PDT Logic

A Probabilistic Doxastic Temporal Logic for Reasoning about Beliefs in Multi-agent Systems

Dissertation Presentation

Dipl.-Ing. Karsten Martiny

Chairman:	Prof. Dr. Stefan Fischer
Reviewers:	Prof. Dr. Ralf Möller
	Prof. Dr. Rüdiger Reischuk

October 4, 2016

UNIVERSITÄT ZU LÜBECK
INSTITUTE OF INFORMATION SYSTEMS

PDT Logic

- A representation formalism to reason about probabilistic beliefs over time in multi-agent systems
- Agents' beliefs are quantified with imprecise probabilities (i.e., probability intervals)
- Time is modeled in discrete steps for a finite set of time points
- Agents' subjective beliefs change upon observing facts

Main Contribution

- Combine and extend results from different fields of formal logic
  - Temporal Logic [SPSS11]
  - Epistemic Logic [FHVM95]
  - Probabilistic Dynamic Epistemic logic [Koo03]
- Create a semantically rich representation formalism for beliefs [MM15a]
- Develop specialized decision procedures [MM16a]

[FHMV95] R. Fagin, J. Halpern, Y. Moses, M. Vardi: **Reasoning About Knowledge** *MIT Press, 1995*

[Koo03] B. Kooi: **Probabilistic Dynamic Epistemic Logic** *Journal of Logic, Language and Information, Volume 12, pages 381-408, September 2003*

[SPSS11] P. Shakarian, A. Parker, G. Simari, V. S. Subrahmanian: **Annotated Probabilistic Temporal Logic**, *ACM Transactions on Computational Logic, Volume 13, pages 1-33, April 2012*

[MM15a] K. Martiny, R. Möller: **A Probabilistic Doxastic Temporal Logic for Reasoning about Beliefs in Multi-agent Systems** *7th International Conference on Agents and Artificial Intelligence (ICAART), Lisbon, Portugal, 2015*

[MM16a] K. Martiny, R. Möller: **PDT Logic: A Probabilistic Doxastic Temporal Logic for Reasoning about Beliefs in Multi-agent Systems**, *Journal of Artificial Intelligence Research (JAIR), Volume 57, pages 39-112, September 2016*

Representation

- Describing possible worlds
  - A propositional language describes *ontic facts*
  - Observation Atoms $Obs_{\mathcal{G}}(F)$ specify that a group of agents $\mathcal{G}$ observes some ontic fact F
- Time
  - Temporal evolution $\Leftrightarrow$ sequence of possible worlds (*thread* Th)
  - Probabilistic temporal relations expressed as *temporal rules* using *frequency functions*
- Probabilistic Beliefs
  - Each thread Th has a prior probability (*"interpretation"*) $\mathcal{I}(Th)$
  - Probabilistic beliefs depend on observations of the respective agent $\Rightarrow$ different threads yield different belief evolutions

Interpretation Updates

Example: two agents 1, 2, six threads $Th_1, \ldots, Th_6$:

$\mathcal{I}(Th_1) = 0.1$

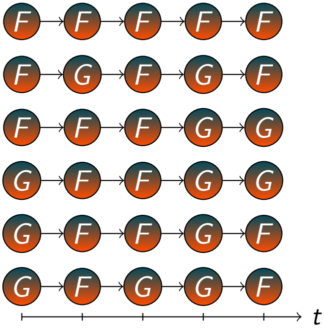
$\mathcal{I}(Th_2) = 0.2$

$\mathcal{I}(Th_3) = 0.1$

$\mathcal{I}(Th_4) = 0.3$

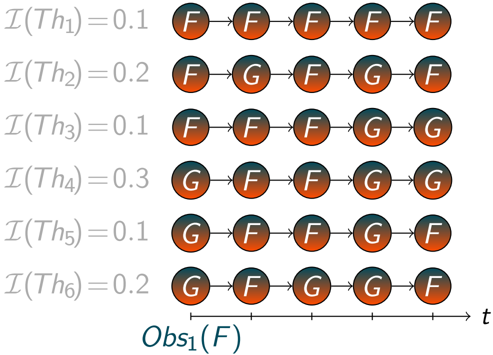
$\mathcal{I}(Th_5) = 0.1$

$\mathcal{I}(Th_6) = 0.2$



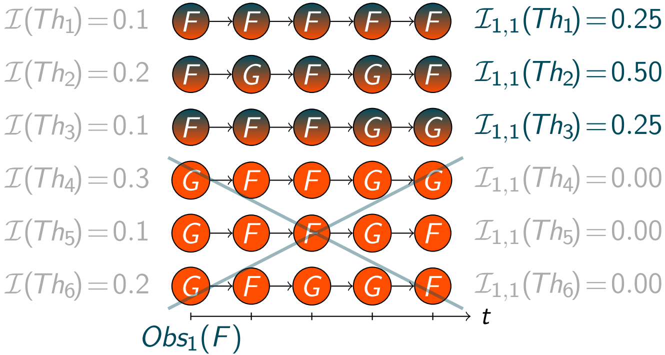
Interpretation Updates

Example: two agents 1, 2, six threads $Th_1, \ldots, Th_6$:



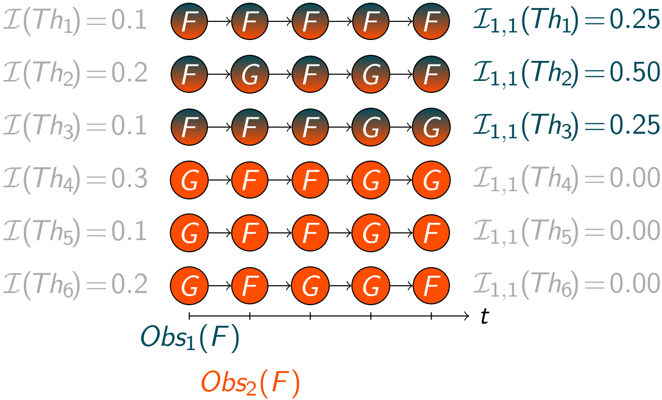
Interpretation Updates

Example: two agents 1, 2, six threads $Th_1, \ldots, Th_6$:



Interpretation Updates

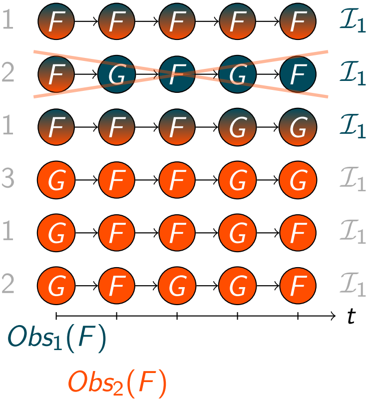
Example: two agents 1, 2, six threads $Th_1, \dots, Th_6$:



Interpretation Updates

Example: two agents 1, 2, six threads $Th_1, \ldots, Th_6$:

$\mathcal{I}(Th_1) = 0.1$	$\mathcal{I}_{1,2}(Th_1) = 0.25$	$\mathcal{I}_{2,2}(Th_1) = 0.125$
$\mathcal{I}(Th_2) = 0.2$	$\mathcal{I}_{1,2}(Th_2) = 0.50$	$\mathcal{I}_{2,2}(Th_2) = 0.000$
$\mathcal{I}(Th_3) = 0.1$	$\mathcal{I}_{1,2}(Th_3) = 0.25$	$\mathcal{I}_{2,2}(Th_3) = 0.125$
$\mathcal{I}(Th_4) = 0.3$	$\mathcal{I}_{1,2}(Th_4) = 0.00$	$\mathcal{I}_{2,2}(Th_4) = 0.375$
$\mathcal{I}(Th_5) = 0.1$	$\mathcal{I}_{1,2}(Th_5) = 0.00$	$\mathcal{I}_{2,2}(Th_5) = 0.125$
$\mathcal{I}(Th_6) = 0.2$	$\mathcal{I}_{1,2}(Th_6) = 0.00$	$\mathcal{I}_{2,2}(Th_6) = 0.250$



Interpretation Updates

Example: two agents 1, 2, six threads $Th_1, \ldots, Th_6$:

$\mathcal{I}(Th_1) = 0.1$

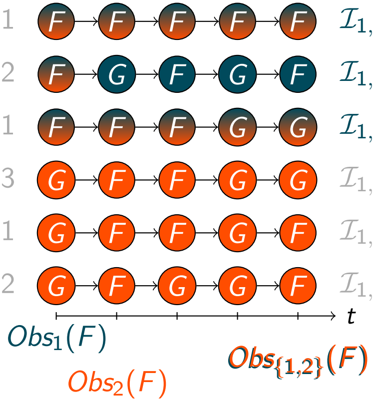
$\mathcal{I}(Th_2) = 0.2$

$\mathcal{I}(Th_3) = 0.1$

$\mathcal{I}(Th_4) = 0.3$

$\mathcal{I}(Th_5) = 0.1$

$\mathcal{I}(Th_6) = 0.2$



$\mathcal{I}_{1,2}(Th_1) = 0.25$ $\mathcal{I}_{2,2}(Th_1) = 0.125$

$\mathcal{I}_{1,2}(Th_2) = 0.50$ $\mathcal{I}_{2,2}(Th_2) = 0.000$

$\mathcal{I}_{1,2}(Th_3) = 0.25$ $\mathcal{I}_{2,2}(Th_3) = 0.125$

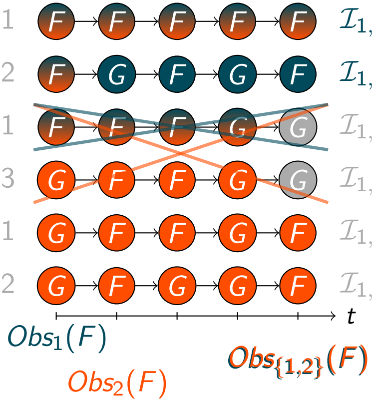
$\mathcal{I}_{1,2}(Th_4) = 0.00$ $\mathcal{I}_{2,2}(Th_4) = 0.375$

$\mathcal{I}_{1,2}(Th_5) = 0.00$ $\mathcal{I}_{2,2}(Th_5) = 0.125$

$\mathcal{I}_{1,2}(Th_6) = 0.00$ $\mathcal{I}_{2,2}(Th_6) = 0.250$

Interpretation Updates

Example: two agents 1, 2, six threads $Th_1, \ldots, Th_6$:



$\mathcal{I}(Th_1) = 0.1$	$\mathcal{I}_{1,5}(Th_1) \approx 0.33$	$\mathcal{I}_{2,5}(Th_1) = 0.25$
$\mathcal{I}(Th_2) = 0.2$	$\mathcal{I}_{1,5}(Th_2) \approx 0.67$	$\mathcal{I}_{2,5}(Th_2) = 0.00$
$\mathcal{I}(Th_3) = 0.1$	$\mathcal{I}_{1,5}(Th_3) = 0.00$	$\mathcal{I}_{2,5}(Th_3) = 0.00$
$\mathcal{I}(Th_4) = 0.3$	$\mathcal{I}_{1,5}(Th_4) = 0.00$	$\mathcal{I}_{2,5}(Th_4) = 0.00$
$\mathcal{I}(Th_5) = 0.1$	$\mathcal{I}_{1,5}(Th_5) = 0.00$	$\mathcal{I}_{2,5}(Th_5) = 0.25$
$\mathcal{I}(Th_6) = 0.2$	$\mathcal{I}_{1,5}(Th_6) = 0.00$	$\mathcal{I}_{2,5}(Th_6) = 0.50$

Belief Operators - Definitions

Agents can have beliefs of three different types, all quantified with a probability interval $[\ell, u]$, seen from thread Th':

- Belief in facts $B_{i,t'}^{\ell,u}(F_t)$:

$$\ell \leq \sum_{Th:\, Th(t) \models F} \mathcal{I}_{i,t'}^{Th'}(Th) \leq u$$

- Belief in temporal rules $B_{i,t'}^{\ell,u}(r_{\Delta t}^{\mathrm{fr}}(F, G))$:

$$\ell \leq \sum_{Th} \mathcal{I}_{i,t'}^{Th'}(Th) \cdot \mathrm{fr}(Th, F, G, \Delta t) \leq u$$

