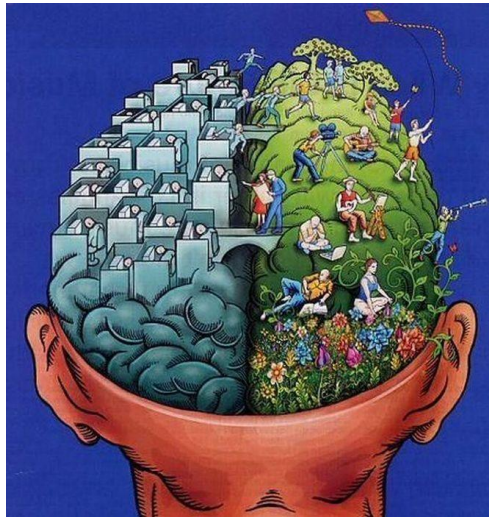




# Budapesti Műszaki és Gazdaságtudományi Egyetem Méréstechnika és Információs rendszerek Tanszék



## Mesterséges Intelligencia - MI

### Logikai ágens, lehetőségek és problémák

Előadó:

Hullám Gábor  
Pataki Béla

Előadás anyaga:

Dobrowiecki Tadeusz



M Ű E G Y E T E M 1 7 8 2

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

[pataki@mit.bme.hu](mailto:pataki@mit.bme.hu)

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



# IBM Watson (2011): Jeopardy kvízzjáték

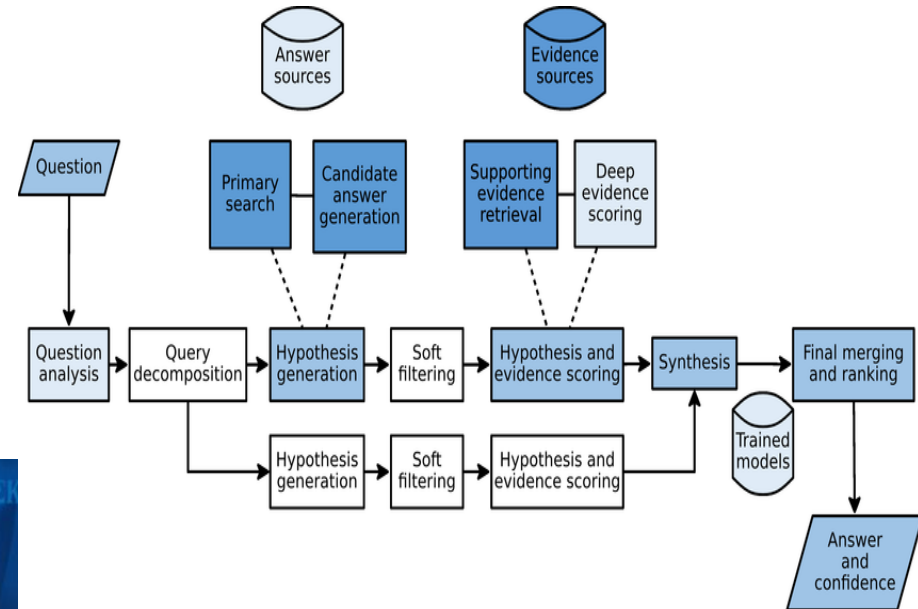
## IBM Grand Challenge

- 1997: **Deep Blue** legyőzi a G. Kaszparov sakkvilágbajnokot.
- 1999-2006<: **Blue Gene**, fehérje struktúra predikció
- 2011: **Watson**

