# Aplikace:

# Znalostní báze

## Znalostní báze

• je systém, který *dostává fakta o  prostředí a dotazy* o něm.

• *Znalostní báze je agentem ve větším systému*, který obsahuje prostředí (také agent), správce (agent), popřípadě další agenty.

• *Správce* ukládá do znalostní báze fakta o prostředí.

• *Prostředí* používáme jako model vnějšího světa.

• Chceme, aby *stav prostředí* poskytoval úplný popis (relevantních rysů) vnějšího světa.

• *Lokální stav znalostní báze* popisuje informace, které báze má o světě a o lokálním stavu správce.

Tento neformální popis poskytuje ještě hodně volnosti jak modelovat globální stavy.

V nejjednodušším případě přijímáme tato omezení

• *vnější svět lze popsat výrokově* pomocí výroků z konečné množiny $\Phi$.

• vnější svět *je stabilní*, tj. pravdivostní hodnoty výroků popisujících svět se s časem nemění.

• *Správce* má úplnou informaci o vnějším světě.

• Vše, *co je uloženo* ve znalostní bázi *je pravdivé*.

• Nejsou *žádné předběžné (a priori) znalosti* o vnějším světě nebo o tom, co bude do znalostní báze uloženo.

Pro reprezentaci systému plyne

• vnější svět lze popsat pravdivostním ohodnocením $\alpha$ prvotních formulí z množiny $\Phi$, které se v žádném běhu nemění.

• stav správce obsahuje ohodnocení $\alpha$ a posloupnost fakt, která do báze uložil.

• lokální stav znalostní báze obsahuje posloupnost $A_1, \ldots, A_n$ faktů, které byly uloženy. Jsou to výrokové formule.

• globální stav označujeme $(\alpha, \langle A_1, \ldots, A_n \rangle, .)$

- *Do znalostní báze* se ukládají a jsou dotazovány *jen fakta o vnějším světě* (vyjádřitelná ve výrokových formulích) a ne fakta o bázi samé.
- Vše, *co je uloženo* ve znalostní bázi *je pravdivé*.
- Nejsou *žádné předběžné (a priori) znalosti* o vnějším světě nebo o tom, co bude do znalostní báze uloženo.

Dostáváme interpretovaný systém $\mathcal{I}^{kb} = (\mathcal{R}^{kb}, \pi^{kb})$,
kde $\mathcal{R}^{kb}$ je množina všech běhů takových, že pro něja - kou posloupnost výrokových formulí $A_1, \ldots, A_n$
platí

KB1. $r(0) = (\alpha, \langle\ \rangle, .)$

KB2. $r(m) = (\alpha, \langle A_1, \ldots, A_n \rangle, .)$ implikuje

($i$) buď $r(m+1) = r(m)$ nebo $r(m+1) = (\alpha, \langle A_1, \ldots, A_k, A_{k+1} \rangle, .)$,

($ii$) $A_1 \wedge \ldots \wedge A_k$ je pravdivá při ohodnocení $\alpha$ a

($iii$) $\pi^{kb}(r, m) = \alpha$, tedy $\pi^{kb}$ je definováno tak,

aby pravdivostní ohodnocení v bodu $(r, m)$

bylo dáno stavem prostředí

Náš předpoklad, že $\mathcal{R}^{kb}$ sestává ze všech běhů, které splňují KB1. a KB2. vylučuje předběžnou informaci o tom, co bude vloženo.

Definujeme-li interpretovaný systém $\mathcal{I}^{kb}$ můžeme říci, jak znalostní báze odpovídá na dotazy.

Předpokládejme, že v bodu $(r,m)$ je báze tázána dotazem $B$, kde $B$ je výroková formule. Protože báze nemá přímý přístup ke stavu prostředí, $B$ nemůže být interpretováno jako dotaz po znalostech báze o vnějším světě.

Znalostní báze by měla odpovědět

$$\begin{cases} \text{ANO} & \text{jestliže } (\mathcal{I}^{kb}, r, m) \models K_{KB} B \\ \text{NE} & \text{jestliže } (\mathcal{I}^{kb}, r, m) \models K_{KB} \neg B \\ \text{NEVÍM} & \text{jinak} \end{cases}$$

Většinou si znalostní báze pamatuje konjunkci toho, co do ní bylo uloženo.

Předpokládejme, že báze je v lokálním stavu $\langle A_1, \ldots, A_k \rangle$, že $k = A_1 \wedge \ldots \wedge A_k$ a znalostní báze ví jenom to, co plyne z $k$.

