

Alkalmazott mesterséges intelligencia (AMI)

<http://www.mit.bme.hu/oktatas/targyak/vimibb01>

4 ea. (2019 ősz)

Logikai következtetést használó MI ágens

<http://mialmanach.mit.bme.hu/aima/ch07>

7. fejezet

Előadó: Pataki Béla

a fóliák

Dobrowiecki Tadeusz és
Hullám Gábor anyagainak
felhasználásával készültek



BME I.E. 414, 463-26-79

pataki@mit.bme.hu,

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

Logika – a reprezentáció és manipuláció eszköze

Szintaktika - legálisan létrehozható szimbolikus mondatok

Szemantika - a világ tényei, amikre a szimbolikus mondatok vonatkoznak

Vonzatreláció és a következtetés

Tények megneveznek valós dolgokat, amelyek egymással kapcsolatban vannak. Ha az egyik igaz a világban, és ezért a másik szükségszerűen szintén az:

vonzat: logikai konzekvencia dolgok között.

Vonzatreláció tudásbázis TB és egy α mondat között:

α **vonzata** TB -nek: $\mathbf{TB} \models \alpha$, ha α minden olyan világban igaz, ahol TB is.

Következtetés: a vonzat „kiszámítása” a mondatok formális manipulálásával.

Bizonyítás = a következtetési algoritmus lépéssorozata

α **bizonyítható** TB -ből: $\mathbf{TB} \vdash \alpha$

Logika – reprezentáció és manipuláció eszköze

Szintaktika - legálisan létrehozható szimbolikus mondatok

Szemantika - a világ tényei, amikre a szimbolikus mondatok vonatkoznak

Mi egy mondat jelentése?

- A mondat írójának egy **interpretációt** kell adnia hozzá, kijelentve, hogy milyen tény tartozik hozzá.
- Egy mondat önmagában nem jelent semmit, és lehet igaz vagy hamis.
- Egy mondat igaz egy bizonyos interpretációban, ha a dolgok állása olyan.
- **Az igazság függ mind a mondat interpretációjától, mind a világ aktuális állapotától.**

Szemantika és interpretáció

Mi lehet pl. a **Fut_Ágens1** mondat jelentése?

Interpretáció: „fut” fizikai mozgás, „Ágens1” egy személy, akkor ebben a világban ez a mondat (most) **igaz**.

Interpretáció: „fut” nem terminált működés, „Ágens1” egy program, akkor ebben a világban ez a mondat (most) **hamis**.



Interpretáció: „fut” fizikai mozgás, „Ágens1” egy személy, akkor ebben a világban a mondat (most) **hamis**.

Interpretáció: „fut” nem terminált működés, „Ágens1” egy program, akkor ebben a világban ez a mondat (most) **igaz**.



Ha adott egy tudásbázis (TB), a következtetés:

- (1) létrehozhat új α mondatot, amely vonzata a TB-nak
- (2) ha adott α mondat, eldönthető, hogy α vonzata-e a TB-nak.

Igazságtartó következtetés, ha **csak igazságfüggvényekkel dolgozik** (igazságfüggvény logikai értéke csakis a részeinek és a logikai műveletek definíciójától függ, **nem függ az interpretációtól**).

Formális bizonyítás – igazságtartó következtetési eljárás lépései.

Következtetési eljárás teljes: ha minden vonzatmondathoz talál egy bizonyítást, avagy ami igaz, az bebizonyítható. $A \models B$ -ből $A \vdash B$

Következtetési eljárás helyes: ha minden bizonyított mondat vonzatrelációban áll a felhasznált tényekkel, avagy amit bebizonyított, az igaz is.

$A \vdash B$ -ből $A \models B$

Egy történet: „Ádám, Béla és Csaba közül valaki kitörhette egy üvegház ablakát.

- Ádám azt állítja, hogy: 'Béla tette, Csaba nem bűnös.'
- Béla azt állítja, hogy: 'Ha Ádám bűnös, akkor Csaba is'.
- Csaba azt állítja, hogy: 'Nem én tettem; többiek közül valaki'."

Lehetséges kérdések:

- Konzisztensek (egyszerre igazak) ezek az állítások?
- Ha egyik sem bűnös, van-e, aki hazudik?
- Ha mindenki igazat mond, akkor ki a bűnös?
- Stb.