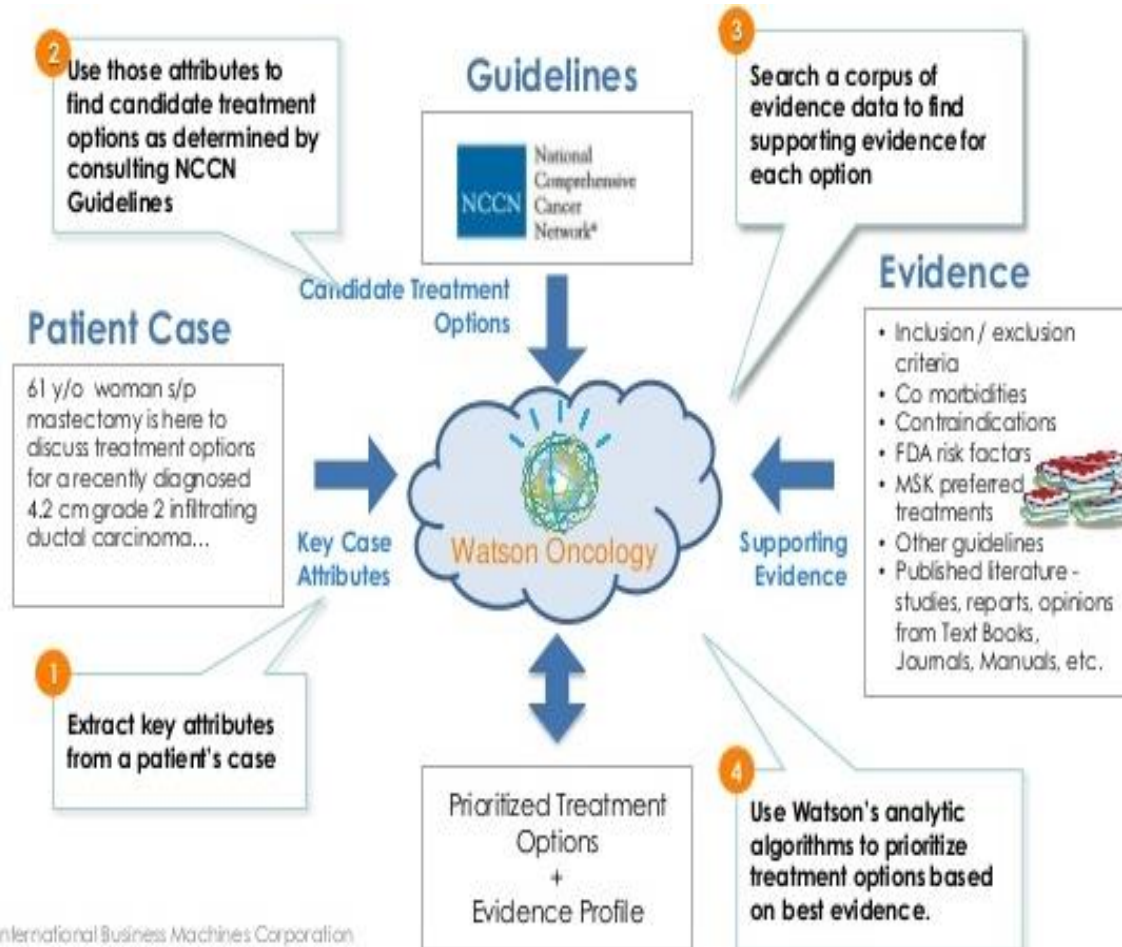
# IBM Watson (2011): Jeopardy kvízzjáték

2011: **Watson**

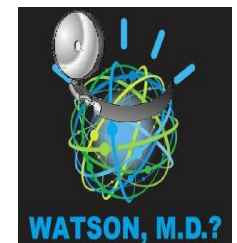
- Beszédértés
- Következtetés
- Játék



# Orvosi döntéstámogató és szakértői rendszerek



© 2014 International Business Machines Corporation



# Szakértői rendszerek az MI történetében

~1930 Univerzális számítási modell: Turing-gép (1936), Univerzális Turing-gép, Church-Turing hipotézis, Zuse, Neumann,...: „vezérlő program is adat”:

1943 McCulloch & Pitts: Bináris kapcsolati agymodell

1950 Turing: "Computing Machinery and Intelligence"

**1956** Dartmouth találkozó: "Artificial Intelligence" megnevezés elfogadása

1950s Korai MI programok: sakk, [tételbizonyítás](#)

## **Számítás(keresés) alapú MI**

- A Fizikai Szimbólumrendszer hipotézise: A.Newel&H.A.Simon (1976)

1966—73 Számítási komplexitási korlátok a keresésben  
Elméleti korlátok a neurális hálózatokban

**1969—79 Tudásalapú szakértői rendszerek (szabály alapú szakértői rendszerek)**

## **Tudásalapú MI**

1986-- Neurális hálózatok újbóli megjelenése

1988-- Valószínűségi szakértői rendszerek

1995-- Gépi tanulás gyors fejlődése

## **Adatvezérelte MI**

Autonóm tanulás alapú MI, Általános Intelligencia, SuperIntelligencia

Szabály alapú rendszerek = Tudásalapú rendszerek  
= Szakértő rendszerek

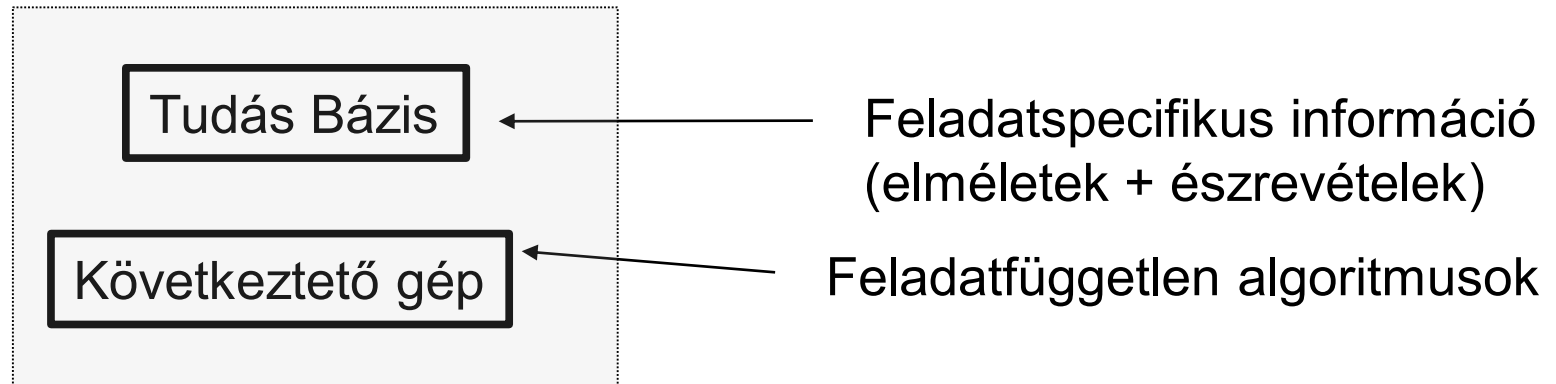
**Mycin**, Stanford, ca. 1972, kb. 500 szabály, szakorvos szintje

**IF** The site of the culture blood, and  
The gram stain of the organism is gramneg, and  
The morphology of the organism is rod, and  
The portal of entry of the organism is urine, and  
The patient has not had a genito-urinary manipulative procedure, and  
Cystitis is not a problem for which the patient has been treated  
**THEN** There is suggestive evidence (.6) - the identity of the organism is e.coli

**IF** The identity of the organism is bacteroides  
**THEN** Recommend therapy chosen from the following drugs:  
1. clindamycin (0.99)      2. chloramphenicol (0.99)  
3. erythromycin (0.57)      4. tetracycline (0.28)  
5. carbenicillin (0.27)

**R1, XCON**, CMU, 1978, DEC VAX rendszerek rendelés utáni konfigurálása,  
kb. 2500 szabály, 1980-1986 között kb. 80,000 rendelés, 95-98% pontosság,  
évi \$25M profit (szerelés sebessége, kevesebb kár, nagyobb elégedettség)