Potom báze může odpovědět dvojím způsobem

$$\text{ANO právě když} \quad \begin{cases} B \ \text{ je důsledkem } \ k \\ \text{nebo} \\ K_{KB}B \ \text{ je důsledkem } \ K_{KB}k \end{cases}$$

obojí vyjde na stejno.

## Věta 1.

Předpokládejme, že

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

a $B$ je výroková formule.

Potom následující tvrzení jsou ekvivalentní

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

$(ii)$ $k \to B$ je výroková tautologie

$(iii)$ $M_n^{rst} \models K_{KB}k \to K_{KB}B$

## **Lemma 1.** (Validnost ve třídách $M_n$ a $M_n^{rst}$)

Nechť $M_n$ a $M_n^{rst}$ jsou třídy Kripkeho struktur , kde posibi-litní relace $K_i$ jsou obecné, respektive relace ekvivalence. Potom platí

(a) pro libovolné dvě formule $A$ , $B$

$M_n$ |= $A$ -> $B$ právě když $M_n$ |= $K_iA$ -> $K_iB$

(b) pro libovolné dvě výrokové formule $A$ , $B$

$M_n^{rst}$ |= $A$ -> $B$ právě když $M_n^{rst}$ |= $K_iA$ -> $K_iB$ navíc

$M_n^{rst}$ |= $A$ -> $B$ právě když $A$ -> $B$ je výroková tautologie

(c) pro libovolné dvě formule $A$ , $B$ platí

$M_n^{rst} \models$ A -> B implikuje $M_n^{rst} \models K_i A \to K_i B$ , ale

existují formule $A$ , $B$ takové, že
$M_n^{rst} \models K_i A \to K_i B$ , ale $M_n^{rst} \not\models A \to B$

Důkaz.

(a) (->) Nechť $M_n \models A \to B$ . Potom
$M_n \models K_i(A \to B)$ (pravidlo generalizace) (1)

Ve třídě $M_n$ je pravdivý Kripkeho distribuční axiom
$K_i(A \to B) \to K_i A \to K_i B$ (2)
odkud z (1) a (2) pravidlem Modus ponens (které je také pravdivé v téže třídě) odvodíme

$M_n \models K_i A \to K_i B$

Předpokládejme, že $\mathcal{M}_n \models K_i A \rightarrow K_i B$

potom $\mathcal{M}_n \not\models K_i A$ nebo $\mathcal{M}_n \models K_i B$

nechť $M_n$ je libovolná Kripkeho struktura ze třídy $\mathcal{M}_n$ a $s$ je její libovolný stav, dostáváme

$(M, s) \not\models K_i A$ nebo $(M, s) \models K_i B$ odkud pro nějaké $t$, $(s, t) \in K_i$ je $(M, t) \not\models A$ nebo pro libovolné $u$, $(s, u) \in K_i$ $(M, u) \models B$ tedy $\mathcal{M}_n \models A \rightarrow B$ protože $M$ a $s$ byly zvoleny libovolně.

(b) První část tvrzení je důsledkem toho, že třída $\mathcal{M}_n^{rst}$ je podtřídou $\mathcal{M}_n$.

Nechť $A$, $B$ jsou výrokové formule a nechť $p_1, \ldots, p_k$ jsou všechny primitivní výrokové konstanty z $\Phi$ vyskytující se v implikaci $A \rightarrow B$.

Protože třída $M_n^{rst}$ sestává ze všech Kripkeho struktur $M = < S , \Phi, \pi , ... >$ obsahuje také všechna pravdivostní ohodnocení $\pi$ uvedených výrokových konstant. Protože implikace $A$ -> $B$ je pravdivá ve všech takových strukturách, je to výroková tautologie.

(c) První část tvrzení plyne z (a). Důkaz druhé části tvrzení. Ve třídě $M_n^{rst}$ je pravdivý Axiom positivní introspekce: pro libovolnou formuli $A$ je pravdivá formule

$$K_i A \text{ -> } K_i K_i A$$

ale formule

$$A \text{ -> } K_i A$$

nemusí být pravdivá (co je pravda, nemusí agent vědět). Tím je důkaz lemmatu 1 dokončen.

# Důkaz Věty 1.

Podle lemmatu 1, (b) platí

. $k \text{ -> } A$ je tautologie právě když $M_n^{rst} \models K_i k \text{ -> } K_i A$

odtud plyne ekvivalence (b) a (c).

