Predikátová logika, situační kalkulus

Jiří Kléma

Katedra počítačů,
FEL, ČVUT v Praze

http://cw.felk.cvut.cz/doku.php/courses/a4b33zui/start

Logika jako nástroj pro reprezentaci znalostí

- Jaké vhodné vlastnosti demonstrovala výroková logika?
  - deklarativnost
    - znalosti a usuzování jsou přirozeně odděleny,
    - procedurální jazyky – postrádají obecný mechanismus pro usuzování,
  - možnost vyjádřit neúplnou informaci
    - pomocí disjunkce (a negace),
    - "Zítra odpoledne půjdu do kina nebo budu sportovat."
    - není samozřejmostí (databáze, datové struktury).
  - je kompozičním a kontextově nezávislým jazykem
    - význam věty je funkcí významu jejich částí,
    - význam $A \wedge B$ má vztah k pravdivosti A a B,
    - na rozdíl od přirozeného jazyka: "Pak to viděla."

Predikátová logika (PL)

- Výroková logika má omezenou expresivitu
  - problémy: bez mechanismu zobecnění nebo škálování (viz wumpusův svět)
    - Všechna pole sousedící s wumpusem zapáchají.
    - $W_{1,2} \Rightarrow (S_{1,1} \wedge S_{2,2} \wedge S_{1,3})$, $W_{2,1} \Rightarrow (S_{1,1} \wedge S_{2,2} \wedge S_{3,1})$ atd.
  - dále vztahy mezi objekty, čas, sylogismy,
- PL zavádí objekty a relace mezi nimi
  - inspiruje se v síle přirozeného jazyka, zachovává ale jednoznačnost.
- **PL 1.řádu (first-order logic, FOL)**
  - záměrné omezení expresivity:
    - relace samotné nelze považovat za objekty,
    - nelze zapisovat tvrzení o všech relacích,
- logiky vyšších řádů
  - větší expresivita, obtížnější teorie i efektivní použití.

Predikátová logika 1.řádu – syntaxe, elementy jazyka

- **konstanty** odkazují na konkrétní objekty
  - petr, pavel, skotsko, základy-umělé-inteligence (Prolog notace),
- **predikáty** vyjadřují vztahy mezi objekty či jejich vlastnosti
  - muž(petr), otec(petr,pavel), žije(petr,skotsko) (arita = počet prvků relace),
- **funkce** jako nepřímé odkazy na objekt
  - otec(petr), žije(petr) (jak je vidět, funkce lze nahradit predikáty),
- **proměnné** odkazují na množiny objektů,
  - X, Y, Z (Prolog notace: začínají velkým písmenem),
- **kvantifikátory** určující interpretaci proměnných
  - $\forall$ (pro všechny) – při jakékoli substituci objektu za proměnnou věta platí,
  - $\exists$ (existuje) – lze nalézt objekt zajišťující platnost věty,
  - $\forall$ X $\exists$ Y učí(X,Y) – každý předmět má alespoň jednoho učitele,
- **logické spojky** a závorky – stejné jako ve výrokové logice.

Predikátová logika 1.řádu – syntaxe

- Syntaxe – definice korektních výrokových formulí
  - **term** je každá konstanta, proměnná nebo funkce na ně aplikovaná
    - odkazuje na objekt, nic jiného termem není (tj. predikáty nejsou termy).
  - **dobře utvořenou formulí** (well-formed formula (WFF)) jsou
    - **atomické formule** $p(t_1, \ldots, t_n)$,
    - kde p je predikátový symbol a t_i termy,
    - formule $\neg p$ a $p \Rightarrow q$, pokud p a q jsou WFF,
    - formule $\forall X p$, pokud X je proměnná a p WFF.
- Syntaxe poznámky
  - WFF může obsahovat volné proměnné (nevázané kvantifikátorem)
    - to působí problémy při interpretaci WFF,
    - **věta** je WFF pouze s vázanými proměnnými,
  - z hlediska definice jsou spojky $\vee, \wedge, \Leftrightarrow$ a kvantifikátor $\exists$ odvozené.

Predikátová logika 1.řádu – sémantika, interpretace

- Jak určovat platnost vět v PL 1.řádu?
  - je dána kontextem, ale ...
- ověřování pravdivosti pro všechny modely?
  - pracujeme s velkými nebo neomezenými doménami (např. reálná čísla),
  - v důsledku může existovat i neomezený počet modelů.
- **Interpretace** definuje:
  - objekty, predikáty a fce odpovídající příslušným symbolům a jejich význam
    - **doména** $\Delta = \{$Jiří Kléma, Filip Železný, STM Y33ZUI, OI A4B33ZUI$\}$,
    - interpretace konstant: $l_C : C \rightarrow \Delta$,
      jirka odkazuje na Jiřího Klému, zui1 na předmět STM Y33ZUI, atd.
    - interpretace predikátů s aritou n: $l_P : P \rightarrow P(\Delta^n)$,
      učí odkazuje na vztah mezi učitelem a předmětem,
      učí/2 $= \{\{$jirka,zui$\}$,$\{$filip,kui$\}\}$.