## Logika

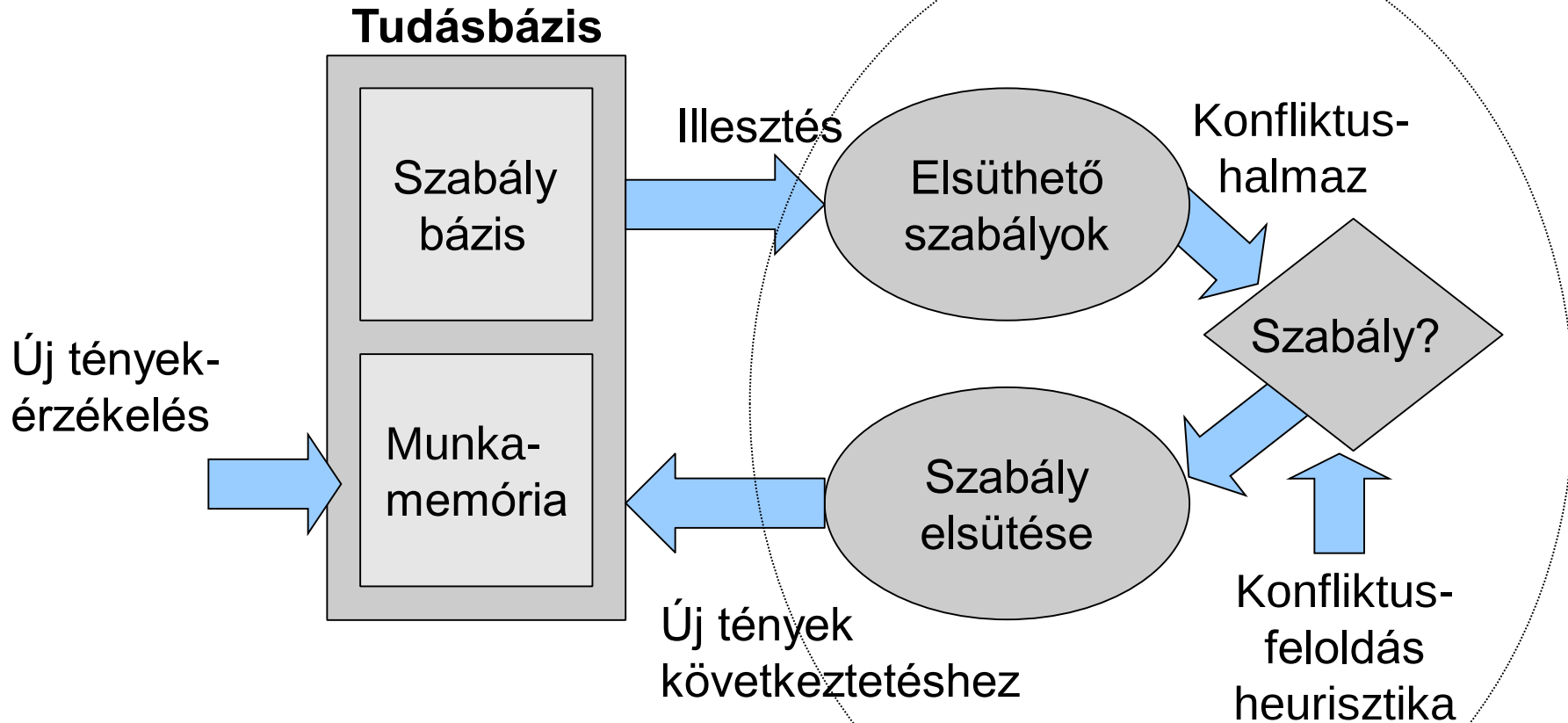


Mire használhatja az Ágens logikai képességeit?

Hogyan fejezze ki ehhez a tudását?

