Towards Human-Aware AI via Planning with Epistemic Preferences

Toryn Q. Klassen, [1,2,3] Christian Muise, [2,4] Sheila A. McIlraith [1,2,3]

[1]Department of Computer Science, University of Toronto, Toronto, Canada
[2]Vector Institute for Artificial Intelligence, Toronto, Canada
[3]Schwartz Reisman Institute for Technology and Society, Toronto, Canada
[4]School of Computing, Queen's University, Kingston, Canada

HAXP workshop
July 10, 2023

Epistemic Preferences

- Two areas of study:
    - Planning with **preferences**: optimize properties of the plan instead of just trying to reach a fixed goal
    - **Epistemic planning**: planning for agents to have particular states of knowledge or belief
- We investigate planning with **epistemic preferences**. For example,
    - An agent might want to move an object from the living room to the kitchen and **prefer** that other agents in the environment **know** the new location,
        - or prefer that Bob knows the new location but not Alice.
    - More generally, for **safety** purposes it might be preferred that some epistemic changes don't occur (e.g., creation of **false beliefs**).

Contributions

The contributions of this work include

- proving the **correctness of an encoding** of planning with epistemic preferences as a traditional (non-epistemic) planning problem,
- and demonstrating the feasibility of this approach through **experimentation**.

We also identify a number of **generic sorts** of preferences, such as to maximize true beliefs, some of which may have safety applications.

Outline

Background: Epistemic Logic and RP-MEP

Epistemic Preferences

Experiments

Background: Epistemic Logic

Given a finite set of propositional symbols $\mathcal{P}$ and a finite set of agents *Ag*, the language of epistemic logic is given by

$$\psi ::= p \mid \neg\psi \mid \psi \wedge \psi \mid \Box_i \psi \mid \top \mid \bot$$

where $p \in \mathcal{P}$ and $i \in Ag$.

- $\Box_i \psi$: agent *i* **believes** ψ
- $\Diamond_i \psi$: agent *i* considers ψ **possible** (defined with $\Diamond_i \psi \overset{\text{def}}{=} \neg \Box_i \neg\psi$)

The **semantics** of epistemic logic are given with models that include sets of **possible worlds**.

Background: Epistemic Logic

Given a finite set of propositional symbols $\mathcal{P}$ and a finite set of agents *Ag*, the language of epistemic logic is given by

$$\psi ::= p \mid \neg\psi \mid \psi \wedge \psi \mid \Box_i \psi \mid \top \mid \bot$$

where $p \in \mathcal{P}$ and $i \in Ag$.

- $\Box_i \psi$: agent *i* **believes** ψ
- $\Diamond_i \psi$: agent *i* considers ψ **possible** (defined with $\Diamond_i \psi \overset{\text{def}}{=} \neg \Box_i \neg\psi$)

The **semantics** of epistemic logic are given with models that include sets of **possible worlds**.

Background: Epistemic Logic

Given a finite set of propositional symbols $\mathcal{P}$ and a finite set of agents Ag, the language of epistemic logic is given by

$$\psi ::= p \mid \neg\psi \mid \psi \wedge \psi \mid B_i\,\psi \mid \top \mid \bot$$

where $p \in \mathcal{P}$ and $i \in Ag$.

- $\Box_i\,\psi$: agent i **believes** ψ
- $\Diamond_i\,\psi$: agent i considers ψ **possible** (defined with $\Diamond_i\,\psi \overset{\text{def}}{=} \neg\Box_i\neg\psi$)

The **semantics** of epistemic logic are given with models that include sets of **possible worlds**.

Background: Epistemic Logic

Given a finite set of propositional symbols $\mathcal{P}$ and a finite set of agents Ag, the language of epistemic logic is given by

$$\psi ::= p \mid \neg\psi \mid \psi \wedge \psi \mid \Box_i \psi \mid \top \mid \bot$$

where $p \in \mathcal{P}$ and $i \in Ag$.

