Making Agents Less Godlike

Christian Muise

Computing and Information Systems,
Melbourne School of Engineering
The University of Melbourne

August 17, 2015

Research Team for "Foundations of Human-Agent Collaboration: Situation-Relevant Information Sharing"

Tim Miller

Adrian Pearce

Liz Sonenberg

Paolo Felli

Christian Muise

Removing The Omni-* Properties

(~~Omniscience~~) Multi-agent Epistemic Planning

We formally characterize a notion of multi-agent epistemic planning, and demonstrate how to solve a rich subclass of these problems using classical planning techniques.

(~~Omnipotence~~) Multi-agent Planning as FOND

We extend a non-deterministic planner to plan in a multi-agent setting, given the goals and possible actions of other agents to plan for what is *plausible*.

Example Goal: Deception

Make Bob believe Sue believes the switch is off, when in fact Bob believes that it is on: $\{B_{Bob}B_{Sue}\neg switch\_on, B_{Bob}switch\_on\}$

Example Action: Gossiping

Precondition for $share(Bob, secret, roomA)$ includes that Bob believes the secret: $B_{Bob}secret$. **Effects** indicate who perceives the gossip (and who is aware of this): $in(Sue, roomA) \rightarrow B_{Sue}secret$, $B_{Joe}in(Sue, roomA) \wedge in(Joe, roomA) \rightarrow B_{Joe}B_{Sue}secret, \ldots$

Trade-offs for Choice of Knowledge Base

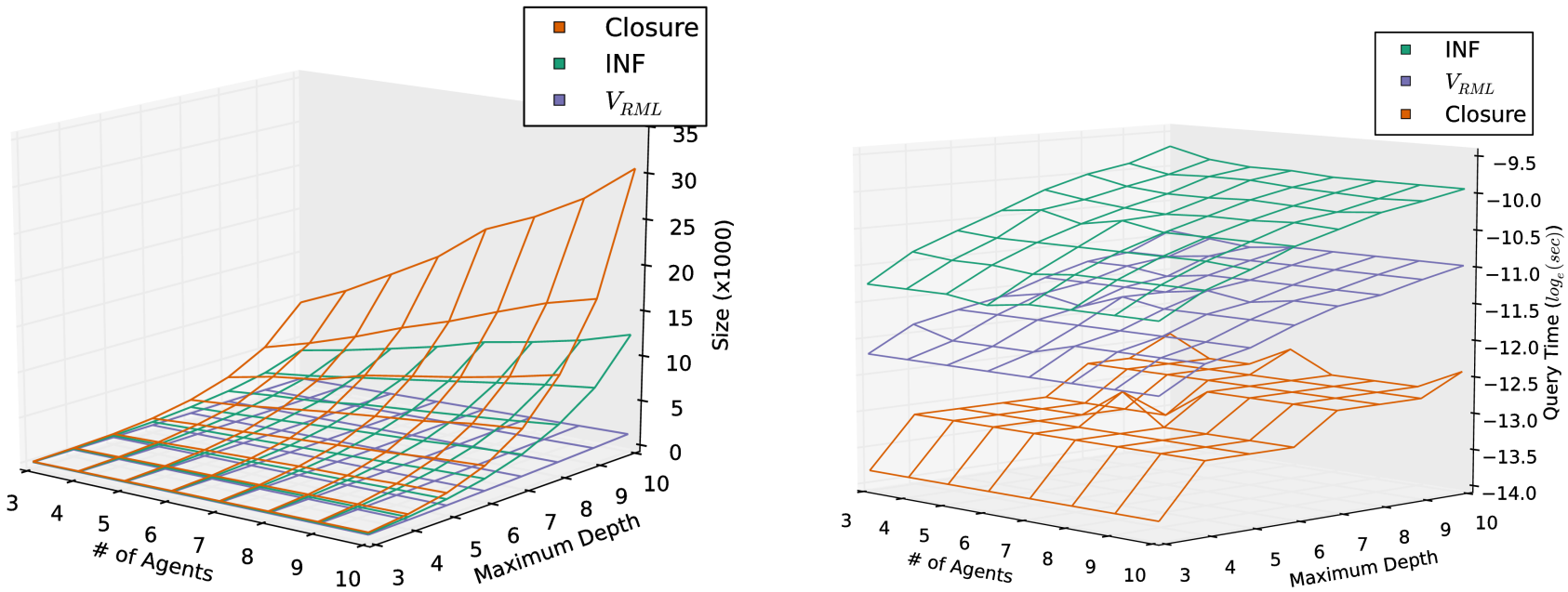
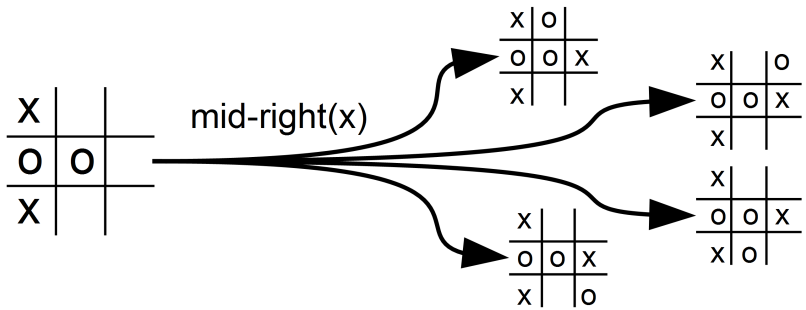


