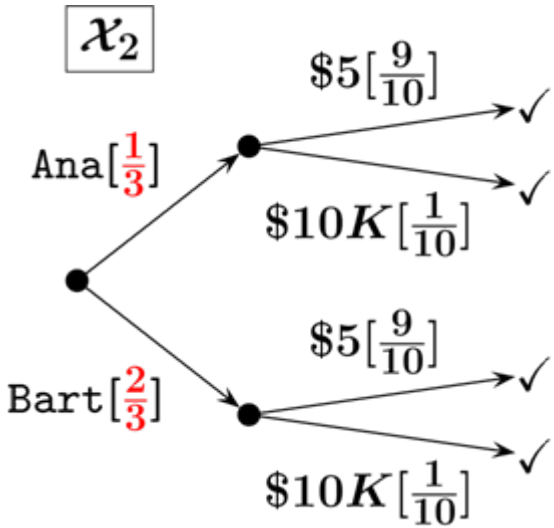
Part II:

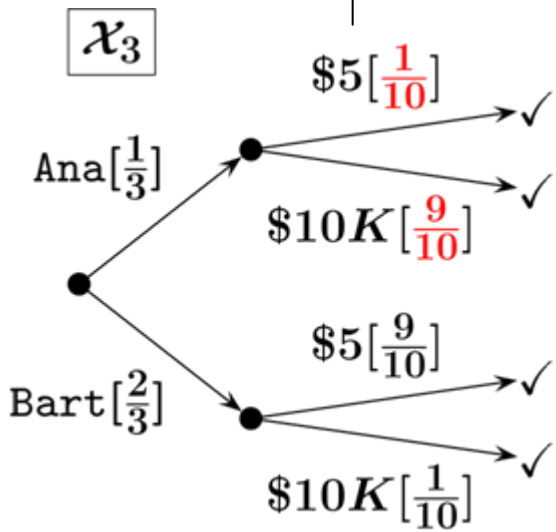
Probabilistic anonymity and

(∃ simulation ➔ anonymity)

Probabilistic anonymity







Anonymous! ***"Anonymous"!*** ***Not anonymous!***

- Are these protocols "anonymous"?
- These are all anonymous in a ***possibilistic*** sense.

  ➔ definition of "probabilistic anonymity"?

Probabilistic anonymity

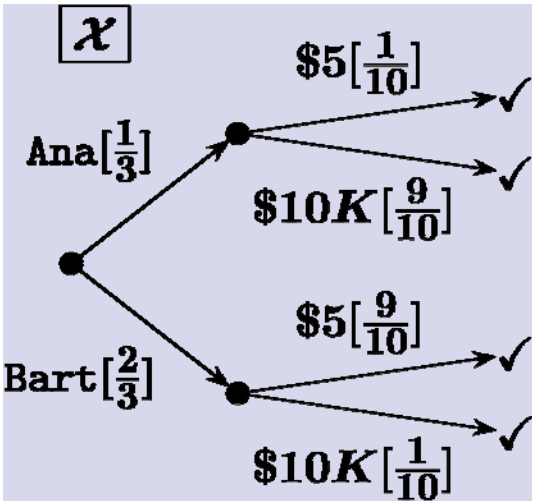
[Bhargava&Palamidessi, CONCUR'05]

Definition (Probabilistic anonymity)

$\mathcal{X}$ is *anonymous*

such as ⟨ $5 ⟩

$$\Leftrightarrow \quad \forall \vec{o}: \text{ observation.}$$
$$\textbf{Prob}(\vec{o} \mid \texttt{Ana donates}) = \textbf{Prob}(\vec{o} \mid \texttt{Bart donates}) = \textbf{Prob}(\vec{o} \mid \texttt{Chris donates}) = \cdots$$



conditional probability

$$\textbf{Prob}(\vec{o} \mid \texttt{Bart donates}) = \frac{\textbf{Prob}(\vec{o} \wedge \texttt{Bart donates})}{\textbf{Prob}(\texttt{Bart donates})}$$

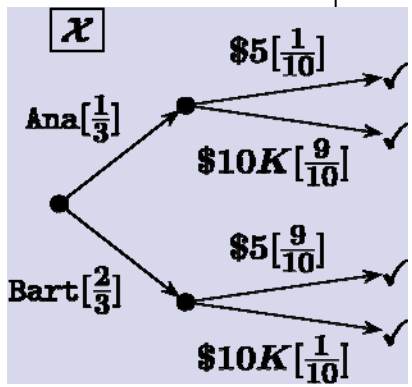
More on probabilistic anonymity

Definition (Probabilistic anonymity)

$\mathcal{X}$ is *anonymous* $\Leftrightarrow$

$\forall \vec{o}$: observation.

