Knowledge in multi-agent systems

Olga Štěpánková

laboratory Gerstner

As a tradition, knowledge and reasoning used to be studied in case of a single individual (philosophers and logicians).

Is it sufficient for analysis of daily situations?

- NL discussion
- Business negotiation
- Decision about the next step in complex trafic situation ?

In all these cases we need to reason about interaction among several agents.

Agents can be people, robots, complex computer systems, …, machines

Examples.

Mr. Bird and Mr. Ladybird.

Watergate case. *Dean does not know, if Nixon knows, that Dean knows, that Nixon knows, that McCord slipped in secretly into O'Brien's office.*

It is not easy to keep the track in such complex reasoning, namely if we are not familiar with the context.

When acting in real world, **an agent must consider** both

- **facts and rules valid** in the considered world,
- **knowledge of his partners** (other agents).

Common knowledge example

Often we assume that all actors in complex situations share meaning of some notions, e.g.

Each driver knows that **red light** means ”stop”, **green** “go” and how these lights are located on a crossroad.

- Is it sufficient to feel safe on a crossroad?
- Consider “turning left”

To make traffic safe, we have to be sure, that everyone

- knows the rules (meaning of the signs),
- **follows it** and
- **knows that all the others follow it as well**, …!

In many situations it is necessary to assume that all the following observations are true simultaneously:

- *Everyone knows the fact* **F**,
- *Everyone knows, that everyone knows the fact* **F**,
- *Everyone knows that, everyone knows, that everyone knows the fact* **F**,
- …

Such a fact **F** is referred to as **common knowledge**. It is a prerequisite for

- Meaningful discussion
- Rational decision making, …

Muddy children puzzle.

Imagine n children playing together. Their mother warned them that if they get dirty there will be severe consequences. During their play some of the children, say k of them exactly, get mud on their forehead.

Along comes their father, who says "*At least one of you has mud on your forehead.*"

Provided $k > 1$, this is no surprise for any child !

The father asks "*Does any of you know whether you have mud on your forehead?*"

over and over again.

Can the children come to some conclusion?

Assume that all the children are perceptive, intelligent, truthful and answer simultaneously. ***k*** dětí je ušmudaných.

Let us denote the father's claim "*At least one of you has mud on your forehead.*" by the symbol ***p*****.**

If $k > 1$, it may seem that father provides no new information. All over this information is useful - why?

Before the father says ***p***, no child can come to an answer to the question ***„Does any of you know if you have mud on your forehead?"***

By induction on k it can be proven that for every round of q questions where $q < k$, all the children have to answer NO.

THUS: ***k* - *1*** times we will hear the answer NO and in the ***k***-th round all the children will answer YES.

Father mediates COMMON KMOWLEDGE !

Formal apparatus to work with knowledge

Kripke's idea of possible worlds semantics for modal logic:

- Besides the true state of affairs, there are a number of alternative states or "worlds" the agent can consider as possible.
- For example we cannot know what will be the weather tomorrow – we can exclude some states (minus 20^{o}C) but other plausible states have to be considered.

Definition. An agent *knows the fact* ***p*** , if ***p*** is true in all worlds the agent considers possible taking into account all information he has available.

Example. 3 logicians in a pub

Example.

Agent1 walks the streets of Prague, where it is sunny. He has no information about the weather in Berlin.

Thus

- **Agent1** has to consider only worlds where there is sunshine in Prague.
- But he can assume nothing about the sky in Berlin – it can be either gray or blue.

Agent1 knows in this case that there is sunny in Prague. But he does not know that there is sunny in Berlin.

Intuitive observation:

The number of considered possible worlds corresponds to vagueness!

The smaller is this number (of possible worlds the agent considers) the more accurate is his knowledge.

As soon as the agent gains some additional info from a reliable resource (e.g. it is sunny in Berlin), he can cancel all possible worlds contradicting the obtained fact.

We need **tools** that will help us **to do relevant reasoning**.

Modal logics provides a language for such reasoning

Let us consider a group of ***n*** agents named *1, 2, . . . , n,* who want to reason in a context that can be described using a set primitive propositions **Φ** denoted as

$$p, p', q, q', \ldots$$

These primitive propositions express the basic facts about the intended context, e.g. „*it is raining in Prague*", „*Mary has mud on her forehead*".

