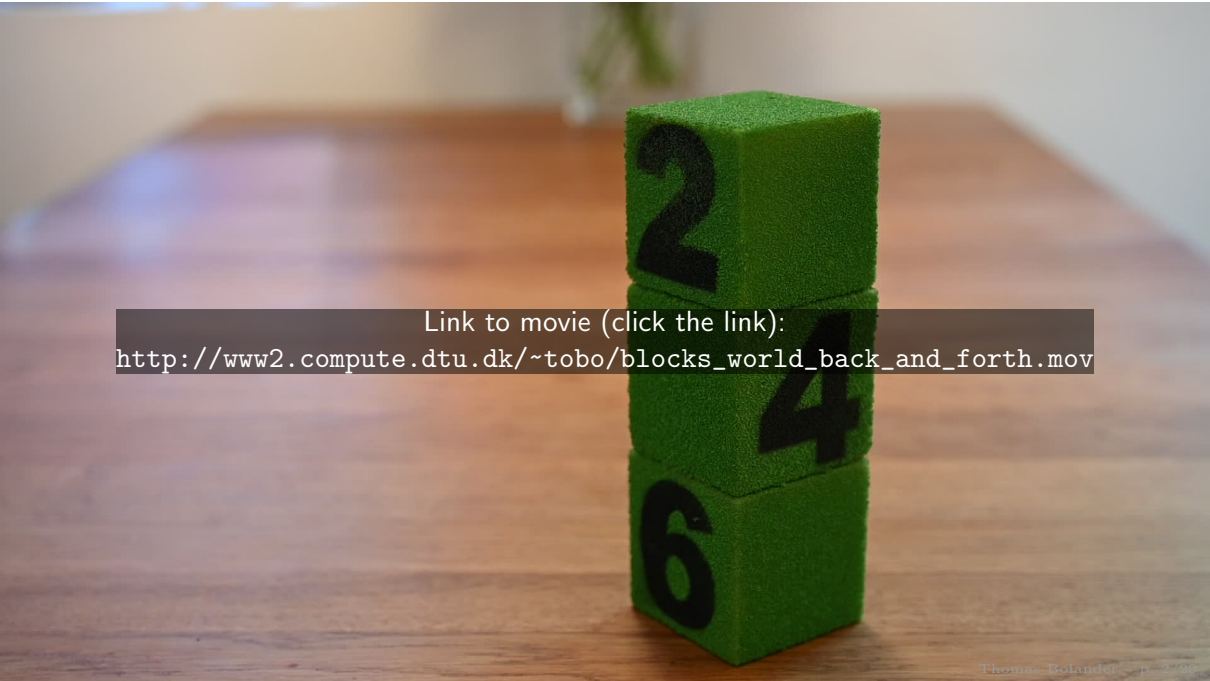
Implicitly Coordinating Agents

Thomas Bolander, DTU Compute

PR-BGI 2026

DTU Compute
Institut for Matematik og Computer Science



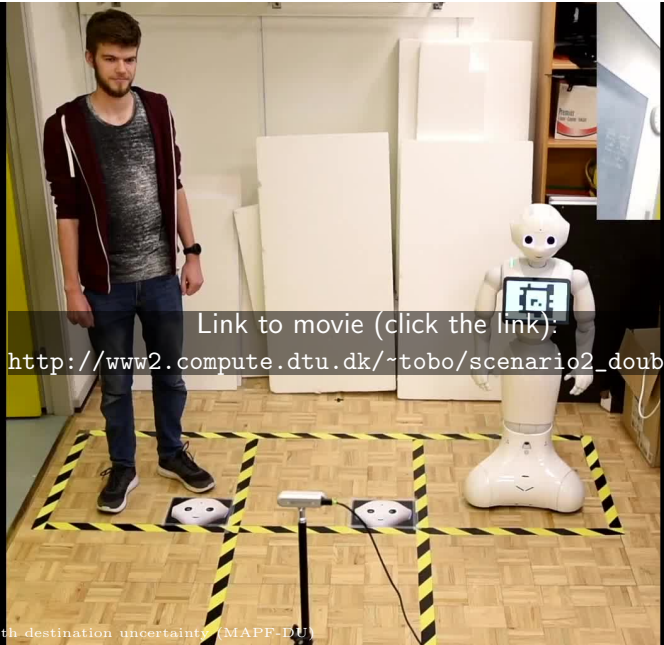
Link to movie (click the link):
http://www2.compute.dtu.dk/~tobo/blocks_world_back_and_forth.mov

How to coordinate without prior agreement?