$\mathbf{Prob}(\vec{o} \mid \texttt{Ana donates}) = \mathbf{Prob}(\vec{o} \mid \texttt{Bart donates}) = \mathbf{Prob}(\vec{o} \mid \texttt{Chris donates}) = \cdots$



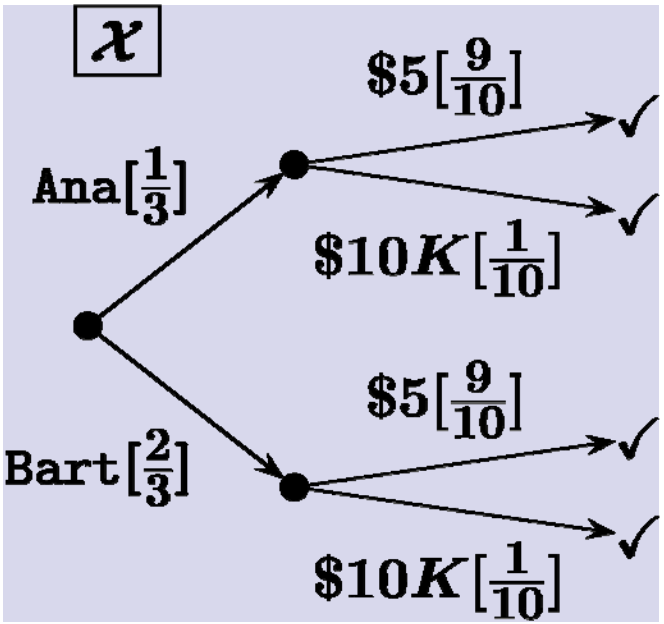
- Intuition:

  "Observation of $\vec{o}$ does not carry any info. on who is donating"

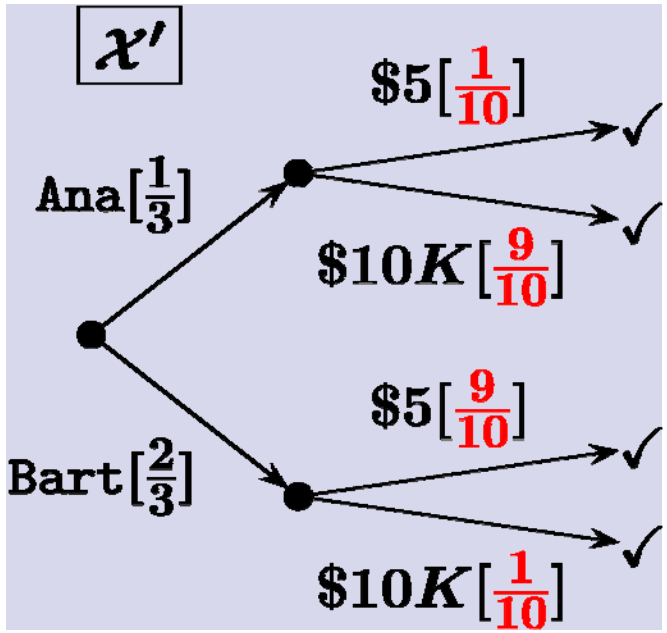
$$\begin{bmatrix} \texttt{Ana} & \mapsto & \frac{1}{3} \\ \texttt{Bart} & \mapsto & \frac{2}{3} \end{bmatrix}$$

  - **A priori** *distribution of suspicion* need not be uniform
  - However, after any observation, any agent must be exactly as suspicious as before

More on probabilistic anonymity

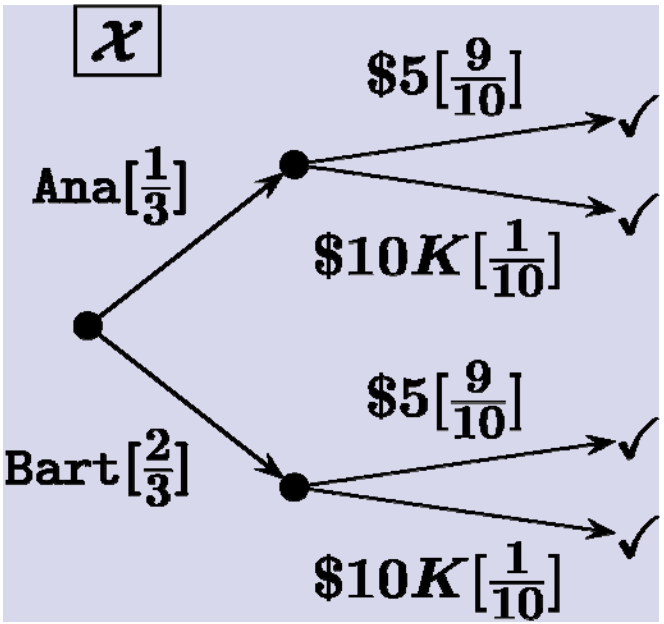


I Know Bart is more likely to donate



I Know Bart is more likely to donate

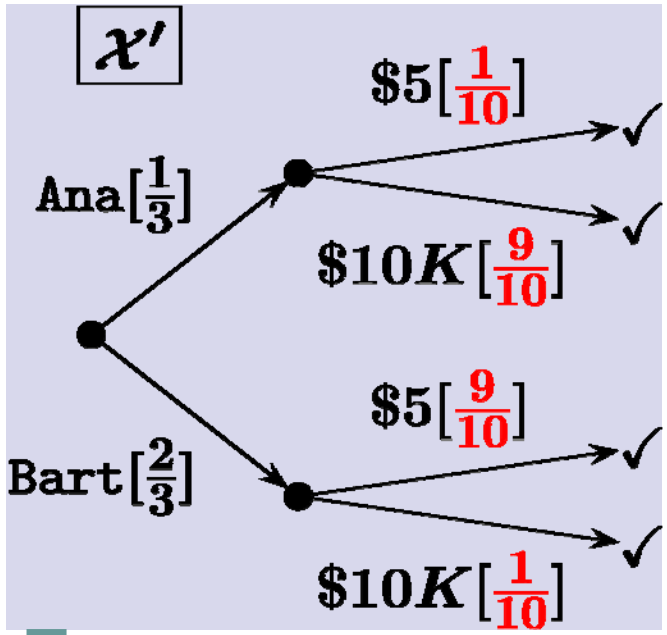
More on probabilistic anonymity



$\$10K$

Someone donated $10K, which is more likely by Bart (as I suspected)…

(anonymous)



$\$10K$

Bart wouldn't donate such an amount. This time it's likely that Ana did!

