Quantitative Analysis of Distributed Probabilistic Systems

Ratul Saha
National University of Singapore

Collaborators

Current Collaborators:

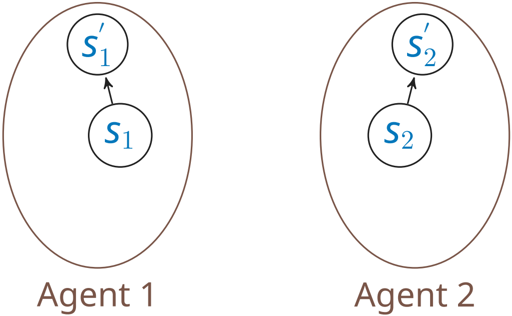
- S Akshay
  IIT Bombay, India
- Javier Esparza
  TU Munich, Germany
- Grant Ingersoll
  CTO, LucidWorks®
- Ansuman Banerjee
  ISI Kolkata, India
- Madhavan Mukund
  CMI, India

Previous Collaborators:

- Sumit K Jha
  Univ. of Central Florida, USA
- P S Thiagarajan
  Harvard Medical School, USA

Distributed

Network of Agents

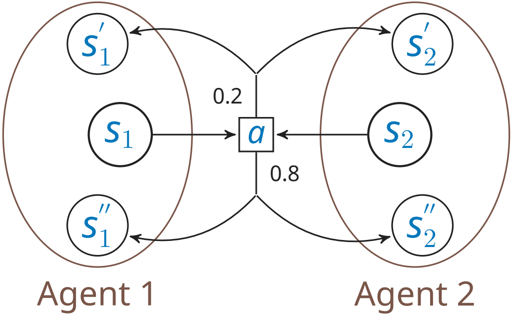


Probabilistic

Distributed

The Synchronization

→ Joint probabilistic move after the synchronization action

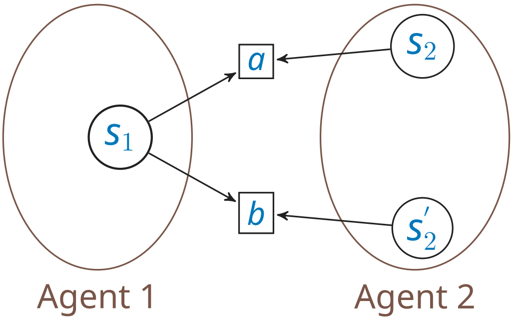


Probabilistic

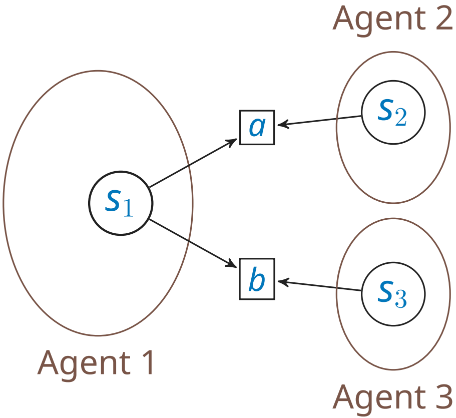
Distributed

Deterministic

Restriction: This is allowed



Restriction: This is not allowed

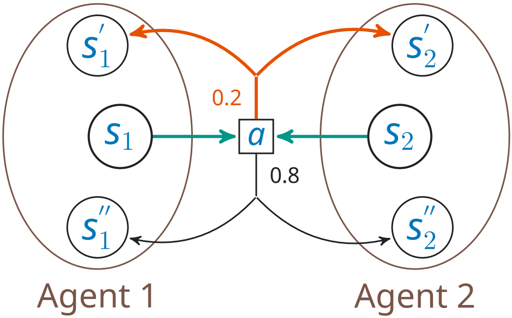


Distributed Markov Chains (DMC)

- Network of communicating probabilistic transition systems
  - Synchronize on shared actions
  - Followed by joint probabilistic move
- Key restriction: no two enabled synchronizations will involve the same agent
  - Syntactically, local state uniquely determines its communicating partners

DMC: Events

→ Event: One synchronization executed at a time, followed by a probabilistic move