Milyen algoritmusok lesznek jók, ügyesek, ...?

# Következtetőgép



## Illesztés:

Premisszára = **előre**-következtetés, forward chaining

Következményre = **hátra**-következtetés, backward chaining



# A Cyc projekt 30 éve (1984-2016)



Cél: Alaptudás/józan ész formális leírása (ontológia)

1984-es becslés:

250 ezer szabály

350 emberév

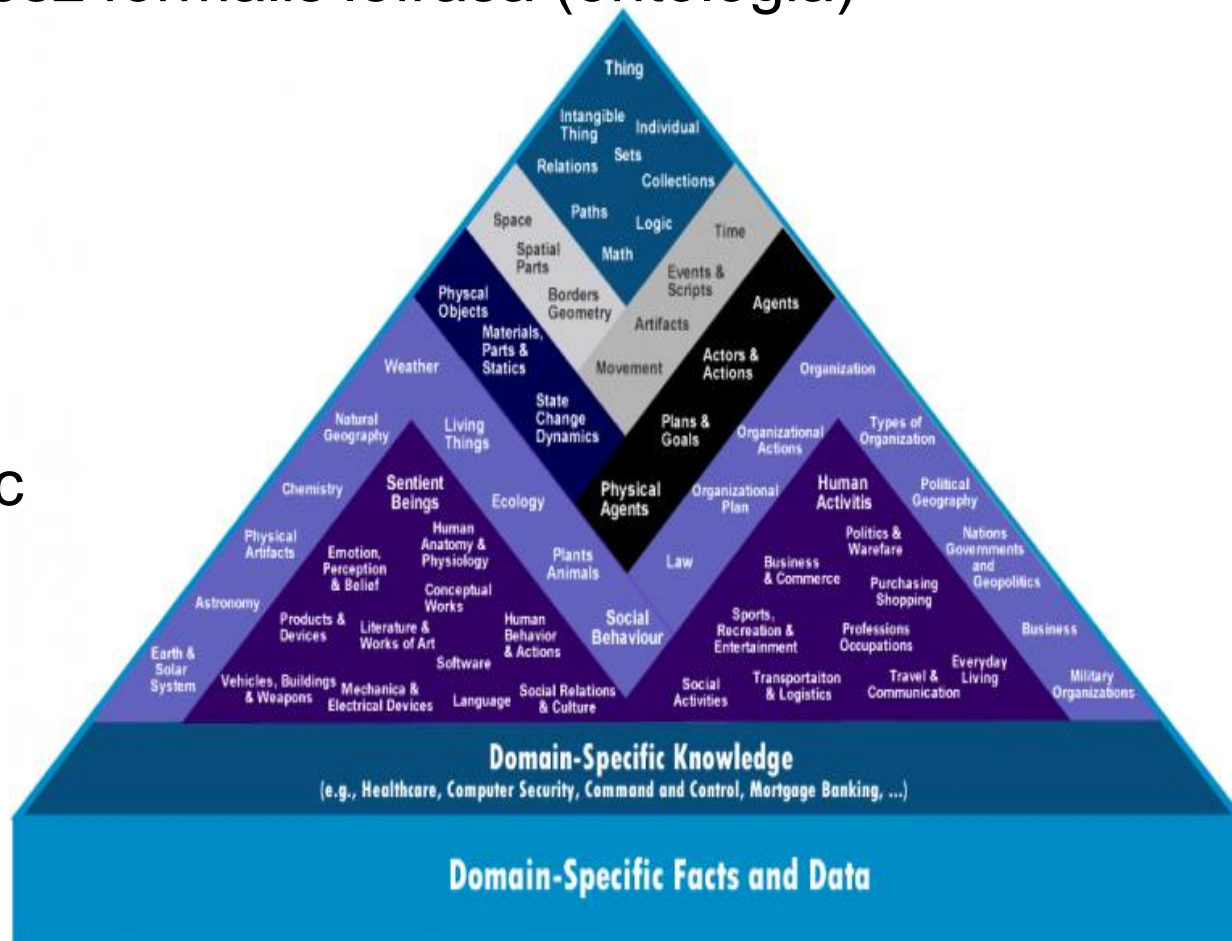
Nyelv: CycL

Elérhetőség: OpenCyc

Jelenlegi állapot

239,000 fogalom

2,093,000 tényállítás



# Szabályalapú szakértői rendszerek manapság

- Jogi szakértői rendszerek
