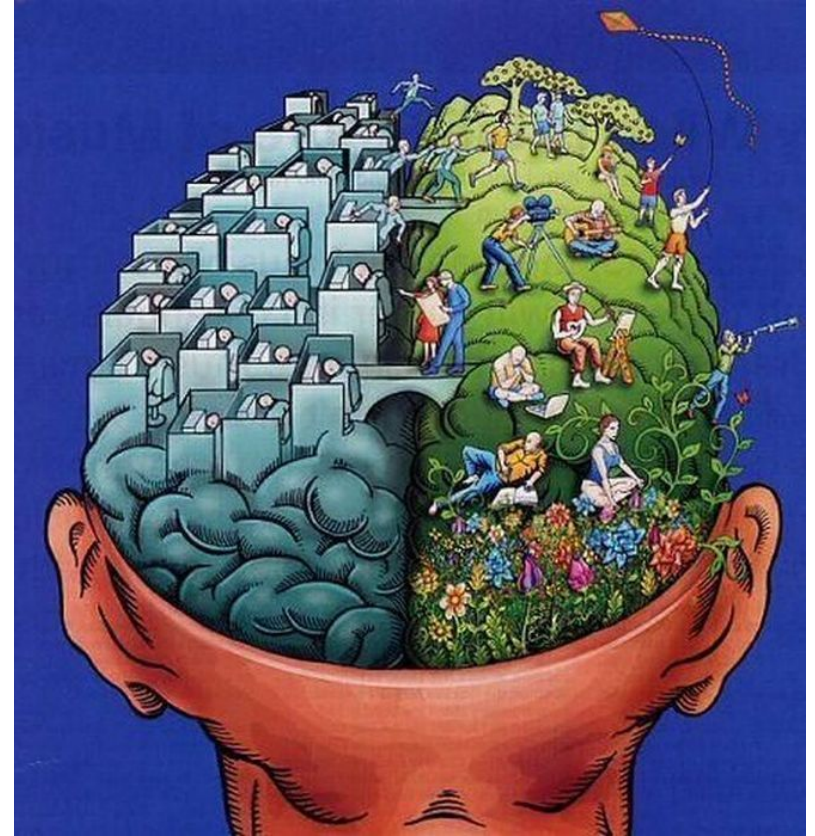


Mesterséges Intelligencia MI

Logikai ágens
ügyesebben

Dobrowiecki Tadeusz
Eredics Péter, és mások



BME I.E. 437, 463-28-99

dobrowiecki@mit.bme.hu,

<http://www.mit.bme.hu/general/staff/tade>

Mit tudunk már?

Logika = szintaktika + szemantika.

A logika használatának alapja a vonzat reláció és a gépi elemzése.

Több matematikai logika van.

Több típusú logikai állítás van.

Bizonyítani is többféle módon lehet (n. fontos a kielégíthetőségi vizsgálat).

Ha a feladat bonyolult, a modellvizsgálatnál jobb a formális bizonyítás.

Logikát jellemezni lehet olyan fogalmakkal, mint monoton, teljes, eldönthető.

Formális bizonyítást viszont olyanokkal, mint teljes, vagy helyes.

Ítélet kalkulus teljes, eldönthető, egzakt bizonyítások exponenciálisak, bonyolult problémákban, nem biztos, hogy elegendő a kifejező ereje.

Természetes dedukció nem, de a rezolúció alapú bizonyítás algoritmizálható.

Kis kapu: logika megszorítása (a kifejező erejében), ezen az áron a bizonyítások felgyorsítása (akár lineárisá).

Formális bizonyítások deduktív sémák, de vannak fontos nem deduktív következtetések is.

Ítéletlogika – feladatbonyolultság határai

Sok fontos gyakorlati alkalmazás (kielégíthetőségi vizsgálattal):

- formális modell-ellenőrzés
- automatikus tesztminta-generálás
- 0-ad (kombinatorikus) ekvivalencia ellenőrzése
- MI tervkészítés (később)
- automatikus tételbizonyítás
- szoftver-verifikálás
- RNA, DNA alakzat kutatás

Pl. Ipari processzor-verifikáció:

14 ciklusra terjedő viselkedés

kb. 1 millió változó, 30 millió literál,

4 millió klóz, 1.5 millió döntés

3h futási idő

(Ez Wumpus N = csak kb. 100-200-nak felel meg!