A történet **ítélet szimbólumai:**

A: Ádám nem bűnös,

B: Béla nem bűnös,

C: Csaba nem bűnös

és az állításaik:

$$\mathbf{SA} = \neg B \wedge C$$

$$\mathbf{SB} = \neg A \Rightarrow \neg C$$

$$\mathbf{SC} = C \wedge (\neg B \vee \neg A)$$

az un. „elmélet” pedig (az igaznak tartott mondatok összessége),
ami egyben a TB tartalma: $\mathbf{TB} = SA \wedge SB \wedge SC$

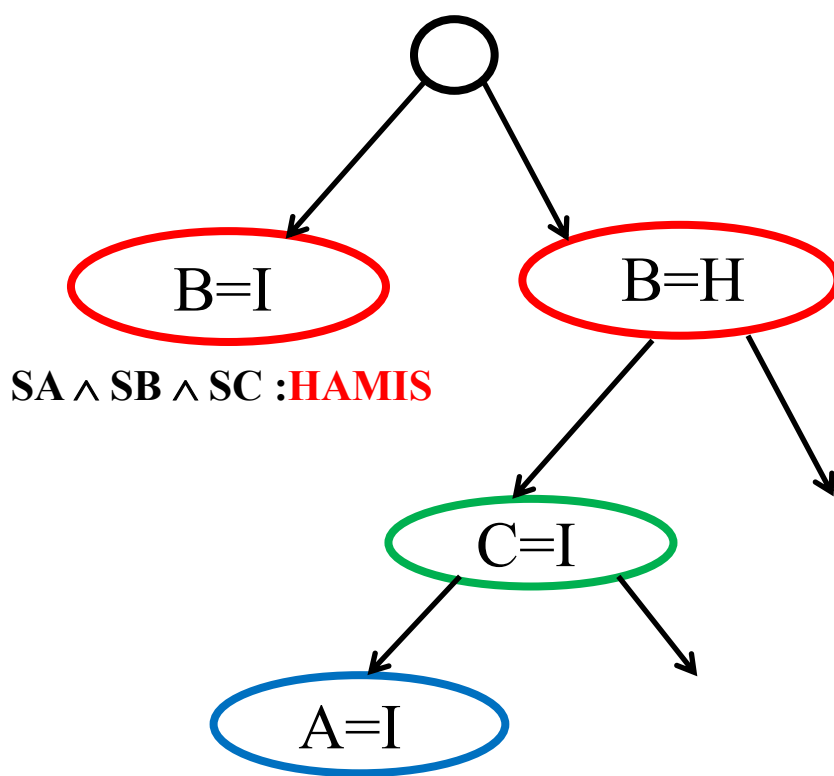
Felsorolásos bizonyítás

A	B	C	SA	SB	SC	TB = SA ∧ SB ∧ SC
H	H	H	H	I	H	H
H	H	I	I	H	I	H
H	I	H	H	I	H	H
H	I	I	H	H	I	H
I	H	H	H	I	H	H
I	H	I	I	I	I	I
I	I	H	H	I	H	H
I	I	I	H	I	H	H

- Konzisztensek (egyszerre igazak) ezek az állítások?
- Ha egyik sem bűnös, van-e, aki hazudik?
- Ha mindenki igazat mond, akkor ki a bűnös?

2^N sor, $O(2^N)$ komplexitás (NP)

Visszalépéses keresés modellek



Keressük

$$TB = SA \wedge SB \wedge SC =$$

$\neg B \wedge C \wedge (C \wedge (\neg B \vee \neg A)) \wedge (\neg A \rightarrow \neg C)$
 olyan modelljét ahol a TB kielégíthető,
 azaz igaz lesz (konzisztensek?)

