Metrics and Approximation

Franck van Breugel

DisCoVeri Group, Department of Computer Science and Engineering
York University, Toronto

June 23, 2010

Metrics

Franck van Breugel

DisCoVeri Group, Department of Computer Science and Engineering
York University, Toronto

June 23, 2010

Behavioural Pseudometrics

Franck van Breugel

DisCoVeri Group, Department of Computer Science and Engineering
York University, Toronto

June 23, 2010

Who is your favourite painter?

- Michelangelo di Lodovico Buonarroti Simoni
- Claude Monet
- Pieter Bruegel
- Franklin Carmichael
- Kartika Affandi-Koberl
- Ebele Okoye
- Frida Kahlo
- Jiao Bingzhen

Mark Tansey

Mark Tansey was born 1949 in San Jose, CA, USA. He is best known for his monochromatic works. His paintings can be found in numerous museums including the New York Metropolitan Museum of Art and the Smithsonian American Art Museum in Washington. His painters have been exhibited at many places including MIT's List Visual Art Center and the Montreal Museum of Fine Arts.

Triumph of the New York School

The Innocent Eye Test

Mark Tansey

Who is this?

Scott Smolka

What are they measuring?

Distances between states of probabilistic concurrent systems.

Alessandro Giacalone, Chi-chang Jou and Scott A. Smolka.
Algebraic reasoning for probabilistic concurrent systems.

In, M. Broy and C.B. Jones, editors, *Proceedings of the IFIP WG 2.2/2.3 Working Conference on Programming Concepts and Methods*, pages 443-458, Sea of Gallilee, April 1990. North-Holland.

Not so easy to find

```
From: Scott Smolka <sas@cs.sunysb.edu>
To: Franck van Breugel <franck@cse.yorku.ca>

Dear Franck,

...  The first thing I will need to do however is
find a copy of the paper.  I do not think I have
one presently ...

All the best,
Scott
```

Why are they measuring those distances?

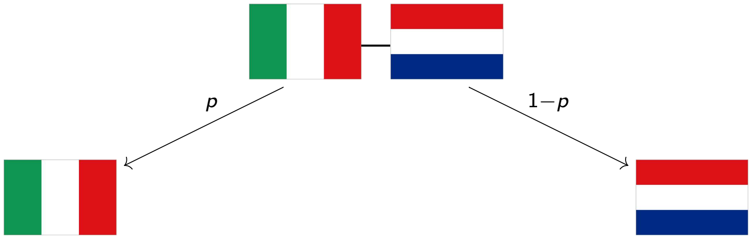
To address that question, let us first try to answer another

Question

Who will win the world cup?

Who will win the world cup?

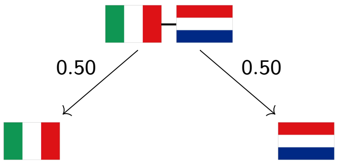
The tournament can be modelled as a probabilistic system. Consider, for example, one possible match in the round of 16.



What is the probability that Italy wins?

Who will win the world cup?

The best we can do is approximate the probability that Italy wins.



These probabilistic systems are not behaviourally equivalent, since the probabilities do not match exactly.

What is a behavioural equivalence?

An equivalence relation that captures which states give rise to the same behaviour.

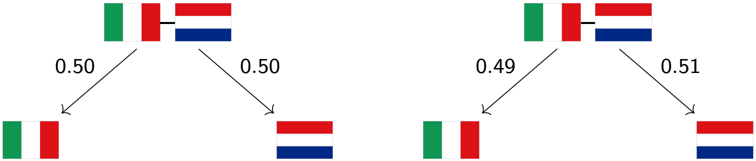
Examples: trace equivalence, bisimilarity, weak bisimilarity, probabilistic bisimilarity, timed bisimilarity, . . .

Why are we interested in behavioural equivalences?

- They answer the fundamental question "Do these two states give rise to the same behaviour?"
- They are used to minimize the state space by identifying those states that are behaviourally equivalent.
- They are used to prove transformations correct.
- . . .

However

For systems with approximate quantitative data, behavioural equivalences make little sense since they are not robust.



Why are GJS measuring those distances?

Problem

Behavioural equivalences for systems with approximate quantitative data are not robust.

Solution

Replace the Boolean valued notion (equivalence relation) with a real valued notion (distance).

Outline

- Part I: back to 1990
- Part II: the rest of the nineties
- Part III: the 21st century

What is a pseudometric?

Definition

Let X be a set. A *pseudometric on* X is a function $d_X : X \times X \to [0, \infty]$ satisfying for all x, y, $z \in X$,

- $d_X(x, x) = 0$,
- $d_X(x, y) = d_X(y, x)$ and
- $d_X(x, z) \leq d_X(x, y) + d_X(y, z)$.

Example

Let X be a set. The *discrete metric on* X is defined by

$$d_X(x, y) = \begin{cases} 0 & \text{if } x = y \\ \infty & \text{otherwise} \end{cases}$$

The *Euclidean metric on* $\mathbb{R}$ is defined by

$$d_{\mathbb{R}}(x, y) = |x - y|$$

From equivalence relation to pseudometric

Proposition

Let X be a set and let $\mathcal{R}$ be an equivalence relation on X. Then

$$d_X(x,y) = \begin{cases} 0 & \text{if } x \,\mathcal{R}\, y \\ \infty & \text{otherwise} \end{cases}$$

is a pseudometric on X.

Example

With the identity relation corresponds the discrete metric.

From pseudometric to equivalence relation

Proposition

Let X be a set and let d_X be a pseudometric on X. Then

$$x \mathcal{R} y \text{ if } d_x(x, y) = 0$$

is an equivalence relation.

Example

The discrete metric and the Euclidean metric both correspond to the identity relation.

Probabilistic transition system

Definition

A *probabilistic transition system* (PTS) is a tuple $\langle S, \mathcal{A}, T \rangle$ consisting of

- a set S of states,
- a set $\mathcal{A}$ of actions, and
- a function $T : S \times A \times S \to [0,1]$ such that for all $s \in S$,

$$\sum_{a \in \mathcal{A} \wedge s' \in S} T(s, a, s') \in \{0, 1\}.$$

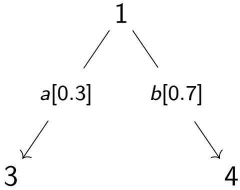
This is a generative model (as mentioned in Roberto's lecture), whereas Prakash presented a reactive model.