Wumpus-világ mégsem ilyen egyszerű játék, amilyennek látszik)

Ítéletlogika – feladatbonyolultság határai

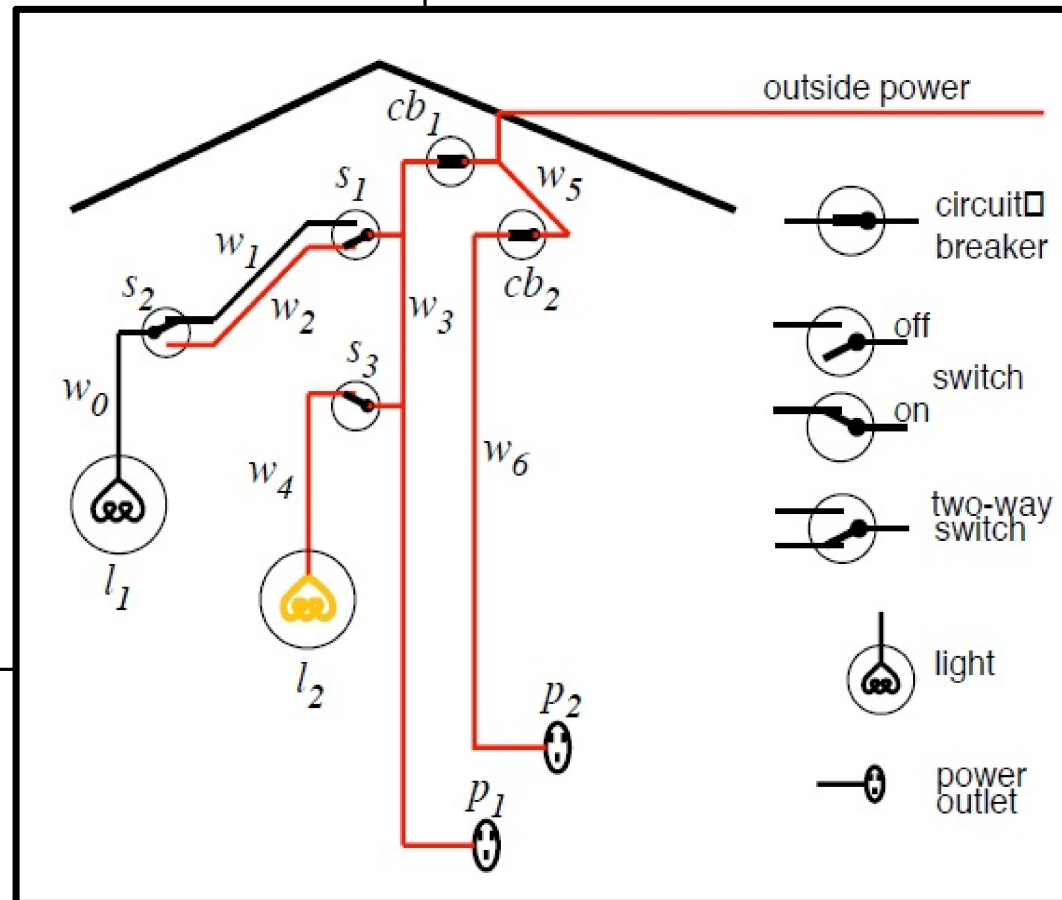
TB = ellentmondás - ... sikeres bizonyítás

TB = konzisztens - ... sikeres modellkeresés

Prover9

Mace4

```
((aw1 & s2) | (aw2 & -s2)) <-> aw0.  
op <-> aw5.  
(aw5 & cb2) <-> aw6.  
aw6 <-> p2.  
(op & cb1) <-> aw3.  
aw3 <-> p1.  
(aw3 & s3) <-> aw4.  
aw4 <-> l2.  
(aw3 & s1) <-> aw1.  
(aw3 & -s1) <-> aw2.  
aw0 <-> l1.  
+ evidencia
```



Elsőrendű logika: milyen a világ valójában?

- **objektumok**: más objektumoktól megkülönböztetett identitással és **tulajdonságokkal** rendelkező dolgok,
- objektumok között különböző **relációk**.
A relációk közül néhány tip. **függvény** – olyan reláció, ahol csak egy „értéke” van egy adott „bemenet” esetén.

Pl.:

objektumok: emberek, elméletek, Róbi Bácsi, színek, sakkjátszma, ...

relációk: testvére, nagyobb mint, belseje, része, színe, birtokol ...

tulajdonságok: piros, kerek, színlelt, páratlan, sokemeletes....

függvények: apja, legjobb barátja, eggyel több mint

Ezeket mind külön-külön tudjuk majd ábrázolni és velük következtetni!

(erősen strukturált tudás erősen strukturált ábrázolása jobban megy)

Út az elsőrendű logika - predikátum kalkulus - felé

450 i.e. Sztoikusok: „ítéletlogika”, következtetés (talán)
322 i. e. Arisztotélesz: szillogizmusok (következtetési minták), kvantorok
1565 Cardano: val. elmélet (ítéletlogika + bizonytalanság)
1847 Boole: ítéletlogika

1879 Frege:

elsőrendű logika (Pierce, Peano)

1922 Wittgenstein:

bizonyítás igazságtáblákkal

1930 Gödel:

FOL - létezik teljes algoritmus

1930 Herbrand:

FOL – konkrét teljes algoritmus (redukció)

1931 Gödel:

aritmetika számára nincs teljes algoritmus

1960 Davis/Putnam

„praktikus” algoritmus ítéletlogikához

1965 Robinson

„praktikus” algoritmus FOL-hoz (rezolúció)

1970 -

ezen az alapon gépi bizonyító rendszerek

Szintaktika és szemantika

konstans szimbólumok: A, B, C, János, Bodri, K21

Interpretáció = az egyes konstans szimbólumok a világ mely objektumaira vonatkoznak.

predikátum szimbólumok: kerek, bátyja,...