To express the claim "*Karin knows that it is raining in Berlin*" we need to enrich the language by modal operators

$$K_1, K_2, \ldots, K_n,$$

(each agent *i* has his own operator K_i), where the expression $K_n \boldsymbol{p}$ is read as „ *the agent **i** knows **p***". Further, we use connectives **¬** (*negation*) and conjunction **&** (often denoted **as** ∧)

Set of ***Formulas*** that can be constructed in the modal logics language.

$$p \in \Phi \quad \Rightarrow \quad p \in \textbf{\textit{Formulas}}$$

$$A, B \in \textbf{\textit{Formulas}} \quad \Rightarrow \quad \neg A, (A \wedge B) \in \textbf{\textit{Formulas}}$$

$$A \in \textbf{\textit{Formulas}} \text{ and } 1 \leq i \leq n \quad \Rightarrow \quad K_i A \in \textbf{\textit{Formulas}}$$

Standard abbreviations from propositional logics

$$A \vee B \quad \text{for} \quad \neg(\neg A \wedge \neg B)$$

$$A \rightarrow B \quad \text{for} \quad \neg A \vee B$$

$$A \leftrightarrow B \quad \text{for} \quad ((A \rightarrow B) \wedge (B \rightarrow A))$$

$$\mathit{true} \text{ for } p \vee \neg p$$

$$\mathit{false} \text{ for } \neg \mathit{true}$$

Expressivity.

a)

$$K_1 K_2 p \wedge \neg K_2 K_1 K_2 p$$

Agent1 knows, that **Agent2** knows ***p*** , but **Agent2** does not know, that **Agent1** knows, that **Agent2** knows ***p*** .

b) How do we express **possibility**?

Possibility is understood as a dual notion to **knowledge**.

Agent1 considers *A* to be possible if the Agent1 does not know ¬*A* for sure, i.e.

$$\neg K_1 \neg A$$

Semantics of modal logics

Kripke's semantics of possible worlds.

Kripke's structure M for n agents over a set $\boldsymbol{\Phi}$ of primitive propositions is the $(n+2)$-tuple $(\boldsymbol{S}, \boldsymbol{\pi}, K_1, K_2, \ldots, K_n)$

- $\boldsymbol{S}$ is a **set** of all worlds (that can be considered in the given context) or **states**,
- $\boldsymbol{\pi}$ is interpretation of states corresponding to a truth function evaluating all primitive propositions from $\boldsymbol{\Phi}$ for each state $\boldsymbol{s}$ separately, ie. $\pi(s) : \Phi \rightarrow \{true,\ false\}$
- binary possibility relations $K_1, K_2, \ldots, K_n$, on $\boldsymbol{S}$ interpreting the modal operators $K_1, K_2, \ldots, K_n$ (rel. K_i connects the state $\boldsymbol{s}$ with any state of $\boldsymbol{S}$ the agent $\boldsymbol{i}$ considers to be **possible alternatives** to $\boldsymbol{s}$).

Kripke's structure can be depicted as a **labeled oriented graph**:

Its nodes are the states ***s*** from ***S***. Each node ***s*** is labeled by the set of primitive propositions that are true in the state ***s*.**

Oriented edges are labeled by sets of the agents as follows:

An oriented edge from the node ***s*** to ***t*** is labeled by the index ***i*** iff the possibility relation K_i of the agent ***i*** contains the pair (***s***, ***t***).

If the relation K_i is symetric it is no more important to highlight orientation of the edges and the arrows are „skipped".

Possible worlds semantics. Given a formula A, we want to specify conditions under which „the formula A is true in the structure M and the state s" (denoted by).

$(M, s) \models A$ is sometimes read as "A holds in (M, s)" or "(M, s) satisfies A".

$(M, s) \models A$ is defined by **induction by the structure** of A :

$$(i) \quad (M,s) \models p \quad \text{iff} \quad \pi(s)(p) = true \quad \{p \in \Phi\}$$

$$(ii) \quad (M,s) \models \neg A \quad \text{iff} \quad (M,s) \not\models A$$

$$(iii) \quad (M,s) \models A \wedge B \quad \text{iff} \quad (M,s) \models A \text{ a } (M,s) \models B$$

$$(iv) \quad (M,s) \models K_i A \quad \text{iff} \quad (M,t) \models A$$

for all t, s.that $(s,t) \in K_i$

Example.

