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

Agent 1

\[ s'_1 \]
\[ s_1 \]

Agent 2

\[ s'_2 \]
\[ s_2 \]
Probabilistic
Distributed
The Synchronization

→ Joint probabilistic move after the synchronization action
Probabilistic
Distributed
Deterministic
Restriction: This is allowed
Restriction: This is **not** allowed

Agent 1

Agent 2

Agent 3
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
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

→ Two players. Each toss a fair coin ($a_1$ and $a_2$)

→ Both tails: they toss again ($tt$)

→ Both heads:
  (i) they toss again with prob $0.9$ ($hh$), or
  (ii) go to an uncertain state with prob $0.1$ ($u$)

→ Different outcome: who tosses Heads wins ($ht$ and $th$)
DMC: Coin Toss Example

Agent 1

Agent 2
DMC: Coin Toss Example

Agent 1

Agent 2
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

$e_1^1, 0.5$

$(T_1, IN_2) \rightarrow (IN_1, IN_2)$
Global Transition System: Coin Toss

agent 1 tossing T

$\left( T_1, IN_2 \right)$

agent 2 tossing T

$\left( IN_1, T_2 \right)$

$e_t^1, 0.5$

$e_t^2, 0.5$
Global Transition System: Coin Toss

\[ (T_1, T_2) \]
\[ (IN_1, T_2) \]
\[ e_t^1, 0.5 \]
\[ e_t^2, 0.5 \]
\[ e_t^1, 0.5 \]
\[ e_t^2, 0.5 \]
Global Transition System: Coin Toss

\[ (T_1, T_2) \quad e_t^1, 0.5 \quad (IN_1, T_2) \]

\[ (T_1, IN_2) \quad e_t^2, 0.5 \quad (IN_1, T_2) \]

\[ (T_1, IN_2) \quad e_t^1, 0.5 \quad (IN_1, IN_2) \]

\[ (IN_1, IN_2) \quad e_t^2, 0.5 \quad (IN_1, H_2) \]

\[ e_h^2, 0.5 \]

\[ (IN_1, H_2) \]
Global Transition System: Coin Toss

\[
\begin{align*}
&T_1; T_2 
&\mathcal{E}_t, 0.5
&(IN_1; T_2) 
&\mathcal{E}_t, 0.5
&T_1; IN_2 
&\mathcal{E}_t, 0.5
&(IN_1; IN_2) 
&\mathcal{E}_t, 0.5
&T_1; H_2 
&\mathcal{E}_t, 0.5
&(IN_1; H_2) 
&\mathcal{E}_t, 0.5
\end{align*}
\]
Global Transition System: Coin Toss

\begin{align*}
\mathbb{G} &\equiv (\mathbb{T}, \mathbb{E}, \mathbb{I}) \\
\mathbb{T} &\equiv \{T_1, T_2, H_1, H_2\} \\
\mathbb{E} &\equiv \{e_1^1, e_1^2, e_2^1, e_2^2\} \\
\mathbb{I} &\equiv \{I_1, I_2\}
\end{align*}

\begin{align*}
(T_1, T_2) &\xrightarrow{e_1^1, 0.5} (IN_1, T_2) & (IN_1, T_2) &\xrightarrow{e_2^2, 0.5} (IN_1, IN_2) \\
(T_1, T_2) &\xleftarrow{e_1^2, 0.5} (IN_1, T_2) & (IN_1, T_2) &\xleftarrow{e_2^1, 0.5} (IN_1, IN_2) \\
(T_1, T_2) &\xleftarrow{e_1^1, 0.5} (H_1, T_2) & (H_1, T_2) &\xleftarrow{e_2^2, 0.5} (H_1, IN_2) \\
(T_1, T_2) &\xrightarrow{e_1^2, 0.5} (H_1, T_2) & (H_1, T_2) &\xrightarrow{e_2^1, 0.5} (H_1, H_2)
\end{align*}

(both agents tossed)
Global Transition System: Coin Toss

(unmarked events have probability 1)
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

Markov chain

Trajectories

Equiv classes of trajectories

Trajectory space \((\sigma\text{-algebra generated by basic cylinders of equiv class of trajectories})\)

Probability measure for the trajectory space

Path space \((\sigma\text{-algebra generated by basic cylinders})\)

Probability measure for the path space
Equivalence Classes of Trajectories

- Transition system
- Markov chain

Equiv classes of trajectories

Trajectory space ($\sigma$-algebra generated by basic cylinders of equiv class of trajectories)

Probability measure for the trajectory space

Path space ($\sigma$-algebra generated by basic cylinders)

Probability measure for the path space
Independence over Events

$\rightarrow e_t^1 I e_h^2$ — agent 1 tossing tail and agent 2 tossing head are independent
Equivalence over Event Sequences

\[ e_1^1, 0.5 \]
\[ e_2^2, 0.5 \]
\[ e_1^1, 0.5 \]
\[ e_2^2, 0.5 \]

\[ \rightarrow [e_1^1 e_2^2] = \{ e_1^1 e_2^2, e_2^2 e_1^1 \} \] — equivalence class over event sequences
Markov Chain Semantics

Transition system

Trajectories

Equiv classes of trajectories

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

Probability measure for the trajectory space

DMC

Markov chain

Paths

Path space (\(\sigma\)-algebra generated by basic cylinders)

Probability measure for the path space
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

\[
\begin{align*}
(F_1, F_2) & \quad \{e_f\} \\
(T_1, T_2) & \quad \{e_{tt}\} \\
(H_1, T_2) & \quad \{e_{ht}\} \\
(U_1, U_2) & \quad \{e_u\} \\
(L_1, W_2) & \quad \{e_{th}\} \\
(T_1, H_2) & \quad \{e_{t1}, e^{2}_{t}\} \\
(H_1, H_2) & \quad \{e^{1}_{h}, e^{2}_{h}\} \\
(IN_1, IN_2) & \quad \{e_{w1}\} \\
(W_1, L_2) & \quad \{e_{w2}\} \\
\end{align*}
\]
Defining the Probability Measure

DMC

Transition system ✓

Markov chain ✓

Trajectories ✓

Equiv classes of trajectories

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

Path space (σ-algebra generated by basic cylinders)

Probability measure for the trajectory space ✓ ✓

Probability measure for the path space
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?