Interpretáció = egy predikátum szimbólum egy bizonyos relációra vonatkozik a modellben.

függvény szimbólumok: koszinusz, apja, bal-lába

Néhány reláció függvény – ahol minden a relációban szereplő objektum pontosan egy másik objektummal van kapcsolatban.

János = apja(Béla) apja(János, Béla) = Igaz

Konstans-, predikátum- és függvénytípusú szimbólumok megválasztása teljes mértékben a felhasználón múlik.

Ez jó is, de nem is. De vajon miért?

Objektum: Bodri kutya, amely
Tulajdonsága: nagy

Konstans: Bodri

Tulajdonság:
nagytestű(x)

Egy igaz állítás:
nagytestű(Bodri)



Konstans: Kutya1

Tulajdonság:
dőgnagy(x)

Egy igaz állítás:
dőgnagy(Kutya1)

Igaz logikában, hogy: $\text{nagytestű(Bodri)} \Leftrightarrow \text{dőgnagy(Kutya1)}$
Nem!

Ha nem mondjuk ki, hogy: $\forall x \text{nagytestű}(x) \Leftrightarrow \text{dőgnagy}(x)$.
Bodri = Kutya1.

azonos jelölés, más interpretációt is takarhat
azonos interpretáció, más jelöléshez is vezethet

Miből építkezünk?

termek: János, Bodri, bal-lába(apja(János)), ..., x , $f(x)$
egy **term** egy objektumra vonatkozó kifejezés.
(konstans szimbólum a legegyszerűbb term)

atomi mondatok: bátyja(Béla, apja(János)),
házas(apja(Richárd), anyja(János)), ...

már tényeket fejeznek ki (van már igazságértékük).
atomi mondat: predikátum szimbólum és az
argumentumait jelentő termek listája

Argumentum lehet nagyon összetett term is,
pl.
házas(apja(apja(apja(apja(apja(apja(Richárd)))))))

összetett mondatok:

bátyja(Béla,apja(János)) $\rightarrow$ házas(apja(Richárd),anyja(János))

logikai összekötő szimbólumok $\neg \wedge \vee \rightarrow \Leftrightarrow \oplus \otimes \downarrow \mid \dots$

a még összetettebb mondatok építésére.

kvantorok:

tulajdonságok, amelyek objektumok egész halmazaira vonatkoznak, ahelyett hogy megneveznénk minden objektumot a nevével.

Az elsőrendű logika standard kvantorai:

univerzális kvantor ($\forall$)

$\forall x. \text{macska}(x) \rightarrow \text{emlős}(x)$

egzisztenciális kvantor ($\exists$)

$\exists x. \text{macska}(x) \wedge \text{úszik}(x)$

konstansok, változók, objektum-objektum függvények

$C, x, f(x)$

\ _____ /



termek + predikátumok

kutya(C)

\ _____ /



negálás + **atomi formulák**

$\neg$ kutya(C)

\ _____ /



$\vee \wedge \rightarrow \forall x \exists y$ **literálok** + logikai operátorok + kvantorok

\ _____ /



Jól definiált formulák

$\forall x. \text{kutya}(x)$

Univerzális kv. BME-n mindenki okos: $\forall x \text{ bme-n}(x) \rightarrow \text{okos}(x)$

$\forall x p(x)$ igaz egy m modellben, a.cs.a, ha $p(x)$ igaz minden modellbeli x objektumra. Univerzális kvantor ekvivalens a $p(x)$ példányosítások konjunkciójával:

$(\text{bme-n}(\text{Béla}) \rightarrow \text{okos}(\text{Béla})) \wedge (\text{bme-n}(\text{Ákos}) \rightarrow \text{okos}(\text{Ákos})) \wedge \dots$

Egzisztenciális kv. Van, aki okos BME-n: $\exists x \text{ bme-n}(x) \wedge \text{okos}(x)$

$\exists x p(x)$ igaz egy m modellben, a.cs.a, ha $p(x)$ igaz valamilyen modellbeli x objektumra. Egzisztenciális kvantor ekvivalens a $p(x)$ példányosítások diszjunkciójával:

$(\text{bme-n}(\text{Béla}) \wedge \text{okos}(\text{Béla})) \vee (\text{bme-n}(\text{Ákos}) \wedge \text{okos}(\text{Ákos})) \vee \dots$

Az $\forall$ és az $\exists$ kapcsolata: a két kvantor szorosan kapcsolódik egymáshoz negáláson keresztül:

$$\forall x \neg \text{szeret}(x, \text{Répa}) = \neg \exists x \text{ szeret}(x, \text{Répa})$$

$$\forall x \text{ szeret}(x, \text{Fagylalt}) = \neg \exists x \neg \text{szeret}(x, \text{Fagylalt})$$

Az egyenlőség

term1 = term2 igaz adott interpretáció mellett, a.cs.a,
ha term1 és term2 ugyanarra az objektumra vonatkoznak.