Let us consider $\Phi = \{p\}$, $n = 2$ and the following Kripke structure

$$M = (S, \pi, K_1, K_2)$$

(i) $S = \{s, t, u\}$

(ii) $\boldsymbol{p}$ is true in the states s and u , not in t, ie.

$$\pi(s)(\boldsymbol{p}) = \pi(u)(\boldsymbol{p}) = \mathit{true} \quad \text{a} \quad \pi(t)(\boldsymbol{p}) = \mathit{false}$$

(iii) The agent1 cannot distinguish the state s from t , ie.

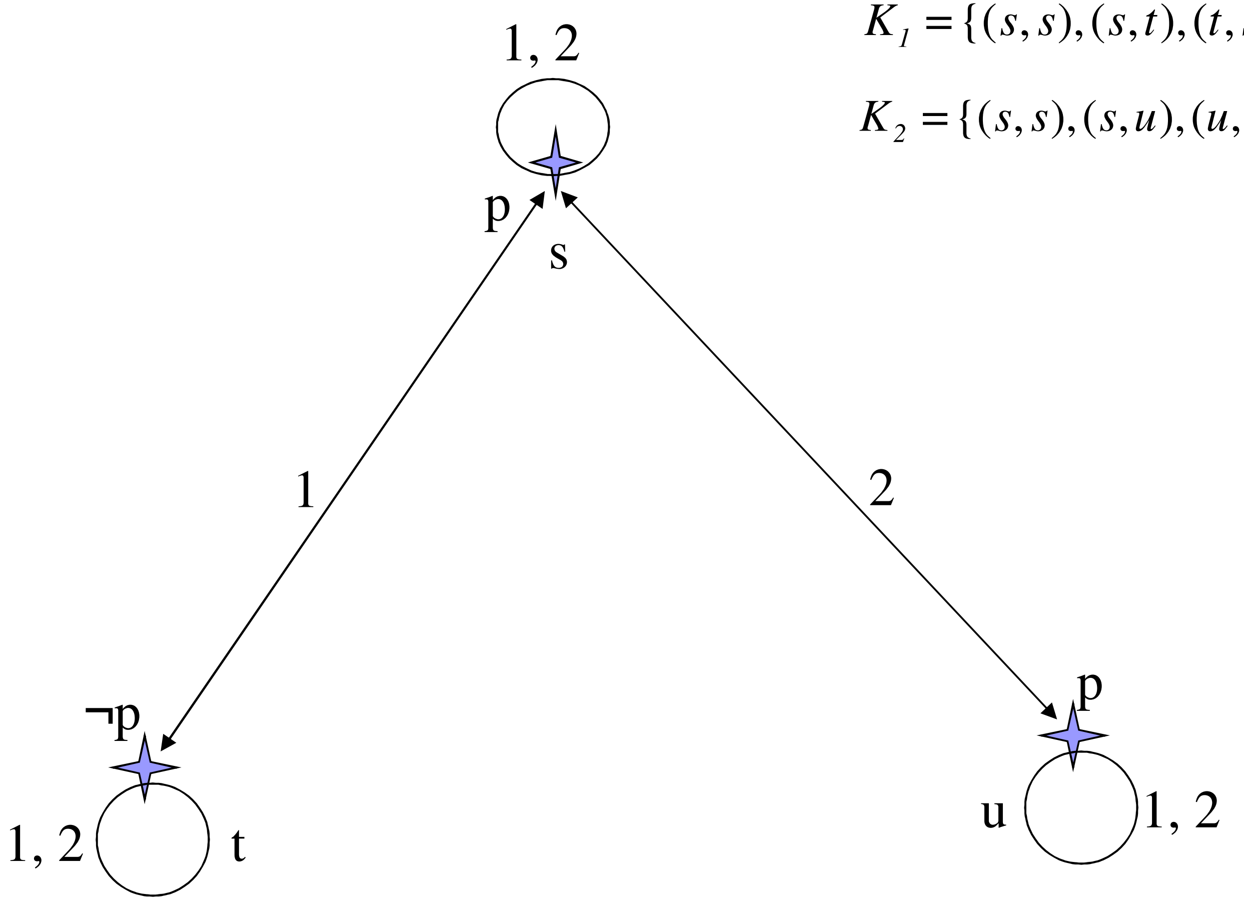
$$K_1 = \{(s,s),(s,t),(t,s),(t,t),(u,u)\}$$

For the agent2 there holds $K_2 = \{(s,s),(s,u),(u,s),(t,t),(u,u)\}$

$$\pi(s)(\boldsymbol{p}) = \pi(u)(\boldsymbol{p}) = \mathit{true} \quad \text{a} \quad \pi(t)(\boldsymbol{p}) = \mathit{false}$$