Predikátová logika 1.řádu – sémantika, interpretace

- Je formule $\forall X(p(X) \vee q(X)) \Rightarrow (\forall X p(X) \vee \forall X q(X))$ tautologií?
  - pro vyvrácení stačí nalézt interpretaci v níž formule neplatí,
  - např. v interpretaci Δ={a,b}, $p_D = \{a\}$, $q_D = \{b\}$,
  - levá strana implikace platí:
    - X=a: $p(a) \vee q(a) = T \vee F = T$, X=b: $p(b) \vee q(b) = F \vee T = T$,
  - pravá strana ne:
    - X=b: $p(b) = F$, $\forall X p(X) = F$ podobně pro X=a a $\forall X q(X) = F$.

Inference v PL 1.řádu – unifikace

- Jak **substituovat** za volné proměnné resp. termy při srovnávání výrazů?
  - KB1: Jirka zná Filipa. Václava zná každý. Každý zná svoji matku. Monika někoho zná.
  - dotaz: Koho zná Jirka?
  - vrať všechny korektní substituce za volnou proměnnou vzhledem ke KB,
  - KB2: $zna(jirka, filip)$, $\forall Y zna(Y, vaclav)$, $\forall Z zna(Z, matka(Z))$, $\exists W zna(monika, W)$,
  - KB3: $zna(jirka, filip)$, $zna(Y, vaclav)$, $zna(Z, matka(Z))$, $zna(monika, s1)$,
  - dotaz: ?zna(jirka,X)
  - substituce: $\theta_1 = \{X|filip\}$, $\theta_2 = \{jirka|Y, X|vaclav\}$, $\theta_3 = \{jirka|Z, X|matka(jirka)\}$,
  - hledáme co **nejobecnější** substituci, aby výrazy byly unifikovatelné.

Inference v PL 1.řádu – unifikace

- Substituce a unifikace nutná pro jakoukoli inferenci
  - viz generalizovaný modus ponens

$$\frac{A, A \Rightarrow B}{B} \rightarrow \frac{p'_1, \ldots, p'_n, (p_1 \wedge p_2 \wedge \cdots \wedge p_n \Rightarrow q)}{Subst(\theta, q)}$$

$$Subst(\theta, p'_i) = Subst(\theta, p_i)$$

  - Když se dva lidé znají, tak se zdraví.
    * $\forall XY(zna(X,Y) \Rightarrow zdravi_se(X,Y))$.,
  - aplikace MP na KB (KB $\vdash \alpha$)
    * $zdravi_se(jirka, filip)$, $zdravi_se(Y, vaclav)$,
    * $zdravi_se(U, matka(U))$, $zdravi_se(monika, s1)$.

Rezoluce v PL 1.řádu – převod do CNF (1)

- Obtížnější převod do CNF než u výrokové logiky – proměnné a kvantifikátory,
- Př: Všichni Římané, kteří znají Markuse, buď nenávidí Caesara nebo si myslí, že každý kdo někoho nenávidí je blázen.

$$\forall X \big[(riman(X) \wedge zna(X, markus)) \Rightarrow \\ (nenavidi(X, caesar) \vee \forall Y \exists Z (nenavidi(Y, Z) \Rightarrow ma_za_blazna(X, Y))) \big]$$

- Obecné kroky převodu:

1. odstranění implikací

$$\forall X \big[\neg (riman(X) \wedge zna(X, markus)) \vee (nenavidi(X, caesar) \vee \\ \vee \forall Y \neg (\exists Z nenavidi(Y, Z)) \vee ma_za_blazna(X, Y))) \big]$$

2. přesun negací k atomickým formulím

$$\forall X \big[\neg riman(X) \vee \neg zna(X, markus) \vee nenavidi(X, caesar) \vee \\ \vee \forall Y \forall Z (\neg nenavidi(Y, Z) \vee ma_za_blazna(X, Y)) \big]$$

Rezoluce v PL 1.řádu – převod do CNF (2)