(not anonymous)

Again the same scenario: (∃ simulation ➔ anonymity)

probabilistic automaton

Theorem

∃ forward/backward simulation from an($\mathcal{X}$) to $\mathcal{X}$

➔ $\mathcal{X}$ is anonymous

Can we do this probabilistically?

- What is an(**X**)?
- What is "probabilistic simulation"?

probabilistic anonymity

Proof.

$\mathcal{X}$ is anonymous

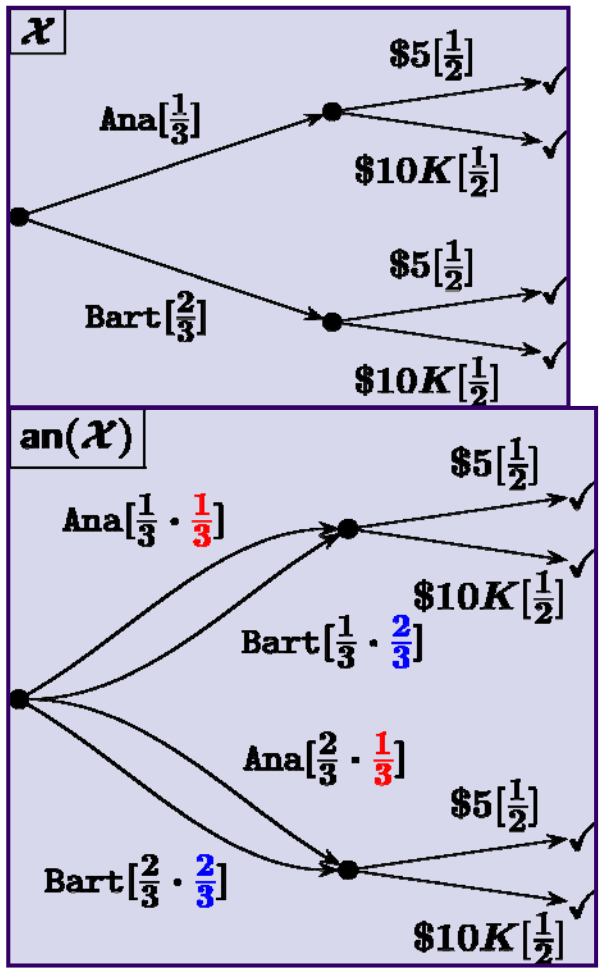
trace *distribution*

➔ tr($\mathcal{X}$) = tr(an($\mathcal{X}$)) (anonymity is

➔ tr($\mathcal{X}$) ⊇ tr(an($\mathcal{X}$)) probabilistic simulation al)

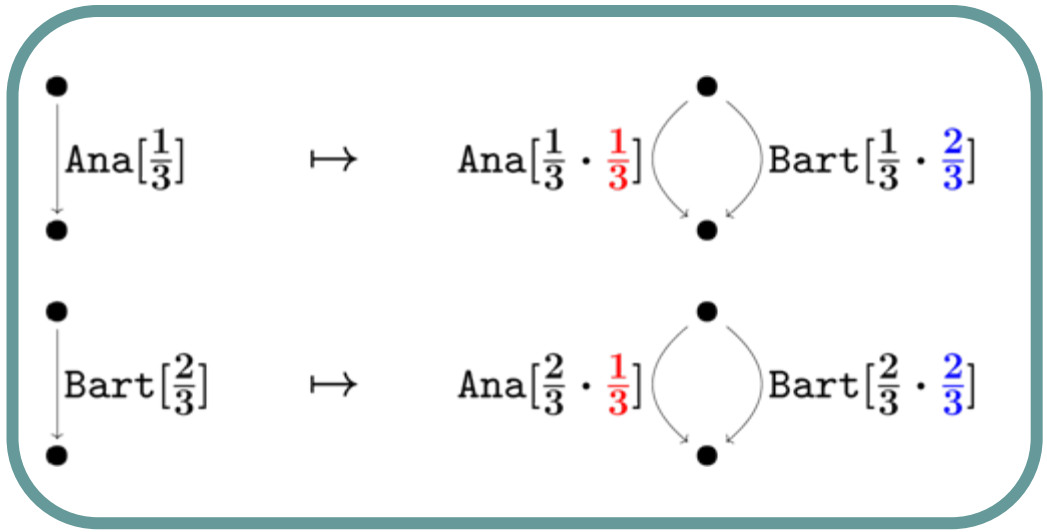
➔ ∃ simulation from an($\mathcal{X}$) to $\mathcal{X}$ (**soundness theorem**!)

an($\mathcal{X}$), in a probabilistic setting

- Idea: distribute probability according to a-priori suspicion
- 

A-priori suspicion: $\begin{bmatrix} \texttt{Ana} & \mapsto & \frac{1}{3} \\ \texttt{Bart} & \mapsto & \frac{2}{3} \end{bmatrix}$

Hence



Again the same scenario: (∃ simulation ➔ anonymity)

probabilistic automaton

Theorem

∃ forward/backward simulation from an($\mathcal{X}$) to $\mathcal{X}$

➔ $\mathcal{X}$ is anonymous

Can we do this probabilistically?