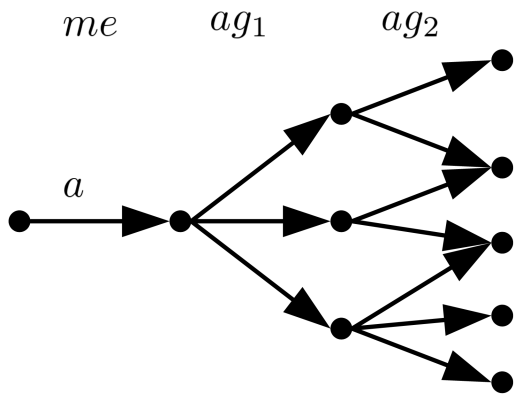
Figure: Size of knowledge base (largest on top) and average query time (slowest on top) for the three types of knowledge bases.

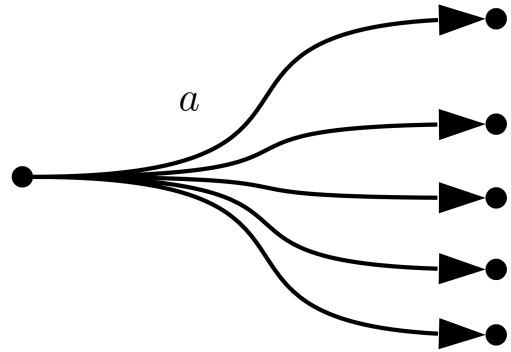
Multi-agent Planning as FOND

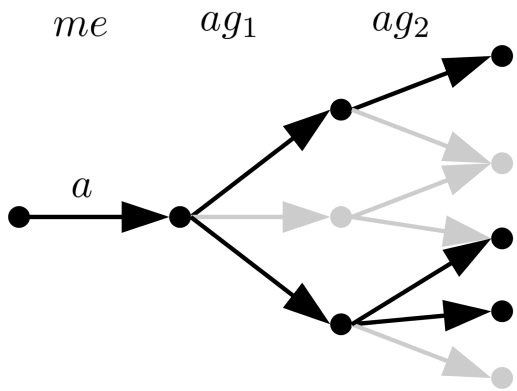


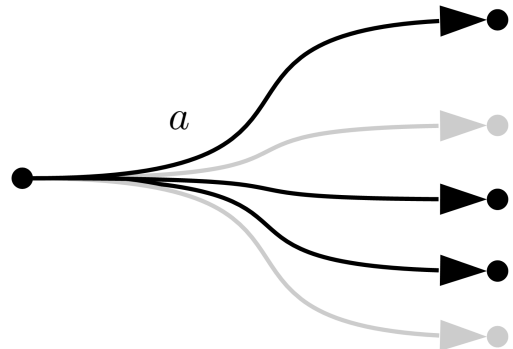
- Perspectival view on multi-agent planning
- Leverage the power of modern FOND planning
- Seamlessly handle a mix of agnostic, collaborative, and combative agents