Nyní ukážeme, že (b) implikuje (a). Předpokádejme tedy, že $k \text{ -> } A$ je tautologie.

Je-li $(r,m) \sim_{KB} (r',m')$ , potom

. $r'(m') = (\alpha', <A_1 , \dots , A_k >, \bullet)$ pro nějaké $\alpha'$

Protože vše, co bylo do znalostní báze KB vloženo je pravdivé, je i konjunkce $k$ pravdivá při pravdivostním ohodnocení $\alpha'$, takže musí platit

. . .
$$(I^{kb},r,m) \models k$$

Protože
.
$$M_n^{rst} \models k \to A$$
.
platí
$$(I^{kb},r',m') \models A$$
.
a tedy také
$$(I^{kb},r,m) \models K_{KB} A$$
.

Tím je (a) dokázáno.

Zbývá dokázat, že (a) implikuje (b). Postupujeme sporem.

Předpokládejme, že (b) neplatí, tedy že $k \to A$ není tautologie.
To znamená, že existuje pravdivostní ohodnocení α', takové, že je pravdivá formule $k \,\&\, \neg A$ .

Protože množina běhů sestává ze všech běhů, které splňují podmínky KB1 a KB2, snadno se dá ukázat, že existuje bod $(r',m')$ takový, že

.
$$r'(m') = (\alpha', <A_1, \ldots, A_k>, \bullet)$$

Protože

.
$$(r,m) \sim_{KB} (r',m') \quad \text{a} \quad (I^{kb},r',m') \models \neg A$$

dostáváme $(I^{kb},r,m) \not\models K_{KB} A$ takže (a) neplatí.

Tím je důkaz Věty 1 dokončen. Tato věta charakterizuje způsob jakým znalostní báze odpovídá na dotazy, které jsou výrokové formule.

Jak by měla odpovídat na dotazy, které nejsou výrokové, například $B \equiv (p \to K_{KB}\, p)$ "jestliže platí $p$, ví to znalostní báze?"

Zde bychom také rádi měli odpovědi na dotaz $B$:

$$\begin{cases} \text{ANO} & \text{jestliže } (\mathcal{I}^{kb}, r, m) \models K_{KB} B \\ \text{NE} & \text{jestliže } (\mathcal{I}^{kb}, r, m) \models K_{KB} \neg B \\ \text{NEVÍM} & \text{jinak} \end{cases}$$

Kdy platí formule $B \equiv K_{KB}(p \to K_{KB}\, p)$ ? (1)

Ukážeme, že formule (1) je ekvivalentní formuli

$$K_{KB}\, p \vee K_{KB} \neg p.$$

Předpokládáme, že relace $K_{KB}$ je relací ekvivalence, tedy pracujeme ve třídě Kripkeho struktur $M_n^{rst}$, která je úplnou axiomatizací modálního systému $S5_n$. Tvrzení platí obecně pro operátory $K_i$.

Nechť $M$ je libovolná Kripkeho struktura ze třídy $M_n^{rst}$, a $s$ je libovolný její stav. Potom platí

$(M,s) |= K_i (p \text{ -> } K_i p)$ právě když *pro každé* $t$, $(s,t)$ ε $K_i$, . .
.
$(M,t) |= p \text{-> } K_i p$

právě když *pro každé t*, $(s,t)$ ε $K_i$, $(M,t)|= p\text{->} K_i p$

. - „ - $(M,t)|= \neg p$ v $K_i p$

. - „ - (symetrie) $(M,t)|= \neg p$ *nebo* $(M,s)|= K_i p$

. $( M,s)|= K_i\neg p$ *nebo* $(M,s)|= K_i p$

tedy $(M,s)|= K_i\neg p$ v $K_i p$

Odpověď bude ANO, pokud $p$ nebo $\neg p$ plyne z toho, co bylo do báze uloženo a NEVÍM jinak.

Přitom odpověď NE není možná, protože pro formuli

$$B \equiv (p \to K_{KB}\, p)$$

platí

$$K_{KB} \neg B \quad \leftrightarrow \quad K_{KB}(p \wedge \neg K_{KB}\, p)$$

a na pravé straně ekvivalence je formule sporná s S5

(cvičení)

Zajímá nás, co lze říci o dotazech týkajících se jen znalostí báze. Pro tento účel můžeme využít výsledku pro Kripkeho struktury ze třídy $M_1^{rst}$ tedy věty ze systému $S5_1$ .