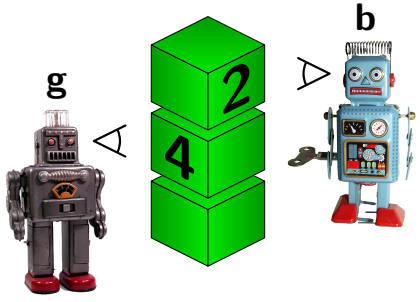
How to coordinate without explicit communication, negotiation or fixed protocol behavior?

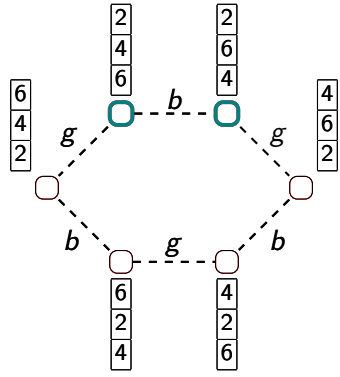
Required: representation of the mental state of other agents. E.g.:

- What's their goal?
- Which actions do they believe will lead to the goal (level of rationality).
- What do they believe about you and your contributions?


Link to movie (click the link):
http://www2.compute.dtu.dk/~tobo/scenario2_double.mp4

Epistemic states: (Multipointed) S5 Kripke models

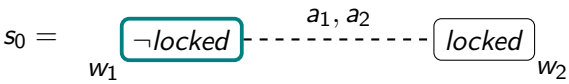


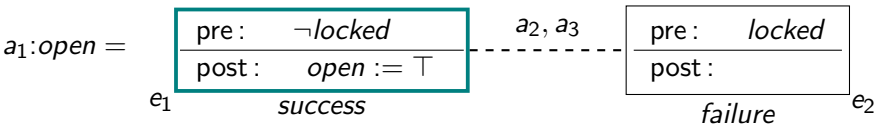


Epistemic states: Multi-pointed epistemic models of multi-agent S5. Nodes are **worlds**, edges are **indistinguishability relations**.
Designated worlds: O (those considered possible by planning agent).

Agent *b*: "Which letter does the middle block have?"

Dynamic epistemic logic (DEL) by example: Product update





$a_1{:}open$ is an **event model** (representing an action). In these, nodes are **events**, and each event has a **precondition** (epistemic formula) and **postconditions** for all atoms (also epistemic formulas).

[Baltag et al., 1998, van Ditmarsch and Kooi, 2008]

Planning based on DEL: Epistemic planning tasks

Definition. An **(epistemic) planning task** $T = (s_0, A, \varphi_g)$ consists of

- A multipointed Kripke model s_0 called the **initial state**.
- A finite set of multipointed event models A called **actions**.
- A **goal formula** φ_g of epistemic logic.

Definition. A (sequential) **plan** for a planning task $T = (s_0, A, \varphi_g)$ is a sequence of actions $\alpha_1, \alpha_2, \ldots, \alpha_n$ from A such that for all $1 \leq i \leq n$, α_i is applicable in $s_0 \otimes \alpha_1 \otimes \cdots \otimes \alpha_{i-1}$ and

$$s_0 \otimes \alpha_1 \otimes \alpha_2 \otimes \cdots \otimes \alpha_n \models \varphi_g.$$

Defining $(\!(\alpha)\!)\varphi := \langle\alpha\rangle\top \wedge [\alpha]\varphi$, this can be reformulated as

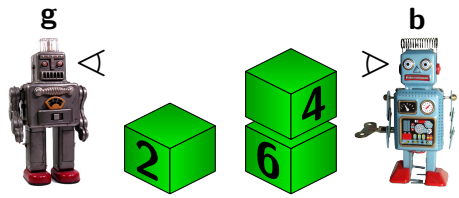
$$s_0 \models (\!(\alpha_1)\!)(\!(\alpha_2)\!) \cdots (\!(\alpha_n)\!)\varphi_g.$$

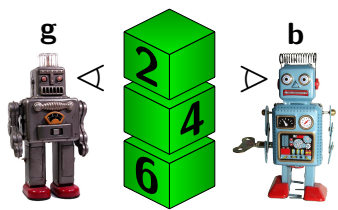
Definition. A plan $i_1{:}\alpha_1, \ldots i_n{:}\alpha_n$ (using notation $i{:}\alpha$ for agent i performing action α) is **implicitly coordinated** if it furthermore holds that :

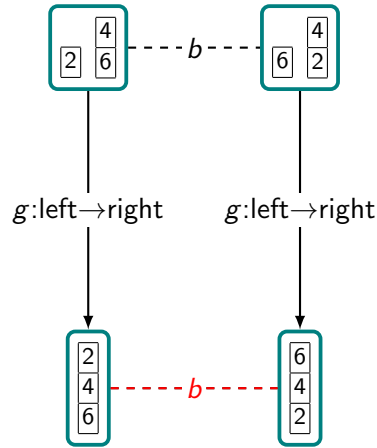
$$s_0 \models K_{i_1}(\!(i_1{:}\alpha_1)\!)K_{i_2}(\!(i_2{:}\alpha_2)\!) \cdots K_{i_n}(\!(i_n{:}\alpha_n)\!)\varphi_g.$$

[Bolander and Andersen, 2011, Engesser et al., 2017, Bolander et al., 2020]

But there's a problem...