4. skolemizace – eliminace existenčních kvantifikátorů
   - zde žádný výskyt, ale: $\forall X \exists Y otec(Y, X) \rightarrow \forall X otec(s1(X), X)$
   - s1 je v daném případě Skolemovou funkcí (každému X přiřazuje otce),
5. oddělení jmen proměnných

$$\forall X riman(X) \vee \forall X rek(X) \rightarrow \forall X riman(X) \vee \forall Y rek(Y)$$

6. přesun univerzálních kvantifikátorů zcela vlevo – **prenexní** tvar,

$$\forall XYZ\big[\neg riman(X) \vee \neg zna(X, markus) \vee \\ nenavidi(X, caesar) \vee \neg nenavidi(Y, Z) \vee ma_za_blazna(X, Y)\big]$$

7. přesun disjunkcí k literálům
   - aplikace distributivních zákonů $a \vee (b \wedge c) \rightarrow (a \vee b) \wedge (a \vee c)$
8. eliminace univerzálních kvantifikátorů – bezkvantifikátorové jádro.

$$\neg riman(X) \vee \neg zna(X, markus) \vee \\ nenavidi(X, caesar) \vee \neg nenavidi(Y, Z) \vee ma_za_blazna(X, Y)$$

Klade Zuzana vajíčka? – motivační příklad 3

:: Pokud vám řeknu, že:

- (S1) Ptakopysk a ježura jsou jediní savci kladoucí vejce.
- (S2) Pouze ptáci a savci jsou teplokrevní.
- (S3) Zuzana, můj pásovec, je teplokrevná a nemá peří.
- (S4) Každý pták má peří.

:: a zeptám se: (D) Klade Zuzana vejce?

Klade Zuzana vajíčka? – motivační příklad 3

:: Pokud vám řeknu, že:

- (S1) Ptakopysk a ježura jsou jediní savci kladoucí vejce.
- (S2) Pouze ptáci a savci jsou teplokrevní.
- (S3) Zuzana, můj pásovec, je teplokrevná a nemá peří.
- (S4) Každý pták má peří.

:: a zeptám se: (D) Klade Zuzana vejce?

:: Inference v přirozeném jazyce:

- Zuzana nemá peří a proto není pták.
- Zuzana je teplokrevná a není pták, proto musí být savec.
- Zuzana je savec a pásovec, není ježura ani ptakopysk a nemůže klást vejce.

:: Jak realizovat strojově? Řetězce jsou problematické pro odvozování ...

Klade Zuzana vajíčka – důkaz rezolucí (1)

- (S1) $\forall X(savec(X) \wedge vejce(X) \Rightarrow jezura(X) \vee ptakopysk(X))$
- (S2) $\forall X(teplokrevny(X) \Rightarrow ptak(X) \vee savec(X))$
- (S3) $pasovec(zuzana) \wedge teplokrevny(zuzana) \wedge \neg peri(zuzana)$
- (S4) $\forall X(ptak(X) \Rightarrow peri(X))$
- (F) $\forall X(pasovec(X) \Rightarrow \neg(ptakopysk(X) \vee jezura(X)))$
- (D) $vejce(zuzana)$