**Lemma 2.** (S$_1$5)

Pro libovolnou formuli v jazyku $L_1$ tvaru $K_1 A$ lze efektivně sestrojit ekvivalentní formuli $B$ , která je Booleovskou kombinací formulí $K_1 B_1$ , ... , $K_1 B_k$ , kde formule $B_1$ , ... , $B_k$ , jsou výrokové.

Důkaz. Indukcí podle složitosti formule $A$ .

**Definice.** (Booleovská kombinace)

Je-li $\Sigma$ třída formulí v nějakém jazyku $L$, říkáme, že *formule $A$ je booleovskou kombinací formulí ze třídy* $\Sigma$, jestliže formule $A$ náleží do třídy $\Sigma$' , která vznikne ze třídy $\Sigma$ uzavřením na negaci a konjunkci (resp. disjunkci).

Důkaz lemmatu. Předpokládejme, že $\Sigma$ je třída všech formulí tvaru $K_1 B$ , kde $B$ je výroková formule.

(a) $A$ je tvaru $K_1 B$ , kde $B$ je výroková formule. Potom $A$ prvkem třídy $\Sigma$ a je tedy booleovskou kombinací formulí tvaru $K_1 B$ , kde $B$ je výroková formule.

(b) Formule $A$ je tvaru $\neg B$ a pro $K_1 B$ již umíme sestrojit booleovskou kombinaci (píšeme $BK( \ldots )$ )

$$|= K_1 B \ <\text{-}> \ BK\,(K_1 B_! \,, \ldots , K_1 B_k) \qquad (1)$$

Potom pro libovolnou Kripkeovou strukturu $M$ a její libovolný stav $s$

$(M,s)|= K_1 \neg B <=>$ *pro lib.* $t$, $(s,t)\ \varepsilon\ K_1 \quad (M,t)|= \neg B$

tedy v $(M,s)$ neplatí $K_1 B$ , a proto $(M,s)|= \neg K_1 B$

Podle (1) $\quad |= \neg K_1 B \ <\text{-}> BK'(K_1 B_1 \,, \ldots )$ tedy

$K_1 A$ je booleovskou kombinací $BK' (K_1 B_1 , \ldots )$ .

(c) $A$ je tvaru $B$ v $C$ , potom není těžké nahlédnout, že $|= K_1 A <-> K_1 B$ v $K_1 C$ a formule $K_1 A$ je booleovskou kombinací formulí požadovaného tvaru.

(d) $A$ je tvaru $K_1 B$ , kde $K_1 B$ je booleovskou kombinací formulí tvaru $K_1 C$ , kde $C$ je výroková formule, potom užitim (b) a (c) dostaneme, že $K_1 A$ je také booleovskou kombinací formulí požadovaného tvaru.

V případu znalostních bází můžeme považovat bázi za jediného agenta. Správce jako agent má informaci o stavu vnějšího prostředí a o výrokových formulích, které byly do báze vloženy, znalostní báze ve svém stavu zná vložené formule a zajímá-li nás jen jak bude báze odpovídat otázky týkající se jejích znalostí, můžeme $K_{1KB}$

odpovídat otázky týkající se jejích znalostí, můžeme $K_{KB}$ po važovat za jedinou modalitu.

**Definice.** ($KB$-formule)

(i) Říkáme, že $B$ je $KB$-formule, jestliže $K_{KB}$ je jedinou modalitou v $B$.

(ii) $KB$-dotaz je dotaz, který je $KB$-formulí.

Za těchto předpokladů lze podle Lemmatu 2 pro každou $KB$ - formuli tvaru $K_{KB}\ A$, efektivně sestrojit ekvivalentní formuli, která je booleovskou kombinací formulí tvaru $K_{KB}\ B$, kde $B$ je výroková.

Z Věty 1 potom plyne, jakým způsobem, báze odpovídá na *KB*-dotazy z toho jak odpovídá na výrokové dotazy.

Odůvodnění.

(a) [syntakticky] Abychom mohli rozhodnout, jaká bude odpověď na dotaz $A$, musíme určit zda platí $K_{KB} A$ nebo $K_{KB} \neg A$.

Přitom použijeme faktu, že každá formule $K_{KB} A$ je ekvivalentní s nějakou booleovskou kombinací formulí $K_{KB} B$, kde $B$ je výroková formule.

Podle Věty 1 pak můžeme vypočist odpovědi na *KB*-dotazy.