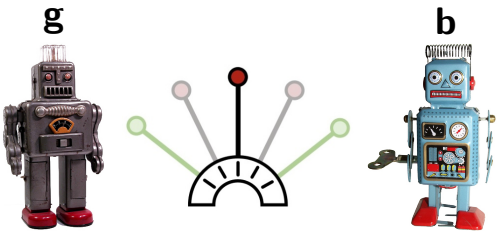






The blue agent didn't learn anything. Why not? Well, since the simple notion of implicit coordination only assumes *rationality in the future*, not *rationality in the past*.

Rationality via standard backward induction

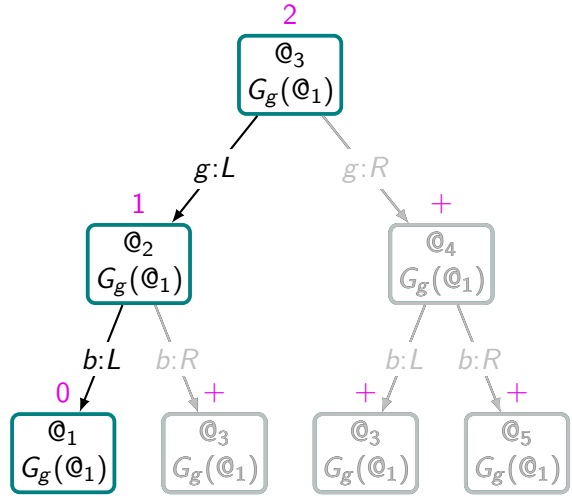


$@_i$: The lever is in position i ($i = 1, \ldots, 5$ from left to right)

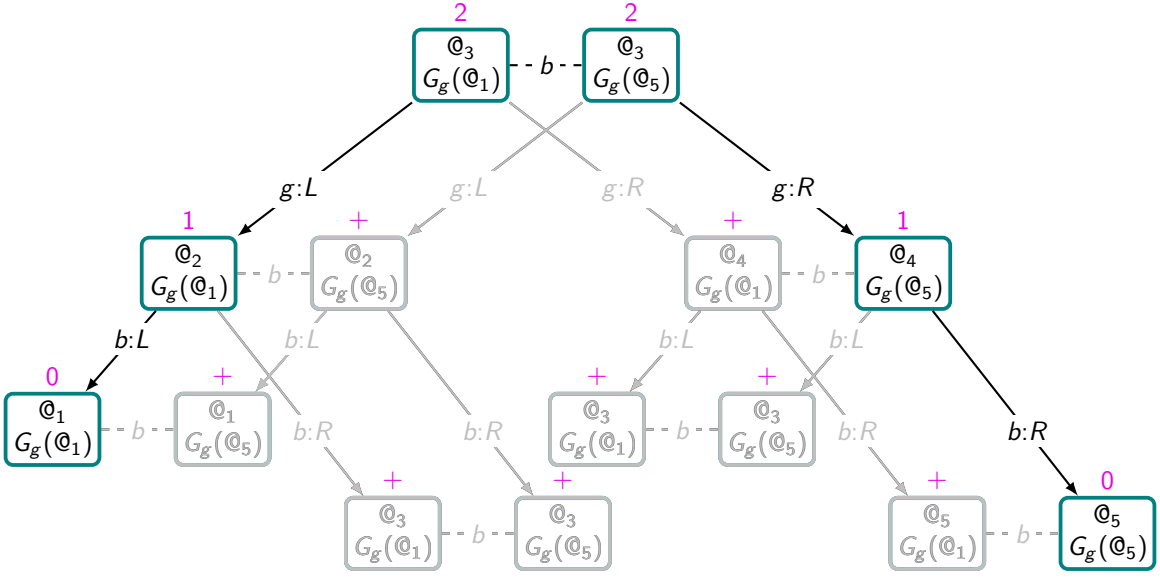
$G_a(\varphi)$: The goal of agent a is φ.

Assumptions: Turn-based, b is altruistic.