1. Ha **B IGAZ**, akkor

$\neg B \wedge C \wedge (\neg B \vee \neg A) \wedge (\neg A \rightarrow \neg C)$
 biztosan HAMIS, tehát szükséges, hogy
 $\neg B$ igaz legyen (bár lehet, hogy akkor
 se teljesül)

2. Maradt:

$C \wedge (\neg A \rightarrow \neg C)$ igazához szükséges,
 hogy **C** legyen **igaz**

3. $\neg A \rightarrow$ Hamis igazához elég (bár
 egyben szükséges is),
 hogy **A IGAZ** ($\neg A$ legyen hamis)

A	B	C	SA	SB	SC	TB
H	H	H	H	I	H	H
H	H	I	I	H	I	H
H	I	H	H	I	H	H
H	I	I	H	H	I	H
I	H	H	H	I	H	H
I	H	I	I	I	I	I
I	I	H	H	I	H	H
I	I	I	H	I	H	H

Érvényes (analitikus mondat, tautológia): ha minden világban minden lehetséges interpretációja igaz, függetlenül attól, hogy mit akar jelenteni és mi a világ állapota.

Pl. „Vagy esik az eső vagy nem esik az eső”

Kielégíthető: ha létezik olyan interpretáció, hogy valamely világban igaz. Ami nem kielégíthető, az **kielégíthetetlen**.

Kvíz következik!

Kvíz (04.1.)

Négy logikai mondatunk van:

- (1) $A \vee B$
- (2) $A \wedge (B \wedge \neg B)$
- (3) $A \wedge B$
- (4) $A \vee (B \vee \neg B)$

Az alábbiak közül melyik igaz?

- A. Az (1) és (3) kielégíthető, a (2) kielégíthetetlen
- B. Az (1) és (3) kielégíthető, a (4) kielégíthetetlen
- C. A (2) kielégíthető, a (4) érvényes
- D. A (2) érvényes, a (4) kielégíthető

A	B	$A \vee (B \vee \neg B)$	$A \wedge (B \wedge \neg B)$	$A \wedge B$	$A \vee B$
0	0	1	0	0	0
0	1	1	0	0	1
1	0	1	0	0	1
1	1	1	0	1	1

Melyik állítás érvényes, kielégíthető, ill. kielégíthetetlen?

Logikai következtetésen alapuló intelligens ágens

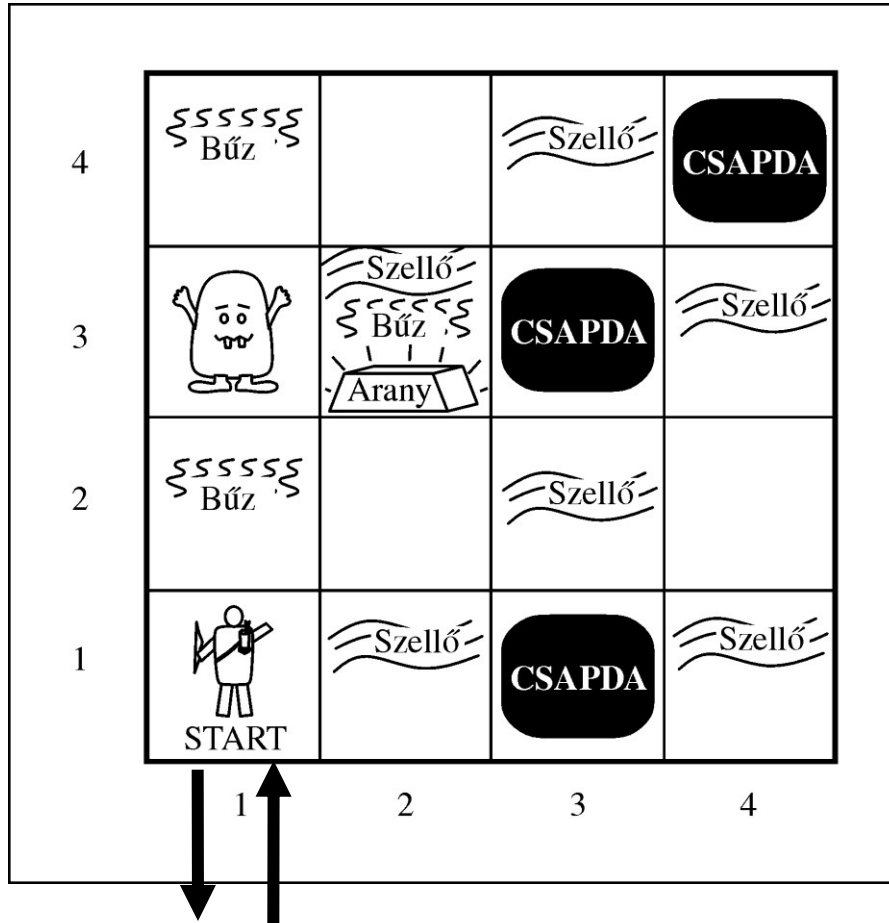
Szabályalapú (logikai, szakértői) rendszerek