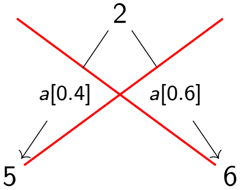
Deterministic probabilistic transition system

Definition

A PTS is *deterministic* if for all $s \in S$ and $a \in \mathcal{A}$,

$$|\{ s' \in S \mid T(s, a, s') > 0 \}| \leq 1.$$





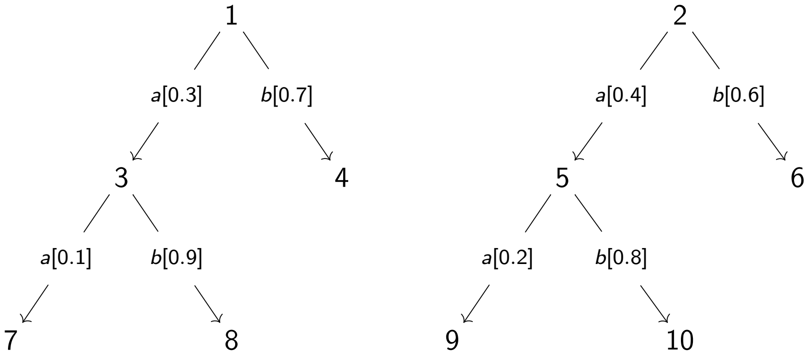
ϵ-Bisimulation

Definition

Let $\epsilon \in [0, 1]$. A relation $\mathcal{R} \subseteq S \times S$ is an *ϵ-bisimulation* if for all $s_1 \mathrel{\mathcal{R}} s_2$ and $a \in \mathcal{A}$

- if $T(s_1, a, s_1') > 0$ then $T(s_2, a, s_2') > 0$ and $|T(s_1, a, s_1') - T(s_2, a, s_2')| \leq \epsilon$ for some s_2' such that $s_1' \mathrel{\mathcal{R}} s_2'$ and
- if $T(s_2, a, s_2') > 0$ then $T(s_1, a, s_1') > 0$ and $|T(s_1, a, s_1') - T(s_2, a, s_2')| \leq \epsilon$ for some s_1' such that $s_1' \mathrel{\mathcal{R}} s_2'$

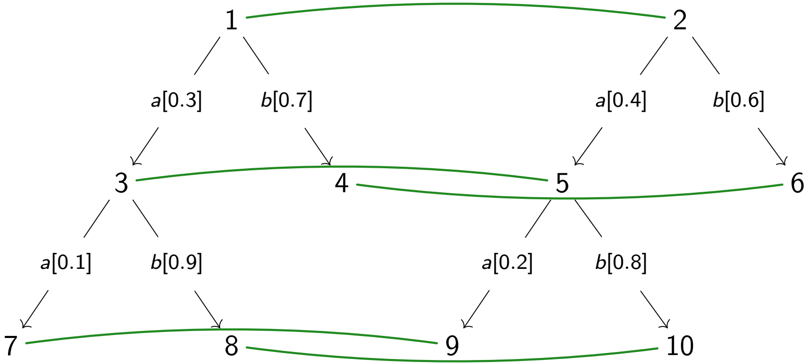
ϵ-Bisimulation



Question

What is the smallest ϵ such that there exists an ϵ-bisimulation $\mathcal{R}$ with $1 \mathcal{R} 2$?

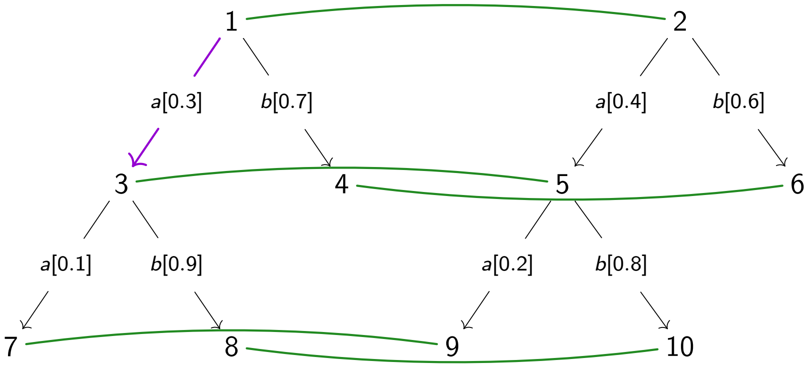
ϵ-Bisimulation



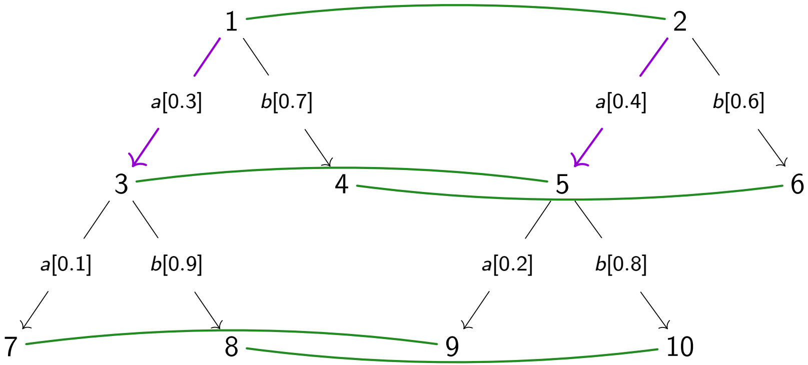
Answer

0.1

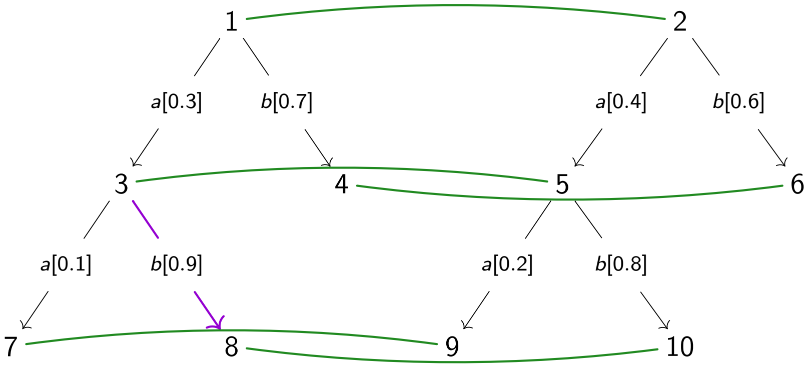
ϵ-Bisimulation



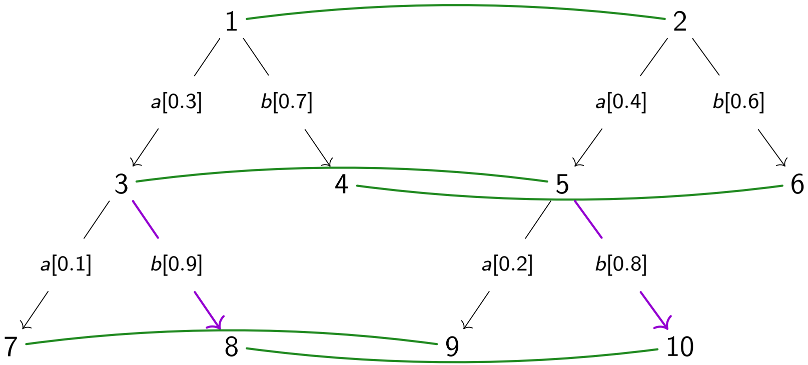
ϵ-Bisimulation



ϵ-Bisimulation



ϵ-Bisimulation



ϵ-Bisimularity

Definition

Let $\epsilon \in [0,1]$. The ϵ-*bisimilarity* relation $\overset{\epsilon}{\sim}$ is defined by