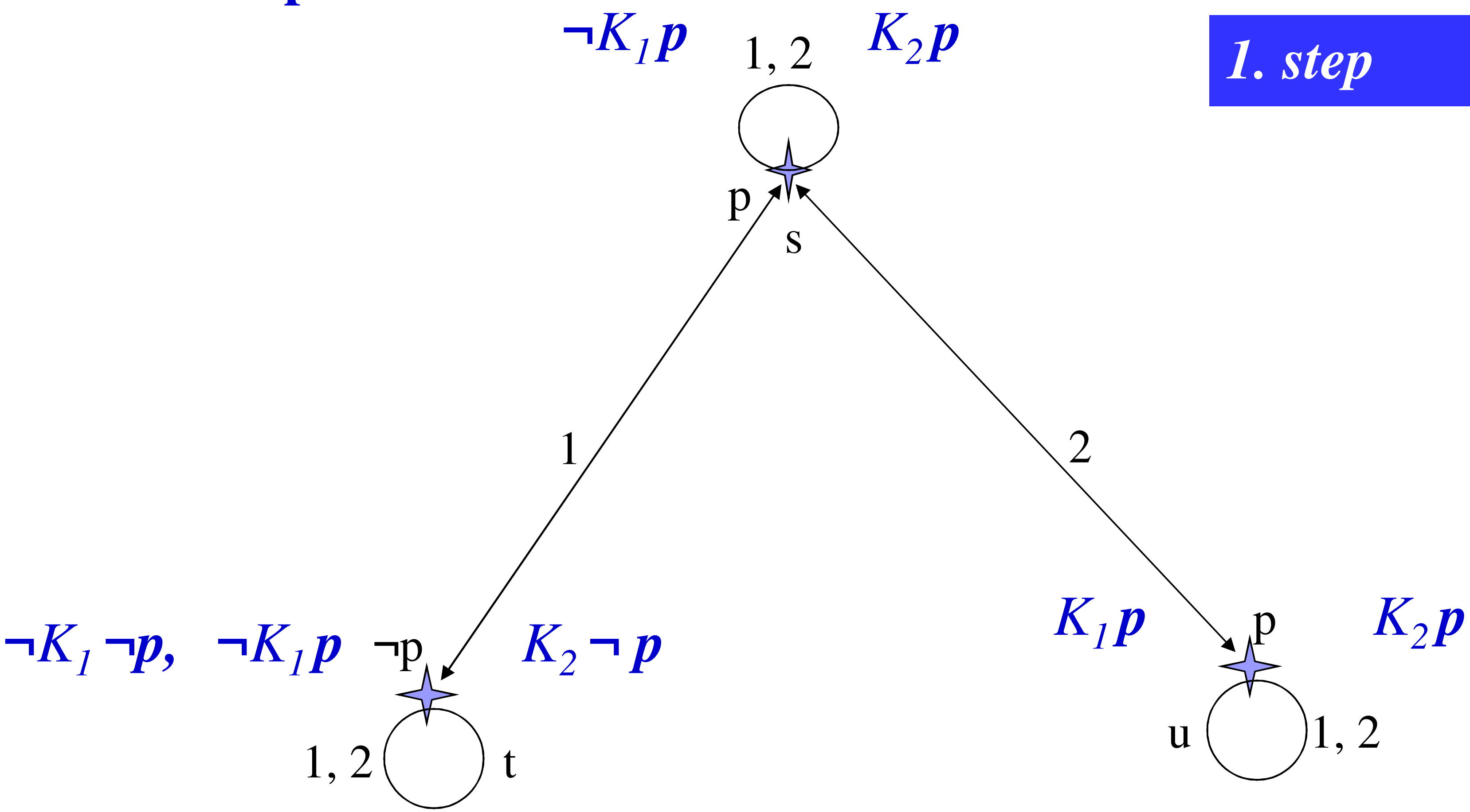
$$K_1 = \{(s,s),(s,t),(t,s),(t,t),(u,u)\}$$