- Az eszközünknek van egy belső modellje a világról – ennél az eszköznél logikai reprezentáció, logikai állítások halmaza. (Más eszközöknek más modellje is lehet a világról.)
- A tények (állítások, ítéletek) és szabályok alapján képes új – közvetlenül nem adott – tényeket kikövetkeztetni.
- A felhasználó kérdéseket tehet fel, és az eszköz megpróbálja kikövetkeztetni a választ. (Pl. az adott jelek alapján XY beteg? Pl. az adott válaszok alapján a vásárló elégedett?)

A szakértői rendszeren alapuló **intelligens ágens észleléseket végezhet**, így is nyerhet új információt, és be is avatkozhat, **megváltoztathatja a saját és/vagy a világ állapotát**.

Logikusan gondolkozó ágens – Hogyan működik ez?

Esettanulmány - a „Wumpus világ” környezetében (jegyzet)



Ha ágens, akkor érzékelései, cselekvései és céljai vannak

Cél: **arany** - megtalálni és kihozni

Érzet(ek):

[Bűz, Szellő, Csillog, Ütközés, Sikoly]

Cselekvések:

Föl, Le, Jobbra/Balra, Megragad, Lő, Szomszédos mezőre Lép

A térképet nem ismeri

... az ágens meghal, ha csapdára vagy Wumpusra rálép.

Hogyan biztosítható a siker (logikai gondolkodással)?

Az állítások szerkezete:

Kvíz következik

- $B(1,1)$: az $(1,1)$ mezőn bűzt érez
- $Sz(4,2)$: a $(4,2)$ mezőn szellőt érez
- $C(3,2)$: a $(3,2)$ mezőn csapda van
- $A(3,3)$: az ágens a $(3,3)$ mezőn van
- $Au(1,4)$: az $(1,4)$ mezőn arany van
- $M(2,2)$: a $(2,2)$ mezőt már meglátogatta
- S : a wumpust eltalálta, sikoly és elpusztul
- $\ddot{U}$: falnak ütközött az ágens
- $Cs(4,2)$: a $(4,2)$ mezőn csillog az arany
- $OK(1,1)$: az $(1,1)$ mező biztonságos

Egy intelligens logikai ágens a Wumpus világ számára

Tudásbázis ciklus: ágens érzékelései $\rightarrow$ mondatok $\rightarrow$ tudásbázis v. cselekvés

1,4	2,4	3,4	4,4
1,3	2,3	3,3	4,3
1,2	2,2	3,2	4,2
OK			
1,1	2,1	3,1	4,1
Á OK	OK		

Á = Ágens
 S = Szellő
 R = Ragyogás, Arany
 OK = Biztonságos négyzet
 C = Csapda
 B = Bűz
 M = Meglátogatott
 W = Wumpus

(a)

04.2. Kvíz

Milyen infó kell ahhoz, hogy a (4,1) mezőről megállapítsa az ágens, hogy biztonságos?

- A. $\neg B(4,1) \wedge \neg Sz(4,1)$
- B. $\neg B(3,1) \wedge \neg B(4,2)$
- C. $\neg Sz(3,1) \wedge \neg Sz(4,2)$
- D. $(\neg B(3,1) \vee \neg B(4,2)) \wedge (\neg Sz(3,1) \vee \neg Sz(4,2))$

A kiinduló helyzet tényei és néhány szabály

Kvíz következik

$\dot{A}(1,1)$

$\neg B(1,1)$

$\neg Sz(1,1)$

$\neg W(1,1)$

$\neg B(1,1) \Rightarrow \neg W(2,1)$

$\neg B(1,1) \Rightarrow \neg W(1,2)$

$\neg Sz(1,1) \Rightarrow \neg C(2,1)$

$\neg Sz(1,1) \Rightarrow \neg C(1,2)$

$\neg C(2,1) \wedge \neg W(2,1) \Rightarrow OK(2,1)$

$\neg C(1,2) \wedge \neg W(1,2) \Rightarrow OK(1,2)$

$A(1,1) \wedge OK(1,2) \Rightarrow \text{FEL}$

$A(1,1) \wedge OK(2,1) \Rightarrow \text{JOBBRA}$

$B(1,1) \wedge OK(1,2) \Rightarrow \text{JOBBRA_LŐ}$

04.3. Kvíz

Szükség lesz-e a kiinduló helyzetben konfliktusfeloldásra a logikai rendszer működési ciklusa során?