Reinterpreting Multi-agent Actions

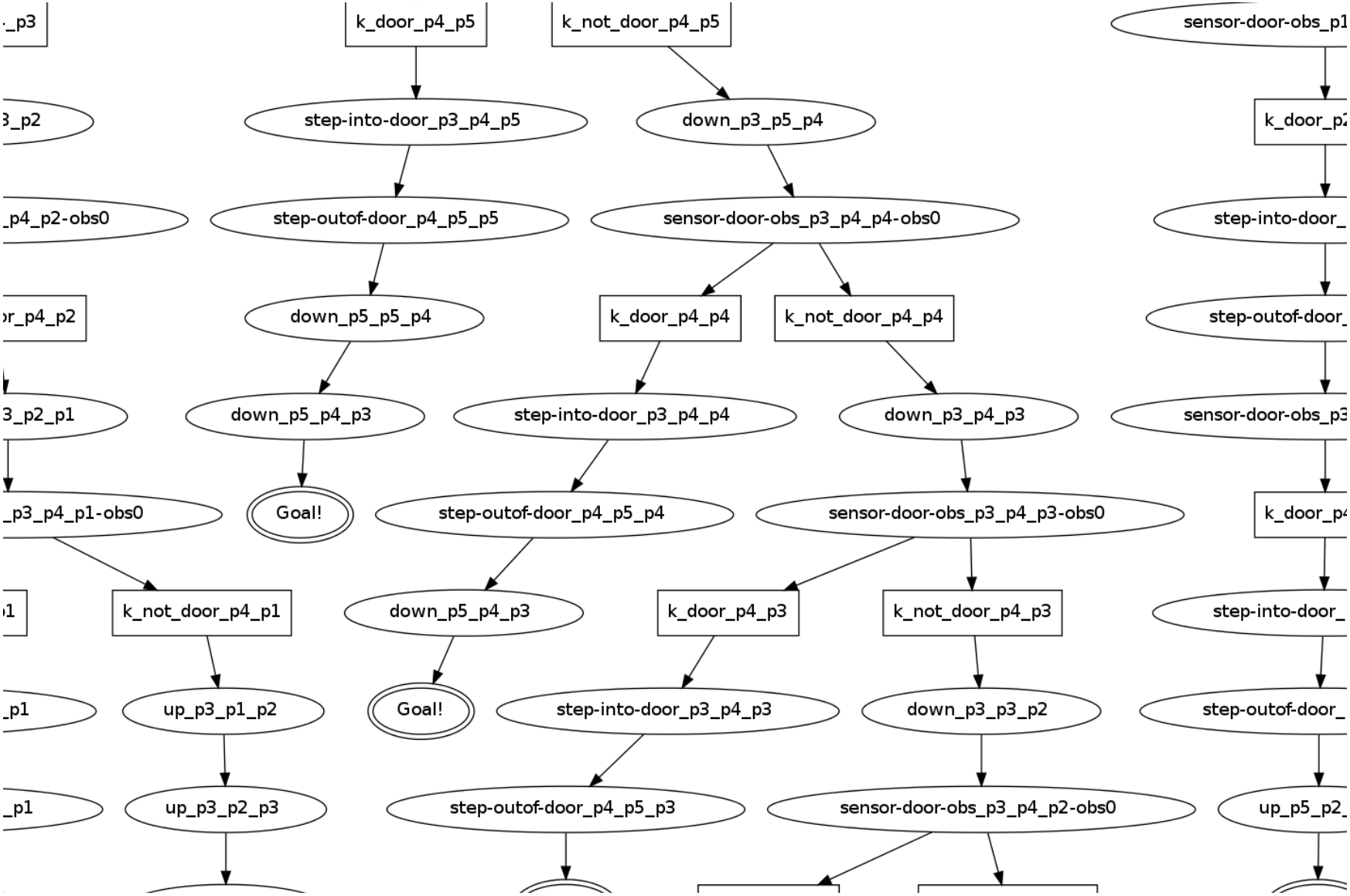




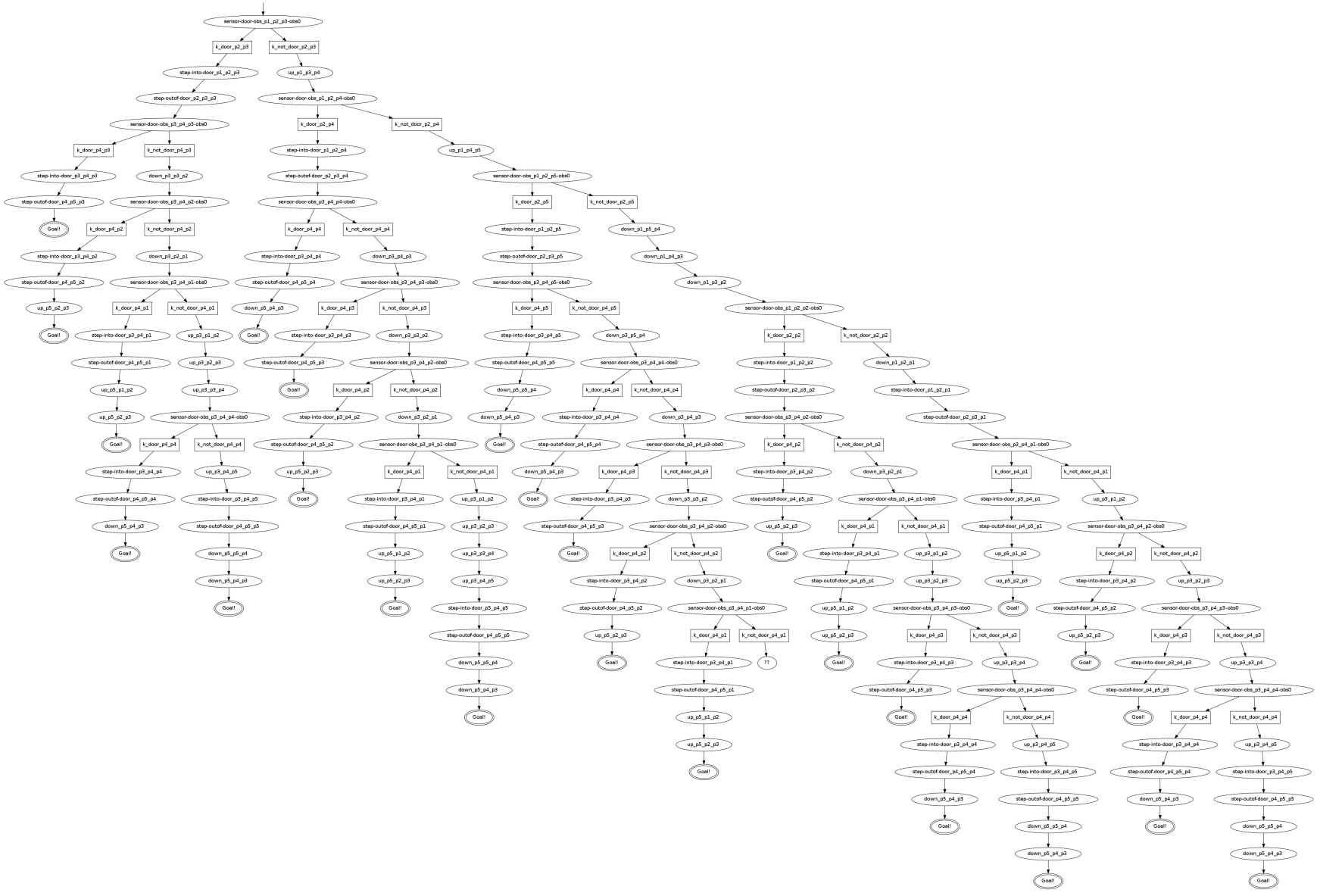




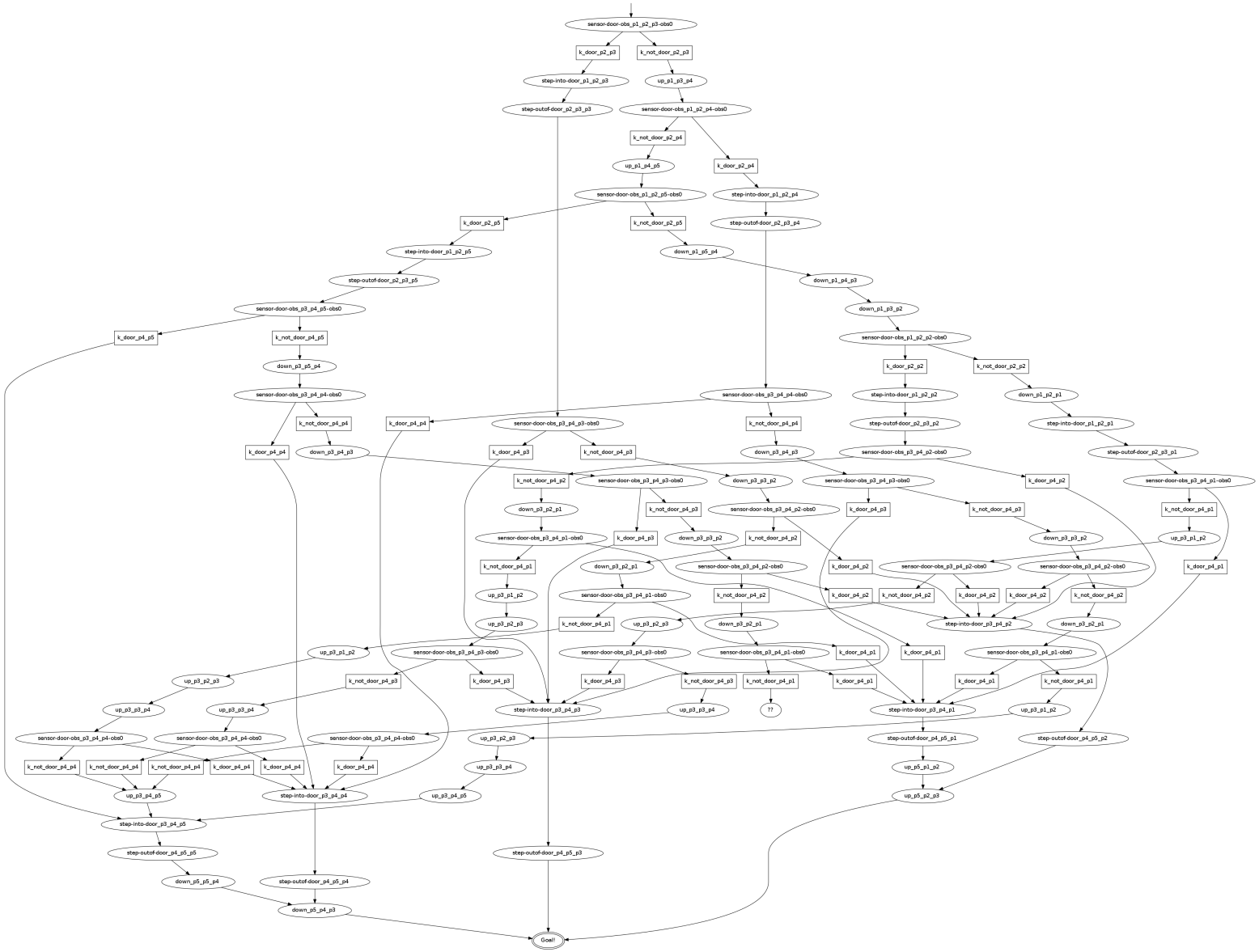