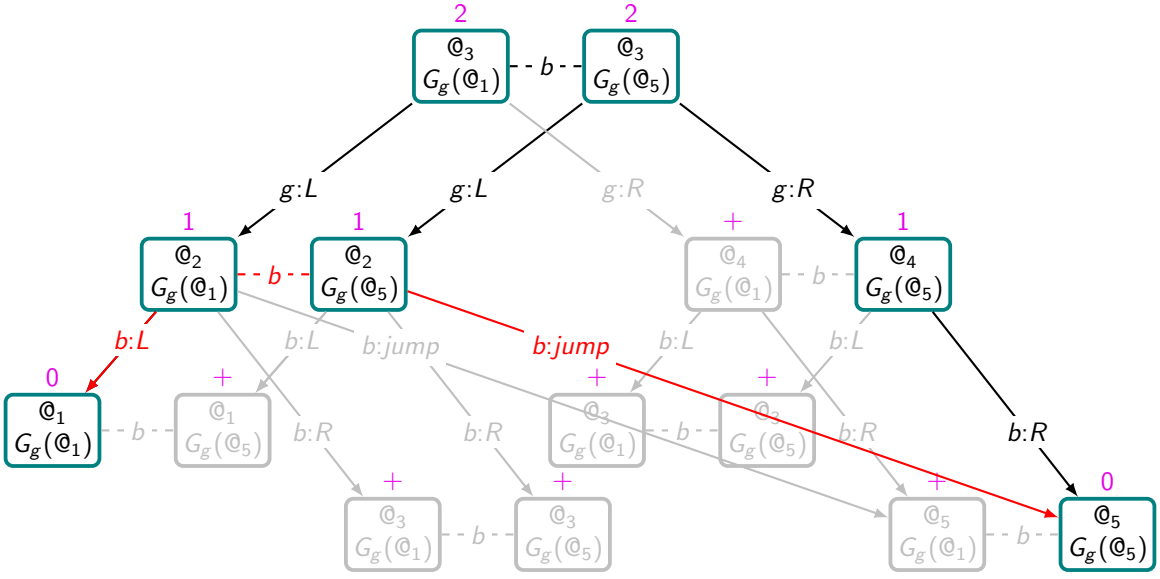
n: (Objective) cost to goal.
$+$: Cost to goal is higher than height of node (world).



Adding goal uncertainty



When forward induction becomes necessary



Coordination via forward induction

How can an action carry **more information** than is encoded in its (observed) consequences?

It only can if certain actions are **expected** and not others.

These expectations then need to be encoded: a **Theory of Mind** of the agent's expected behavior (agent type).

Forward induction (game theory): Assume that the other agents' past actions have been rational.

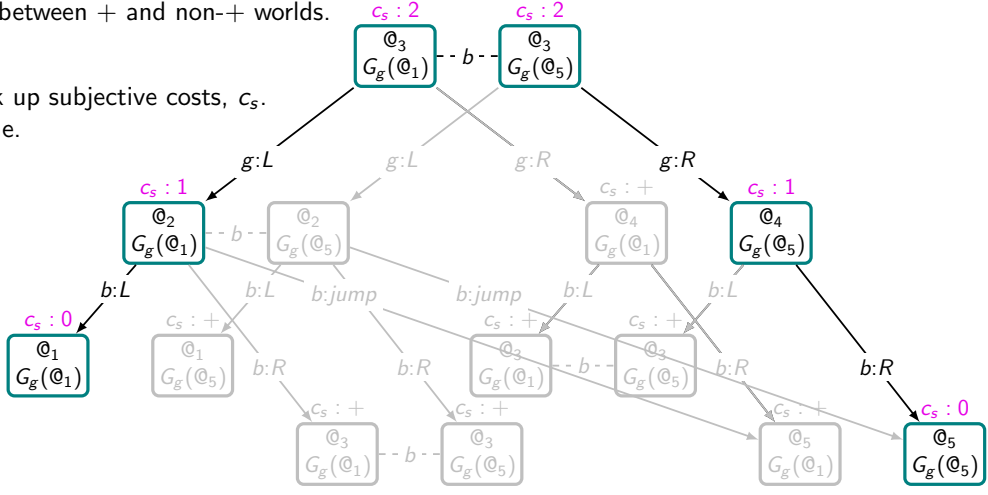
Bad news: Standard forward induction solution concepts in (epistemic) game theory [Battigalli, 1997] involve quantifying over all policies = exponential in size of the game tree!

We (Victoria Núñez Romero and me, in preparation) develop a simpler multipass algorithm: slightly weaker solution concept, but exponential speedup.

Algorithm for simple forward induction

Algorithm (simplified):

1. Back up objective costs, c_o.
2. Cut links between + and non-+ worlds.
3. Repeat:
   - 3.1 Back up subjective costs, c_s.
   - 3.2 Prune.