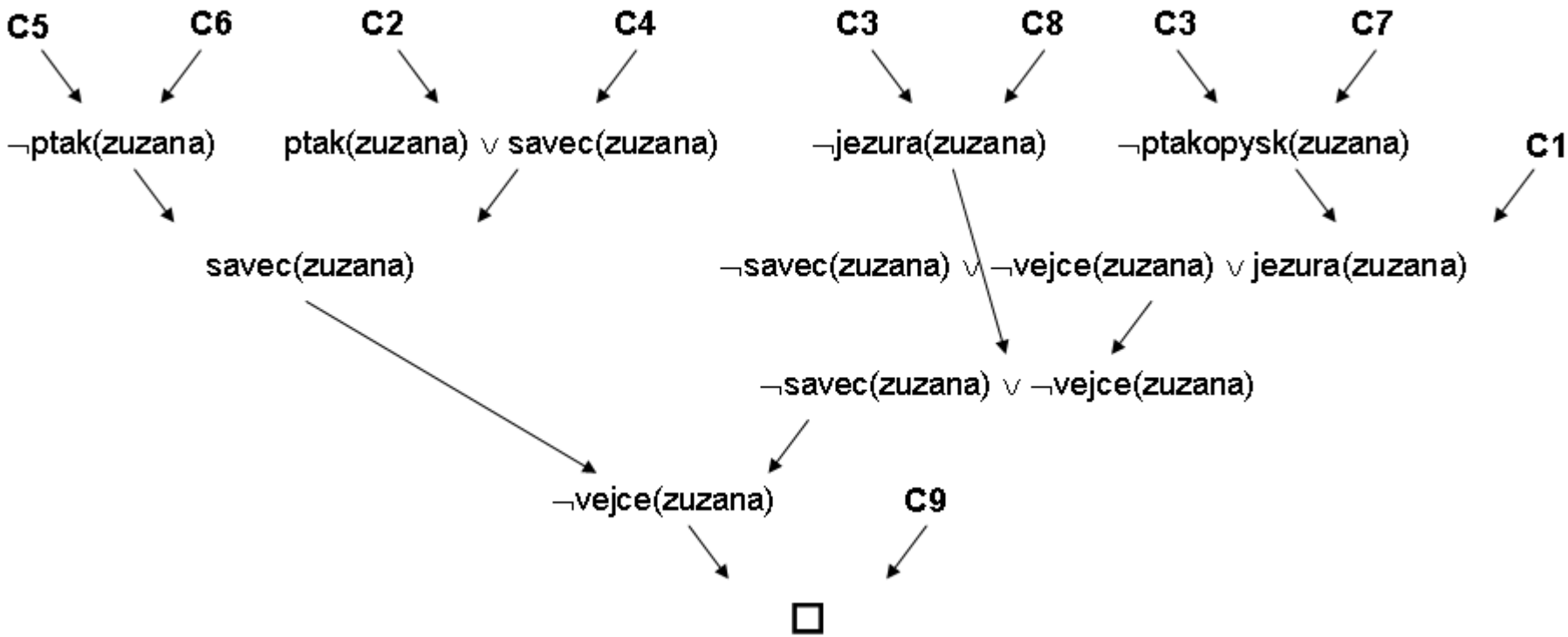
Klade Zuzana vajíčka – důkaz rezolucí (1)

- (S1) $\forall X(savec(X) \wedge vejce(X) \Rightarrow jezura(X) \vee ptakopysk(X))$
- (S2) $\forall X(teplokrevny(X) \Rightarrow ptak(X) \vee savec(X))$
- (S3) $pasovec(zuzana) \wedge teplokrevny(zuzana) \wedge \neg peri(zuzana)$
- (S4) $\forall X(ptak(X) \Rightarrow peri(X))$
- (F) $\forall X(pasovec(X) \Rightarrow \neg(ptakopysk(X) \vee jezura(X)))$
- (D) $vejce(zuzana)$

- (C1) $\neg savec(X) \vee \neg vejce(X) \vee jezura(X) \vee ptakopysk(X)$
- (C2) $\neg teplokrevny(X) \vee ptak(X) \vee savec(X)$
- (C3,4,5) $pasovec(zuzana) \wedge teplokrevny(zuzana) \wedge \neg peri(zuzana)$
- (C6) $\neg ptak(X) \vee peri(X)$
- (C7,8) $(\neg pasovec(X) \vee \neg ptakopysk(X)) \wedge (\neg pasovec(X) \vee \neg jezura(X))$
- (C9) $vejce(zuzana)$

Klade Zuzana vajíčka – důkaz rezolucí (2)

:: Strom rezolučního zamítnutí:



:: Tvrzení, že Zuzana klade vejce je ve sporu s teorií. **Zuzana vejce neklade.**

FOL ve wumpusově světě

- FOL je dostatečně expresivní pro popis wumpusova světa a inferenci v něm
- reprezentace prostředí
  - sousednost jeskyní

    $$\forall XYAB\big[sousedi([X,Y],[A,B]) \Leftrightarrow \\ [A,B] \in \{[X+1,Y],[X-1,Y],[X,Y+1],[X,Y-1]\}\big]$$
  - diagnostická pravidla
    * od pozorovatelných jevů k jejich příčinám – pro výskyt vánku:

      $\forall S\,(vanek(S) \Rightarrow \exists R\,(sousedi(R,S) \wedge sachta(R)))$,

      $\forall S\,(\neg vanek(S) \Rightarrow \neg\exists R\,(sousedi(R,S) \wedge sachta(R)))$,
  - kauzální pravidla
    * vlastnosti prostředí generují vnější projevy – pro jeskyně:

      $\forall R\,(sachta(R) \Rightarrow \forall S\,(sousedi(R,S) \Rightarrow vanek(S)))$,

      $\forall S\,[\forall R\; sousedi(R,S) \Rightarrow \neg sachta(R)] \Rightarrow \neg vanek(S)$,
  - výše uvedená diagnostická a kauzální pravidla jsou ekvivalentní zápisu:

    $\forall S\,(vanek(S) \Leftrightarrow \exists R\,(sousedi(R,S) \wedge sachta(R)))\,.$