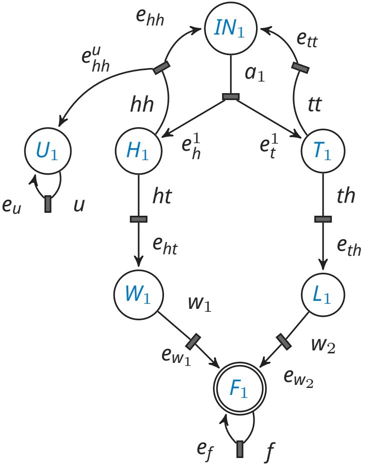


$e = ((s_1, s_2), a, (s_1', s_2'))$ is an event, $p_e = 0.2$

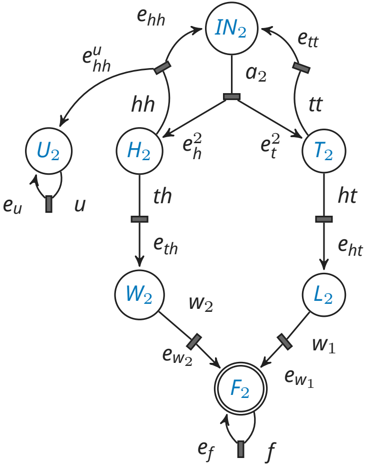
DMC: Coin Toss Example

- $\rightarrow$ Two players. Each toss a fair coin (a_1 and a_2)
- $\rightarrow$ Both tails: they toss again (*tt*)
- $\rightarrow$ Both heads:
  (i) they toss again with prob 0.9 (*hh*), or
  (ii) go to an uncertain state with prob 0.1 (*u*)
- $\rightarrow$ Different outcome: who tosses Heads wins (*ht* and *th*)

DMC: Coin Toss Example

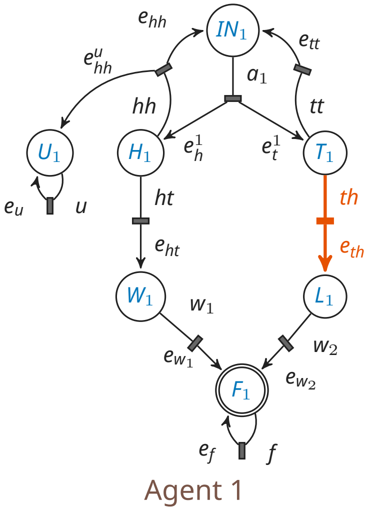


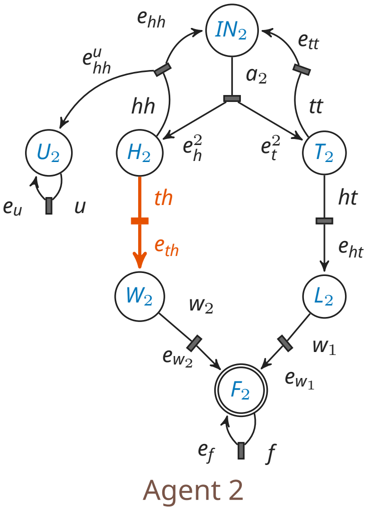
Agent 1



Agent 2

DMC: Coin Toss Example





Global Transition System

→ Associate a global transition system based on event occurrences

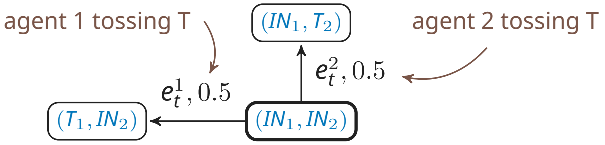
→ This is interleaved semantics

Global Transition System: Coin Toss

agent 1 tossing T

$$(T_1, IN_2) \xleftarrow{e_t^1,\ 0.5} (IN_1, IN_2)$$

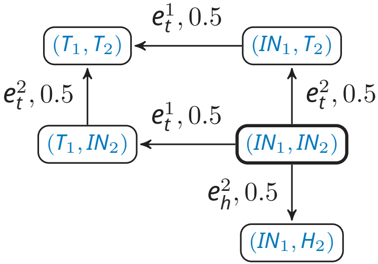
Global Transition System: Coin Toss



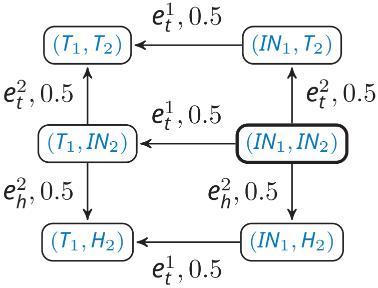
Global Transition System: Coin Toss

$$
\begin{tikzcd}
(T_1, T_2) & (IN_1, T_2) \arrow[l, "{e_t^1, 0.5}"'] \\
(T_1, IN_2) \arrow[u, "{e_t^2, 0.5}"] & \boxed{(IN_1, IN_2)} \arrow[l, "{e_t^1, 0.5}"'] \arrow[u, "{e_t^2, 0.5}"']
\end{tikzcd}
$$