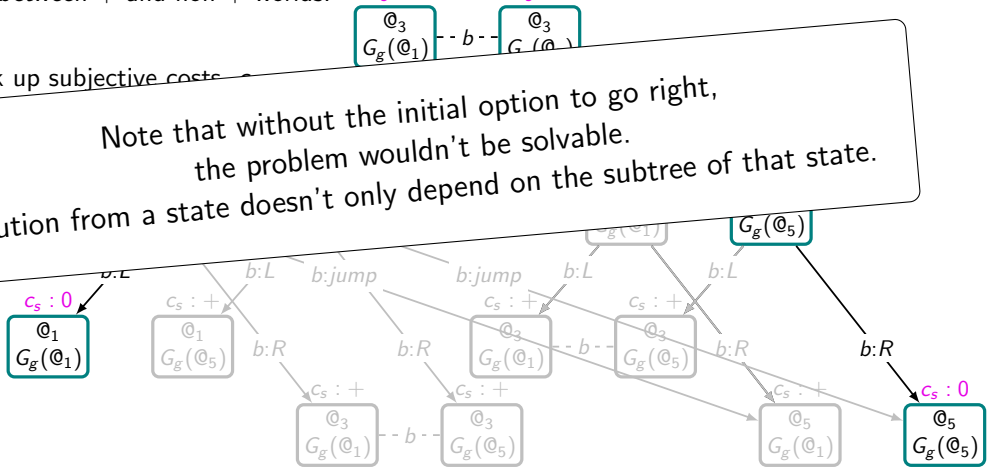
Algorithm for simple forward induction

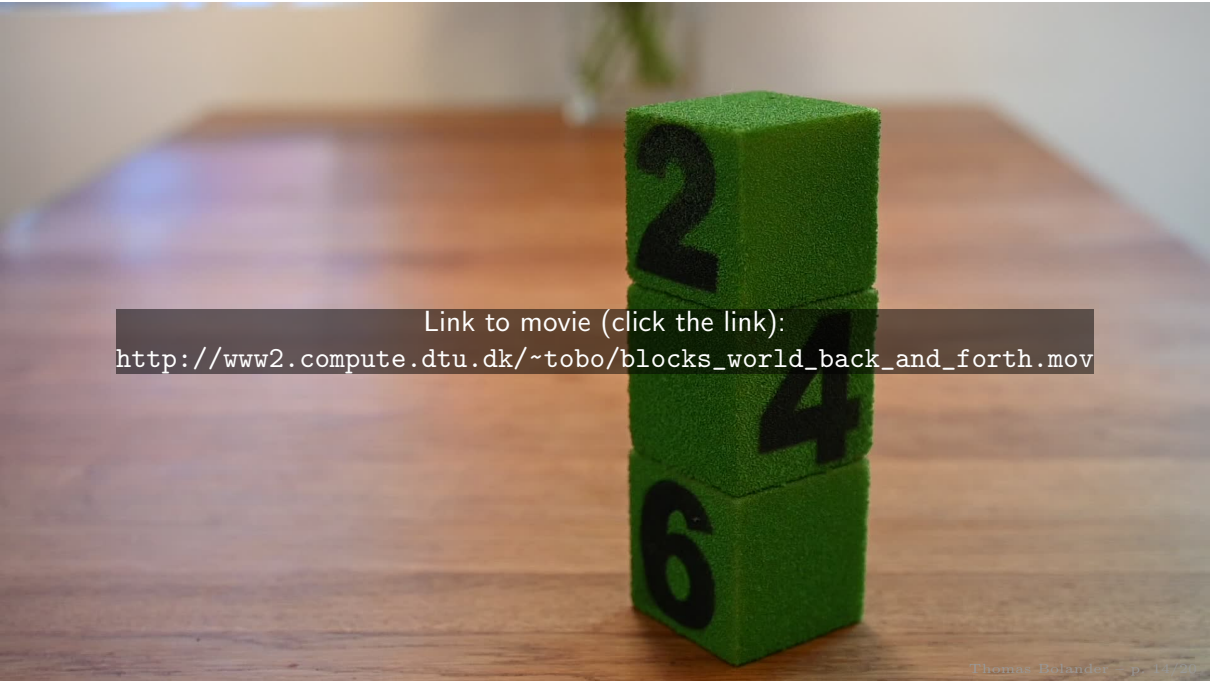
Algorithm (simplified):

