Knowledge and some real-life problems

laboratory Gerstner

Knowledge base

• is a system *obtaining facts about its environment* that is supposed to *answer some queries*.

• ***Knowledge base KB** is one of the agents in a larger system*, where at least the following agents are present:

•***Environment*** is an agent representing the model of the external world.

•***Teller*** is the agent telling the ***KB*** the facts about the external world.

Environment´s state is expected to provide complete description (of the relevant features) of the external world.

• *Local state of **KB*** describes information the ***KB*** has obtained about the external world up tp now.

•*Local state of **Teller*** contains all he knows about „the external world“ + „the knowledge up to now provided to ***KB***“ ...

This informal description is still rather vague – it offeres many possibilities for modelling the global states.

The simples choice could be based on the following constraints:

• The external world can be described using a set of primitive propositions Φ.

• The external world is *stable*, ie. the truth values of the primitive propositions about the external world do not change in time.

• The *Teller* has full information about the external world.

• All information *stored in **KB*** is *valid*.

• There is *no a priori knowledge* about the external world or about the content of ***KB***.

For simplicity let us assume, that

• The **external world** can be described by an evaluation **α** of primitive propositions from **Φ** (and this **α** remains fixed!).

• The **Teller** has access to the evaluation **α** and to the sequence of facts, he has provided as input into ***KB***.

• ***KB´s*** **local state** contains the sequence $A_1, \ldots, A_n$ of formulas provided by the **Teller** as input (it could be either ***propositional*** *or* ***modal formulas***).

• **Global state** is denoted by $(\alpha, \langle A_1, \ldots, A_n \rangle, .)$

1. ***KB*** that stores **propositional formulas, only!**

- ***KB*** stores ***propositional formulas*** created from the primitive propositions **Φ.**

- The ***queries*** subjected to ***KB*** have the form of propositional formulas created from propositions **Φ .**

- All formulas *stored* in ***KB*** *are true* under the evaluation **α.**

- There are no *a priori assumptions* about the external world and about information provided by the ***Teller*** to ***KB***.

  These assumptions represent the constraints that will be used to construct the corresponding Kripke structures.

Let I^{kb} describe gradual addition of information into the ***KB*** in the form of a sequence *r* of the global states. Let us define

$(I^{kb}, r, m) \models K_{KB}\varphi$ iff φ holds in **all the next time points**.

How the *KB* should answer a query *B* ?

Choice **a)**

Suppose ***KB*** is asked at a certain moment (r,m) the query ***B***, where ***B*** is a propositional formula. Our knowledge base has no access to the external world – that is why ***B*** cannot be interpreted as a query about the external world! It can be understood as a question „What the ***KB*** knows about the external world, ie. $\boldsymbol{K_{KB}(B)}$**?** **“**.

The KB should answer

$$\begin{cases} \text{YES} & \text{if } (\mathcal{I}^{kb}, r, m) \models K_{KB} B \\ \text{NO} & \text{if } (\mathcal{I}^{kb}, r, m) \models K_{KB} \neg B \\ \text{Do Not Know} & \text{else} \end{cases}$$

Choice b)

The ***KB*** remembers conjunction of all information provided up to now.

Suppose ***KB*** is in the local state

$\langle A_1, \dots, A_k \rangle$, denote $\kappa = A_1 \wedge \dots \wedge A_k$

and let the knowledge base know everything what is a consequence of κ.

In such a case the query ***B*** can get a positive answer in 2 cases:

$$\text{YES iff} \begin{cases} B \quad \text{is a consequence of } \kappa \\ \text{or} \\ K_{KB}B \ \text{ is a consequence of } K_{KB}\kappa \end{cases}$$

*If **KB** contains only propositional formulas about the external world* and no facts about ***KB*** itself, then the answers to propositional queries are the same for both choices a) and b):

KB **theorem 1.**

Let us assume that KB contains propositional formulas only, ie.

$$r_{KB}(m) = \langle A_1, \ldots, A_k \rangle, \quad \kappa = A_1 \wedge \ldots \wedge A_k$$

and B is a propositional formula.