Pl.

$(1 = 2)$ és $(\forall x \text{ szorzás}(n\text{-gyök}(x), n\text{-gyök}(x)) = x)$
kielegíthető állítások

$(2 = 2)$ érvényes állítás

a testvér logikai definíciója:

$$\forall x, y \text{ testvér}(x, y) \leftrightarrow [\neg(x = y) \wedge \\ \exists m, f \neg(m = f) \wedge \\ \text{szülő}(m, x) \wedge \\ \text{szülő}(f, x) \wedge \\ \text{szülő}(m, y) \wedge \\ \text{szülő}(f, y)]$$

Bizonyítások

Mi a helyzet a modell alapú következtetéssel?

Adott TB-hoz felsorolni a modelleket azt jelenti, hogy a modelleket meg kell fogalmazni:

a problématerület objektumainak minden $n = 1 \dots \infty$ száma mellett,
a szókészlet minden k attribútumú $p(x_1, \dots, x_k)$ predikátumra,
a szókészlet minden n elemű k -nárís relációra,
a szókészlet minden C logikai konstansra,
a C logikai konstans minden problématerület objektum hozzárendelésére n db. objektumból ...

T.i. nem tudjuk, mely kombináció lesz jó?

A legjobb (legegyszerűbb) esetben is legalább **megszámlálható**.
A függvények (relációk) ágyazása a termekben a lényegi nehézség.

Redukció ítéletlogikai következtetésre

(egy változó nélküli term - **alapterm**) (grounding)

Változók nélküli predikátum kalkulus = ítélet kalkulus!

kutya(Bodri)

nagytestű(Bodri)

kutya(Bodri) $\wedge$ nagytestű(Bodri) $\rightarrow$ fél(Béla,Bodri)

KutyaBodri

NagytestűBodri

KutyaBodri $\wedge$ NagytestűBodri $\rightarrow$ FélBélaBodri

KB

NB

KB $\wedge$ NB $\rightarrow$ FBB

X1

X2

X1 $\wedge$ X2 $\rightarrow$ X3

Redukció ítéletlogikai következtetésre

Új TB: minden univerzális mondat minden lehetséges példányosítása. Egy alapmondat vonzata az új TB-nak, ha vonzata volt az eredeti TB-nak. Minden FOL TB átalakítható ítéletlogikai alapalakba a vonzatok megtartásával.

Probléma: ha függvények is vannak, akkor ∞ -sok alapterm létezik,
pl. *apja(apja ... (apja(János)))*

Herbrand (1930): Ha egy mondat egy FOL TB vonzata, akkor vonzata a példányosított TB egy **véges** részhalmazának is.

Eljárás: minden $n = 0, 1, 2, \dots$

példányosított TB létesítése n -mélységű termekkel
a kérdéses mondat vonzata-e?

Probléma: működik jól, ha a vonzat létezik, ∞ -hurok, ha nincs.

Turing (1936), Church (1936): FOL-ban a vonzat félig eldönthető.

Probléma: példányosítás sok irreleváns mondatot generál!

p db. k -áris predikátum és n konstans: $p \times n^k$ példányosítás.

Függvényekkel annál rosszabb!

Következtetési lépések kibővítése

Modus Ponens:

$$\frac{A \quad A \rightarrow B}{B} \quad \frac{p(A) \quad \forall x, p(x) \rightarrow q(x)}{q(A)}$$

(x/A illesztés = **unifikálás (egyesítés) + behelyettesítés**

Unifikál(p, q) = θ , ha if $p\theta = q\theta$

<u>p</u>	<u>q</u>	<u>θ</u>
ismer(János, x)	ismer(János, Ági)	{x / Ági}
ismer(János, x)	ismer(y, BME)	{x / BME, y / János}
ismer(János, x)	ismer(y, anya(y))	{y / János, x / anya(János)}
ismer(János, x)	ismer(x, BME)	kudarc

de további lehetőségek is vannak:

Univerzális kvantor eliminálása:

$$\frac{\forall x, p(x, A)}{p(B, A)}$$

Egzisztenciális kvantor eliminálása: $\frac{\exists x q(x, A)}{q(B_s, A)}$
 (un. skolemizálás)

feltéve, hogy B_s -nek másutt szerepe a TB-ban nincs!

B_s az un. **Skolem konstans**, bizonyos tulajdonságokkal igen, de a feladatban önálló (interpretált) léttel **NEM rendelkező objektum**.
 Lehet $f(x)$ **Skolem függvény** is, amely minden x -hez egy külön Skolem konstanst rendel hozzá ...

$$\frac{\forall x \exists y q(x, y)}{\forall x q(x, f_s(x))}$$

$$\frac{\forall ember \exists szív. van-szíve(ember, szív)}{\forall ember van-szíve(ember, f_{szív}(ember))}$$

$$\frac{\forall x p(x, A)}{p(B, A)}$$

ezzel szemben az univerzális kvantor eliminálásánál akármilyen létező objektumot megjelölő konstanst lehetett használni!