$$K_2 = \{(s,s),(s,u),(u,s),(t,t),(u,u)\}$$



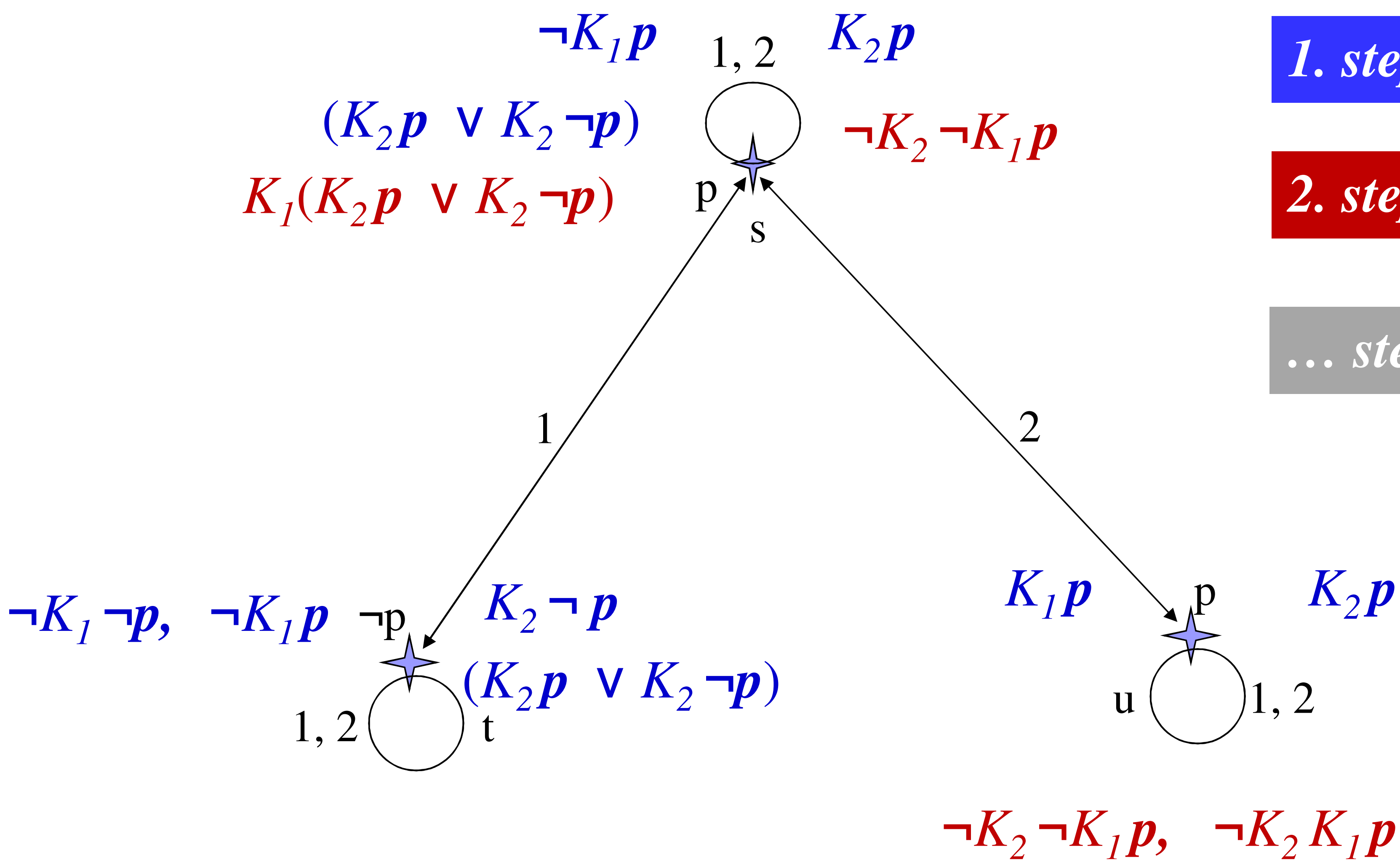
Truth values of some compound formulas **containing just 1 modal operator**

1. step



laboratory Gerstner

$$(M,s) \models p \wedge \neg K_1 p \wedge K_2 p \wedge K_1(K_2 p \vee K_2 \neg p) \wedge \neg K_2 \neg K_1 p$$