Global Transition System: Coin Toss



Global Transition System: Coin Toss

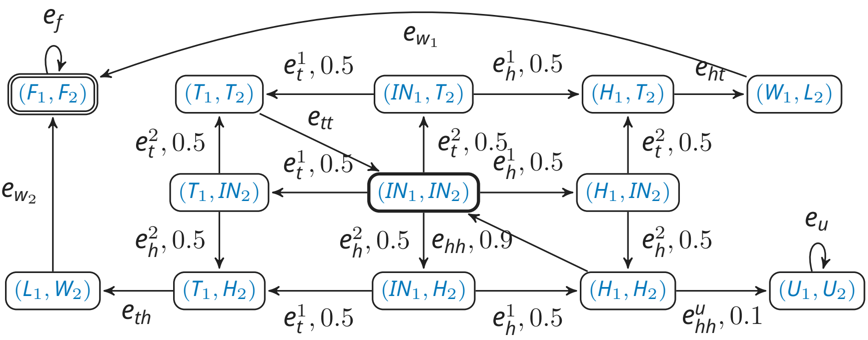


Global Transition System: Coin Toss

$$
\begin{tikzcd}
(T_1, T_2) & (IN_1, T_2) \arrow[l, "{e_t^1, 0.5}"'] \arrow[r, "{e_h^1, 0.5}"] & (H_1, T_2) \\
(T_1, IN_2) \arrow[u, "{e_t^2, 0.5}"] \arrow[d, "{e_h^2, 0.5}"'] & \boxed{(IN_1, IN_2)} \arrow[l, "{e_t^1, 0.5}"'] \arrow[r, "{e_h^1, 0.5}"] \arrow[u, "{e_t^2, 0.5}"'] \arrow[d, "{e_h^2, 0.5}"] & (H_1, IN_2) \arrow[u, "{e_t^2, 0.5}"'] \arrow[d, "{e_h^2, 0.5}"] \\
(T_1, H_2) & (IN_1, H_2) \arrow[l, "{e_t^1, 0.5}"] \arrow[r, "{e_h^1, 0.5}"'] & (H_1, H_2)
\end{tikzcd}
$$

(both agents tossed)

Global Transition System: Coin Toss



(unmarked events have probability 1)

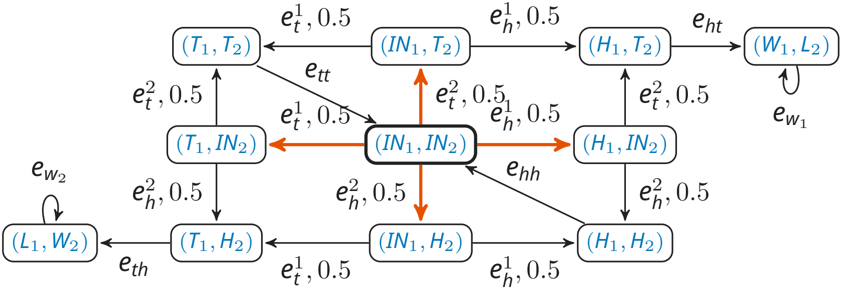
The Trajectory Space

→ We wish to reason about the behavior of the system using the interleaved semantics

Problem: It is *hard* to define a probability measure over the set of maximal trajectories

The Trajectory Space

Due to mix of concurrency and stochasticity, *TS* is not a Markov chain in general

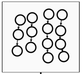


The Solution

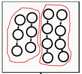
DMC

Transition system

Trajectories



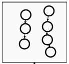
Equiv classes of trajectories



Trajectory space (σ-algebra generated by basic cylinders of equiv class of trajectories)