- What is an(**X**)?
- What is "probabilistic simulation"?

probabilistic anonymity

Proof.

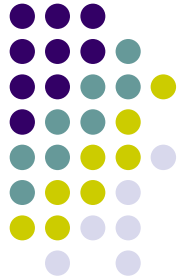
$\mathcal{X}$ is anonymous

trace *distribution*

➔ tr($\mathcal{X}$) = tr(an($\mathcal{X}$)) (anonymity is

➔ tr($\mathcal{X}$) ⊇ tr(an($\mathcal{X}$)) probabilistic simulation al)

➔ ∃ simulation from an($\mathcal{X}$) to $\mathcal{X}$ (**soundness theorem**!)



Intermezzo:

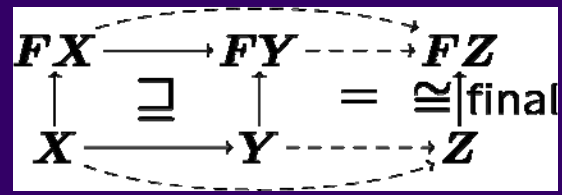
Coalgebraic theory of *traces* and *simulations*

Theory of coalgebras

- ***Categorical*** theory of "state-based systems"

  Everything as *objects* and *arrows*

  - Focus on "essence"
  - Abstraction, genericity

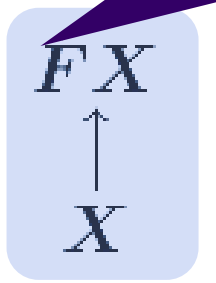
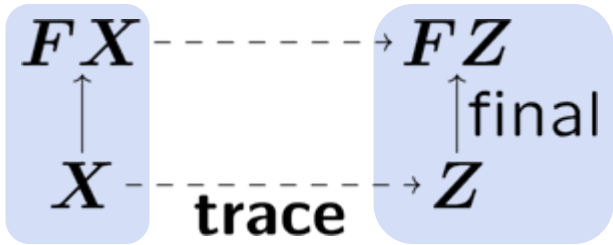
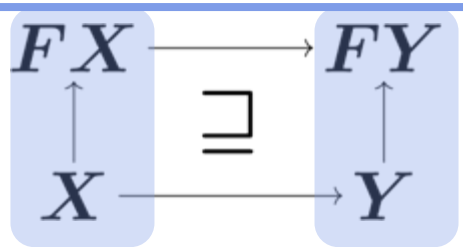
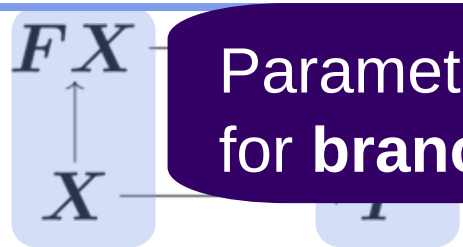


- Base categories
  - **Sets** ➔ theory of *bisimilarity* [Rutten,TCS'00]
  - **Kl(T)** ➔ theory of *traces* and *simulations*

    [IH,Jacobs,Sokolova,CMCS'06][IH,CONCUR'06]

Traces and simulations coalgebraically

Parameter F : for **transition-type**

System	a coalgebra 
Trace semantics	 in $\mathcal{K}\ell(T)$
Simulations	 **forward** sim.  **backward** sim.

Parameter T : for **branching-type**

General soundness theorem : ∃ simulation ➔ trace incl.

Genericity

T and F

By changing parameters, the framework covers

- different **branching-types** by different ***T***
  - non-determinism
  - probabilistic branching

We exploit this!

- Anything we can do in a non-det. setting,
- we can also do in a probabilistic setting

- different **transition-types** by di
  - LTS: $x \mapsto (a,x')$
  - Context-free grammar: $x \mapsto \langle$ "$\neg$", $x \rangle$

    $x \mapsto \langle x,$ "$\wedge$", $x \rangle$

Systems as coa

F: a functor
for **transition-type**

A system is: $FX \uparrow X$ in $\mathcal{K}\ell(T)$

T: a monad
for **branching-type**

- T is a **monad**, for branching-type. Examples:
  - $\mathcal{P}$, powerset monad
    - for non-determinism
  - $\mathcal{D}$, subdistribution monad
    - for (generative) probabilistic branching

$$\mathcal{D}X = \{\text{probability subdistributions on } X\} = \{d : X \to [0,1] \mid \sum_{x \in X} d(x) \leq 1\}$$

Systems as coalgebras

A system is:

$$\frac{\begin{array}{c} FX \\ \uparrow \\ X \end{array} \text{ in } \mathcal{K}\ell(T)}{TFX}$$

, a function

A category where **branching** is implicit

$\mathcal{K}\ell(T)$: the **Kleisli** category for T

- Main point:

  for T = $\mathcal{P}$,

$$\frac{X \longrightarrow Y \text{ in } \mathcal{K}\ell(\mathcal{P})}{X \longrightarrow \mathcal{P}Y \text{, a function}}$$

$$x \longmapsto \{y_1, y_2, \ldots\}$$

Parameters

A system is: $\begin{array}{c} FX \\ \uparrow \\ X \end{array}$ in $\mathcal{K}\ell(T)$