- A. Nem mert nincsenek egymásnak ellentmondó információk
- B. Igen, mert több szabály feltételrészére is illeszthetők a tények
- C. Igen, mert $\neg Sz(1,1)$ és $\neg B(1,1)$ egymásnak ellentmond
- D. Nem mert a wumpus nem lehet sem (1,2)-n, sem (2,1)-n

1,4	2,4	3,4	4,4
1,3	2,3	3,3	4,3
1,2	2,2	3,2	4,2
OK			
1,1	2,1	3,1	4,1
Á			
OK	OK		

(a)

Á

 = Ágens
 S = Szellő
 R = Ragyogás,
 Arany
 OK = Biztonságos
 négyzet
 C = Csapda
 B = Bűz
 M = Meglátogatott
 W = Wumpus

1,4	2,4	3,4	4,4
1,3	2,3	3,3	4,3
1,2	2,2	3,2	4,2
OK	C?		
1,1	2,1	3,1	4,1
M	Á	C?	
OK	S		
	OK		

(b)

1,4	2,4	3,4	4,4
1,3 W!	2,3	3,3	4,3
1,2 Á B	2,2 OK	3,2	4,2
1,1 M OK	2,1 S M OK	3,1 C!	4,1

(a)

Á = Ágens
 S = Szellő
 R = Ragyogás,
 Arany
 G = Biztonságos
 négyzet
 C = Csapda
 B = Bűz
 M = Meglátogatott
 W = Wumpus

1,4	2,4 C?	3,4	4,4
1,3 W!	2,3 Á B R	3,3 C?	4,3
1,2 B M OK	2,2 S M OK	3,2	4,2
1,1 M OK	2,1 S M OK	3,1 C!	4,1

(b)

Környezeti titkok (nem hozzáférhető!) kiderítése logikai modellezéssel és gondolkodással, majd jó cselekvések választása.

Rezolúció (1963) Q mondat bizonyítása TB alapján

1. TB átírása klóz formába. (Az állítások VAGY kapcsolata).

$$\begin{array}{c} B \vee A \\ \hline \neg B \vee G \\ \hline A \vee G \end{array}$$

2. Q kérdés negálása. Negált Q kérdés átírása klóz formára.

3. Kiterjesztett TB' = TB $\cup$ $\neg$ Q tudásbázis létesítése.

$$\begin{array}{c} X \\ \hline \neg X \\ \hline \emptyset \end{array}$$

4. Rezolúciós lépés ciklikus elvégzése:

- kilépés üres rezolvensre, akkor TB' egy ellentmondás, tehát mivel TB igaz, Q igaz kell, hogy legyen.

Rezolúció „cáfolat-teljes” lépés, garantáltan megtalálja az ellentmondást. Teljes kereséssel párosítva teljes bizonyítás.

(Ha a TB-ből következik Q, akkor megtalálja az ellentmondást. Ha például a TB-ből a $Q \vee W$ következik és W igaz, akkor nem tudhatja, hogy Q igaz-e.)

Ítéletlogika eldönthető és teljes

Minden jól definiált mondat akár **igaz**, akár **hamis** volta belátható véges algoritmussal (a vonzat eldönthető) → a **következtetés igazságtábla módszere teljes**, mindig lehetséges kiszámolni a tábla 2^n sorát bármely n ítéletszimbólumot tartalmazó bizonyítás esetében.

A számítási idő azonban n -ben **exponenciális**.