1. Back up objective costs, c_o.
2. Cut links between + and non-+ worlds.
3. Repeat:

   3.1 Back up subjective costs

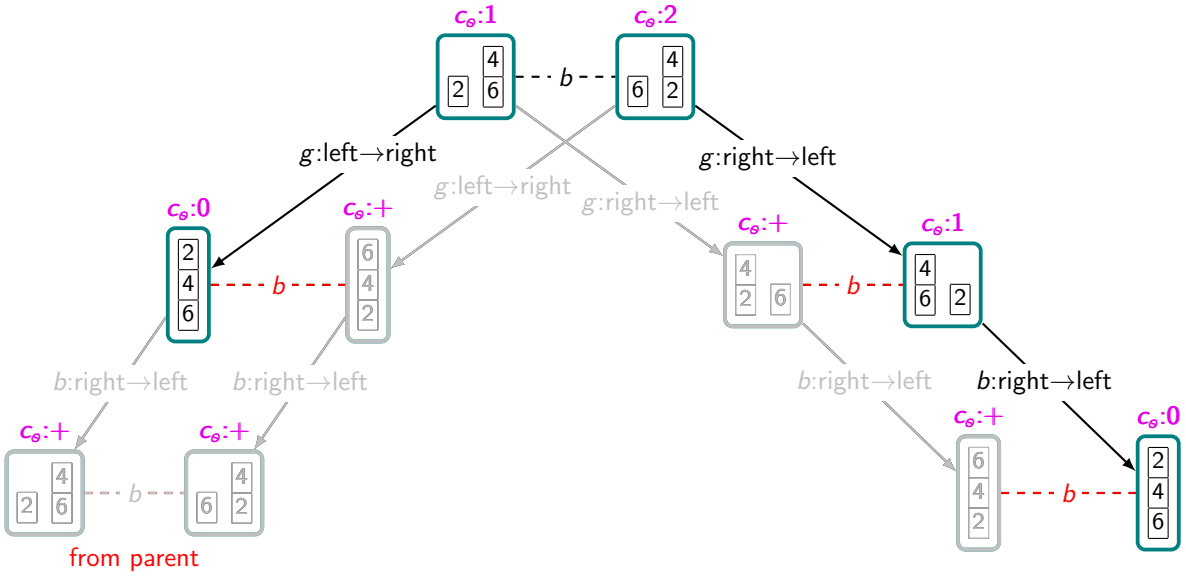
Note that without the initial option to go right,
the problem wouldn't be solvable.
Solution from a state doesn't only depend on the subtree of that state.



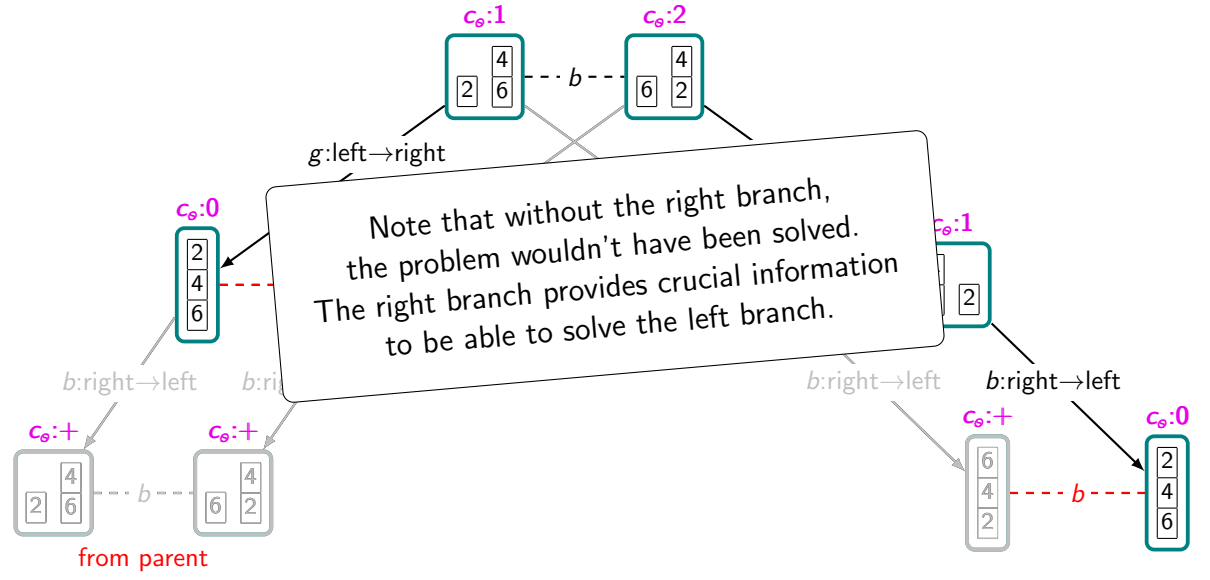


Link to movie (click the link):
http://www2.compute.dtu.dk/~tobo/blocks_world_back_and_forth.mov

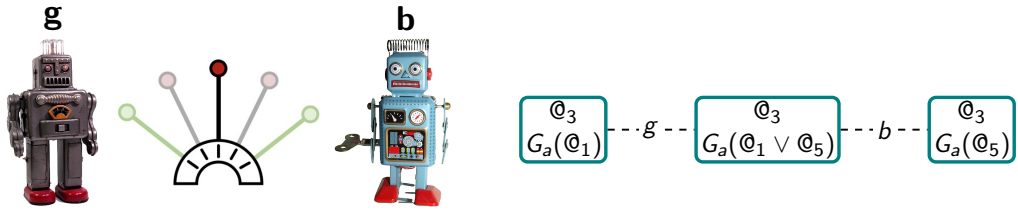
Box off, box on: Our algorithm on the initial blocks world example