$\dfrac{}{\begin{array}{c} TFX \\ \uparrow \\ X \end{array}}$, a function

branching-type: non-determinism

transition-type: terminate or (output, next)

- T = $\mathcal{P}$, F = 1 + Σ × -

$\begin{array}{c} \mathcal{P}(1 + \Sigma \times X) \\ \uparrow \\ X \end{array}$ such as $\{\checkmark, (a_1, x_1), (a_2, x_2)\}$ (with x below)

$x \to \checkmark$
$x \xrightarrow{a_1} x_1$
$x \xrightarrow{a_2} x_2$

→ LTS with explicit termination

Parameters

A system is: $\dfrac{FX \uparrow X}{TFX \uparrow X}$ in $\mathcal{K}\ell(T)$, a function

- T = $\mathcal{P}$, F = 1 + Σ × -
  → LTS with explicit termination
- T = $\mathcal{D}$, F = 1 + Σ × -
  → Generative probabilistic system
- T = $\mathcal{P}$, F = (Σ + -)*
  → Context-free grammar

Trace semantics

[IH, Jacobs, Sokolova. CMCS'06]

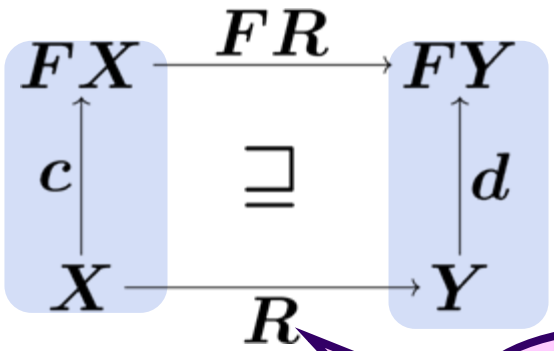
1. $\exists$ final coalgebra $\begin{array}{c} FZ \\ \cong\uparrow \\ Z \end{array}$ in $\mathcal{K}\ell(T)$

For any system $\begin{array}{c} FX \\ \uparrow \\ X \end{array}$, $\exists!\, f$ such that

$$\begin{array}{ccc} FX & \xrightarrow{Ff} & FZ \\ \uparrow & & \cong\uparrow \\ X & \xrightarrow[f]{} & Z \end{array}$$