(b) [sémanticky] Jiný způsob charakterizace, jak báze odpoví na KB-dotazy.

Je - li dána výroková formule $A$, nechť $S^A$ sestává ze všech pravdivostních ohodnocení $\alpha$, které jsou modelem $A$.

Nechť

$$M^A = (S^A, \pi, K_{KB})$$

kde $\pi(\alpha) = \alpha$ a $K_{KB}$ je universální relace na $S^A$

$(t.j. K_{KB} = S^A \times S^A)$.

Můžeme si myslet, že $M^A$ je maximální model $A$, protože obsahuje všechny modely $A$.

Následující tvrzení ukazuje, že

je-li $k$ konjunkce všech formulí, které byly do báze uloženy, potom pro libovolnou formuli $B$, dostáváme

$$K_{KB}B \quad \textit{právě když} \quad M^{k} \models K_{KB}B$$

Intuitivně, bylo-li do báze uloženo $k$ pak vše, co báze ví je $k$. Maximální model pro $k$ to zachycuje.

**Věta 2.**

Předpoklá dejme, že

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

Potom pro všechny KB - formule $B$ platí

$$(\mathcal{I}^{kb}, r, m) \models B \quad \textit{právě když} \quad (M^{k}, r_e(m)) \models B$$

Věta 2 ukazuje, že Znalostní báze může odpovědět na $KB$-dotaz $A$ tím, že vyhodnotí zda

. . .

$$(M^k, r_e(m)) \models K_{KB} A$$

.

Přitom pravdivost $K_{KB} A$ v $(M^k, r_e(m))$ nezávisí na $r_e(m)$. Proto můžeme stále chápat konjunkci $k$ jako representaci toho, co znalostní báze ví.

Dá se ukázat, že Větu 2 lze rozšířit tak, aby mohla řešit libovolné dotazy, nejenom $KB$-dotazy.

Dosavadní diskuse ukazuje, že v našem přístupu je možné modelovat standardní typy znalostních bází.

Jaké nám to dává možnosti ?

• dostáváme explicitní předpoklady pro standardní representaci,

• podle Věty 2 můžeme mluvit o tom, co znalostní báze ví vzhledem ke svým znalostem,

• navíc, jak ukážeme v dalším, dovoluje nám to uvažovat o některých variantách našich předpokladů. Flexibilita modelu uvažovaného ve Větě 2 nám usnadňuje zvažovat východiska, která se nabízejí, když budeme modifikovat naše předpoklady.

## Předběžné znalosti.

Začneme rozborem situací, kdy je dána nějaká předběžná znalost o tom, jaké budou vkládány informace.

To, že uvažujeme *všechny* běhy splňující podmínky KB1 a KB2 zaručuje, že není dána žádná předběžná znalost o informacích vkládaných do báze znalostí.

V praxi se setkáváme s předpoklady, které jsou mlčky vkládány do konvencí podle nichž je informace zadávána.

Příkladem toho je konvence, že je-li pravdivé (elementární) tvrzení $p$, je znalostní bázi zadáváno přednostně. Takovou konvenci můžeme v našem přístupu snadno popsat pomocí omezení množiny běhů.

Vyjádříme to omezující podmínkou, že pro každý bod $(r,m)$ takový, že $r(m) = (\alpha, <A_1 , \dots , A_k >, .)$ pro nějaké $k \geq 1$ , platí $A_1 = p$ , právě když $p$ je pravdivé při ohodnocení $\alpha$ . Připomeňme, že lokální stav znalostní báze obsahuje i pořadí, ve kterém byly informace ukládány.

Z toho bezprostředně vyplývá, že znalostní báze ví zda platí $p$ nebo $\neg p$ jakmile do ní byl vložen první fakt.

Předběžné znalosti o vnějším světě lze vyjádřit vhodnou modifikací běhů v informovaném systému $\mathrm{I}^{kb}$ .

Předpokládejme například, že je známo, že primitivní výrok $p$ musí platit. V takovém případě uvažujeme jen takové běhy, kdy $r_e(0) = \alpha$ pro nějaké ohodnocení $\alpha$, při kterém platí $p$.

Situaci, kdy správce nemá úplnou informaci o vnějším světě (ale stále má k disposici úplnou informaci o bázi), můžeme řešit tím, že správcova stavu místo jednoho pravdivostního ohodnocení vložíme neprázdnou množinu pravdivostních ohodnocení $T$. Ta potom representuje množinu vnějších světů, které správce považuje za možné.