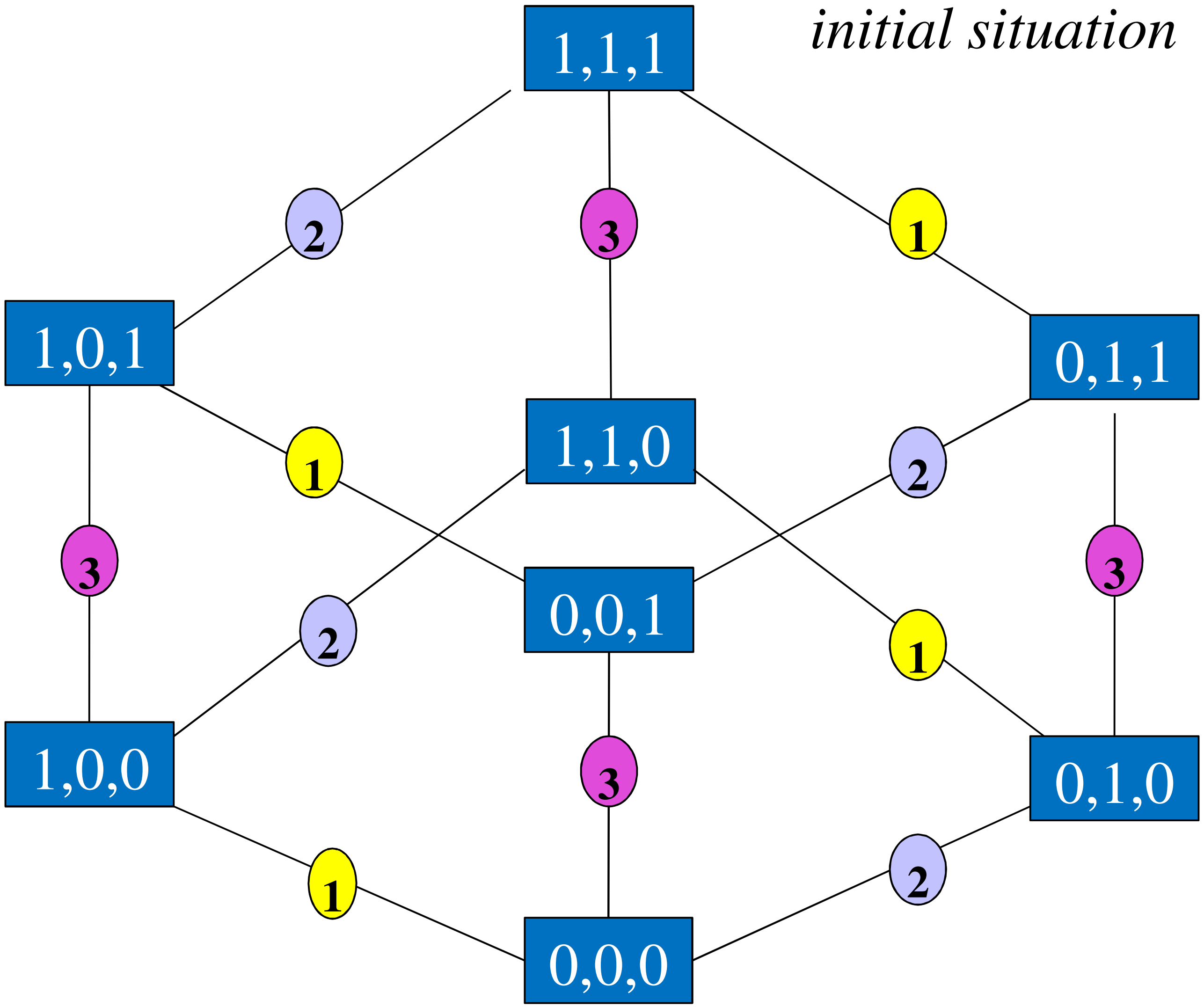


1. step

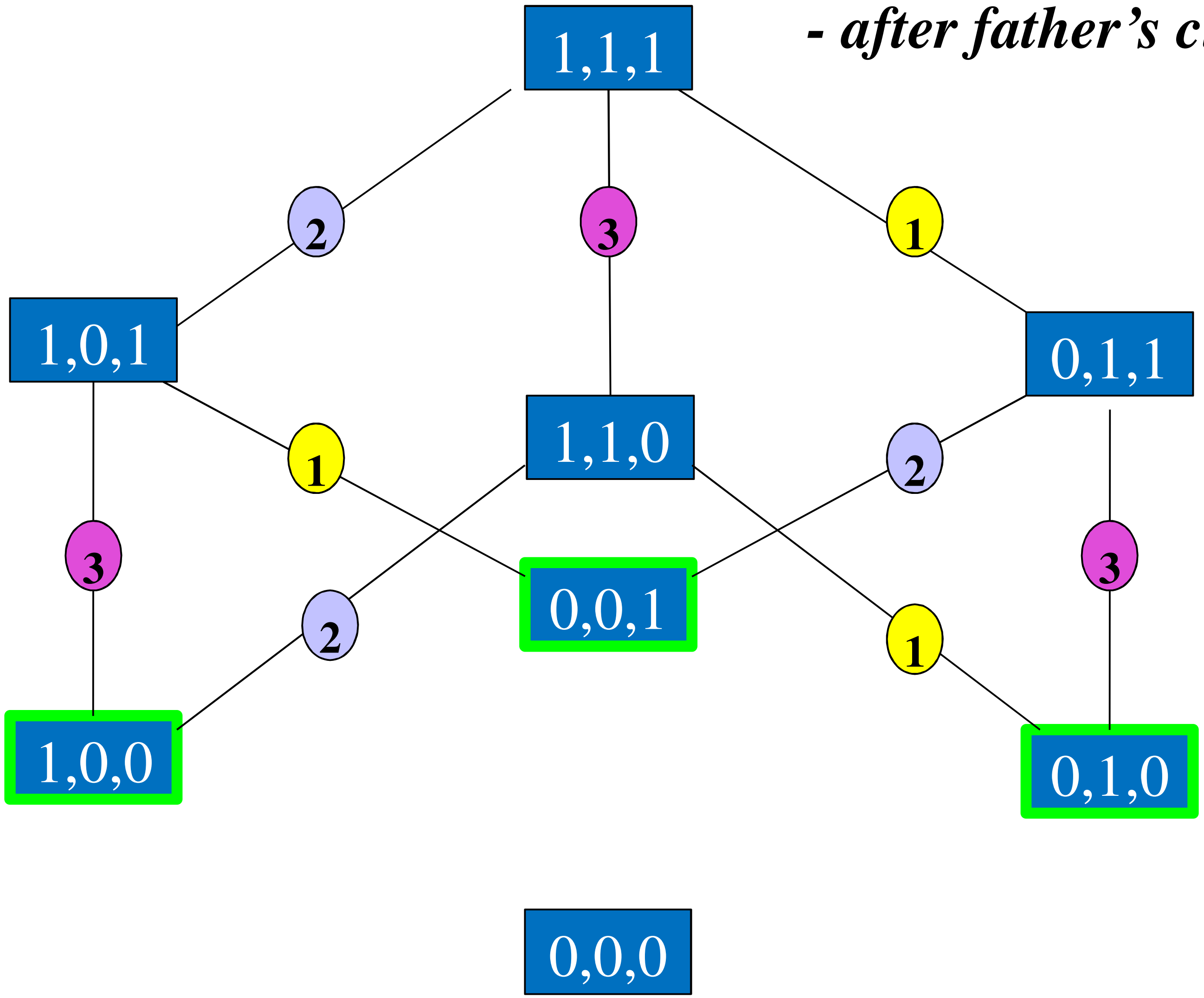
2. step

… steps

Kripke structure for 3 muddy children – *initial situation*



Kripke structure for 3 muddy children - after father's claim p



⇔

Let us assume that the **possibility relation** K_i has the properties of **equivalence**, namely K_i is reflexive, symetric and transitive.

Does this fact have an influence on the structure of the set of states **S** ?

For any $v \in \mathbf{S}$ let us define the set $\mathbf{S}_v^i$ of all the possible alternative states to v the agent i consideres possible $\mathbf{S}_v^i = \{t \in \mathbf{S}: (v,t) \in K_i\}$.

Suppose s, t are states from **S** such that $(s,t) \in K_i$, then the agent i has in both states s and t the same sets of possible alternative states $\mathbf{S}_s^i$ and $\mathbf{S}_t^i$, ie. $\mathbf{S}_s^i = \mathbf{S}_t^i$ and both these sets of states are for the agent i in this case undistinguishable!

This is a consequence of the assumption that K_i is an equivalence relation (namely it is symmetric and tranzitive).

Common and distributed knowledge

To express these notions three additional modal operators are introduced for any nonempty subset $\boldsymbol{G}$ of the set of all agents $\{1,2, \ldots .\ n\}$, namely

$$E_G \quad \{\text{"}\textit{everyone in the group G knows}\text{"}\}$$

$$C_G \quad \{\text{"}\textit{it is a common knowledge among the agents in G}\text{"}\}$$

$$D_G \quad \{\text{"}\textit{it is a distributed knowledge among the agents in G}\text{"}\}$$

If $\boldsymbol{A}$ is a formula, then $\boldsymbol{E_G A}$, $\boldsymbol{C_G A}$ and $\boldsymbol{D_G A}$ are formulas as well.