- Adó- és pénzügyi tanácsadó rendszerek
- Ügyfélszolgálat és technikai támogató rendszerek

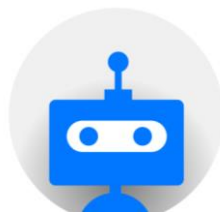
- Természetes nyelvű interfészek



- Digitális asszisztensek



- Csetbotok



**Drift**

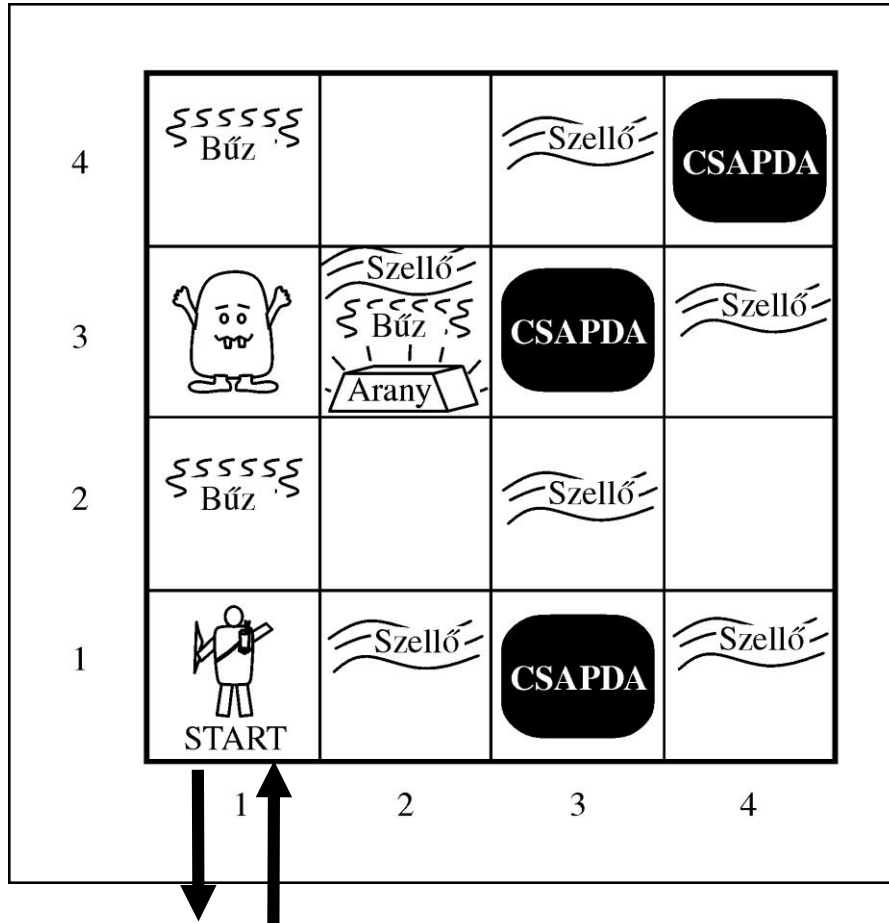


Workbot  
FOR SALESFORCE



# Logikusan gondolkozó ágens - Mit jelentsen ez?

Esettanulmány - a „Wumpus világ” környezetében (jegyzet)



**Ha ágens, akkor érzékelések, cselekvések és célok.**

Érzet(ek):

[Búz, Szellő, Csillog, Ütközés, Sikoly]  
hol vagyunk, nem tudni.