- $\Box_i \psi$: agent i **believes** ψ
- $\Diamond_i \psi$: agent i considers ψ **possible** (defined with $\Diamond_i \psi \overset{\text{def}}{=} \neg \Box_i \neg\psi$)

The **semantics** of epistemic logic are given with models that include sets of **possible worlds**.

Background: Proper Epistemic Knowledge Bases (PEKBs)

RP-MEP represents planning states as **proper epistemic knowledge bases**[1].

A proper epistemic knowledge base **(PEKB)** is a set (or conjunction) of **restricted modal literals (RMLs)**, which are formulas given by this grammar:

$$\varphi ::= p \mid \neg p \mid \Box_i \varphi \mid \Diamond_i \varphi \mid \top \mid \bot$$

The **depth** of an RML φ is the number of belief operators in it; e.g.,

- p has depth 0
- and $\Box_i \Diamond_j \Box_i \neg p$ has depth 3.

The absence of **disjunction** in PEKBs simplifies some computations.

---

[1] Lakemeyer, G. & Lespérance, Y. *Efficient Reasoning in Multiagent Epistemic Logics.* in *ECAI 2012* (IOS Press, 2012), 498–503.

Background: Multi-agent Epistemic Planning (MEP) problems[2]

$\langle \mathcal{P}, \mathcal{A}, Ag, \mathcal{I}, \mathcal{G} \rangle$:

- $\mathcal{P}$ is the set of propositions;
- $\mathcal{A}$ is a finite set of **actions**, where an action a is a pair $\langle Pre_a, Eff_a \rangle$, in which
  - Pre_a is a PEKB and
  - Eff_a is a set of **conditional effects** of the form $\langle \gamma_i, \varphi_i \rangle$ where the PEKB γ_i is the condition and the RML φ_i is the effect;
- Ag is a finite set of agents;
- $\mathcal{I}$ is a PEKB representing the **initial state**;
- and $\mathcal{G}$ is another PEKB, representing the **goal**.

---

[2] Muise, C., Belle, V., Felli, P., McIlraith, S. A., Miller, T., Pearce, A. R. & Sonenberg, L. Efficient Multi-agent Epistemic Planning: Teaching Planners About Nested Belief. *AIJ* **302,** 103605 (2022).

Background: RP-MEP

A **restricted perspectival** MEP (RP-MEP)[3] problem with depth bound d is a MEP problem with the restriction that there is some **root agent** $\star \in Ag$ such that

- any RML in the initial state, goal, or an action precondition is of the form $\Box_\star \varphi$, any RML in a conditional effect is either of the form $\Box_\star \varphi$ or $\Diamond_\star \varphi$,
- and any RML anywhere in the problem has depth at most $d + 1$.

Computation:

- Muise et al. showed how to encode an RP-MEP problem R as a **classical**$^{+}$ planning problem C(R).

---

[3] Muise, C., Belle, V., Felli, P., McIlraith, S. A., Miller, T., Pearce, A. R. & Sonenberg, L. Efficient Multi-agent Epistemic Planning: Teaching Planners About Nested Belief. *AIJ* **302,** 103605 (2022).

Outline

Background: Epistemic Logic and RP-MEP

Epistemic Preferences

Experiments

Formalizing Epistemic Preferences

- Given an RP-MEP problem, we can add **preferences (with associated weights)** of the form $\langle \psi_i, r_i \rangle$ where ψ_i is a PEKB and $r_i \in \mathbb{R}$.
  - ψ_i contains only RMLs of the form $\Box_\star \varphi$ of depth $\leq d + 1$
- A plan π is **optimal** for an RP-MEP problem with preferences if it **maximizes the sum** of the weights of the preferences satisfied.

Computation:

- We can extend the RP-MEP compilation to also encode preferences, by encoding each preference formula in the same manner as the goal.
- **Theorem:**
  A plan is optimal for an RP-MEP problem with preferences iff it corresponds to an optimal plan for the compiled classical$^+$ problem with preferences.