- Nested beliefs $B_{i,t'}^{\ell,u}(B_{j,t}^{\ell_j,u_j}(\cdot))$:

$$\ell \leq \sum_{Th,\ \mathcal{I}_{j,t}^{Th} \models B_{j,t}^{\ell_j,u_j}(\cdot)} \mathcal{I}_{i,t'}^{Th'}(Th) \leq u$$

Belief Operators - Example

$\mathcal{I}(Th_1) = 0.1$ — $\mathcal{I}_{1,2}(Th_1) = 0.25$

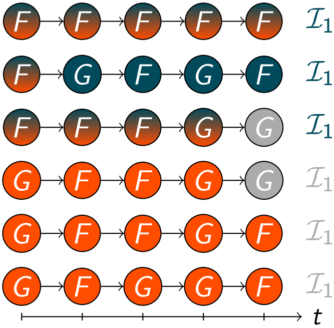
$\mathcal{I}(Th_2) = 0.2$ — $\mathcal{I}_{1,2}(Th_2) = 0.50$

$\mathcal{I}(Th_3) = 0.1$ — $\mathcal{I}_{1,2}(Th_3) = 0.25$

$\mathcal{I}(Th_4) = 0.3$ — $\mathcal{I}_{1,2}(Th_4) = 0.00$

$\mathcal{I}(Th_5) = 0.1$ — $\mathcal{I}_{1,2}(Th_5) = 0.00$

$\mathcal{I}(Th_6) = 0.2$ — $\mathcal{I}_{1,2}(Th_6) = 0.00$



$$B_{1,2}^{.1,.3}(G_5)$$

Belief Operators - Example

$\mathcal{I}(Th_1) = 0.1$

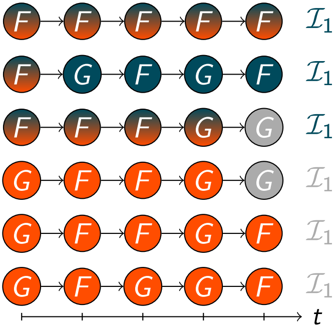
$\mathcal{I}(Th_2) = 0.2$

$\mathcal{I}(Th_3) = 0.1$

$\mathcal{I}(Th_4) = 0.3$

$\mathcal{I}(Th_5) = 0.1$

$\mathcal{I}(Th_6) = 0.2$



$\mathcal{I}_{1,2}(Th_1) = 0.25$

$\mathcal{I}_{1,2}(Th_2) = 0.50$

$\mathcal{I}_{1,2}(Th_3) = 0.25$

$\mathcal{I}_{1,2}(Th_4) = 0.00$

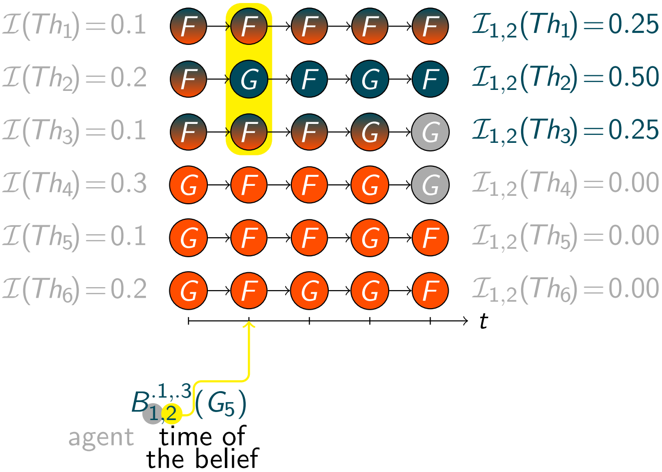