Cselekvések:

Előre, Jobbra/Balra, Megragad,  
Lő, Mászik, ... ágens meghal,  
ha csapdára vagy Wumpusra rálép.

**cél: arany** - megtalálni és kihozni

**Hogyan biztosítható a siker**  
(logikai gondolkodással)?

|              |     |     |     |
|--------------|-----|-----|-----|
| 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 |
| <div>Á</div> |     |     |     |
| OK           | OK  |     |     |

(a)

Á

 = Ágens  
 S = Szellő  
 R = Ragyogás,  
 Arany  
 G = 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  | <div>C?</div> |               |     |
| 1,1 | 2,1           | 3,1           | 4,1 |
| M   | <div>Á</div>  | <div>C?</div> |     |
| OK  | S             |               |     |
|     | OK            |               |     |

(b)

|              |     |     |     |
|--------------|-----|-----|-----|
| 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 |
| <div>Á</div> |     |     |     |
| OK           | OK  |     |     |

(a)

Á

 = Ágens  
 S = Szellő  
 R = Ragyogás,  
 Arany  
 G = 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 |
| <div>OK</div> | <div>C?</div> |               |     |
| 1,1           | 2,1           | 3,1           | 4,1 |
|               | <div>Á</div>  | <div>C?</div> |     |
| M             | S             |               |     |
| OK            | OK            |               |     |

(b)

|                |                     |           |     |
|----------------|---------------------|-----------|-----|
| 1,4            | 2,4                 | 3,4       | 4,4 |
| 1,3<br>W!      | 2,3                 | 3,3       | 4,3 |
| 1,2<br>Á<br>B  | 2,2<br>OK           | 3,2       | 4,2 |
| 1,1<br>M<br>OK | 2,1<br>S<br>M<br>OK | 3,1<br>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<br>C?           | 3,4       | 4,4 |
| 1,3<br>W!           | 2,3<br>Á<br>B R     | 3,3<br>C? | 4,3 |
| 1,2<br>B<br>M<br>OK | 2,2<br>S<br>M<br>OK | 3,2       | 4,2 |
| 1,1<br>M<br>OK      | 2,1<br>S<br>M<br>OK | 3,1<br>C! | 4,1 |

(b)

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

# Logika – reprezentáció és manipuláció eszköze

**Szintaktika** - legálisan létrehozható szimbolikus mondatok

**Szemantika** - a világ tényei, amikre a szimbolikus mondatok vonatkoznak

## Vonzatreláció és a következtetés

**Tények** neveznek valós dolgokat, amelyek egymással kapcsolatban vannak.

Ha az egyik igaz a világban, és a másik szükségszerűen is az:

**vonzat: logikai konzekvencia dolgok között.**

Vonzatreláció tudásbázis TB és egy  $\alpha$  mondat között:

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

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

**Bizonyítás** = a következtetési algoritmus lépéssorozata

$\alpha$  **bizonyítható TB** -ből: **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 ilyen.
- **Az igazság függ mind a mondat interpretációjától, mind a világ aktuális állapotától.**



# Szemantika és interpretáció

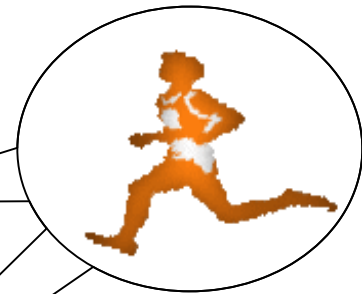
mi lehet pl. a **Fut\_Ágens1** mondat jelentése?

Ha „fut” fizikai mozgás, „Ágens1” egy személy, és a világ most: akkor ez a mondat (most) **igaz**.

Ha „fut” nem terminált működés, „Ágens1” egy program, és a világ most ugyanilyen, akkor ez a mondat (most) **hamis**.

Ha „fut” fizikai mozgás, „Ágens1” egy személy és a világ most: akkor ez a mondat (most) **hamis**.