Types of Epistemic Preferences

Some categories of epistemic preferences (where ℓ is an arbitrary literal):

- **truth**: (the root agent believes that) agent *i* correctly believes that ℓ is true: $(\Box_\star \ell \wedge \Box_\star \Box_i \ell)$.
- **misconception**: (the root agent believes that) agent *i* incorrectly believes that ℓ is true: $(\Box_\star \bar{\ell} \wedge \Box_\star \Box_i \ell)$
- **oblivious**: (the root agent believes that) agent *i* considers ℓ possible: $\Box_\star \Diamond_i \ell$
- **conscious**: (the root agent believes that) agent *i* believes ℓ: $\Box_\star \Box_i \ell$

Outline

Background: Epistemic Logic and RP-MEP

Epistemic Preferences

Experiments

Experiments: domains

We used these domains:[4]

Corridor: An agent who has a secret can walk around and make (possibly false) announcements that are believed by other nearby agents.

Grapevine: All (non-root) agents can move and make (possibly false) announcements; each starts with its own secret. Agents only believe announcements that don't contradict their own beliefs.

---

[4] Muise, C., Belle, V., Felli, P., McIlraith, S. A., Miller, T., Pearce, A. R. & Sonenberg, L. Efficient Multi-agent Epistemic Planning: Teaching Planners About Nested Belief. *AIJ* **302,** 103605 (2022).

Experiments: preferences

- We added preferences generated from the **truth**, **misconception**, **oblivious**, and **conscious** categories.

- In each experiment, **all** preferences from the relevant category (e.g., **truth**) that involve any non-root agent *i* and any literal ℓ are generated
  - (excluding some literals that agents always know the truth value of).

- All the preferences are given weight 1.

Experiments: computation

To compute plans for each problem, we

1. **Compile** the problem into a classical$^+$ planning problem, using the **RP-MEP** program.
2. Transform each classical$^+$ planning problem with preferences by applying the **compilation from preferences into action costs** from Keyder and Geffner[5].
3. **Solve** the resulting classical$^+$ planing problem with costs using LAMA[6].

---

[5] Keyder, E. & Geffner, H. Soft Goals Can Be Compiled Away. *JAIR* **36,** 547–556 (2009).

[6] Richter, S. & Westphal, M. The LAMA Planner: Guiding Cost-Based Anytime Planning with Landmarks. *JAIR* **39,** 127–177 (2010).

Experiments: results

	Preference Type													
Problem	none		truth			misconception			oblivious			conscious		
	$\lvert\pi\rvert$	time	$\lvert\pi\rvert$	time	prefs	$\lvert\pi\rvert$	time	prefs	$\lvert\pi\rvert$	time	prefs	$\lvert\pi\rvert$	time	prefs
Corridor-3	5	0.36	6	0.35	**2**/6	6	0.36	**1**/6	6	0.36	**3**/6	6	0.37	**3**/6
Corridor-5	5	0.40	8	0.63	**4**/10	6	0.64	**2**/10	6	0.43	**6**/10	6	0.46	**5**/10
Corridor-7	5	3.72	8	14.76	**5**/14	9	12.76	**4**/14	6	5.04	**8**/14	7	2.18	**7**/14
Grapevine-4-2	4	4.66	11	37.45	**15**/32	7	51.42	**11**/32	5	33.34	**26**/32	8	33.84	**16**/32
Grapevine-4-4	6	2.91	15	33.32	**14**/32	10	33.30	**10**/32	8	49.93	**24**/32	9	32.60	**16**/32
Grapevine-4-8	11	45.43	14	37.37	**12**/32	13	33.46	**8**/32	13	32.73	**20**/32	12	34.67	**16**/32
Grapevine-8-2	4	31.13	19	61.22	**63**/128	16	57.01	$51^\dagger$/128	6	59.55	118/128	9	61.53	$40^\dagger$/128
Grapevine-8-4	5	24.63	17	59.78	$37^\dagger$/128	15	61.16	$28^\dagger$/128	15	60.71	116/128	13	59.36	$26^\dagger$/128
Grapevine-8-8	7	32.82	32	61.13	60/128	27	60.73	52/128	13	59.93	112/128	18	61.00	$28^\dagger$/128