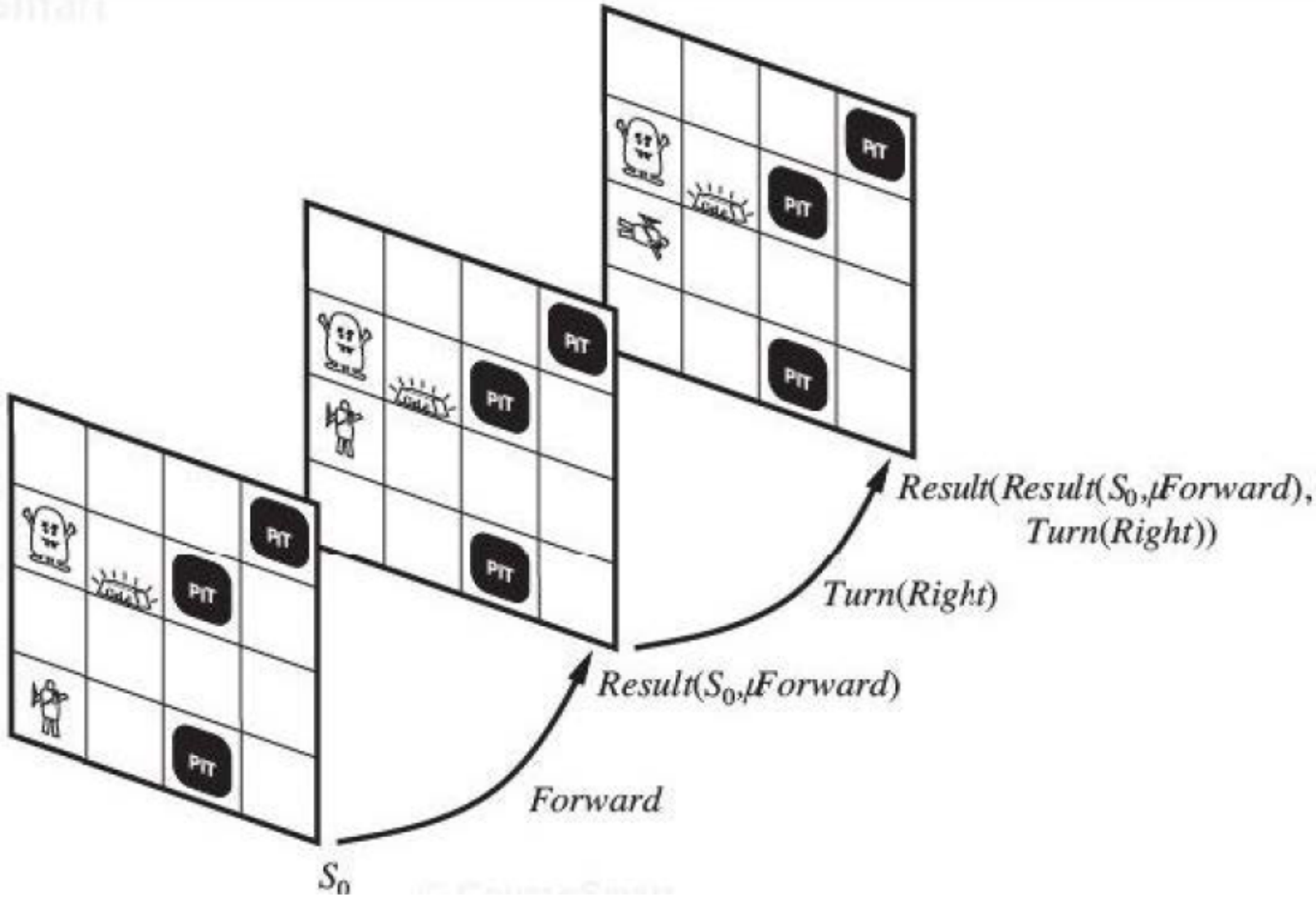
FOL ve wumpusově světě

- obtížnější je zachycení vnímání, reflexe, modifikace prostředí
  - vnímání (T vyjadřuje čas, resp. číslo kroku)
    $\forall TSB\,(senzory([S, B, trpyt], T) \Rightarrow trpyt(T))$
  - reflexe
    $\forall T\,(trpyt(T)) \Rightarrow nejlepsi_akce(uchop, T))$

Sledování změn

- fakta platí v konkrétních **situacích**, spíše než věčně,
- **situační kalkul** je jedním ze způsobů reprezentace změn ve FOL
  - predikáty dělí na neměnné (rigid, eternal) a flexibilní (fluents)
    - bez nebo s možnou změnou platnosti,
  - přidává situační argument ke každému flexibilnímu predikátu
    - např. $drzi(zlato, nyni)$, term $nyni$ označuje situaci,
  - neměnný např. $sachta(R)$, $sousedi(R, S)$,
  - situace jsou vázány $result$ funkcí,
    - s je situace, $result(a, s)$ je také situace,
    - do $result(a, s)$ dospějeme provedením akce a v situaci s.