$$\overset{\epsilon}{\sim} = \bigcup \{ \mathcal{R} \mid \mathcal{R} \text{ is an } \epsilon\text{-bisimulation} \}.$$

Proposition

- $\overset{\epsilon}{\sim}$ is an ϵ-bisimulation.
- If $\epsilon \leq \epsilon'$ then $\overset{\epsilon}{\sim} \subseteq \overset{\epsilon'}{\sim}$.
- $\overset{0}{\sim}$ is probabilistic bisimilarity.

A pseudometric for deterministic PTSs

Definition

The function $E_S : S \times S \rightarrow 2^{[0,1]}$ is defined by

$$E_S(s_1, s_2) = \{ \epsilon \mid s_1 \mathcal{R} s_2 \text{ for some } \epsilon\text{-bisimulation } \mathcal{R} \}.$$

The function $d_S : S \times S \rightarrow [0, 1]$ is defined by

$$d_S(s_1, s_2) = \begin{cases} \inf E_S(s_1, s_2) & \text{if } E_S(s_1, s_2) \neq \emptyset \\ 1 & \text{otherwise} \end{cases}$$

Proposition

- d_S is a pseudometric.
- For all s_1, $s_2 \in S$, $d_S(s_1, s_2) = 0$ iff s_1 and s_2 are probabilistic bisimilar.

ϵ-Bisimulation

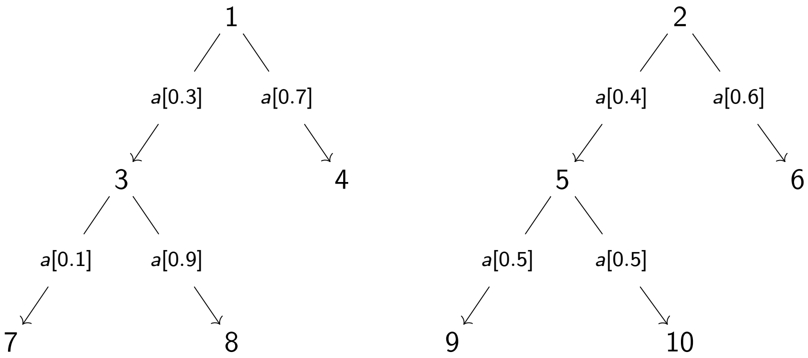
Let us adapt the notion of ϵ-bisimulation for all PTSs (not necessarily deterministic).

Definition

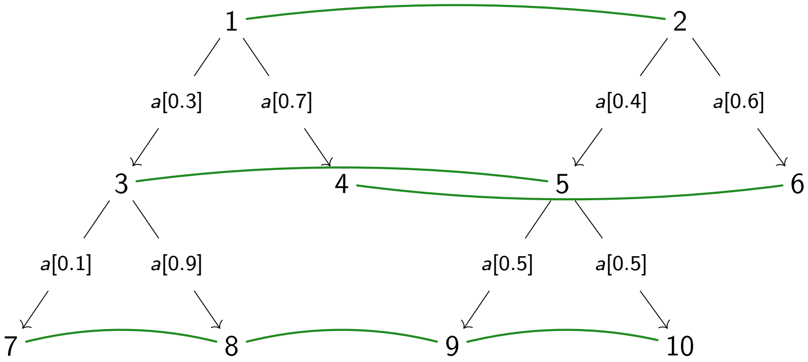
Let $\epsilon \in [0, 1]$. An equivalence relation $\mathcal{R} \subseteq S \times S$ is an *ϵ-bisimulation* if for all $s_1 \, \mathcal{R} \, s_2$, $a \in \mathcal{A}$ and $B \in S/\mathcal{R}$

- if $\sum_{s \in B} T(s_1, a, s) > 0$ then $\sum_{s \in B} T(s_2, a, s) > 0$ and $|\sum_{s \in B} T(s_1, a, s) - \sum_{s \in B} T(s_2, a, s)| \leq \epsilon$ and
- if $\sum_{s \in B} T(s_2, a, s) > 0$ then $\sum_{s \in B} T(s_1, a, s) > 0$ and $|\sum_{s \in B} T(s_1, a, s) - \sum_{s \in B} T(s_2, a, s)| \leq \epsilon$.