$\mathcal{I}_{1,2}(Th_5) = 0.00$

$\mathcal{I}_{1,2}(Th_6) = 0.00$

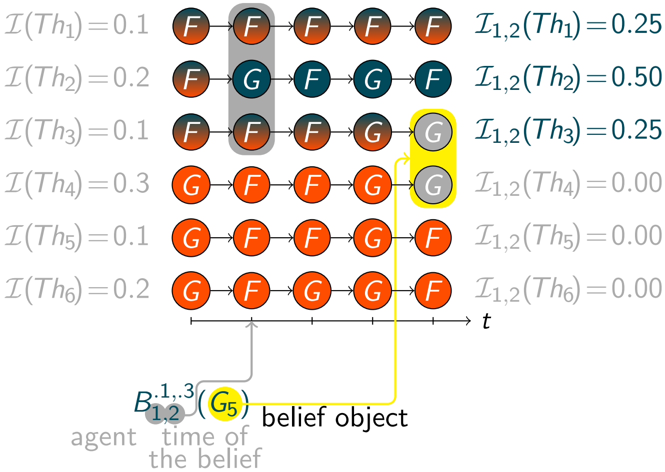
$B^{.1,.3}_{1,2}(G_5)$

agent

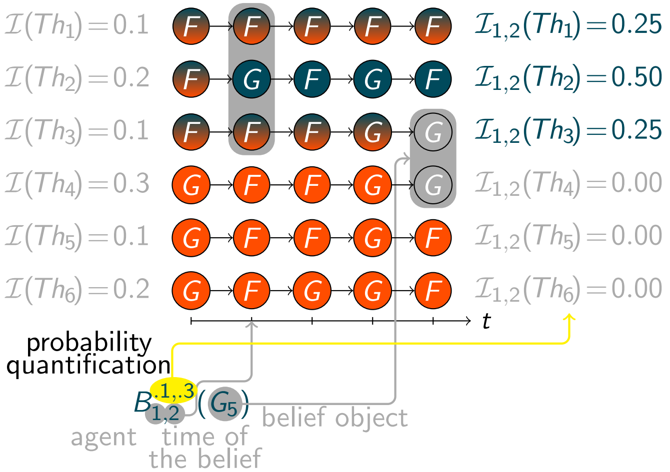
Belief Operators - Example



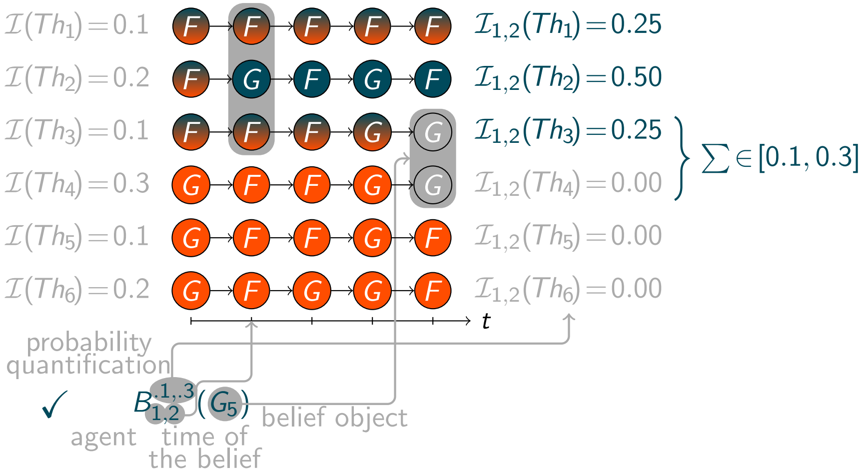
Belief Operators - Example



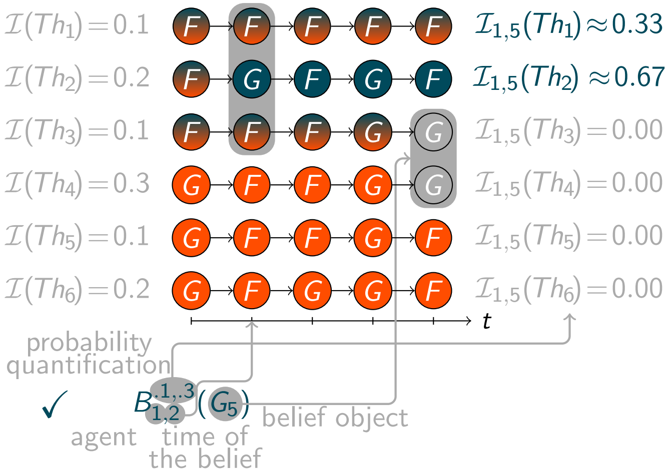
Belief Operators - Example



Belief Operators - Example

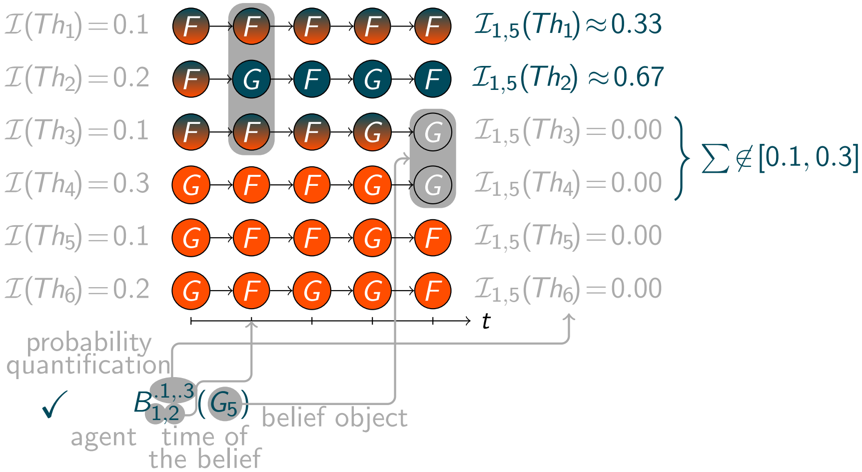


Belief Operators - Example



$B_{1,5}^{.1,.3}(G_5)$

Belief Operators - Example



✓ $B_{1,2}^{.1,.3}(G_5)$

✗ $B_{1,5}^{.1,.3}(G_5)$

Satisfiability Checking

- Given: a set of belief formulae $\mathfrak{B}$
- Goal: check satisfiability of $\mathfrak{B}$ (w.r.t. a specified problem)
- A possible problem specification: [MM15a]
  - Exhaustive set of threads $\mathcal{T}$
  - Prior probabilities $\mathcal{I}$

\+ Easy to perform (*PTIME*)

− Specification is very large ($\Rightarrow$ restricted applicability)

[MM15a] K. Martiny, R. Möller: **A Probabilistic Doxastic Temporal Logic for Reasoning about Beliefs in Multi-agent Systems**
7th International Conference on Agents and Artificial Intelligence (ICAART), Lisbon, Portugal, 2015

Satisfiability Checking

- Alternative problem specification: encode all information in $\mathfrak{B}$ [MM16a]
- Determine possible threads $\mathcal{T}$ automatically
- Transform $\mathcal{T}$ and $\mathfrak{B}$ to a *0-1 Mixed Integer Linear Program (LP)*
- LP has a solution $\Leftrightarrow$ $\mathfrak{B}$ is satisfiable

\+ Specification is small

− Poor *worst-case* complexity (*EXPSPACE*)

[MM16a] K. Martiny, R. Möller: **PDT Logic: A Probabilistic Doxastic Temporal Logic for Reasoning about Beliefs in Multi-agent Systems**, *Journal of Artificial Intelligence Research (JAIR), Volume 57, pages 39-112, September 2016*

Satisfiability Checking

Optimization

- Existence of a model determines satisfiability
- Explore the search space step by step
- Test corresponding LPs for each step
- Major challenge: The semantics prevents pruning
- Use dependency-directed search heuristics for exploration
- Limit the search space to *intended models*

What could not be addressed here...