Sledování změn



Russell, Norvig: AIMA, 3rd edition, 2010.

Popis a použití akcí

- "důsledkový" axiom – popisuje změny, které jsou výsledkem akce
  - $\forall s\ seZlatem(s) \rightarrow drzi(zlato, result(uchop, s))$
- axiom "rámce" – popisuje, co akce **nezměnila**
  - $\forall s\ maSip(s) \rightarrow maSip(result(uchop, s))$
- **problém rámce**: jak elegantně zvládnout fakta beze změny

  (a) reprezentační

  vyhni se axiomům rámce,

  F flexibilních predikátů, A akcí $\rightarrow \mathcal{O}(FA)$ axiomů rámce,

  (b) inferenční

  vyhni se opakovanému přepisování k udržení informace o stavu,

Popis a použití akcí

- **kvalifikační problém**
  - věrný popis reálných akcí vyžaduje nekonečnou péči,
  - co když zlato bude kluzké nebo přibité k zemi nebo . . . ?
- **problém důsledku**
  - skutečné akce mají mnoho skrytých důsledků,
  - s agentem se přesouvá vše co drží,
  - se zlatem se přesouvá i prach,
  - rukavice se opotřebovávají.

Popis a použití akcí

- axiomy "následných stavů" řeší reprezentační problém rámce,
- každý axiom se dotýká predikátu (a nikoli akce samotné)

  P platí po provedení akce $\Leftrightarrow$ [akce vedla ke splnění P
  $\vee$ P již platilo a akce P nezneplatnila]

- pro manipulaci se zlatem

$$\forall a, s\ drzi(zlato, Result(a, s)) \Leftrightarrow$$
$$[(a = uchop \wedge seZlatem(s)) \vee (drzi(zlato, s) \wedge a \neq uvolni)]$$

- dostaneme F axiomů
  - s celkovým počtem literálů $\mathcal{O}(AE)$ (E je počet efektů na akci).
- alternativní zápis pro $drzi$ přes axiomy rámce a důsledkové axiomy?

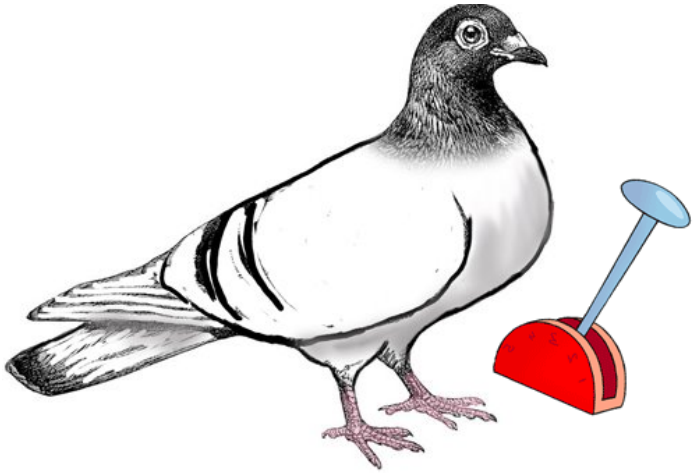
Problém rámce a lidský mozek

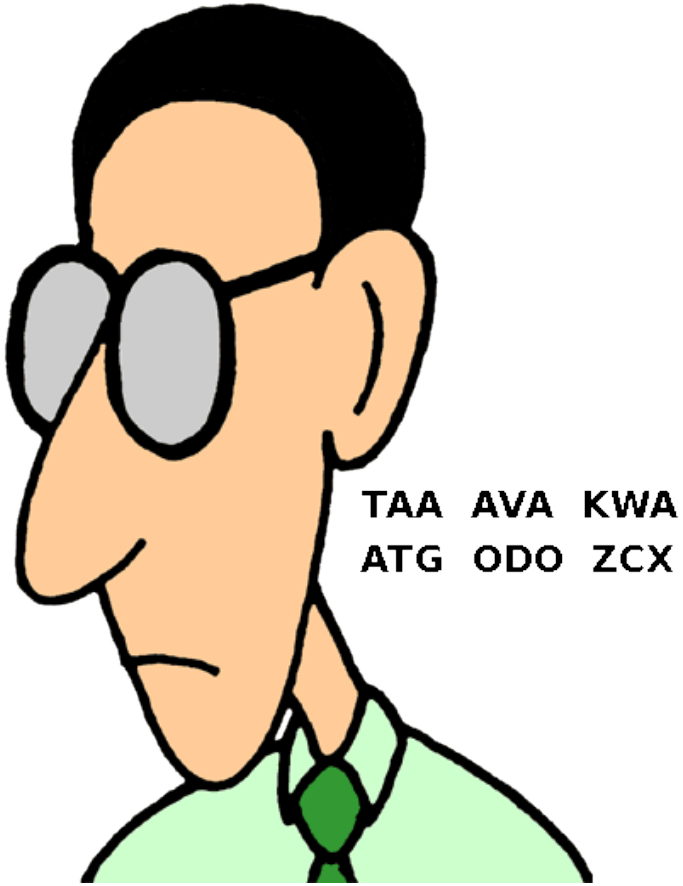
- Sherlock Holmes a pes, který neštěkal

```
Is there any point to which you would wish to draw my attention?
To the curious incident of the dog in the night-time.
The dog did nothing in the night-time.
That was the curious incident, remarked Sherlock Holmes.
```