Predikátum kalkulus tulajdonságai

Teljes: minden igaz állítás belátható annak

(Gödel, 1930, egzisztenciális bizonyítás)

(Herbrand, 1930, ítéletlogikai redukció)

A vonzat csak félig eldönthető:

hamis állítás hamis volta nem mutatható ki!

(a bizonyítási eljárásnak **állítás = Igaz** esetén van kilépési pontja, **állítás = Hamis** esetén az eljárásnak nincs kilépési pontja, és munka közben, logikai alapon, nem dönthető el, hogy melyik esettel foglalkozunk.)

Gyakorlati következtetés: gépen egy állítás logikai értéke, vagy általánosan egy állításhalmaz konzisztenciája elvben nem mutatható ki a **véges erőforrások (idő)** miatt!

Egy tanulságos történet

“A jog azt mondja, hogyha egy amerikai fegyvert ad el egy ellenséges nemzetnek, akkor bűnös. **Nono** egy ország, az **Amerika** ellensége, rendelkezik rakétákkal, amiket **West** ezredes adott el nekik, aki viszont egy amerikai. Be lehet-e bizonyítani, hogy **West** ezredes bűnös?”

IRAN-CONTRA botrány

NSC (National Security Council) titkos külügyi műveletei Ronald Reagan elnöksége idején, kongresszusi határozatok megkerülésével.

Célok: fegyvert eladni Iránnak, a Libanonban tartott amerikai túsók elengedése reményében + a profit illegális átirányítása nicaraguai Contra-knak, akik a Sandinista kormány ellen harcoltak.

Botrány: 1986 - 1993, irányító Vice Admiral John Poindexter, és beosztottja (a bűnbak) Lt. Colonel Oliver North.

"Ha egy amerikai fegyvert ad el egy ellenséges nemzetnek, akkor bűnös."

$$1. \forall x \forall y \forall z \text{ amerikai}(x) \wedge \text{fegyver}(y) \wedge \text{nemzet}(z) \\ \wedge \text{ellenséges}(z) \wedge \text{elad}(x, y, z) \rightarrow \text{bűnös}(x)$$

" Nono nemzet rendelkezik rakétákkal."

$$2. \exists x \text{ van}(\text{Nono}, x) \wedge \text{rakéta}(x)$$

" ... rakétákkal, amiket West ezredes adott el nekik."

$$3. \forall x \text{ van}(\text{Nono}, x) \wedge \text{rakéta}(x) \rightarrow \text{elad}(\text{West}, x, \text{Nono})$$

Rakéta egy fegyver.

$$4. \forall x \text{ rakéta}(x) \rightarrow \text{fegyver}(x)$$

Amerika ellensége egy ellenséges valaki.

$$5. \forall x \text{ ellensége}(x, \text{Amerika}) \rightarrow \text{ellenséges}(x)$$

West, aki amerikai ...

$$6. \text{amerikai}(\text{West})$$

Nono egy nemzet

$$7. \text{nemzet}(\text{Nono})$$

Nono, az Amerika ellensége ...

$$8. \text{ellensége}(\text{Nono}, \text{Amerika})$$

$$9. \text{nemzet}(\text{Amerika})$$

(ezek eddig az axiómák!)

Az elmélet és az axiómák: matematikában és MI-ben

axiómák = TB

kérdés

bizonyítás és utána?

Matematika

minimális

....

hosszú

dicsőség

redundanciamentes

MI

maximális

redundáns

...

minél
rövidebb

cselekvés
végrehajtása

10. (2, egziszt. kvantor el.) $\Rightarrow \text{van}(\text{Nono}, R1) \wedge \text{rakéta}(R1)$
11. (10, ÉS-el.) $\Rightarrow \text{van}(\text{Nono}, R1)$
12. $\text{rakéta}(R1)$
13. (4, univ. kvantor el.) $\Rightarrow \text{rakéta}(R1) \rightarrow \text{fegyver}(R1)$
14. (12, 13, Modus Ponens) $\Rightarrow \text{fegyver}(R1)$
15. (3, univ. kv. el.) $\Rightarrow \text{van}(\text{Nono}, R1) \wedge \text{rakéta}(R1) \rightarrow \text{elad}(\text{West}, R1, \text{Nono})$
16. (15, 10, MP) $\Rightarrow \text{elad}(\text{West}, R1, \text{Nono})$
17. (1, 3x univ. kv. el.) $\Rightarrow$
 $\text{amerikai}(\text{West}) \wedge \text{fegyver}(R1) \wedge \text{nemzet}(\text{Nono})$
 $\wedge \text{ellenséges}(\text{Nono}) \wedge \text{elad}(\text{West}, R1, \text{Nono}) \rightarrow \text{bűnös}(\text{West})$
18. (5, univ.kv. el.) $\Rightarrow \text{ellensége}(\text{Nono}, \text{Amerika}) \rightarrow \text{ellenséges}(\text{Nono})$
19. (8, 18, MP) $\Rightarrow \text{ellenséges}(\text{Nono})$
20. (6, 7, 14, 16, 19, ÉS bevezetése) $\Rightarrow$
 $\text{amerikai}(\text{West}) \wedge \text{fegyver}(R1) \wedge \text{nemzet}(\text{Nono}) \wedge \text{ellenséges}(\text{Nono})$
 $\wedge \text{elad}(\text{West}, R1, \text{Nono})$
21. (17, 20, MP) $\Rightarrow \text{bűnös}(\text{West})$