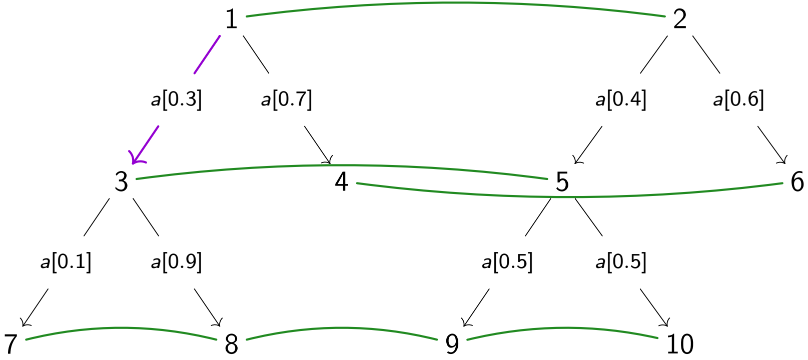
ϵ-Bisimulation



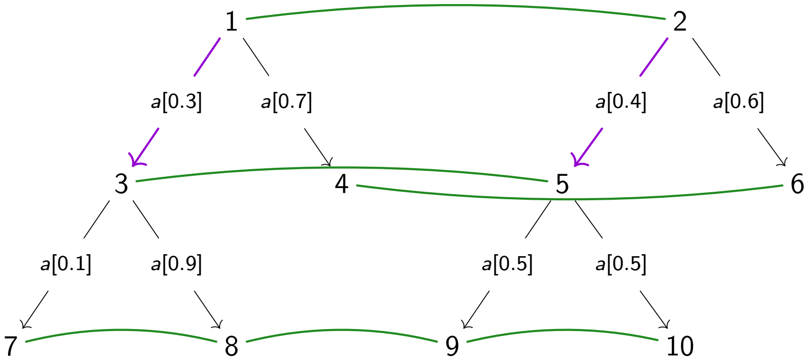
ϵ-Bisimulation



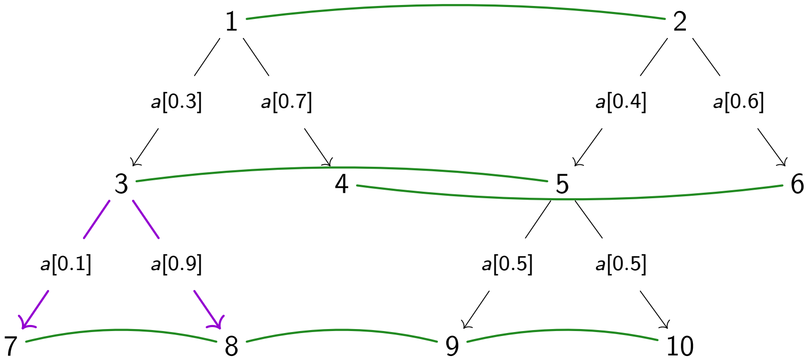
ϵ-Bisimulation



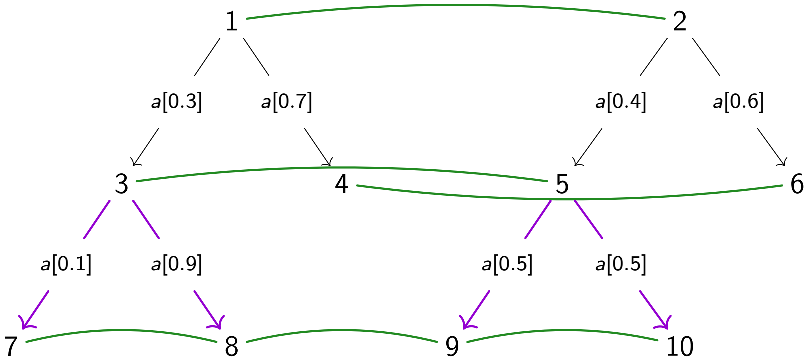
ϵ-Bisimulation



ϵ-Bisimulation



ϵ-Bisimulation



A distance for PTSs

The definitions of E_S and d_S remain unchanged.

Proposition

For all s_1, s_2, $s_3 \in S$,

- $d_S(s_1, s_1) = 0$, and
- $d_S(s_1, s_2) = d_S(s_2, s_1)$.

Proposition

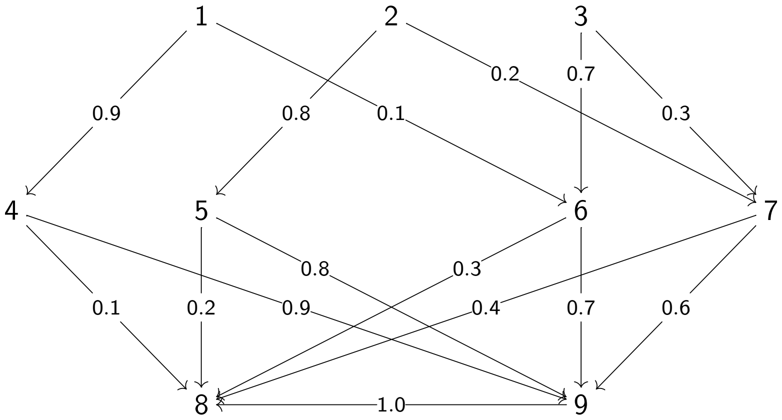
There exist s_1, s_2, $s_3 \in S$ such that

- $d_S(s_1, s_3) \not\leq d_S(s_1, s_2) + d_S(s_2, s_3)$.

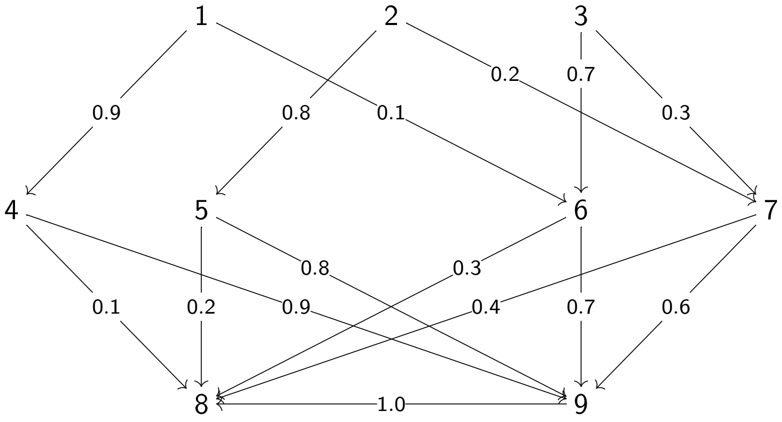
Corollary

d_S is not a pseudometric.