Examples of formulas with operators $\boldsymbol{E_G}$, $\boldsymbol{C_G}$ or $\boldsymbol{D_G A}$ and their reading

$K_3 \neg C_{[1,2]} p$ — Agent3 knows, that p is not common knowledge between the agents 1 a 2.

$D_G q \wedge \neg C_G q$ — q is distributed knowledge of agents in the group $\boldsymbol{G}$, but it is not a common knowledge there.

In order to define semantics of these operators we have to introduce (nested) iteration of the operator $\boldsymbol{E_G}$:

$$E_G^0 A \equiv A \qquad\qquad E_G^{n+1} A \equiv E_G E_G^n A$$

Let us define

$$(M,s) \models E_G A \quad \Leftrightarrow \quad (M,s) \models K_i A \; for\ all \;\; i \in G$$

$$(M,s) \models C_G A \quad \Leftrightarrow \quad (M,s) \models E_G^k A \; for\ all \; 1 \leq k$$

Both notions have an interesting graphical interpretation:

Let **G** be a nonempty set of agents. We say that the state ***t*** is ***G*-reachable** from the state ***s*** in $0 < k$ steps, if there is a sequence of states

$$s \equiv s_0, s_1, \ldots, s_k \equiv t$$

Such that, for any $j, 0 \leq j < k$ there exists $i \in G$ such that

$$(s_j, s_{j+1}) \in K_i.$$

We say that ***t*** is **G**-reachable from ***s***, if ***t*** is ***G***-reachable in finite number of steps.

Lemma.

$$(i)\ (M,s) \models E_G^k A \quad \Leftrightarrow \quad (M,t) \models A \quad \text{for any } t\text{, that is } G-\text{reachable in } k \text{ steps from } s$$

$$(ii)\ (M,s) \models C_G A \quad \Leftrightarrow \quad (M,t) \models A \quad \text{for any } t \text{ that is } G-\text{reachable from } s.$$

Proof.

(i) Can be proved by induction on k, (ii) is a consequence of (i).

Both claims are valid for any possibility relations K_i - no specific property (e.g. equivalence) is requested from K_i .

Let us denote by $\boldsymbol{p}$ the statement „one of the children is muddy“

If we consider just the set of 2 children ($k = 2$) and the situation when both children are muddy. It is not hard to show, that each child is in a state, where it is true that „this child knows p“, but the claim „everyone knows that everyone knows p“ is not true.

Similarly, if $k = 3$ and all are muddy, the claim „everyone knows, that everyone knows p“ holds. But the following statement does not hold „ everyone knows, that everyone knows, that everyone knows p“

Excercise. Suppose there are exactly k muddy children. Before the father´s first claim each child is in a state where $\boldsymbol{E}^{(k-1)}\boldsymbol{p}$ holds but where $\boldsymbol{E}^{k}\boldsymbol{p}$ does not hold.

Task 4. Who is the best?

Deadline for submition: 9.5. 09:00 am.

Three Masters of Logic want to find out who was the wisest amongst them and asked their Grand Master to resolve their dispute. "Easy," the old sage said:

"I will blindfold you and paint either red, or blue dot on each man's forehead. When I take your blindfolds off, if you see at least one red dot, **raise your hand**. The one, who guesses the color of the dot on his forehead first, wins.„ And so it was done …
When he took their blindfolds off, all three men raised their hands as the rules required, and sat in silence pondering. Finally, one of them said: "I have a red dot on my forehead."

A: How did the winner guess?

Task 4b. Who is the best?

After losing the "Spot on the Forehead" contest, the two defeated Puzzle Masters complained that the winner had made a slight pause before raising his hand … . And so the Grand Master vowed to set up a ***truly fair test*** to reveal the best logician amongst them.

He showed the three men 5 hats - two white and three black. Then he turned off the lights in the room and put a hat on each Puzzle Master's head. After that the old sage hid the remaining two hats, but before he could turn the lights on, one of the Masters (as chance would have it, the winner of the previous contest) announced the color of his hat. And he was right once again.

B: What color was the winner´s hat? What could have been his reasoning?

Recommended resources

- Ronald Fagin, Joseph Y. Halpern, Yoram Moses, Moshe Y. Vardi: *Reasoning About Knowledge*, MIT Pres 1995, 2003
- Chapter 8 in the volume *Umělá inteligence*(6), Academia 2012