(igaz deduktíve az axiómákból)

Honnan tudjuk, melyik konstanst írunk be változók helyére?
--

Általánosított Modus Ponens

$$\frac{p_1', p_2', \dots, p_n', (p_1 \wedge p_2 \wedge \dots \wedge p_n) \rightarrow q}{q_\theta}$$

ahol $p_i'\theta = p_i\theta$, minden i -re

Pl. $p_1' = \text{király}(\text{János})$ $p_1 = \text{király}(x)$
 $p_2' = \text{mohó}(y)$ $p_2 = \text{mohó}(x)$
 $\theta = \{x / \text{János}, y / \text{János}\}$
 $q = \text{gonosz}(x)$ $q_\theta = \text{gonosz}(\text{János})$

$$\frac{\text{király}(\text{János}), \text{mohó}(y), (\text{király}(x) \wedge \text{mohó}(x) \rightarrow \text{gonosz}(x))}{\text{gonosz}(\text{János})} \quad (\forall)$$

Általánosított Modus Ponens helyes.

Előrecsatolt alkalmazása FOL határozott klózokra teljes és polinomiális komplexitású (ha terminálódik és nincsenek függvények).

A premisszák általános illesztése NP-nehéz.

Hátracsatolt alkalmazása esetén a tár komplexitása lineáris, de a séma nem teljes, nem hatékony.

Gépi bizonyítás kérdése – illesztés redundáns apparátusban

Valaki mindig szereti azt, aki minden állatot szeret.

Tegyük fel, hogy a fenti természetes nyelvű mondat logikai reprezentációja:

$$\forall x [\forall y \text{ állat}(y) \rightarrow \text{szeret}(x, y)] \rightarrow [\exists y \text{ szeret}(y, x)]$$

Mi itt a probléma?

Modus Ponens? Horn-klózik? Más?

Mintaillesztési probléma premisszára, következményre?

Redundanciák lekezelése?

1. $\forall x [\forall y \text{ állat}(y) \rightarrow \text{szeret}(x, y)] \rightarrow [\exists y \text{ szeret}(y, x)]$
2. $\neg \text{szeret}(\text{János}, \text{Bodri}) \wedge \text{állat}(\text{Bodri})$
3. $\forall y \neg \text{szeret}(y, \text{János})$

Következik-e, pl.: $1 + 2 \rightarrow 3$? Pl. Modus Ponens sémával?

Modus Ponens alapú bizonyítás **nem teljes**, de 1' rendű logika **teljes** (Gödel, 1930, teljességi tétel), avagy igaz tételek bizonyítása kell, hogy létezzen (de hogyan?).

Az első praktikusán jól algoritmizált „hogyan”, a **rezolúció** (Robinson, 1963), de a félig eldönthetőség, mint probléma megmarad!

Rezolúció predikátum kalkulusban:

	$q(x) \vee p(x, \text{Atti})$
	$\neg q(\text{Béla}) \vee r(x)$
$x/\text{Béla}$	$p(\text{Béla}, \text{Atti}) \vee r(\text{Béla})$

Egyesítés (unifikálás)

ez viszont akkor lehetséges, ha:

	$q(x)$ és $\neg q(y)$, ha x/y
x	y
változó-1	változó-2
változó	konstans
konstans	változó

nem megy, ha:

konstans-1	konstans-2	Id.	$\neg \text{kutya}(\text{Bodri})$	$\text{kutya}(\text{János}) ?$
változó	f(változó)			

Transzformáció klóz formára

1. Ekvivalenciát eltüntetni: $A \Leftrightarrow B = (A \rightarrow B) \wedge (B \rightarrow A)$

2. Implikációt eltüntetni: $A \rightarrow B = \neg A \vee B$

3. Negálást az atomi formulák szintjére áthelyezni.

$$\neg (A \vee B) = \neg A \wedge \neg B, \quad \neg \forall x p(x) = \exists x \neg p(x)$$

4. Egzisztenciális kvantorokat eltüntetni: Skolemizálás

5. Ha szükséges, a változókat átnevezni.

$$\forall x p(x) \vee \forall x q(x) \quad \rightarrow \quad \forall x p(x) \vee \forall y q(y)$$

6. Univerzális kvantorokat balra kihelyezni.

$$....\forall x....\forall y.. = \forall x\forall yx...y..$$

A cél:
az eredeti állításforma
és a klózforma
logikailag ekvivalensek!