This is a Contingent Plan



This is a Contingent Plan



This is a Contingent Plan

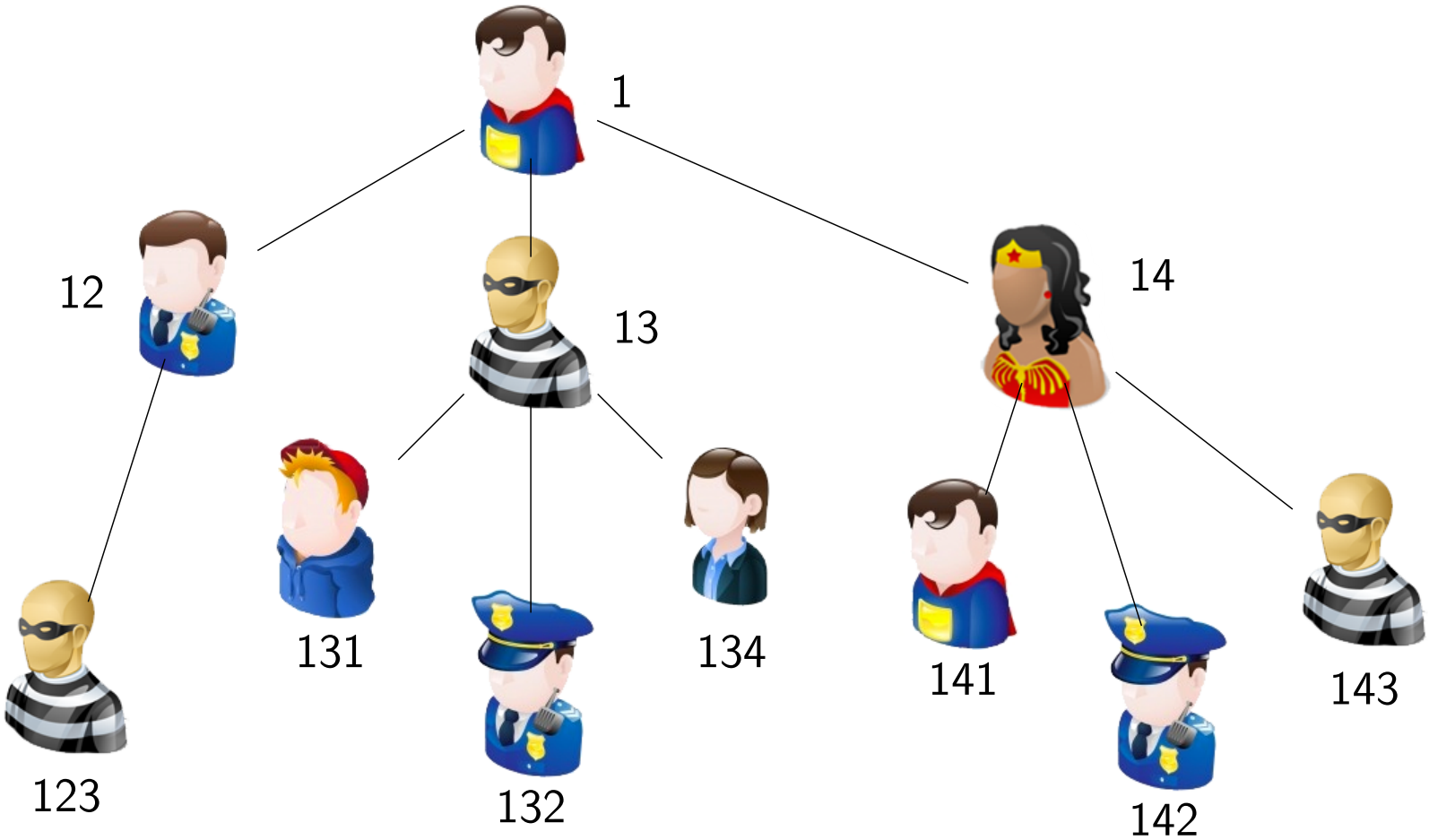


Paolo Felli (PhD from Università di Roma)

1. Modelling: agent models
2. Reasoning: empathetic and stereotypical
3. Social: acceptable behaviours

Nested modelling

A scenario with four physical agents : $\{\underline{\mathbf{1}}, 2, 3, 4\}$.



Example: agent "13" represents the modelling that agent 1 uses for representing agent 3.

Stereotypical and empathetic reasoning.