Protože nám stále jde o znalosti, požadujeme, aby $\alpha \in T$, tedy aby skutečný vnější svět byl mezi světy, které správce považuje za možné.

Abychom se vyhnuli redundanci, označujeme správcovy stavy $<T, .>$. Globální stavy jsou potom tvaru

$$(\alpha, <A_1, \ldots, A_k>, <T, .>)$$

Abychom podrželi podmínku, že vše, co je do báze vloženo je pravdivé, požadujeme, aby správce vkládal do báze jen takové výrokové formule $A$, které jsou pravdivé při všech ohodnoceních v množině $T$.

Dá se ukázat, že správce vloží formuli $A$ jen v případě $K_T A$ , kde $K_T A$ označuje tvrzení “správce ví $A$ “.

Při takovém přístupu Věty 1. a 2. zůstávají v platnosti v podstatě se stejným důkazem.

Více informátorů (správců) báze je také realistický předpoklad. V praxi bývá pro znalostní bázi více zdrojů informací.

V této situaci již není hranice mezi správci a znalostní bází tak ostrá a každý agent může být chápán jako správce i jako báze. K této situaci se vrátíme v dalších kapitolách.

Přesvědčení místo pravdy na straně správce lze modelovat tím, že množina světů $T$, které správce považuje za možné, nemusí obsahovat skutečný svět. Přitom stále trváme na tom, že správce vkládá do báze jen takové výrokové formule $A$, které jsou pravdivé při všech ohodnoceních v množině $T$.

V takovém případě říkáme, že $A$ je pouze *přesvědčení* (*belief*) správce, nikoli pravda.

Dá se ukázat, že tuto situaci je možno vyjádřit jinou posibilitní relací než $\sim_{KB}$.

Formule, které nejsou výrokové, lze také vkládat do znalostních bází, i když je to trochu složitější.

Například předpokládejme, že do báze je vloženo tvrzení .

$$p \text{ -> } K_{KB}\, p$$

. které říká, že je-li parvdivé $p$, potom báze o tom ví. Tako-vá informace může být velmi užitečná, jestliže báze umí ověřit co ví a co neví. V takovém případě báze muže ověřit, zda ví $p$. Jestliže ne, pak může odvodit, že $p$ není prav-divé.

Tento příklad ukazuje, že když umožníme dávat bázi informace o vnejším světě, pak báze pomocí introspekce odvodit důsledky o vnějším světě.

Jsou-li bázi dávána tvrzení, která nejsou výroková, pak už nemůžeme reprezentovat její znalosti jako dosud konjunkcí tvrzení, která do ní byla vložena. V takovém případě můžeme do báze vložit fakt, který byl pravdivý v okamžiku, kdy byl vložen, ale nezůstává platný v žádném dalším časovém bodu.

Například, předpokládejme, že primitivní výrok $p$ je ve vnějším světě pravdivý, ale báze nedostala žádnou počáteční informaci. V takové situaci je formule

$$p \,\&\, \neg K_{KB}\, p \qquad (1)$$

jistě pravdivá, ale v okamžiku, kdy do báze vložíme tento fakt, jistě neplatí, že báze ví (1).

Jak jsme již zjistili dříve, formule

$$K_{KB}(p \mathbin{\&} \neg K_{KB}\, p) \qquad (2)$$

není bezesporná s S5.

Nicméně, znalostní báze přece jen něco získá, když je do ní vložen fakt (1) : potom ví, že $p$ je pravdivé. Jako důsledek (2), $K_{KB}\, p$ má platit po té, co do báze vložíme (1).

Fakta obsahující tvrzení o znalostech báze. V našem přístupu můžeme popsat i systém, který vznikne když připustíme, aby do báze byla vložena fakta obsahující tvrzení o jejích znalostech.

Jako dosud , předpokládáme, že lokální stav báze sestává z posloupnosti formulí, připouštíme navíc i modální formule, které mohou mluvit o znalostech, které báze má.

V takovém případě nastanou obtíže, chceme-li připustit jen běhy, ve kterých jsou do báze vkládány jen pravdivé formule. Protože pracujeme s formulemi, které obsahují znalosti, není jasné, zda můžeme rozhodnout je-li daná formule pravdivá, aniž bychom měli k dispozici celý systém.

Naším problémem však je nejdřív takový systém sestrojit. K tomu o všem potřebujeme více prostředků, než máme teď k disposici. Vrátíme se k tomuto problému v dalších kapitolách.