$d_S(1,3) \not\leq d_S(1,2) + d_S(2,3)$

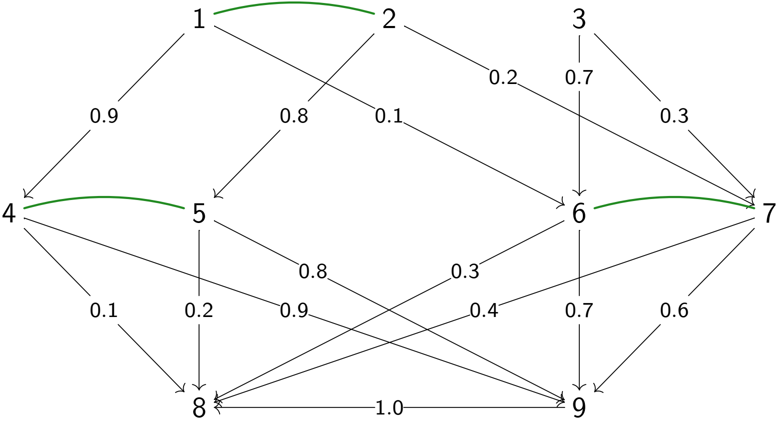


$d_S(1,3) \not\leq d_S(1,2) + d_S(2,3)$



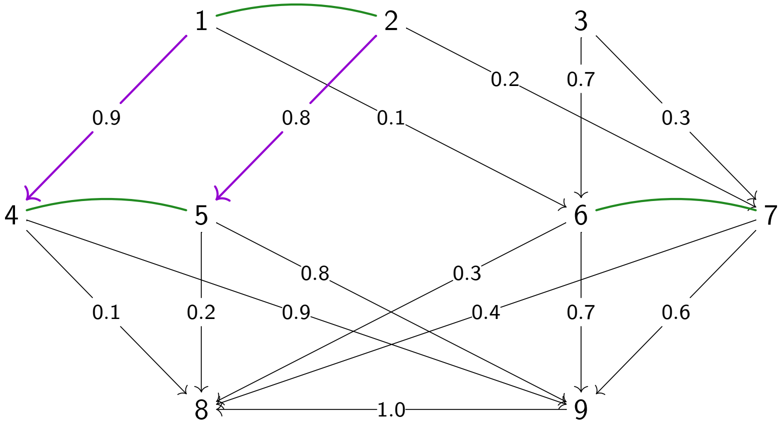
$d_S(1,2) \leq 0.1$

$d_S(1,3) \not\leq d_S(1,2) + d_S(2,3)$



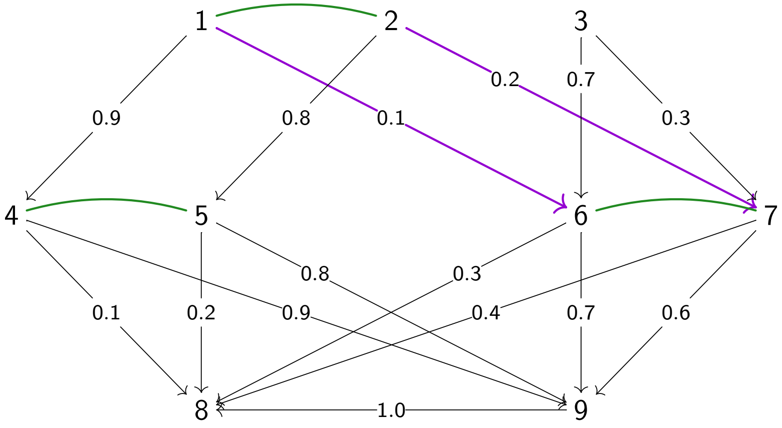
$d_S(1,2) \leq 0.1$

$d_S(1,3) \not\leq d_S(1,2) + d_S(2,3)$



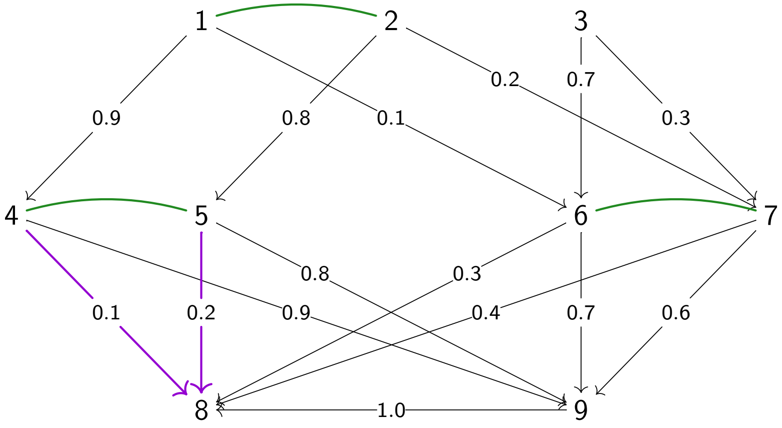
$d_S(1,2) \leq 0.1$

$d_S(1,3) \not\leq d_S(1,2) + d_S(2,3)$



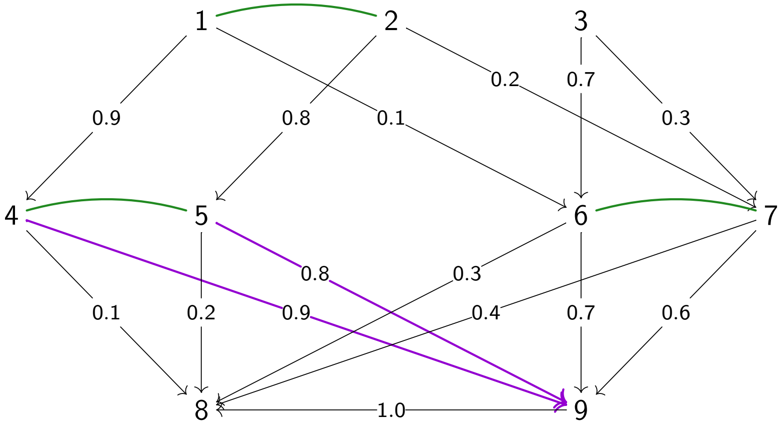
$d_S(1,2) \leq 0.1$

$d_S(1,3) \not\leq d_S(1,2) + d_S(2,3)$



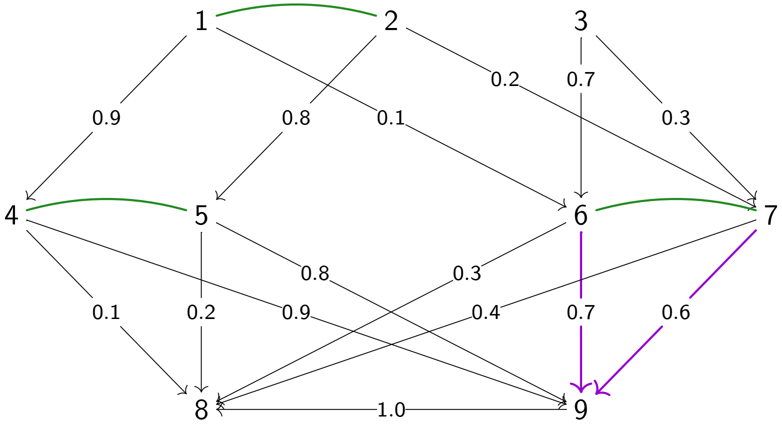
$d_S(1,2) \leq 0.1$

$d_S(1,3) \not\leq d_S(1,2) + d_S(2,3)$