Probability measure for the trajectory space

Markov chain

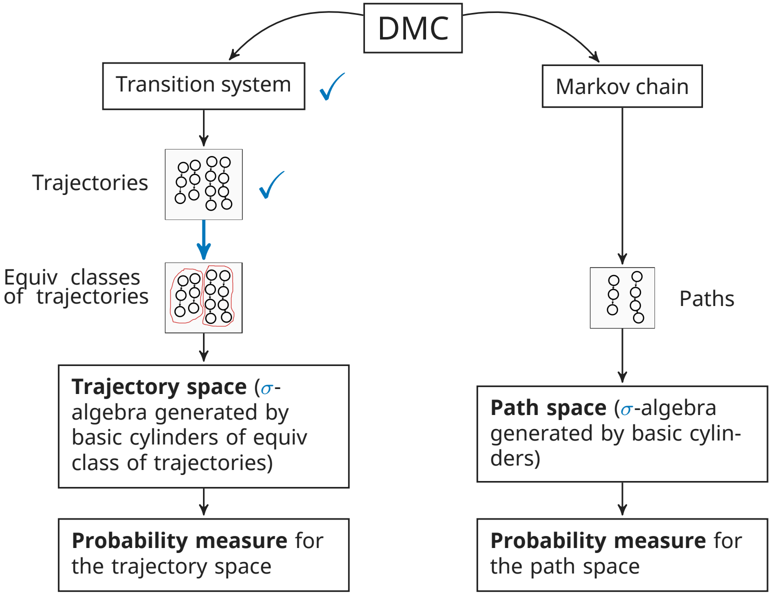


Paths

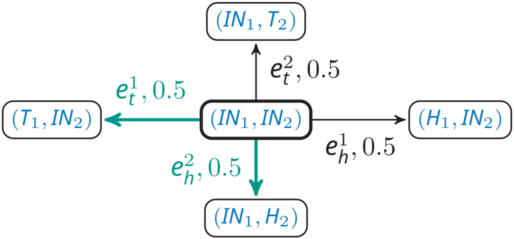
Path space (σ-algebra generated by basic cylinders)

Probability measure for the path space

Equivalence Classes of Trajectories

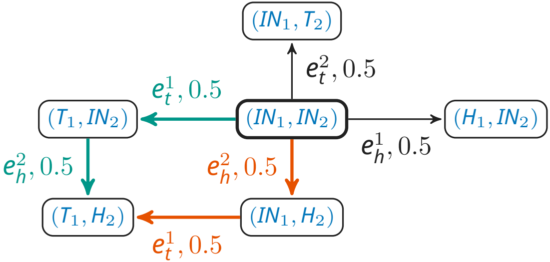


Independence over Events



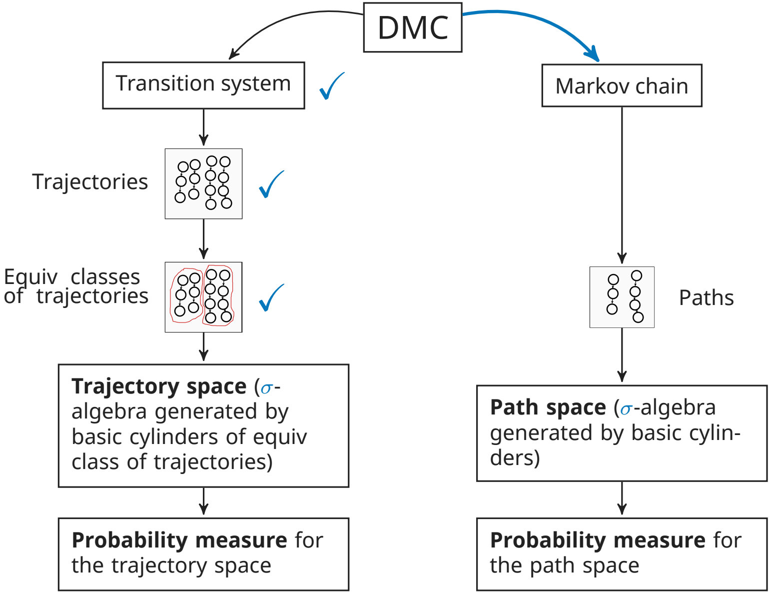
→ $e_t^1 \; I \; e_h^2$ — agent 1 tossing tail and agent 2 tossing head are independent

