О приложениях эпистемической логики в компьютерных науках

ФКН НИУ ВШЭ, 29.06.2023

В.В.Долгоруков, Международная лаборатория логики, лингвистики и формальной философии

Эпистемическая логика

- Логика для знания, убеждения и их динамики
- Семейство модальных логик
- Зародилась внутри философии
- Основное приложение: формальная эпистемология
- Также: экономика (эпистемическая теория игр), лингвистика, cs
- Монографии:
  - Hintikka J. Knowledge and Belief: An Introduction to the Logic of the Two Notions. Ithaca: Cornell University Press. 1962.
  - Fagin R., Halpern J., Moses Y, Vardi M., Reasoning about Knowledge. Cambridge MA: MIT Press, 1995.
  - Meyer J., van der Hoek W. Epistemic Logic for AI and Computer Science. Cambridge: Cambridge University Press, 1995.
  - van Ditmarsch H., van der Hoek W., Kooi B. Dynamic Epistemic Logic. Dordrecht: Springer, 2008.
  - Handbook of Epistemic Logic. L.: College Publications, 2015.

Какая логика соответствует знанию?

- $(K)\ \ K_i(\varphi \to \psi) \to (K_i\varphi \to K_i\psi)$
- $(T)\ K_i\varphi \to \varphi$
- $(4)\ K_i\varphi \to K_iK_i\varphi$
- $(5)\ \neg K_i\varphi \to K_i\neg K_i\varphi$
- Правила вывода: MP, Nec
- Консенсус: логика знания между S4 (KT4) и S5 (KT5)

Парадокс познаваемости (Фитча)

- Аксиомы
  - $\varphi \to \Diamond K\varphi$
  - $K\varphi \to \varphi$
  - $K(\varphi \wedge \psi) \to (K\varphi \wedge K\psi)$
- Теорема: $p \to Kp$

Приложения в CS

- эпистемическое планирование
- верификация эпистемических протоколов
- эпистемическая криптография: русская карточная задача

Языки эпистемической логики

- $(\mathscr{E}\mathscr{L})\ \varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_i\varphi$
- $(\mathscr{E}\mathscr{L}\mathscr{C})\ \varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_i\varphi \mid C_G\varphi$
- $(\mathscr{P}\mathscr{A}\mathscr{L})\ \varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_i\varphi \mid [!\varphi]\psi$

Модель Крипке

- $M = (W, (\sim_i)_{i \in Ag}, V)$, где
- W – непустое множество возможных миров (состояний)
- $\sim_i \subseteq W \times W$ – отношение достижимости
- $V : Prop \to \mathscr{P}(W)$ – функция оценки

Условия истинности в модели

- $M, w \vDash p \iff w \in V(p)$
- $M, w \vDash \neg\varphi \iff M, w \nvDash \varphi$
- $M, w \vDash \varphi \wedge \psi \iff M, w \vDash \varphi$ и $M, w \vDash \psi$
- $M, w \vDash K_i\varphi \iff \forall w'(w \sim_i w' \Rightarrow M, w' \vDash \varphi)$
- $M, w \vDash C_G\varphi \iff \forall w'(w(\bigcup_{i \in G} \sim_i)^* w' \Rightarrow M, w' \vDash \varphi)$
- $M, W_d \vDash \varphi \iff M, w \vDash \varphi$ для всех $w \in W_d$