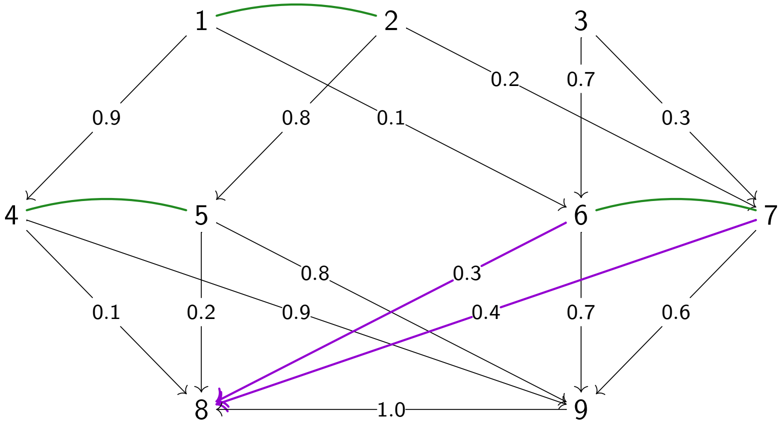
$d_S(1,2) \leq 0.1$

$d_S(1,3) \not\leq d_S(1,2) + d_S(2,3)$



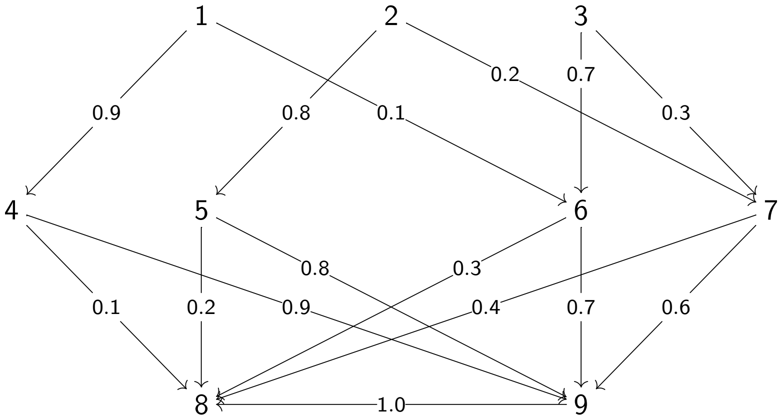
$d_S(1,2) \leq 0.1$

$d_S(1,3) \not\leq d_S(1,2) + d_S(2,3)$



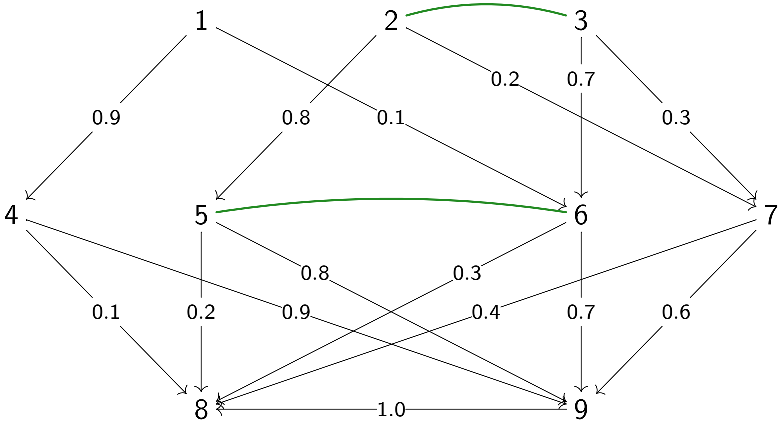
$d_S(1,2) \leq 0.1$

$d_S(1,3) \not\leq d_S(1,2) + d_S(2,3)$



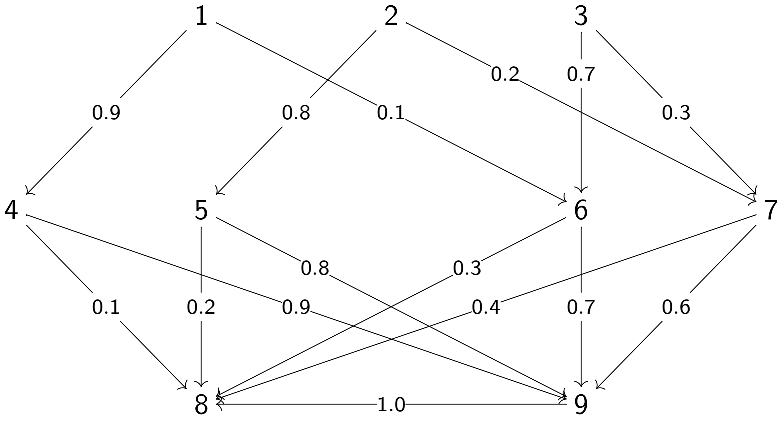
$d_S(1,2) \leq 0.1 \wedge d_S(2,3) \leq 0.1$

$d_S(1,3) \not\leq d_S(1,2) + d_S(2,3)$



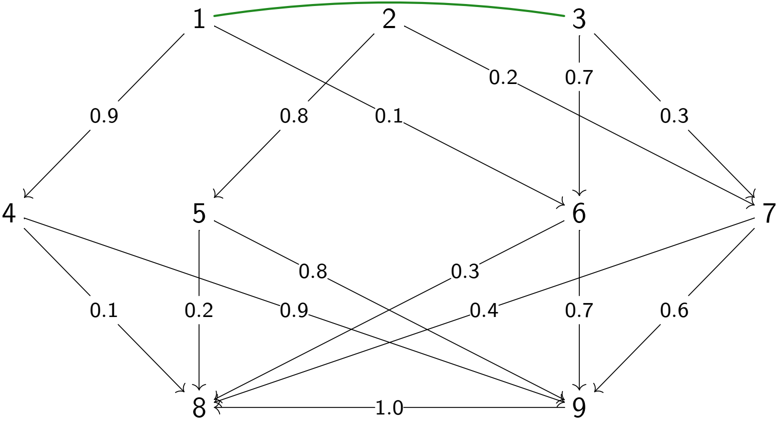
$d_S(1,2) \leq 0.1 \wedge d_S(2,3) \leq 0.1$

$d_S(1,3) \not\leq d_S(1,2) + d_S(2,3)$



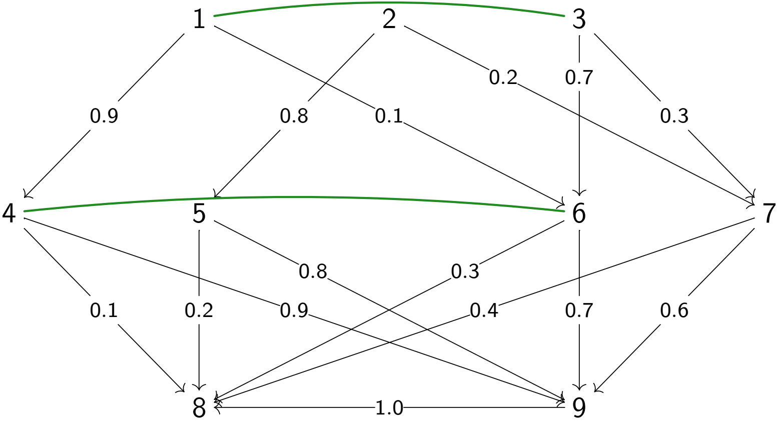
$$d_S(1,2) \leq 0.1 \wedge d_S(2,3) \leq 0.1 \wedge d_S(1,3) > 0.25$$