Equivalence over Event Sequences

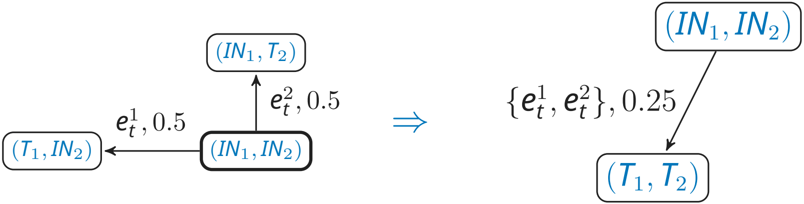


$\rightarrow$ $[e_t^1 e_h^2] = \{e_t^1 e_h^2, e_h^2 e_t^1\}$ — equivalence class over event sequences

Markov Chain Semantics



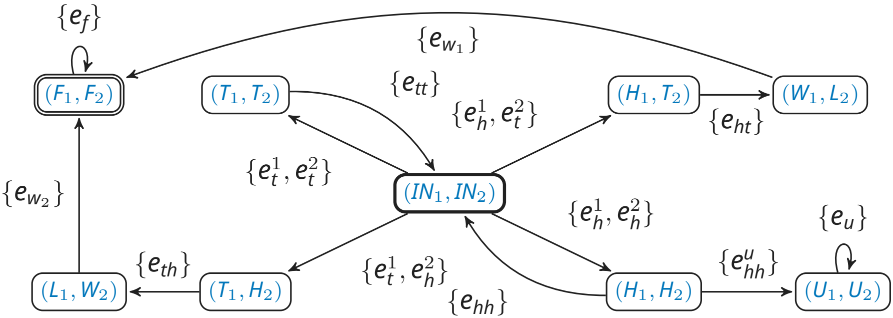
Markov Chain Semantics



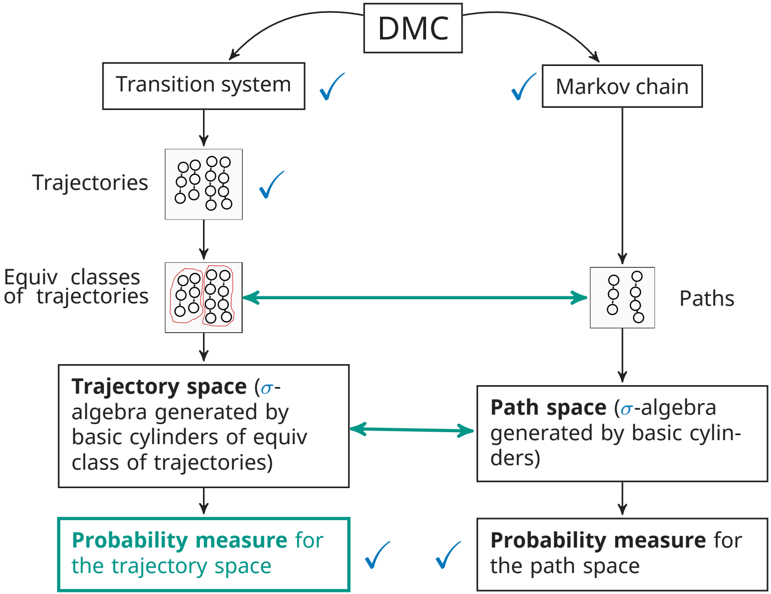
→ $\{e_t^1, e_t^2\}$ is a maximal step at (IN_1, IN_2)

→ The probability of a step is the product of probabilities associated with the events in the step

Coin Toss: Global Markov Chain



Defining the Probability Measure



Theoretical Results

Expressiveness

→ Close connection with Petri nets

→ More expressive than Free-choice

→ Open: But how much more?

Termination Properties

- → Attach non-neg real weights to events
- → Interpret weights: Probability, expected cost and expected time of termination
- → Perform both exact and approximate verification
- → Open: Can we attach time interval to the local/global states?

Syntactic Reduction

→ Reduce the system preserving termination properties

→ Free-choice subclass: can be reduced to summarization

→ Open: Can we identify the reason behind the gap?

Ambitious Open Problems

→ Extend termination properties to full PCTL (or variant)

→ Model partially observable systems

→ Learning parameters with Big Data

Application Domains

Application Domains

→ Stochastic analysis of Business Process Management (BPM) systems

(i) Throughput analysis

(ii) Simulation with statistical guarantee

→ Model distributed cloud computing systems

(i) Model shard-replica systems

(ii) Predict fault-tolerance and eventual consistency

Thank you!

Questions?