7. Diszjunkciókat literál szintjére áthelyezni.

$$(A \wedge B) \vee C = (A \vee C) \wedge (B \vee C) \quad \text{konjunktív normal forma (CNF, Davis, 1960)}$$

8. Konjunkciókat eliminálni. Bontás diszjunktív klózokra

9. Ha szükséges, a változókat átnevezni.

10. Univerzális kvantorokat elhagyni. **ami marad: $\neg$, $\vee$, és most?**

A példa: Valaki mindig szereti azt, aki minden állatot szeret.

$$\forall x [\forall y \text{ állat}(y) \rightarrow \text{szeret}(x, y)] \rightarrow [\exists y \text{ szeret}(y, x)]$$

$$\forall x \neg [\forall y \neg \text{állat}(y) \vee \text{szeret}(x, y)] \vee [\exists y \text{ szeret}(y, x)]$$

$$\forall x [\exists y \neg (\neg \text{állat}(y) \vee \text{szeret}(x, y))] \vee [\exists y \text{ szeret}(y, x)]$$

$$\forall x [\exists y \neg \neg \text{állat}(y) \wedge \neg \text{szeret}(x, y)] \vee [\exists y \text{ szeret}(y, x)]$$

$$\forall x [\exists y \text{ állat}(y) \wedge \neg \text{szeret}(x, y)] \vee [\exists y \text{ szeret}(y, x)]$$

$$\forall x [\exists y \text{ állat}(y) \wedge \neg \text{szeret}(x, y)] \vee [\exists z \text{ szeret}(z, x)]$$

$$\forall x [\text{állat}(F(x)) \wedge \neg \text{szeret}(x, F(x))] \vee \text{szeret}(G(x), x)$$

$$[\text{állat}(F(x)) \wedge \neg \text{szeret}(x, F(x))] \vee \text{szeret}(G(x), x)$$

$$[\text{állat}(F(x)) \vee \text{szeret}(G(x), x)] \wedge [\neg \text{szeret}(x, F(x)) \vee \text{szeret}(G(x), x)]$$

Egy állításból két (lehet több is) klóz:

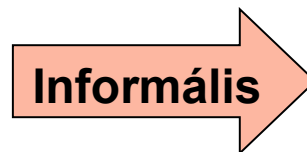
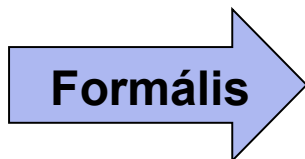
a. $\text{állat}(F(x)) \vee \text{szeret}(G(x), x)$

b. $\neg \text{szeret}(x, F(x)) \vee \text{szeret}(G(x), x)$ $F(x), G(x)$ Skolem függvények

Rezolúciós bizonyítás procedúrája

Adott állítások halmaza F , a bizonyítandó állítás Q

1. Az F halmaz összes állítását konvertáljuk klóz formába F' .
2. Negáljuk a Q -t és konvertáljuk klóz formába. Adjuk hozzá az F' -hez.
3. Ismételjük az alábbi ciklust, amíg:
 - (a) **ellentmondásra** rá nem futunk,
 - (b) **az előrehaladást** már nem tapasztaljuk, vagy
 - (c) **az erőforrások előre meghatározott mennyiségét ki nem használjuk:**
4. **válasszunk meg** két klózt.
5. **alkalmazzunk rezolúciós lépést.**
6. **ha a rezolvens üres**, megvan az ellentmondás.
Ha nincs, adjuk hozzá a többi klóz-hoz és folytatjuk (3.)



Rezolúciós stratégiák (klózek kiválasztási heurisztikái) (a keresés irányadása)

- 1. Egységklóz preferencia:** lényeges gyorsítás, ha klózek egyike egy szimpla literál: $P, \neg P \vee [.....] \Rightarrow [.....]$ rövidebb!
Horn-klóz alakú tudásbázisban az eljárás teljes, különben nem!
- 2. 'Set of Support':** rezolúció (egy klóz 'Set of Support'-ból és egy 'külső' klóz), rezolvens vissza 'Set of Support'-ba.
Eljárás teljes, ha 'Set of Support'-n kívüli klózek teljesíthetők, gyakorlatban: 'Set of Support' = a negált kérdés (a többit úgyis elhisszük)
- 3. Input rezolúció:** az egyik klóz mindig az előbbi rezolvens, az első lépésnél viszont a kérdés. **Horn-klóz alakú tudásbázisban az eljárás teljes**, különben nem!
- 4. Lineáris rezolúció:** P és Q rezolválható, ha P benne van az eredeti tudásbázisban, vagy ha P a Q őse a bizonyítási fában. **Lineáris rezolúció egy teljes eljárás.**
- 5. Egyszerűsítések**
Elimináljunk minden olyan állítást, amely egy tudásbázisban létező állításnál specifikusabb. ... ami tautológia ... ami

1. $\forall x \forall y \forall z a(x) \wedge f(y) \wedge n(z) \wedge ell(z) \wedge elad(x,y,z) \rightarrow b(x)$
2. $\exists x v(N,x) \wedge r(x)$
3. $\forall x v(N,x) \wedge r(x) \rightarrow elad(W,x,N)$
4. $\forall x r(x) \rightarrow f(x)$
5. $\forall x e(x,A) \rightarrow ell(x)$
6. $a(W)$
7. $n(N)$
8. $e(N,A)$
9. $n(A)$