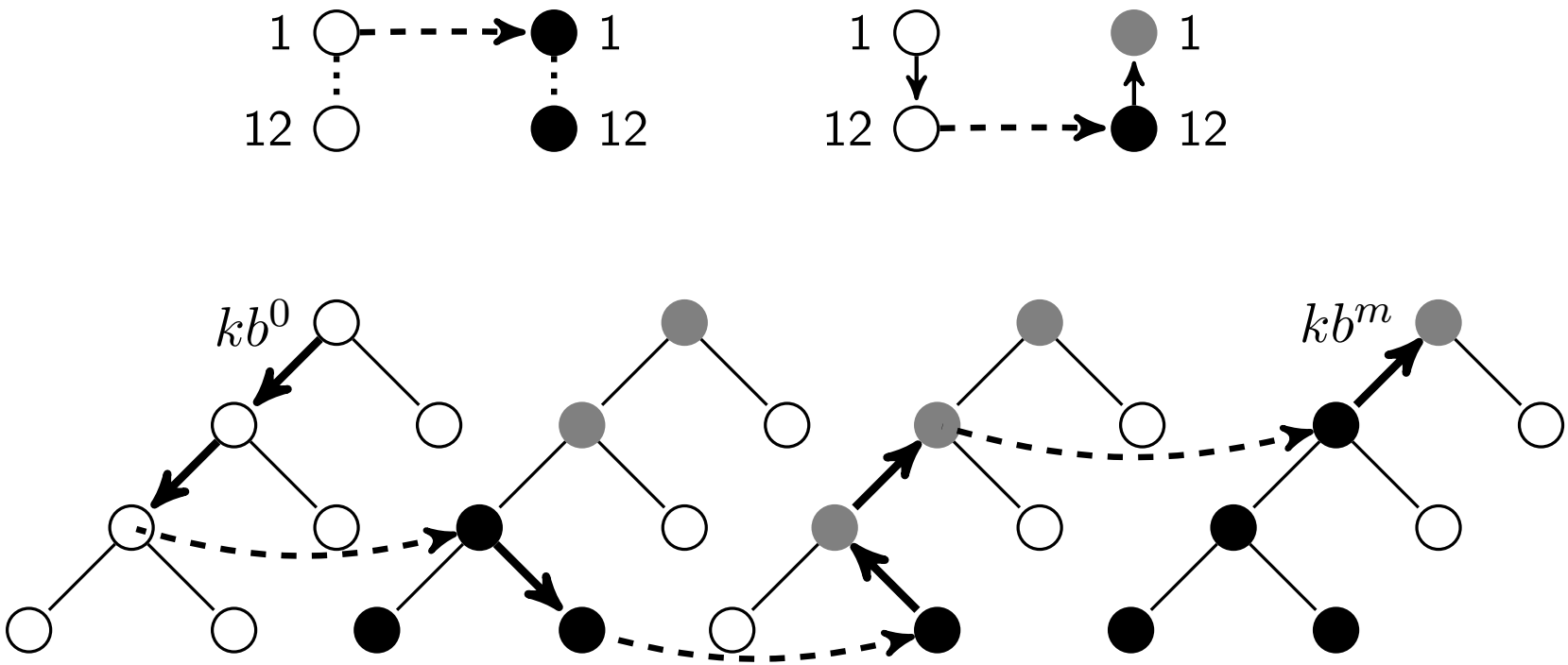


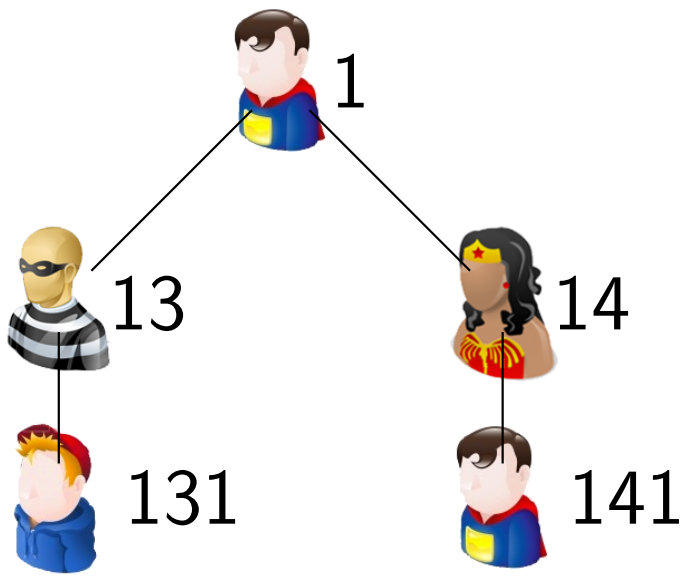
Figure: Expansion: a *path* in the tree of agent models

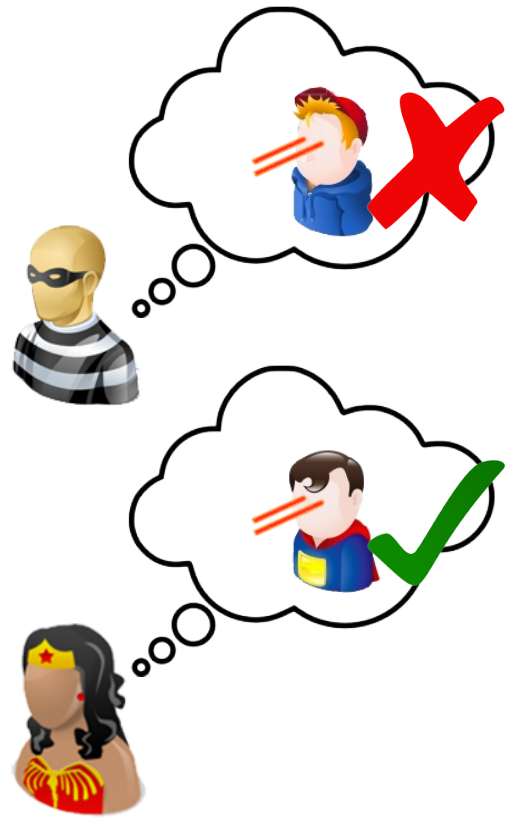
Agent models can be used not only for defining the reasoning of a perspectival agent, but also to define the set of *acceptable* behaviours in the social context. Acceptable $\equiv$ "that makes sense".

$\Rightarrow$ **Simulate** and check possible executions:

$\rightarrow$ deception

$\rightarrow$ etiquette





$\Rightarrow$ Perform model-checking to **synthesize** acceptable strategies.

Any Questions?

Project Page

`http://agentlab.cis.unimelb.edu.au/project-hac.html`

Personal Research Page

`http://www.haz.ca/research.html`