2. This induced f gives (finite) trace semantics

Forward simulation



in $\mathcal{K}\ell(T)$

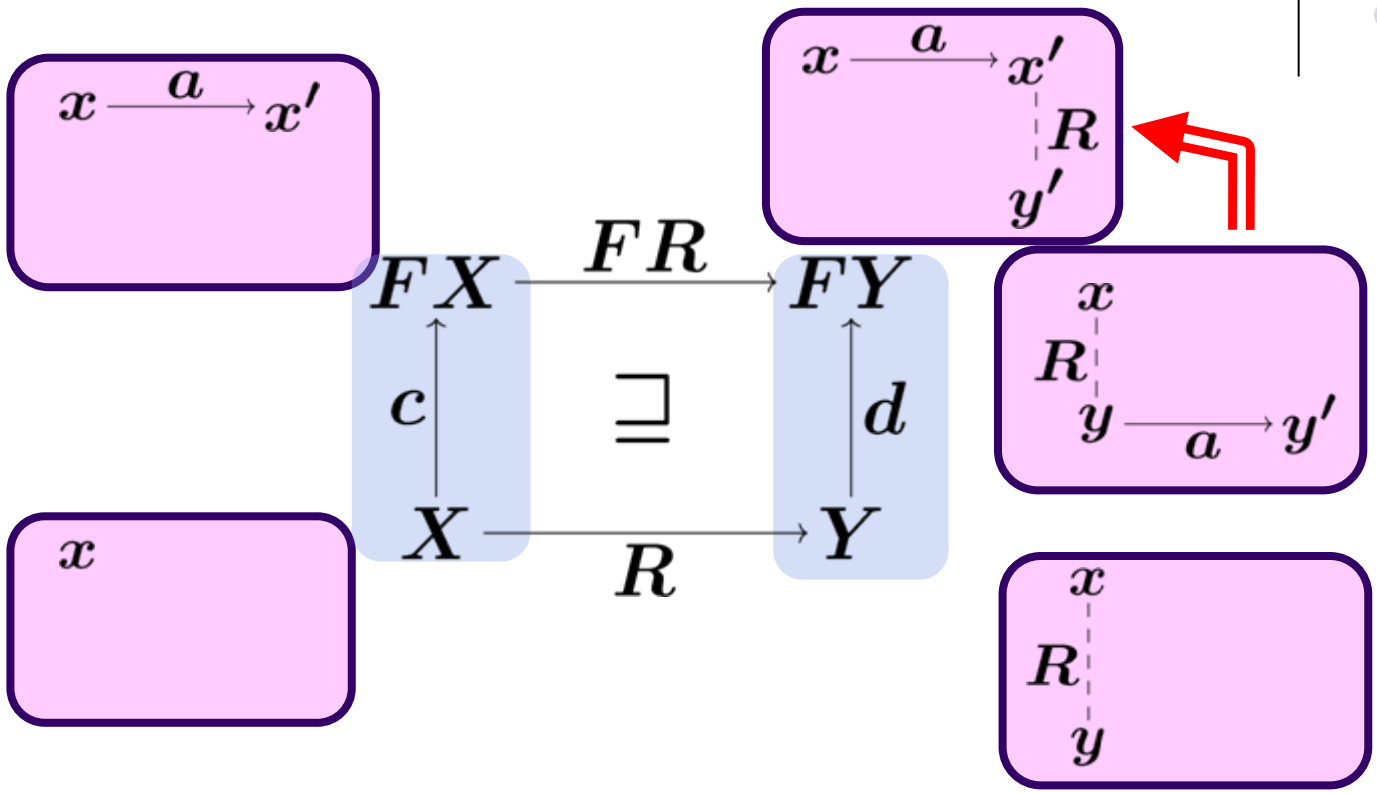
Take T=$\mathcal{P}$

R : a **relation**, because

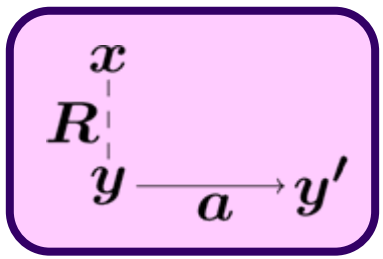
$$\frac{X \longrightarrow Y \text{ in } \mathcal{K}\ell(\mathcal{P})}{X \longrightarrow \mathcal{P}Y \text{, a function}}$$