- $|\pi|$: length of plan
- Times (in seconds): taken by LAMA on the encoded problem (encoding times not included). LAMA was run with a search time limit of 30 seconds.
- x/y: x out of y preferences were satisified (bold $\boldsymbol{x}$ is known to be optimal, $x^\dagger$ with the $\dagger$ is known to be suboptimal)

Experiments: results

Problem	Preference Type													
	none		**truth**			**misconception**			**oblivious**			**conscious**		
	$\lvert\pi\rvert$	time	$\lvert\pi\rvert$	time	prefs	$\lvert\pi\rvert$	time	prefs	$\lvert\pi\rvert$	time	prefs	$\lvert\pi\rvert$	time	prefs
Corridor-3	5	0.36	6	0.35	**2**/6	6	0.36	**1**/6	6	0.36	**3**/6	6	0.37	**3**/6
Corridor-5	5	0.40	8	0.63	**4**/10	6	0.64	**2**/10	6	0.43	**6**/10	6	0.46	**5**/10
Corridor-7	5	3.72	8	14.76	**5**/14	9	12.76	**4**/14	6	5.04	**8**/14	7	2.18	**7**/14
Grapevine-4-2	4	4.66	11	37.45	**15**/32	7	51.42	**11**/32	5	33.34	**26**/32	8	33.84	**16**/32
Grapevine-4-4	6	2.91	15	33.32	**14**/32	10	33.30	**10**/32	8	49.93	**24**/32	9	32.60	**16**/32
Grapevine-4-8	11	45.43	14	37.37	**12**/32	13	33.46	**8**/32	13	32.73	**20**/32	12	34.67	**16**/32
Grapevine-8-2	4	31.13	19	61.22	**63**/128	16	57.01	51†/128	6	59.55	118/128	9	61.53	40†/128
Grapevine-8-4	5	24.63	17	59.78	37†/128	15	61.16	28†/128	15	60.71	116/128	13	59.36	26†/128
Grapevine-8-8	7	32.82	32	61.13	60/128	27	60.73	52/128	13	59.93	112/128	18	61.00	28†/128

- $|\pi|$: length of plan
- Times (in seconds): taken by LAMA on the encoded problem (encoding times not included). LAMA was run with a search time limit of 30 seconds.
- x/y: x out of y preferences were satisified (bold $\boldsymbol{x}$ is known to be optimal, $x^\dagger$ with the $\dagger$ is known to be suboptimal)

Experiments: results

	Preference Type													
Problem	none		truth			misconception			oblivious			conscious		
	$\lvert\pi\rvert$	time	$\lvert\pi\rvert$	time	prefs	$\lvert\pi\rvert$	time	prefs	$\lvert\pi\rvert$	time	prefs	$\lvert\pi\rvert$	time	prefs
Corridor-3	5	0.36	6	0.35	**2**/6	6	0.36	**1**/6	6	0.36	**3**/6	6	0.37	**3**/6
Corridor-5	5	0.40	8	0.63	**4**/10	6	0.64	**2**/10	6	0.43	**6**/10	6	0.46	**5**/10
Corridor-7	5	3.72	8	14.76	**5**/14	9	12.76	**4**/14	6	5.04	**8**/14	7	2.18	**7**/14
Grapevine-4-2	4	4.66	11	37.45	**15**/32	7	51.42	**11**/32	5	33.34	**26**/32	8	33.84	**16**/32
Grapevine-4-4	6	2.91	15	33.32	**14**/32	10	33.30	**10**/32	8	49.93	**24**/32	9	32.60	**16**/32
Grapevine-4-8	11	45.43	14	37.37	**12**/32	13	33.46	**8**/32	13	32.73	**20**/32	12	34.67	**16**/32
Grapevine-8-2	4	31.13	19	61.22	**63**/128	16	57.01	$51^{\dagger}$/128	6	59.55	118/128	9	61.53	$40^{\dagger}$/128
Grapevine-8-4	5	24.63	17	59.78	$37^{\dagger}$/128	15	61.16	$28^{\dagger}$/128	15	60.71	116/128	13	59.36	$26^{\dagger}$/128
Grapevine-8-8	7	32.82	32	61.13	60/128	27	60.73	52/128	13	59.93	112/128	18	61.00	$28^{\dagger}$/128