$d_S(1,3) \not\leq d_S(1,2) + d_S(2,3)$



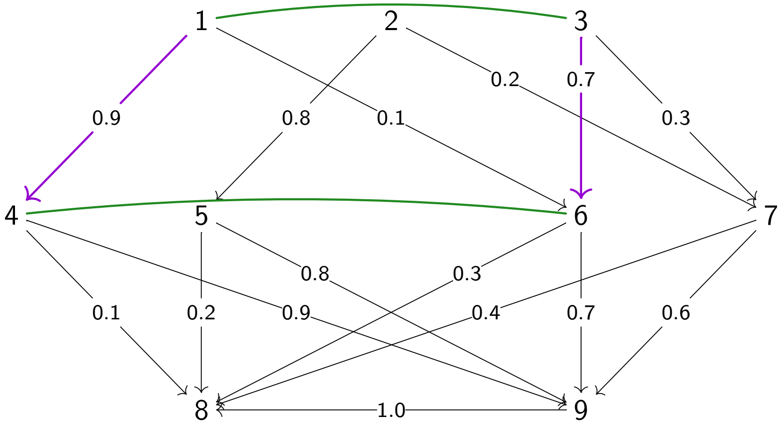
$d_S(1,2) \leq 0.1 \wedge d_S(2,3) \leq 0.1 \wedge d_S(1,3) > 0.25$

$d_S(1,3) \not\leq d_S(1,2) + d_S(2,3)$



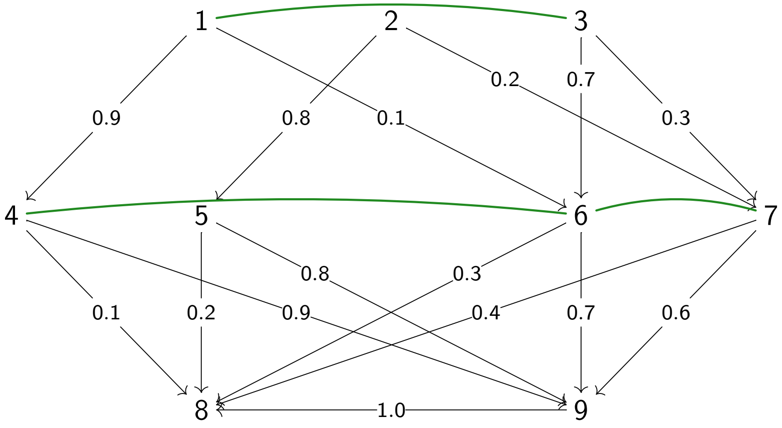
$$d_S(1,2) \leq 0.1 \wedge d_S(2,3) \leq 0.1 \wedge d_S(1,3) > 0.25$$

$d_S(1,3) \not\leq d_S(1,2) + d_S(2,3)$



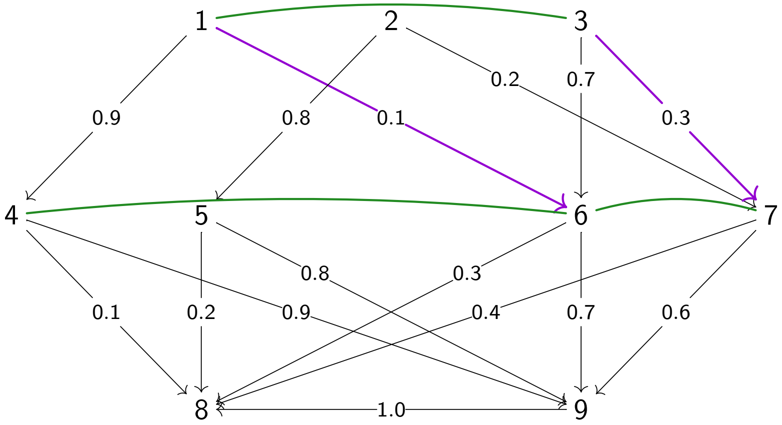
$d_S(1,2) \leq 0.1 \wedge d_S(2,3) \leq 0.1 \wedge d_S(1,3) > 0.25$

$d_S(1,3) \not\leq d_S(1,2) + d_S(2,3)$



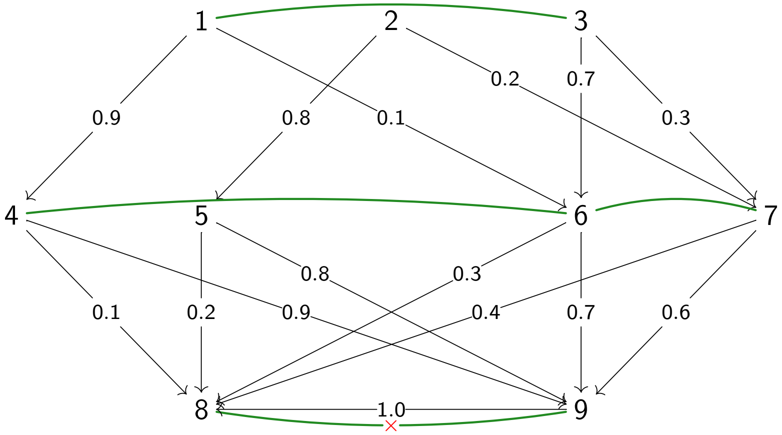
$d_S(1,2) \leq 0.1 \wedge d_S(2,3) \leq 0.1 \wedge d_S(1,3) > 0.25$

$d_S(1,3) \not\leq d_S(1,2) + d_S(2,3)$



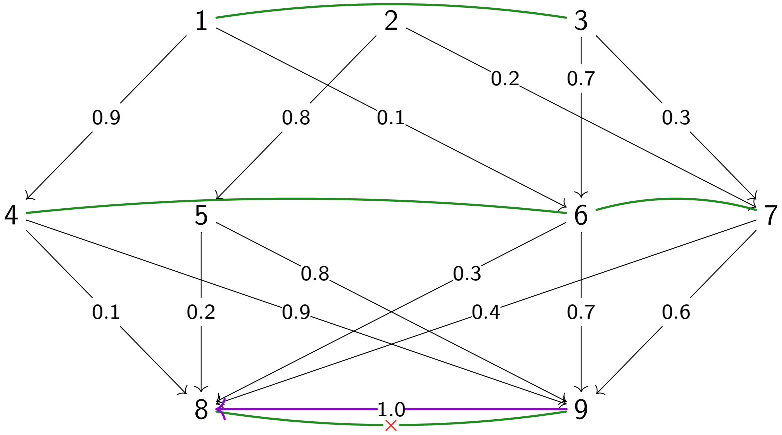
$d_S(1,2) \leq 0.1 \wedge d_S(2,3) \leq 0.1 \wedge d_S(1,3) > 0.25$