$$\frac{}{R \subseteq X \times Y \text{, a relation}}$$

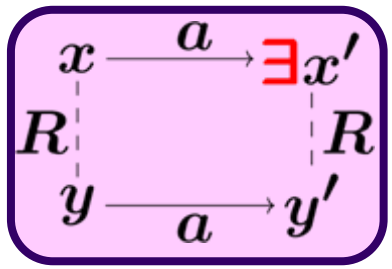
Forward simulation



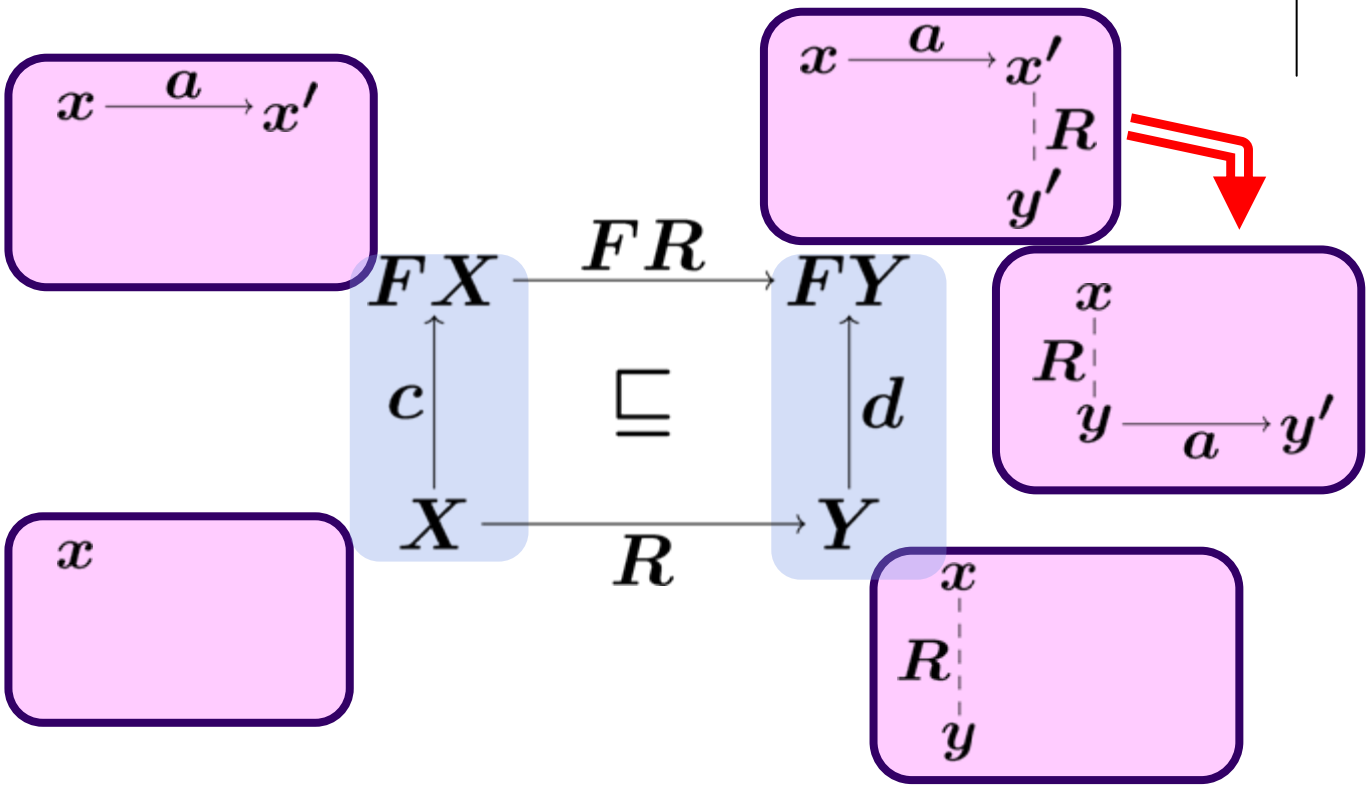
Hence

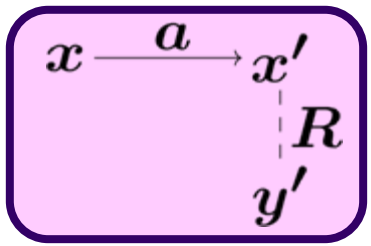


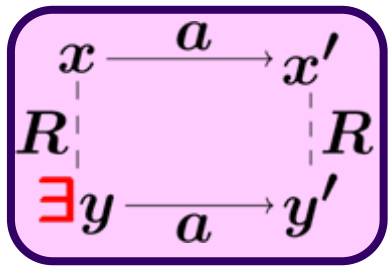
implies



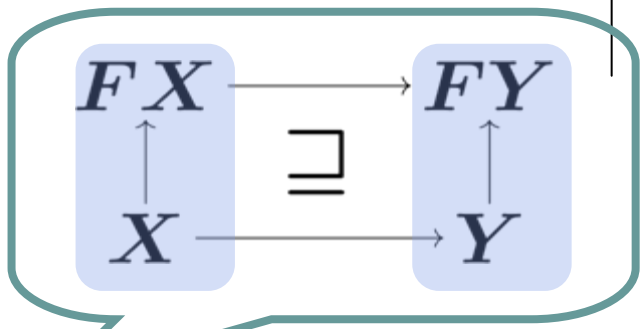
Backward simulation



Hence 

implies 

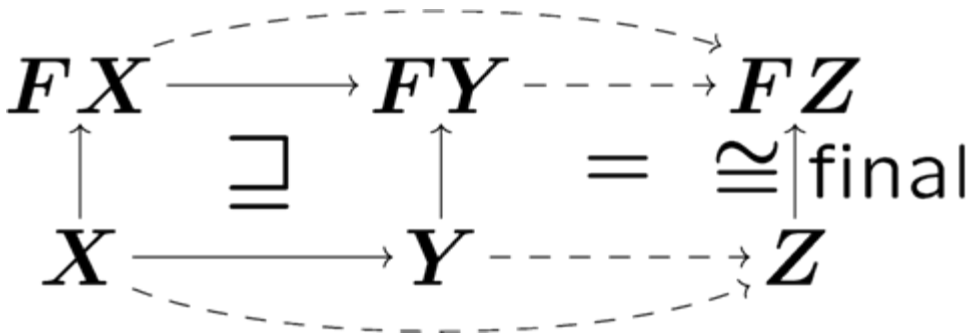
Main result: general soundness theorem



$\exists$ forward/backward simulation

➡ trace inclusion

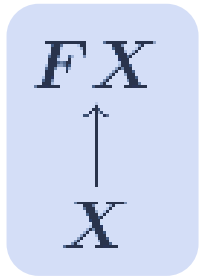
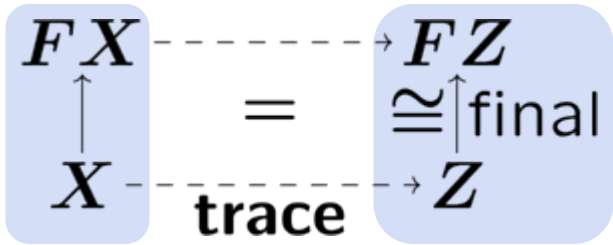
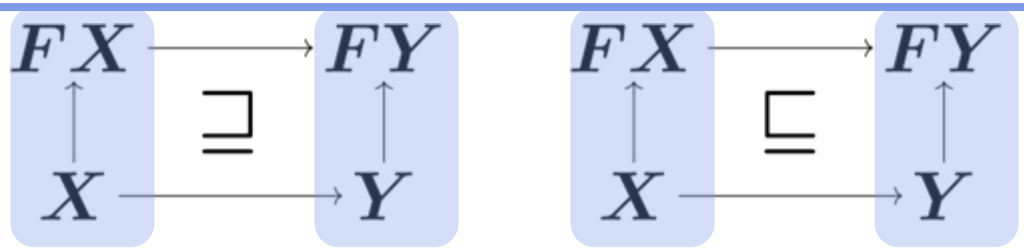
Proof.



in $\mathcal{K}\ell(T)$

- Also completeness result, as easily

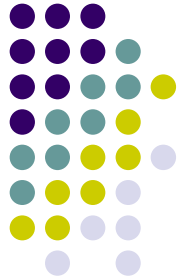
Summary: we have illustrated

System	
Trace semantics	
Simulations	

forward sim. **backward** sim.

in $\mathcal{K}\ell(T)$

General soundness theorem : ∃ simulation ➔ trace incl.



Back to probabilistic anonymity

Back to probabilistic anonymity

probabilistic automaton

Theorem

∃ forward/backward simulation from an($\mathcal{X}$) to $\mathcal{X}$

➔ $\mathcal{X}$ is anonymous

Can we do this probabilistically?

- What is an(**X**)?
- What is "probabilistic simulation"?

probabilistic anonymity

Proof.