In this case the following are **equivalent**

$(i)\ (\mathcal{I}^{kb}, r, m) \models K_{KB} B$

$(ii)\ \kappa \rightarrow B$ is a propositional tautology

$(iii)\ M_n^{rst} \models K_{KB}\kappa \rightarrow K_{KB} B$

What happens if the query is not limited to **propositional formulas ?**

Let us consider a query $B \equiv (p \rightarrow K_{KB}\, p)$

"Is it the case, that if p holds then KB knows it?"

How the answer for the query ***B*** should be obtained? Can we apply the choice a) used earlier? Namely can the answer to ***B*** be

$$\begin{cases} \text{YES} & \text{if } (\mathcal{I}^{kb}, r, m) \models K_{KB} B \\ \text{NO} & \text{if } (\mathcal{I}^{kb}, r, m) \models K_{KB} \neg B \\ \text{Do Not Know} & \text{else} \end{cases}$$

What we know about validity of the formula $K_{KB}(p \rightarrow K_{KB}\, p)$? (1)

According to **Mod_T8 a, Mod_T8b**) we know that $K_{KB}\,(p \rightarrow K_{KB}\, p\,)$ is equivalent to the formula

$$K_{KB}\, p \;\vee\; K_{KB} \neg\, p \qquad (2)$$

Now, we are ready to answer the query $\boldsymbol{B = (p \rightarrow K_{KB}\, p)}$ by considering the equivalent transcriptions

1. $\boldsymbol{K_{KB}\,(p \rightarrow K_{KB}\, p)} = K_{KB}\, p \lor K_{KB} \neg p$
2. $\boldsymbol{K_{KB} \neg(p \rightarrow K_{KB}\, p)} = K_{KB}\,(p \;\&\; \neg K_{KB}\, p)$

We have proven earlier in **Mod_T7** that the formula $K_{KB}\,(p \;\&\; \neg K_{KB}\, p)$ is contradictory (and thus cannot be proven)!
Consequently, the query ***B*** can never result in the answer NO!

To answer the query $\boldsymbol{B = (p \rightarrow K_{KB}\, p)}$ let us consider the equivalent transcription $\boldsymbol{K_{KB}\,(p \rightarrow K_{KB}\, p)} = K_{KB}\, p \lor K_{KB} \neg p$. Thus the answer is

YES, if p or $\neg p$ is a consequence of KB content

Do Not Know else.

Does it make sense to provide *KB* with input described by non-propositional formulas ?

Let us assume, that ***KB*** has current information represented by a sequence of formulas $<F_1, \ldots, F_i>$ and the next provided information is $\boldsymbol{F_{i+1} = (p \rightarrow K_{KB}\, p)}$ saying „if ***p*** is valid, ***KB*** knows it“.

This information can be useful, if ***KB*** can verify what it knows and what it does not know. If there holds $<F_1, \ldots, F_i> \vdash \neg K_{KB}\, p$, then ***KB*** can deduce from the input $\boldsymbol{F_{i+1}}$ that p does not hold.

This shows that thanks to the non-propositional input the ***KB*** can gain (derive) new information about the external world using **introspection.**

If KB obtains non-propositional input**, its knowledge can be** no more represented as a ***conjunction of the input sequence*** (as in the propositional case).

On the next page we are explaining an ***example situation***, when the provided input represents a fact, that

- has been true at the moment of input,
- does not remain true in any later time point.

Under such conditions it is impossible to describe knowledge in the form of ***conjunction of provided input***. ***Why?*** The upper conjunction results in a contradiction (from which anything can be derived) !

Example: Assume ***KB*** obtains as input a fact, that

- is true in the current moment,
- but it does not remain valid in any next state.

Let the primitive proposition ***p*** about the external world be valid, but the ***KB*** does not know ***p*** yet (until the tact ***j*** no such information has been provided). In such a situation there must hold

$$\boldsymbol{p} \;\&\; \neg K_{KB}\, \boldsymbol{p} \qquad (3)$$

This can be provide by the *Teller* in the tact ***j* +1**. But even provided this information (3), we cannot claim that ***KB*** knows (3), ie. (3) holds in al the future time points. If this would be the case, it would have to be true that $K_{KB}(p \;\&\; \neg K_{KB}\, p)$

But this is impossible! This formula contradicts S5.

$$\boldsymbol{p} \,\&\, \neg\, K_{KB}\, \boldsymbol{p} \qquad (3)$$

We have learned already through the analysis of the formula (3) that, if ***KB*** gets the information φ, it **does not have to be** the case that $K_{KB}\,\varphi$ knows it.