FOL állítások

$b(W)?$

1. $\neg a(x1) \vee \neg f(y1) \vee \neg n(z1) \vee \neg ell(z1) \vee \neg elad(x1,y1,z1) \vee b(x1)$

2a. $v(N,Rs)$

2b. $r(Rs)$

3. $\neg v(N,x2) \vee \neg r(x2) \vee elad(W,x2,N)$

4. $\neg r(x3) \vee f(x3)$

5. $\neg e(x4,A) \vee ell(x4)$

6. $a(W)$

7. $n(N)$

8. $e(N,A)$

9. $n(A)$

10. $\neg b(W)$

klózek

$$11 = 2a + 3$$

$x2/Rs$

$$12 = 11 + 2b$$

$$13 = 12 + 1$$

$x1/W, y1/Rs, z1/N$

$$14 = 13 + 10$$

$$15 = 14 + 6$$

$$16 = 15 + 7$$

$$17 = 16 + 4$$

$$18 = 17 + 5$$

$$19 = 18 + 8$$

$$20 = 19 + 2b$$

rezolúció

Input rezolúció

Horn-klóz tudásbázis

Ujra West ezredes ügye

1. $\neg a(x1) \vee \neg f(y1) \vee \neg n(z1) \vee \neg ell(z1) \vee \neg elad(x1,y1,z1) \vee b(x1)$

2a. $v(N, Rs)$

2b. $r(Rs)$

3. $\neg v(N, x2) \vee \neg r(x2) \vee elad(W, x2, N)$

4. $\neg r(x3) \vee f(x3)$

5. $\neg e(x4, A) \vee ell(x4)$

6. $a(W)$

7. $n(N)$

8. $e(N, A)$

9. $n(A)$

10. $\neg b(W)$

klózek

West ezredes ügye
egy csavarral

És ha mégsem biztos előre, hogy ez West volt?
Hogy lehet rábizonyítani, anélkül, hogy a nevét
sugallnánk a bizonyításnak?

$\exists x \text{ bűnös}(x)$?

10. $\neg(\exists x \text{ bűnös}(x))$
 $\forall x \neg \text{bűnös}(x)$
 $\neg \text{bűnös}(x)$
 $\neg b(x5)$

Az elmélet és az axiómák: matematika és MI

<u>axiómák = TB</u>	<u>kérdés</u>	<u>bizonyítás és utána?</u>	<u>Hol az infó?</u>
Matematika minimális redundanciamentes		hosszú	pávatoll üres rezolvens tényében
MI maximális redundans	...	minél rövidebb	cselekvés végrehajtása üres rezolvens tényében ÉS a behelyettesítésekben

A tudásbázis építése

tudásbeszerzés (knowledge acquisition):

tudásmérnök $\Leftrightarrow$ valódi szakértő

tudásszervezés (knowledge engineering): a tudásbázis építése
ontológia szervezése (ontological engineering)

A jó és a rossz tudásbázis tulajdonságai

- egy jó tudásreprezentációs nyelv:

nagy kifejező erejű, tömör, egyértelmű, környezet-érzékeny, hatékony

- tudásbázis: letisztázott, korrekt, csak fontos relációk

kompromisszum: tömörség $\Leftrightarrow$ korrektség

Tudásbázis fogyasztója: az emberi olvasó és a következtető gép

Gyakori hiba:

- embernek értelmesnek tűnő predikátumnevek használata
- értelmesek lesznek-e a következtető eljárásnak is?

A tudásbázis építése – mire lehet szükség

objektumok, relációk ...

csoportosulások, szervezetek, halmazok, képviselők ...

objektumok részei, struktúrái, ...

objektumok méretei ...

objektumok és relációk változása az időben ...

...

(valódi objektumok modellezése ágens agyában ...)

ágens agyában létező objektum(fogalmak) modellezése
ágens agyában?

...

Tudásszervezés (Knowledge Engineering) lépései

- 1. Döntsük el, miről fogunk beszélni:**
objektumok, tények, relációk (**ontológia**)
- 2. Döntsük el predikátumok, függvények és konstansok szótárát:**
logikai szintű nevek (amire szabad kezünk van)
- 3. A tárgytartományra vonatkozó általános tudás kódolása:**
logikai állítások = axiómák, precíz módon,
a következtetési eljárás automatizmusához
- 4. Specifikus problémaegyedek leírása:**
ha az ontológia jó, könnyű, mert a nyelv adott
- 5. Kérdések megfogalmazása a következtető eljárás számára, válaszok fogadása:**
jutalom: helyettünk már csak a gép dolgozik
(formális hitelességgel)

házi állatok, kutyák

nagy(x), hangos(x),
harapós(x), kölyke(x), ...

$\forall x. \text{harapós}(x) \rightarrow$
 $\text{harapós}(\text{kölyke}(x))$

Bodri = kölyke(Cézár)
nagy(Bodri), ...

harapós(Bodri)?