$d_S(1,3) \not\leq d_S(1,2) + d_S(2,3)$



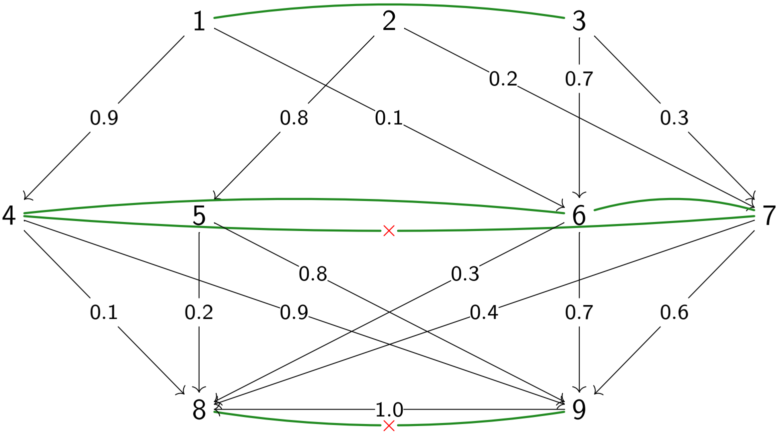
$d_S(1,2) \leq 0.1 \wedge d_S(2,3) \leq 0.1 \wedge d_S(1,3) > 0.25$

$d_S(1,3) \not\leq d_S(1,2) + d_S(2,3)$



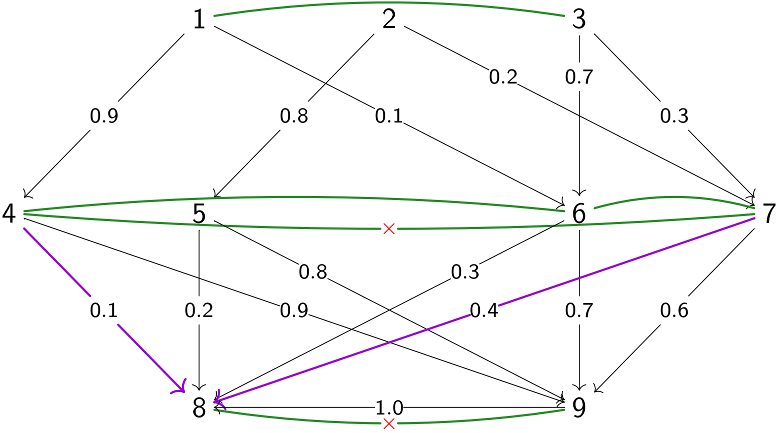
$$d_S(1,2) \leq 0.1 \wedge d_S(2,3) \leq 0.1 \wedge d_S(1,3) > 0.25$$

$d_S(1,3) \not\leq d_S(1,2) + d_S(2,3)$



$d_S(1,2) \leq 0.1 \wedge d_S(2,3) \leq 0.1 \wedge d_S(1,3) > 0.25$

$d_S(1,3) \not\leq d_S(1,2) + d_S(2,3)$



$$d_S(1,2) \leq 0.1 \wedge d_S(2,3) \leq 0.1 \wedge d_S(1,3) > 0.25$$

Compositional reasoning

Observation

If a behavioural equivalence is a congruence then it supports compositional reasoning.

Example

If $s_1 \sim s_2$ then $s_1 + s \sim s_2 + s$.

Question

What is a quantitative analogue of congruence?

A quantitative analogue of congruence

Let $\oplus$ denote a random choice.

- If $s_1 \sim s_2$ then $s_1 \oplus s \sim s_2 \oplus s$.

A quantitative analogue of congruence

Let $\oplus$ denote a random choice.

- If $s_1 \sim s_2$ then $s_1 \oplus s \sim s_2 \oplus s$.
- If $d_S(s_1, s_2) = 0$ then $d_S(s_1 \oplus s, s_2 \oplus s) = 0$.

A quantitative analogue of congruence

Let $\oplus$ denote a random choice.

- If $s_1 \sim s_2$ then $s_1 \oplus s \sim s_2 \oplus s$.
- If $d_S(s_1, s_2) = 0$ then $d_S(s_1 \oplus s, s_2 \oplus s) = 0$.
- For all $\epsilon \geq 0$, if $d_S(s_1, s_2) \leq \epsilon$ then $d_S(s_1 \oplus s, s_2 \oplus s) \leq \epsilon$.

A quantitative analogue of congruence

Let $\oplus$ denote a random choice.

- If $s_1 \sim s_2$ then $s_1 \oplus s \sim s_2 \oplus s$.
- If $d_S(s_1, s_2) = 0$ then $d_S(s_1 \oplus s, s_2 \oplus s) = 0$.
- For all $\epsilon \geq 0$, if $d_S(s_1, s_2) \leq \epsilon$ then $d_S(s_1 \oplus s, s_2 \oplus s) \leq \epsilon$.
- $d_S(s_1 \oplus s, s_2 \oplus s) \leq d_S(s_1, s_2)$.

Let $\oplus$ denote a random choice.

- If $s_1 \sim s_2$ then $s_1 \oplus s \sim s_2 \oplus s$.
- If $d_S(s_1, s_2) = 0$ then $d_S(s_1 \oplus s, s_2 \oplus s) = 0$.
- For all $\epsilon \geq 0$, if $d_S(s_1, s_2) \leq \epsilon$ then $d_S(s_1 \oplus s, s_2 \oplus s) \leq \epsilon$.
- $d_S(s_1 \oplus s, s_2 \oplus s) \leq d_S(s_1, s_2)$.
- $\cdot \oplus s$ is nonexpansive.

Summary of Part I: back to 1990

Giacalone, Jou and Smolka

- advocated the use of pseudometrics instead of equivalence relations to compare the behaviour of states of systems with approximate quantitative data,
- introduced a pseudometric for deterministic PTSs, and
- proposed nonexpansiveness as a quantitative generalization of congruence.

But what about PTSs that are not deterministic?