Intuitively, the ***KB*** should gain something after it has been told (3) : it should know, that $\boldsymbol{p}$ is true! Consequently input of (3) should result in $K_{KB}\boldsymbol{p}$.

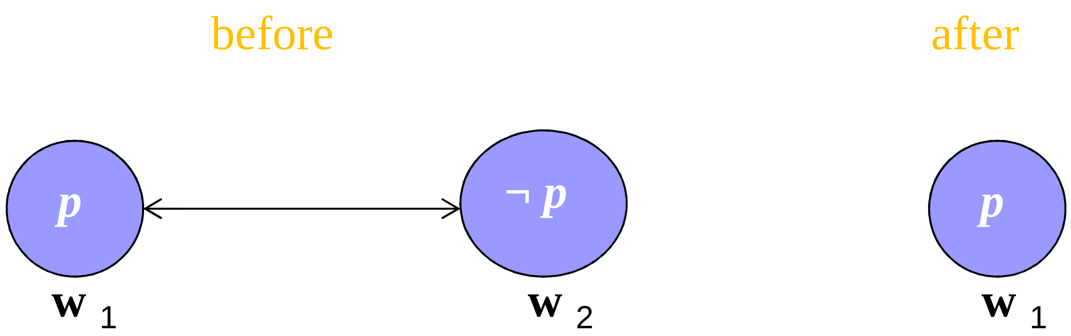
How to make this intuition true? The knowledge base cannot treat modal formulas in the same way as the propositional formulas!

A new approach has to be designed. The viable solution is based on **modal logics**!

Common knowledge

Can a public announcement result in common knowledge?

- Prohlášení směrem k panu *i* : „*You do not know it, but I have to tell you that you have some bug on your back.*“
- Let us use ***p*** for the claim „*You have some bug on your back .*“ and γ for the formula $\neg K_i p \,\&\, p$



Obviously, $K_i \gamma$ does not hold in the state w_1. Consequently, it cannot be the case that $C\gamma$ holds.

Common knowledge in ASMP (asynchronous message passing systems)

Asychronous sending of messeges

- Typical knowledge of the agent 1 in time 0: K_1 „*Agent* 2 *did not obtain any message from me*, *yet*.“ (denoted as $K_1\,\alpha$)
- Agent 1 sends a message „*Hello*.“ to the agent 2.
- This act of sending a message results in a **loss** of certain knowledge, namely $K_1\,\alpha$

It can be shown that there is a close connection between common knowledge and synchronization of agent’s activities!

Communication and common knowledge

Let us consider 2 generals, who communicate by sending letters through human messengers: When can they be sure that both their armies can attack jointly (in coordinated way)?

Common knowledge is closely related to communication means used in the system. It can be proven, that in an asynchronous system no new knowledge or common knowledge can be gained through **sending messeges**:

Theorem: For any structure ***S*** and any formula ψ there holds that in any time development of this system (described by a run ***r***) and any instant ***m*** :

$$(S,r,m) \models C_G\psi \text{ if and only if } (S,r,0) \models C_G\psi$$

Recommended resources

- Modal logic

Ronald Fagin, Joseph Y. Halpern, Yoram Moses, Moshe Y. Vardi: *Reasoning About Knowledge*, MIT Pres 1995, 2003

- Temporal logic

Michael Fisher: *Introduction into Formal Methods Using Temporal Logic*, John Wiley & Sons, 2011

- Modal and temporal logics

Chapter 8 in the volume *Umělá inteligence*(6), Academia 2012

FIRST PhD Autumn School on Modal Logic, November 10-11 2009, materials from the courses *Temporal Logics for Specification and Verification* (V. Goranko), *Computational Modal Logic* (C. Areces, P. Blackburn), http://hylocore.ruc.dk/m4m6school.html