PAL: логика с публичными объявлениями

Обновление модели $M^{!\varphi} = (W', (\sim_i')_{i \in Ag}, V')$

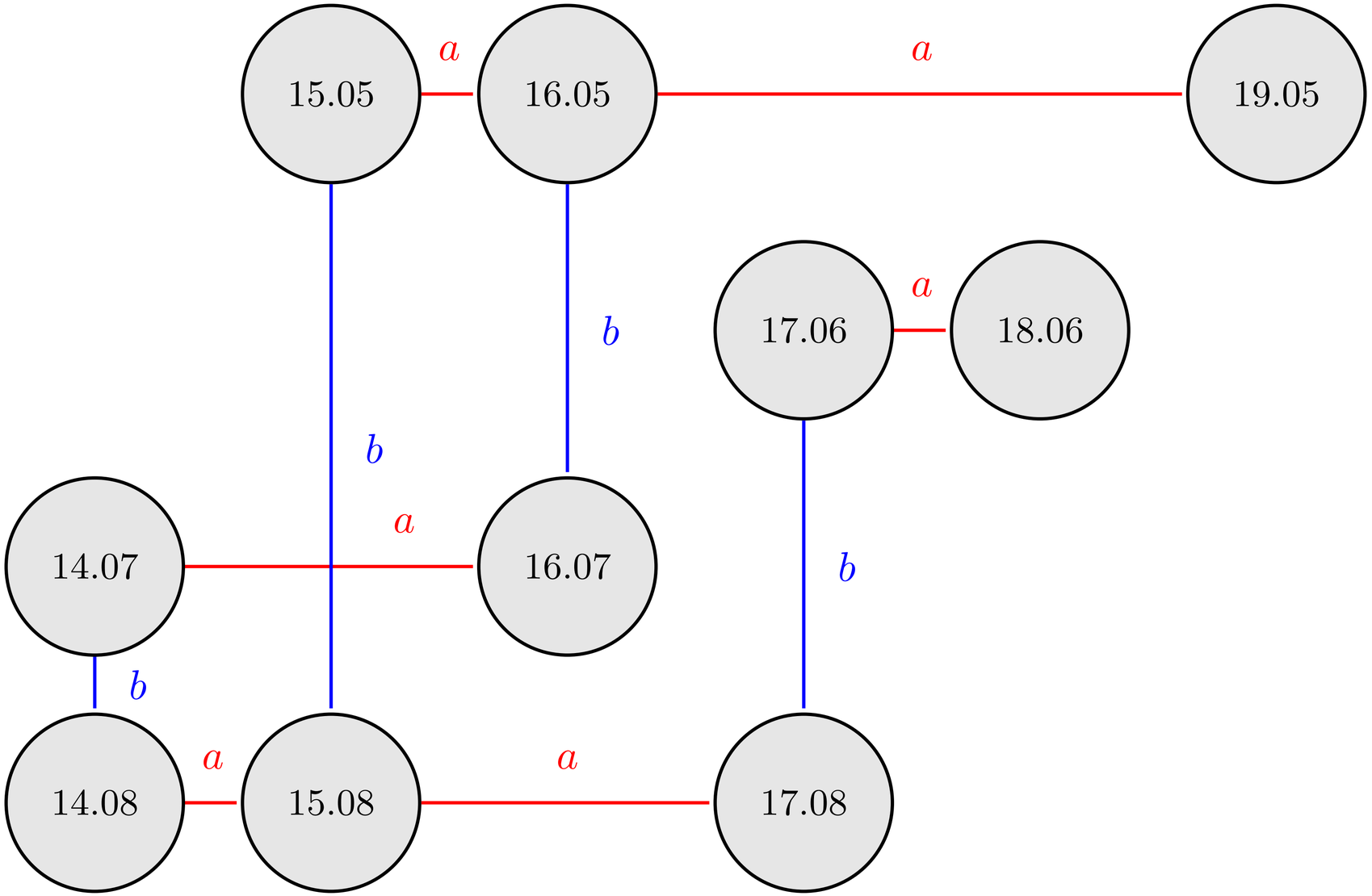
- $W' = \{w \in W \mid M, w \vDash \varphi\}$
- $\sim_i' = \sim_i \cap (W' \times W')$
- $V'(p) = V(p) \cap W'$

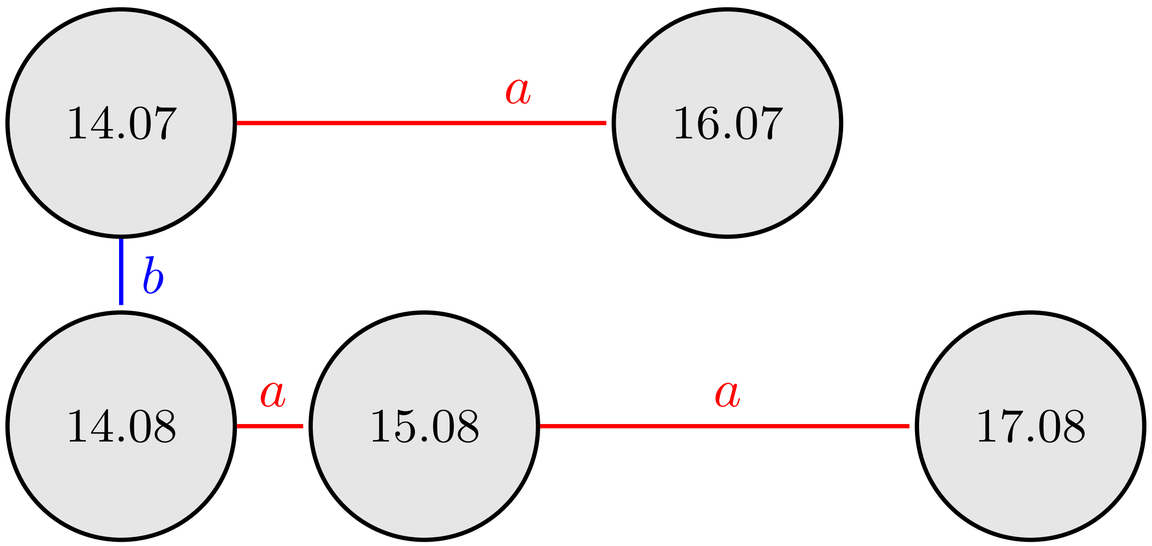
Условие истинности:

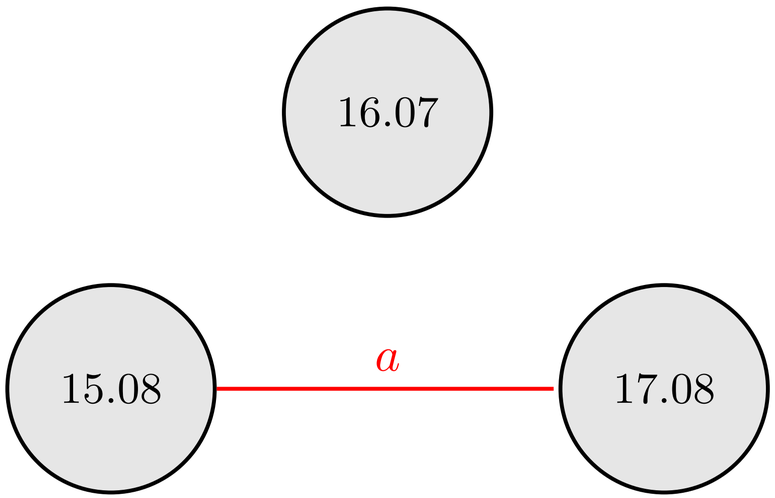
$$M, w \vDash [!\varphi]\psi \iff \varphi \Rightarrow M^{!\varphi}, w \vDash \psi$$

Задача

- Альберт и Бернард только что познакомились с Шерил, и захотели узнать, когда у нее день рождения. Шерил перечислила список из 10 возможных дат: 15 мая, 16 мая, 19 мая, 17 июня, 18 июня, 14 июля, 16 июля, 14 августа, 15 августа и 17 августа.
- Потом Шерил сказала Бернарду только день ее рождения, а Альберту — месяц.
- "Я не знаю, когда у Шерил день рождения, но я точно знаю, что Бернард тоже не знает», — сказал Альберт.
- «Сначала я не знал, когда у Шерил день рождения, но теперь я знаю», — возразил Бернард.
- На это Альберт ответил: «Тогда я тоже знаю, когда у Шерил День рождения».
- Когда же у Шерил день рождения?


15.05
16.05
19.05
17.06
18.06
14.07
16.07
14.08
15.08
17.08
a
b


14.07
a
16.07
b
14.08
a
15.08
a
17.08

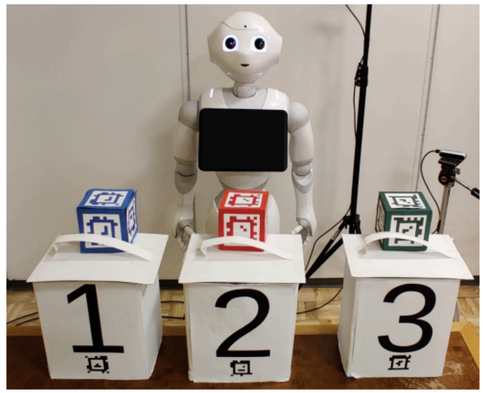

16.07
15.08
a
17.08

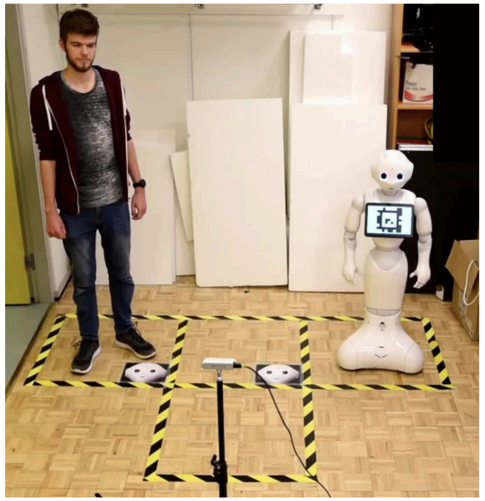
Эпистемическое планирование

- Один агент. Дано текущее эпистемическое состояние и желаемое эпистемическое состояние. Задача: перейти от одного состояния к другом, затратив минимальное количество ресурсов
- Мультиагентный случай. Текущее и желаемое состояния могут отсылать к метазнанию.