- Co tedy umí lidský mozek a paměť?
  - ani lidská paměť neukládá vše,
  - ale někdy se tak tváří,
  - zjednodušení si lidé neuvědomují,
  - soustředí se na akci, rámec je rekonstruován,
  - proto mají i lidé problém povšimnout si toho, *co se nestalo*.

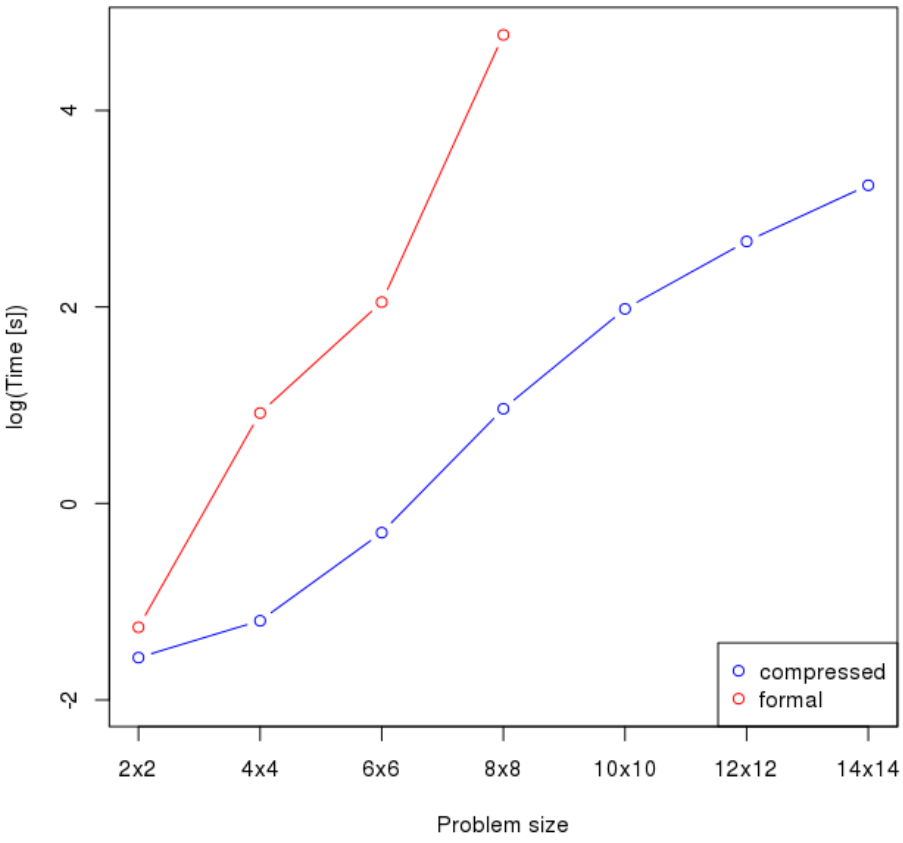
Problém rámce a lidský mozek





Wumpus: srovnání formálního a efektivního řešení

- formální řešení
  - modulární oddělení znalostí,
  - rozšiřitelné a udržovatelné,
- efektivní řešení
  - minimalizuje počet predikátů,
  - v krajním případě 1,
  - implicitní redukce prohledávacího prostoru,
- demo.