Box off, box on: Our algorithm on the initial blocks world example



Results on forward induction



With the original approach to implicit coordination [Bolander et al., 2018], implicit coordination didn't guarantee successful executions! With forward induction, they do:

Theorem (Bolander and Romero, in preparation)

Implicit coordination with forward induction guarantees succesful coordination/executions. (Individual policies are coordinated: induced (non-deterministic) joint policy is guaranteed to always terminate in a goal state.)

Theorem

Standard goal recognition by filtering is subsumed by forward induction.

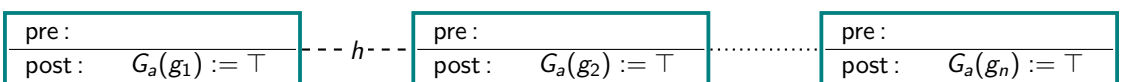

Link to movie (click the link):
http://www2.compute.dtu.dk/~tobo/children_cabinet_cropped.mov
© Warneken & Tomasello

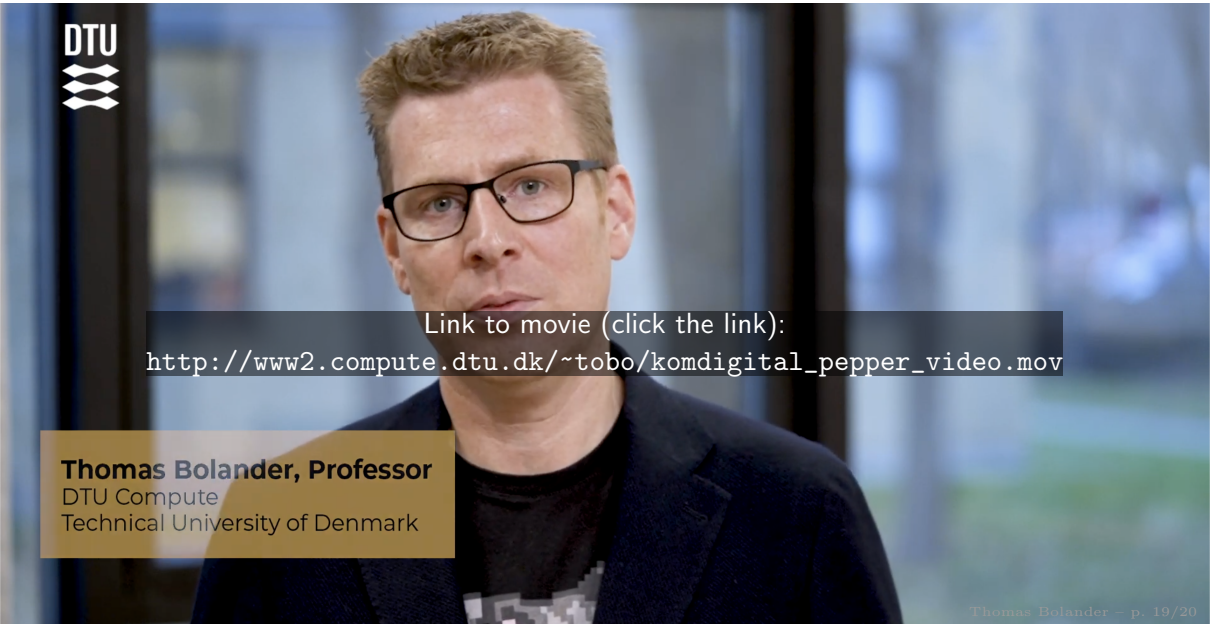
Where do the candidate goals come from to begin with?

Some options:

- Assume a set of candidate goals given as in standard goal recognition.
- Use observed behavior to create candidate goals. E.g. candidate goals via MLLMs.

Adding candidate goals can be done via DEL event models. For candidate goals $G_a = \{g_1, \ldots, g_n\}$ of an *actor agent* a, a *helper agent* h can apply the event model:




DTU
Link to movie (click the link):
http://www2.compute.dtu.dk/~tobo/komdigital_pepper_video.mov
Thomas Bolander, Professor
DTU Compute
Technical University of Denmark

From knowledge to beliefs—and proactivity

KD45 properties not preserved under product update. We can use *plausibility models* [Baltag and Smets, 2008] instead.