Ítéletlogika monoton

Logika **monoton**: amikor új mondatokat adunk hozzá a TB-hoz, minden korábban maga után vonzott mondata az eredeti TB-nak továbbra is vonzata marad az új, nagyobb tudásbázisnak.

$$\text{ha } TB_1 \models a, \text{ akkor } (TB_1 \cup TB_2) \models a$$

Az igaz mondatok száma csak nőni tud! **Jó ez, vagy nem jó?!**

Jó: amit egyszer nehéz volt bebizonyítani, majd „ingyen van”.

Rossz: a változó világ logikai leírása is változik és ami igaz volt, nem biztos, hogy később annak megtartható.

Σ : amíg a probléma tökéletesen statikus, a monotonitással nyerünk, különben ráfizethetünk.

Problémák az ítéletlogikai ágenssel

Általában: - **túl sok ítéletszimbólumot kell kezelni.**
- **nehézkes a változások kezelése.**

Emiatt: **lassú a következtetési eljárás.** (igazságtábla mérete)

De sok fontos gyakorlati alkalmazás (kielégíthetőségi vizsgálattal):

- formális modell-ellenőrzés
- automatikus tesztminta-generálás
- MI tervkészítés
- automatikus tételbizonyítás
- szoftver-verifikálás

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

Elsőrendű logika: milyen a világ valójában?

Objektumok: más objektumoktól megkülönböztetett identitású,
tulajdonságokkal rendelkező dolgok

Relációk: objektumok közötti kapcsolatok

- n-elemű
- egyelemű (unáris): tulajdonságok

Relációk közül néhány **függvény** – csak egy „érték” egy adott „bemenetre”.

Pl.: **objektumok:** emberek, elméletek, színek, sakkjátszmák, ...
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)

Elsőrendű logika, predikátum kalkulus

Szintaktika és szemantika

konstans szimbólumok: X, János, Bodri, K21 (a világ objektumai).

predikátum szimbólumok: kerek, bátyja,... (egy bizonyos reláció).

kerek(X)

bátyja(József, Anna)

függvény szimbólumok: koszinusz, apja, bal-lába (a relációban szereplő objektum pontosan egy másik objektummal van kapcsolatban).

János = apja(Béla)

apja(János, Béla) = Igaz

Építőelemek

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

atom: bátyja(Géza,Árpád), házas(apja(Richárd),anyja(János)),
atom(i mondat): predikátum szimbólum és az argumentumait jelentő
termek

tényeket fejez ki (van igazságértéke).

összetett mondat: bátyja(Misi,János) $\rightarrow$ házas(apja(Misi),anyja(János))

logikai összekötő szimbólumok $\neg \wedge \vee \rightarrow \Leftrightarrow$

a még összetettebb mondatok építésére.

Építőelemek

kvantorok: tulajdonságok, amelyek objektumok 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 ($\exists$) $\exists x. \text{macska}(x) \wedge \text{úszik}(x)$

Egyenlőség: (**term1 = term2**) igaz adott interpretáció mellett, a.cs.a, ha term1 és term2 ugyanarra az objektumra vonatkoznak.

Az $\forall$ és az $\exists$ kapcsolata: $\forall x \neg \text{szeret}(x, \text{Répa}) = \neg \exists x \text{ szeret}(x, \text{Répa})$
 $\neg \forall x \text{ szeret}(x, \text{Fagylalt}) = \exists x \neg \text{szeret}(x, \text{Fagylalt})$

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

Fizikai világ: Objektum: Bodri kutya, amely
Tulajdonsága: nagy

Konstans: Bodri
Predikátum:
nagytestű(x)
Egy igaz állítás:
nagytestű(Bodri)



Konstans: Kutya1
Predikátum:
dőgnagy(x)
Egy igaz állítás:
dőgnagy(Kutya1)

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

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

Predikátum kalkulus tulajdonságai

Teljes: minden igaz állítás belátható (bebizonyítható)

A vonzat csak félig eldönthető: állítás hamis volta nem mutatható ki!

- a bizonyítási eljárásnak **igaz állítás** esetén van kilépési pontja,
- **hamis állítás** esetén nincs kilépési pontja, és munka közben, logikai alapon, nem dönthető el, hogy melyik esettel foglalkozunk.

Gyakorlati következtetés: számítógépen egy állításhalmaz konzisztenciája elvben nem mutatható ki a **véges erőforrások (idő)** miatt!