$\mathcal{X}$ is anonymous

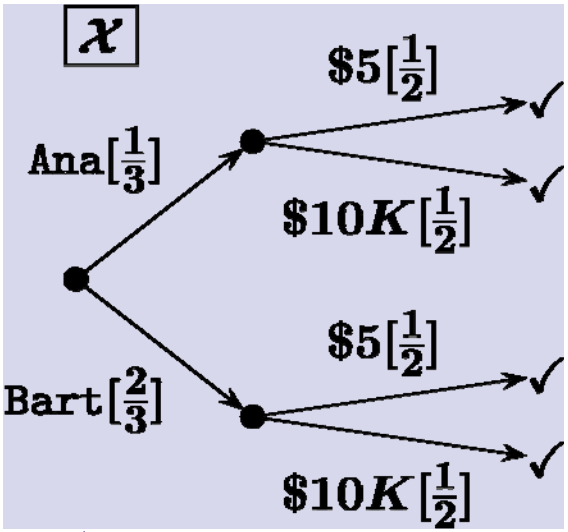
trace *distribution*

🡐 tr($\mathcal{X}$) = tr(an($\mathcal{X}$)) (anonymity is

🡐 tr($\mathcal{X}$) ⊇ tr(an($\mathcal{X}$)) probabilistic simulation al)

🡐 ∃ simulation from an($\mathcal{X}$) to $\mathcal{X}$ (**soundness theorem**!)

Probabilistic simulation via coalgebraic simulation



f : fwd simulation from $\mathcal{Y}$ to $\mathcal{X}$

$\Leftrightarrow$

$$\begin{array}{ccc} FX & \longrightarrow & FY \\ \uparrow & \sqsupseteq & \uparrow \\ X & \xrightarrow[f]{} & Y \end{array}$$

in $\mathcal{K}\ell(T)$

$\Leftrightarrow$

$$\begin{array}{lcll} s(x) & \leq & \sum_{y \in Y} t(y) \cdot f(y)(x) & \text{for any } x \in X, \\ \sum_{x \in X} f(y)(x) \cdot c(x)(e, \checkmark) & \leq & d(y)(e, \checkmark) & \text{for any } y \in Y \text{ and } e \in \mathcal{A}, \\ \sum_{x \in X} f(y)(x) \cdot c(x)(e, x') & \leq & \sum_{y' \in Y} d(y)(e, y') \cdot f(y')(x') & \\ & & & \text{for any } y \in Y,\ e \in \mathcal{A} \text{ and } x' \in X. \end{array}$$

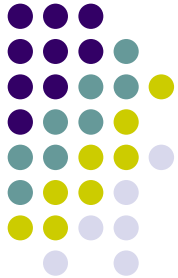
Coalgebra for

- $T = \mathcal{D}$ (prob. branching)
- $F = \mathcal{A} \times \{\checkmark\} + \mathcal{A} \times \_$

Soundness comes for free by

$$\begin{array}{ccccc} FX & \longrightarrow & FY & \dashrightarrow & FZ \\ \uparrow & \sqsupseteq & \uparrow & = & \cong \uparrow \text{final} \\ X & \longrightarrow & Y & \dashrightarrow & Z \end{array}$$

Probabilistic
(∃ simulation ➔ anonymity)
[IH&Kawabe,2006]



Theorem

∃ forward/backward simulation from an($\mathcal{X}$) to $\mathcal{X}$

➔ $\mathcal{X}$ is anonymous

Probabilistic

Probabilistically

Proof.

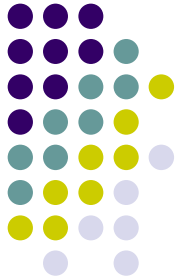
$\mathcal{X}$ is anonymous

🡐 tr($\mathcal{X}$) = tr(an($\mathcal{X}$)) (anonymity is trace-based)

🡐 tr($\mathcal{X}$) ⊇ tr(an($\mathcal{X}$)) (⊆ is by constr. of an($\mathcal{X}$))

🡐 ∃ simulation from an($\mathcal{X}$) to $\mathcal{X}$ (**soundness theorem**!)

Probabilistic
(∃ simulation ➔ anonymity)
[IH&Kawabe,2006]



Theorem

∃ forward/backward simulation from an($\mathcal{X}$) to $\mathcal{X}$

Probabilistic

➔ $\mathcal{X}$ is anonymous

Probabilistically

- Should be useful in ***theorem-proving*** anonymity
  - This is the case in a non-deterministic setting
    - Verification of the FOO voting protocol [Kawabe,Mano,Sakurada,Tsukada'06]
  - Probabilistic verification example is yet to be found
    - Currently our running example is dining cryptographers [Chaum'88]

Conclusion

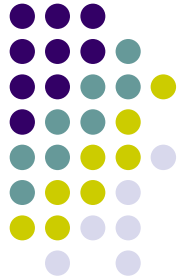
non-deterministic (∃ simulation ➔ anonymity)
[KawabeMST06]

Generic, **coalgebraic** theory of
traces and simulations [IH,Jacobs,Sokolova]

- $T = \mathcal{P}$ ➔ non-determinism
- $T = \mathcal{D}$ ➔ probability

probabilistic **(∃ simulation ➔ anonymity)**
[Current work]

Conclusion and future work



- Future work
  - Bigger **verification example**?
  - Systems with **both non-determinism and probability**?
    - As in [Segala&Lynch'95]
    - Important also for anonymity applications [Palamidessi,MFPS'05]
    - Suitable coalgebraic framework is missing
  - **Weaker notion** of anonymity?
    - Current "anonymity" notion is pretty strong
    - "Simulation-based" method also for weaker notions?

For you to take home:

- ***Probability*** in anonymity
- Category theory in action!

Thank you for your attention!

Ichiro Hasuo, U. Nijmegen, NL

http://www.cs.ru.nl/~ichiro