- $|\pi|$: length of plan
- Times (in seconds): taken by LAMA on the encoded problem (encoding times not included). LAMA was run with a search time limit of 30 seconds.
- x/y: x out of y preferences were satisified (bold $\boldsymbol{x}$ is known to be optimal, $x^{\dagger}$ with the $\dagger$ is known to be suboptimal)

Experiments: results

							Preference Type							
Problem	**none**		**truth**			**misconception**			**oblivious**			**conscious**		
	$\lvert\pi\rvert$	time	$\lvert\pi\rvert$	time	prefs	$\lvert\pi\rvert$	time	prefs	$\lvert\pi\rvert$	time	prefs	$\lvert\pi\rvert$	time	prefs
Corridor-3	5	0.36	6	0.35	**2**/6	6	0.36	**1**/6	6	0.36	**3**/6	6	0.37	**3**/6
Corridor-5	5	0.40	8	0.63	**4**/10	6	0.64	**2**/10	6	0.43	**6**/10	6	0.46	**5**/10
Corridor-7	5	3.72	8	14.76	**5**/14	9	12.76	**4**/14	6	5.04	**8**/14	7	2.18	**7**/14
Grapevine-4-2	4	4.66	11	37.45	**15**/32	7	51.42	**11**/32	5	33.34	**26**/32	8	33.84	**16**/32
Grapevine-4-4	6	2.91	15	33.32	**14**/32	10	33.30	**10**/32	8	49.93	**24**/32	9	32.60	**16**/32
Grapevine-4-8	11	45.43	14	37.37	**12**/32	13	33.46	**8**/32	13	32.73	**20**/32	12	34.67	**16**/32
Grapevine-8-2	4	31.13	19	61.22	**63**/128	16	57.01	51†/128	6	59.55	118/128	9	61.53	40†/128
Grapevine-8-4	5	24.63	17	59.78	37†/128	15	61.16	28†/128	15	60.71	116/128	13	59.36	26†/128
Grapevine-8-8	7	32.82	32	61.13	60/128	27	60.73	52/128	13	59.93	112/128	18	61.00	28†/128

- $|\pi|$: length of plan
- Times (in seconds): taken by LAMA on the encoded problem (encoding times not included). LAMA was run with a search time limit of 30 seconds.
- x/y: x out of y preferences were satisified (bold $\boldsymbol{x}$ is known to be optimal, $x^{\dagger}$ with the $\dagger$ is known to be suboptimal)

Conclusion

- We have considered planning with **epistemic** preferences, i.e., over knowledge or beliefs.
- Our approach to computing plans for such problems makes use of **two existing compilations**:
  1. the **RP-MEP encoding** of epistemic planning problems as classical$^+$ problems,
  2. and **Keyder and Geffner's compilation**[7] of preferences into costs.
- Future work may further explore **applications** of planning with epistemic preferences in areas such as AI safety.

---

[7] Keyder, E. & Geffner, H. Soft Goals Can Be Compiled Away. *JAIR* **36,** 547–556 (2009).