Theory of Mind у роботов

Thomas Bolander, Lasse Dissing,Nicolai Herrmann, DEL-based Epistemic Planning for Human-Robot Collaboration: Theory and Implementation //18th International Conference on Principles of Knowledge Representation and Reasoning (KR-2021)





Эпистемическое планирование с имплицитной координацией

- Улучшение понимания между роботом и человеком
- Задача: помоги найти мне очки
- Если человек первым найдет свои очки?
- Робот должен обладать ToM

Модель действия

- $\mathscr{E} = (E, (Q_i)_{\in Ag}, pre, post)$
- E – непустое множество событий
- $Q_i \subseteq E \times E$ – отношение достижимости
- $pre : E \to \mathscr{L}$ – предусловие события
- $post : E \to \mathscr{L}$ – постусловие события

Результат обновления модели

- $W' = \{(w, e) \in W \times E \mid M, w \vDash pre(e)\}$
- $(w_1, e_1) \sim'_i (w_2, e_2) \iff w_1 \sim_i w_2$ и $e_1 Q_i e_2$
- $(w, e) \vDash p \iff (M, w \vDash p$ и $\neg p \notin post(e))$ или $p \in post(e)$
- $W'_d = \{(w, e) \in W' \mid w \in W_d, e \in E_d\}$

Gossip-протокол

- Задача: У каждого из 6 друзей есть своя новость. Они созваниваются (попарно) и обмениваются новостями. Какое количество звонков нужно, чтобы обменяться всеми новостями?
- Достаточно 8 звонков

Gossip-протокол

N	Агенты	1	2	3	4	5	6
1	12	12	12	3	4	5	6
2	34	12	12	34	34	5	6
3	56	12	12	34	34	56	56
4	13	1234	12	1234	34	56	56
5	45	1234	12	1234	3456	3456	56
6	16	123456	12	1234	3456	3456	123456
7	24	123456	123456	1234	123456	3455	123456
8	35	123456	123456	123456	123456	123456	123456

Gossip-протокол

- В общем виде: достаточно $2n - 4$ звонков ($n \geq 4$)

Русская карточная задача

- Задача. В колоде 7 карт. Первый и второй игрок берут по три карты, оставшаяся карта достается третьему игроку. Могут ли первый и второй игроки обменяться открытыми сообщениями так, чтобы каждый из них знал какие у кого карты, а третий бы не смог вычислить их наборы карт.
- Допустим, что распределение такое: 123.456.7
- Решение:

a : “У меня 012, 034, 056, 135, 246”, b: “у с 6”

Формальная верификация модели

van Ditmarsch H.P. et al. Model Checking Russian Cards // Electronic Notes in Theoretical Computer Science. 2006. Vol. 149, № 2. P. 105–123.

- DEMO (Haskell)
- Model Checking Multi-Agent Systems (ATL, C++)

GAL: group announcement logic

- $\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_i\varphi \mid [!\varphi]\psi \mid [G]\varphi$
- $M, x \vDash [G]\varphi \iff$ для любого множества формул $\{\psi_i \mid \in G\} \subseteq \mathscr{L}_K$

$M, x \vDash [!\bigwedge_{i \in G} K_i\psi_i]\varphi$

- $\langle G\rangle\varphi := \neg[G]\neg\varphi$

Russian Cards: GAL

- $klm_i := k_i \wedge l_i \wedge m_i$
- $\bigwedge_{ijk} (ijk_b \to K_a ijk_b)$ one
- $\bigwedge_{ijk} (ijk_a \to K_b ijk_a)$ two
- $\bigwedge_{q=0}^{6} ((q_a \to \neg K_c q_a) \wedge (q_b \to \neg K_c q_b))$ three
- $M, x \vDash \langle ab \rangle (one \wedge two \wedge three)$

Эпистемическая криптография

Joseph Y Halpern and Rafael Pass. “A Knowledge-Based Analysis of the Blockchain”. In: Proceedings of the 16th Conference on Theoretical Aspects of Rationality and Knowledge. Edited by Jérôme Lang. TARK 2017. 2017. doi: 10.4204/EPTCS.251.22

Gattinger, M, & van Eijck, D.J.N. (2015). Towards Model Checking Cryptographic Protocols with Dynamic Epistemic Logic. In *Proceedings of International Workshop on Learning and Adaptation in MAS 2015 (LAMAS)* (pp. 1–14). AAMAS.

Ограничения в ресурсах

- Belardinelli, G., Rendsvig, R.K. (2021). Epistemic Planning with Attention as a Bounded Resource. In: Ghosh, S., Icard, T. (eds) Logic, Rationality, and Interaction. LORI 2021. Lecture Notes in Computer Science, vol 13039. Springer, Cham. https://doi.org/10.1007/978-3-030-88708-7_2

- Dolgorukov, V., Gladyshev, M. (2023). Dynamic Epistemic Logic for Budget-Constrained Agents. In: Areces, C., Costa, D. (eds) Dynamic Logic. New Trends and Applications. DaLí 2022. Lecture Notes in Computer Science, vol 13780. Springer, Cham. https://doi.org/10.1007/978-3-031-26622-5_4

- $[?_i\varphi]\psi$ - поиск ответа на вопрос требует ресурсов

- общее знание + $[?_G\varphi]\psi$ - группа может объединить усилия в поиске ответа на вопрос