Vytváření plánů

- počáteční podmínky v bázi znalostí
  - $poloha(agent, [1,1], s_0)$, $poloha(zlato, [1,2], s_0)$,
- dotaz:
  - $Zjisti(KB, \exists s\ drzi(zlato, s))$,
  - tj., v jaké situaci budu držet zlato?
- odpověď:
  - $\{s/result(uchop, result(jdi_vpred, s_0))\}$,
  - tj., jdi vpřed a poté uchop zlato,
- předpoklady:
  - agent se zajímá pouze o plány započínající v s_0,
  - s_0 je jedinou situací popsanou v KB.

Vytváření plánů: lepší způsob

- vyjádři plány jako posloupnosti akcí $[a_1, a_2, \ldots, a_n]$,
- $planResult(p, s)$ je výsledkem provedení plánu p ve stavu s
  - projekční úloha – výsledná situace po aplikování dané posloupnosti akcí?
    $pozice(agent, [1,1], s_0) \wedge pozice(zlato, [1,2], s_0) \wedge \neg drzi(O, s_0)$
    $pozice(zlato, [1,1], result([jdi([1,1],[1,2]), uchop(zlato), jdi([1,2],[1,1])], s_0))$
  - plánovací úloha – jaká posloupnost akcí vede k dané situaci?
- dotaz $Zjisti(KB, \exists p\ drzi(zlato, planResult(p, s_0)))$
  - má řešení $\{p/[jdi_vpred, uchop]\}$,
- definice $planResult$ v rámci $result$
  - $\forall s\ planResult([\,], s) = s$,
  - $\forall a, p, s\ planResult([a|p], s) = planResult(p, result(a, s))$,
- motivace pro **plánování** jako disciplínu
  - specializovaný plánovač uspěšnější než obecné usuzování/prohledávání.

Situační kalkul: příklad

- Mějme logickou KB definující vzdálenost mezi českými městy:
  - $vzdal(praha, brno, 202)$, $vzdal(brno, zlin, 100)$,
- stav agenta nechť je dán jeho polohou a benzínem v nádrži
  - $pozice(Mesto, Stav)$, $nadrz(Benzin, Stav)$,
  - spotřeba nechť je 6l/100km,
- zapište axiom definující predikát delkaCesty(P,D)
  - například $delkaCesty([praha, brno, zlin], 302)$,
- zapište situační axiom definující následky akce $projdiCestu(P)$
  - postihněte změnu pozice agenta a objemu benzínu v nádrži,
- jsou potřeba axiomy rámce?

Situační kalkul: příklad

- rekurzivní definice delkaCesty(P,D):

$$\forall X\ delkaCesty([X], 0)$$

$$\forall X, Y, L, D_1, D_2\ vzdal(X, Y, D_1) \wedge delkaCesty([Y|L], D_2) \Rightarrow delkaCesty([X, Y|L], D_1 + D_2)$$

- pro následky akce $projdiCestu(P)$ je třeba sledovat splnění předpokladů

$$\forall X, Y, B, D, P, S\ pozice(X, S) \wedge delkaCesty(P, D) \wedge nadrz(B,S) \wedge (B > D * 6/100) \wedge first(P, X) \wedge last(P, Y) \Rightarrow pozice(Y, Result(projdiCestu(P), S)) \wedge nadrz(B - D * 6/100, Result(projdiCestu(P), S))$$

- akce $projdiCestu(P)$ mění oba flexibilní predikáty
  - axiomy rámce nejsou třeba, nutné například pro hypotetický predikát: $pojizdnostVozu(Pojizdnost, Stav)$.

Shrnutí

- Je obtížnější tvořit a spravovat znalosti o proměnlivém světě
  - obecné problémy – důsledku a zejména rámce,
  - i lidská paměť rámec rekonstruuje pouze přibližně,
  - reprezentace změn ve FOL $\rightarrow$ situační kalkul,
- další způsoby reprezentace znalostí
  - produkční systém,
- praktické využití
  - třeba sémantický web,
- kde se dozvíte více?
  - A4M33RZN – Pokročilé metody reprezentace znalostí,
  - A4M33AU – Automatické usuzování.

Doporučené doplňky – zdroje přednášky

:: Četba

- Russel, Norvig: **AI: A Modern Approach**, Logical Agents, chapter 7
  - reprezentace z pohledu inteligentního agenta,
  - dostupná v pdf – `http://aima.cs.berkeley.edu/newchap07.pdf`.
- Mařík a kol. **Umělá inteligence 1**
  - kapitola Reprezentace znalostí
    - základní formáty, logika, sémantické sítě, rámce,
  - kapitola Řešení úloh a dokazování vět
    - predikátová logika a důkazní prostředky,
- Mařík a kol. **Umělá inteligence 2**
  - kapitola Znalostní inženýrství
    - praktická, znalostní systémy v konkrétních aplikacích.