Thesis contents not covered in the talk:

- Formal analysis of the logic's properties [MM15a],[MM16a]
- Temporal relations (*"frequency functions"*) [MM15a],[MM16a]
- Detailed discussion of application scenarios
  - Cyber security [MMM15]
  - Stock markets [MM15b]
- Extension of the temporal model to infinite streams [MM14]
- Abductive reasoning [MM15b]

Publications

[MM14] Karsten Martiny and Ralf Möller:
PDT Logic for Stream Reasoning in Multi-agent Systems
6th International Symposium on Symbolic Computation in Software Science (SCSS), Tunis, Tunisia, 2014

[MM15a] Karsten Martiny and Ralf Möller:
A Probabilistic Doxastic Temporal Logic for Reasoning about Beliefs in Multi-agent Systems
7th International Conference on Agents and Artificial Intelligence (ICAART), Lisbon, Portugal, 2015

[MMM15] Karsten Martiny, Alexander Motzek, and Ralf Möller:
Formalizing Agents' Beliefs for Cyber-Security Defense Strategy Planning
8th International Conference on Computational Intelligence in Security for Information Systems, Burgos, Spain, 2015

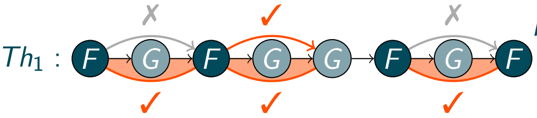
[MM15b] Karsten Martiny and Ralf Möller:
Abduction in PDT Logic
28th Australasian Conference on Artificial Intelligence (AI), Canberra, Australia, 2015

[MM16a] Karsten Martiny and Ralf Möller:
PDT Logic: A Probabilistic Doxastic Temporal Logic for Reasoning about Beliefs in Multi-agent Systems
Journal of Artificial Intelligence Research (JAIR), Volume 57, pages 39-112, September 2016

[MM16b] Karsten Martiny and Ralf Möller:
Reasoning about Imprecise Beliefs in Multi-Agent Systems
accepted for publication in
KI Zeitschrift - Special Issue on Challenges for Reasoning under Uncertainty, Inconsistency, Vagueness, and Preferences

Frequency Functions

- Point frequency function *pfr*
  - Expresses how frequently some event F is followed by another event G in *exactly* Δt time units
  - $pfr(Th, F, G, \Delta t) = \frac{|\{t : Th(t) \models F \wedge Th(t+\Delta t) \models G\}|}{|\{t : (t \leq t_{max} - \Delta t) \wedge Th(t) \models F\}|}$
- Existential frequency function *efr*
  - Expresses how frequently some event F is followed by another event G *within* Δt time units
  - $efr(Th, F, G, \Delta t) = \frac{efn(Th, F, G, \Delta t, 0, t_{max})}{|\{t : (t \leq t_{max} - \Delta t) \wedge Th(t) \models F\}| + efn(Th, F, G, \Delta t, t_{max} - \Delta t, t_{max})}$
    with $efn(Th, F, G, \Delta t, t_1, t_2) = |\{t : (t_1 < t \leq t_2) \wedge Th(t) \models F$
    $\wedge \exists t' \in [t, \min(t_2, t + \Delta t)] \; (Th(t') \models G)\}|$
- Example:



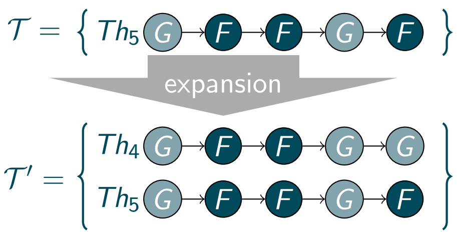
$pfr(Th_1, F, G, 2) = \frac{1}{3}$

$efr(Th_1, F, G, 2) = 1$

Semantic Challenge for Decision Procedures

- Example: Determine satisfiability for
  - $\mathfrak{B} = \{B^{1,1}_{1,0}(r^{pfr}_1(G,F))\}$ *("G is always directly followed by F")*
  - $\mathfrak{B}' = \{B^{0.6,\ 0.9}_{1,0}(r^{pfr}_1(G,F))\}$ *("the probability that G is directly followed by F is between 0.6 and 0.9")*

- step-wise satisfiability checking:



$sat(\mathfrak{B})$ ✓ (with $\mathcal{I}(Th_5) = 1$)
$\neg sat(\mathfrak{B}')$ ✗

$sat(\mathfrak{B}')$ ✓ (e.g., with $\mathcal{I}(Th_4) = 0.5$, and $\mathcal{I}(Th_5) = 0.5$)