Multi-agent planning planning on plausibility models:
[Pieper et al., 2025, Bolander and Grosinger, 2026]

Some definitions and results from [Bolander and Grosinger, 2026] (where E_i is a new *expectation modality*):

- **Definition**. Let U be an undesirability formula for an agent i, and $a_1, \ldots, a_n$ an applicable action sequence in a state s. A formula φ is a *relevant announcement* (wrt. i, U, $a_1, \ldots, a_n$, and s) if the following holds:

$$s \models \neg U \quad \wedge \quad [a_1; \cdots; a_n] \underbrace{(U \wedge E_i \neg U)}_{\neg U \text{ is falsely expected}} \quad \wedge \quad [\underbrace{\Uparrow(i, \varphi)}_{\text{soft announce } \varphi}; a_1; \cdots; a_n] \underbrace{(U \to E_i U)}_{\text{false expectation removed}}$$

- **Theorem**. Let U be a propositional undesirability formula for an agent i, and $a_1, \ldots, a_n$ an applicable action sequence in a state s. Deciding whether relevant announcements exist, and computing them when they do, can both be done in time polynomial in $|s \otimes a_1 \otimes \cdots \otimes a_n| + |U|$.

Appendix: References I

Baltag, A., Moss, L. S. and Solecki, S. (1998).

The Logic of Public Announcements and Common Knowledge and Private Suspicions.
In Proceedings of the 7th Conference on Theoretical Aspects of Rationality and Knowledge (TARK-98), (Gilboa, I., ed.), pp. 43–56, Morgan Kaufmann.

Baltag, A. and Smets, S. (2008).

A Qualitative Theory of Dynamic Interactive Belief Revision.
In Logic and the Foundations of Game and Decision Theory (LOFT7), (Bonanno, G., van der Hoek, W. and Wooldridge, M., eds), vol. 3, of Texts in Logic and Games pp. 13–60, Amsterdam University Press.

Battigalli, P. (1997).

On rationalizability in extensive games.
Journal of Economic Theory *74*, 40–61.

Bolander, T. and Andersen, M. B. (2011).

Epistemic Planning for Single- and Multi-Agent Systems.
Journal of Applied Non-Classical Logics *21*, 9–34.

Bolander, T., Charrier, T., Pinchinat, S. and Schwarzentruber, F. (2020).

DEL-based Epistemic Planning: Decidability and Complexity.
Artificial Intelligence *287*, 1–34.

Bolander, T., Dissing, L. and Herrmann, N. (2021).

DEL-based Epistemic Planning for Human-Robot Collaboration: Theory and Implementation.
In Proceedings of the 18th International Conference on Principles of Knowledge Representation and Reasoning (KR 2021).

Appendix: References II

Bolander, T., Engesser, T., Mattmüller, R. and Nebel, B. (2018).
Better Eager Than Lazy? How Agent Types Impact the Successfulness of Implicit Coordination.
In Proceedings of the 16th International Conference on Principles of Knowledge Representation and Reasoning (KR 2018) AAAI Press.

Bolander, T. and Grosinger, H. J. (2026).
Epistemic Proactivity via Reasoning about Beliefs and Expectations Using Plausibility Models.
Journal of Logic, Language and Information , 1–42.

DTUdk (2020).
KomDigital: R2DTU – A Pepper robot.
Youtube.
https://www.youtube.com/watch?v=9qYCFew58uM [Accessed: 24 November 2021].

Engesser, T., Bolander, T., Mattmüller, R. and Nebel, B. (2017).
Cooperative Epistemic Multi-Agent Planning for Implicit Coordination.
In Proceedings of Methods for Modalities number 243 in Electronic Proceedings in Theoretical Computer Science pp. 75–90,.

Nebel, B., Bolander, T., Engesser, T. and Mattmüller, R. (2019).
Implicitly Coordinated Multi-Agent Path Finding under Destination Uncertainty: Success Guarantees and Computational Complexity.
Journal of Artificial Intelligence Research *64*, 497–527.

Pieper, J., Engesser, T. and Nebel, B. (2025).
Towards Implicit Coordination Planning with Knowledge and Belief.
In Festschrift for Andreas Herzig on the Occasion of his 65th Birthday: Essays in Honor of Andi College Publications.

Appendix: References III

van Ditmarsch, H. and Kooi, B. (2008).
Semantic Results for Ontic and Epistemic Change.
In Logic and the Foundation of Game and Decision Theory (LOFT 7), (Bonanno, G., van der Hoek, W. and Wooldridge, M., eds), Texts in Logic and Games 3 pp. 87–117, Amsterdam University Press.

Warneken, F. and Tomasello, M. (2006).
Cabinet Task.
Max Planck Institute for Evolutionary Anthropology.
https://www.eva.mpg.de/fileadmin/content_files/psychology/videos/children_cabinet.mpg [Accessed: 24 November 2021].