Ha „fut” nem terminált működés, „Ágens1” egy program, és a világ most ugyanilyen, akkor ez a mondat (most) **igaz**.



# Modellek

Bármely (formális is) világ, ahol egy mondat igaz egy bizonyos interpretációban. Egy mondatnak számos modellje lehet.

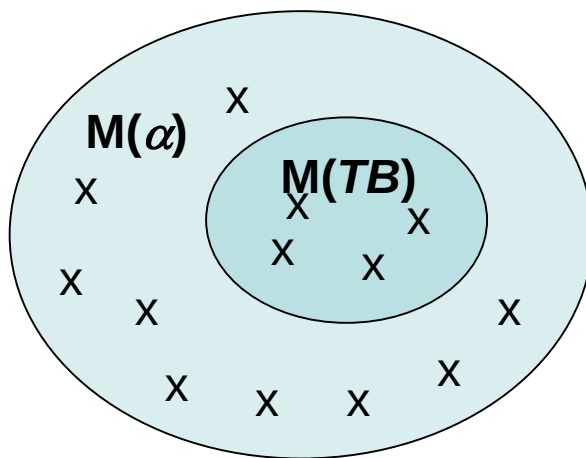
Minél többet állítunk (pl. minél több információt adunk hozzá a tudásbázishoz), annál kevesebb modellünk lesz.

Modellek nagyon fontosak (mert):

**egy  $\alpha$  mondat vonzata a TB tudásbázisnak,  
ha a TB modelljei mind modelljei az  $\alpha$ -nak is.**

Ha ez így van, akkor ha TB igaz, akkor  $\alpha$  is igaz

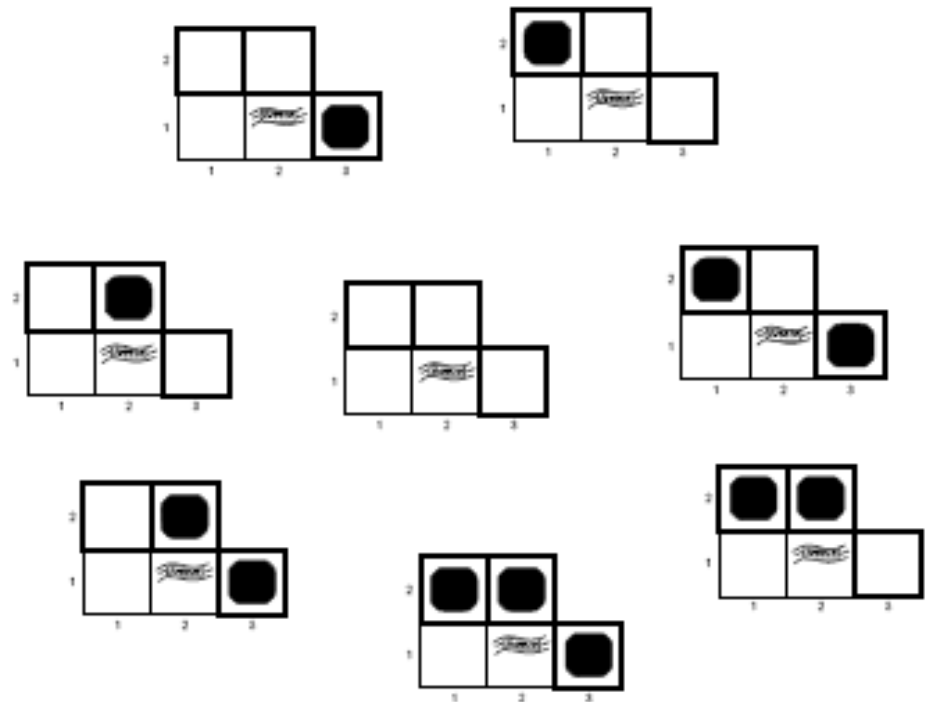
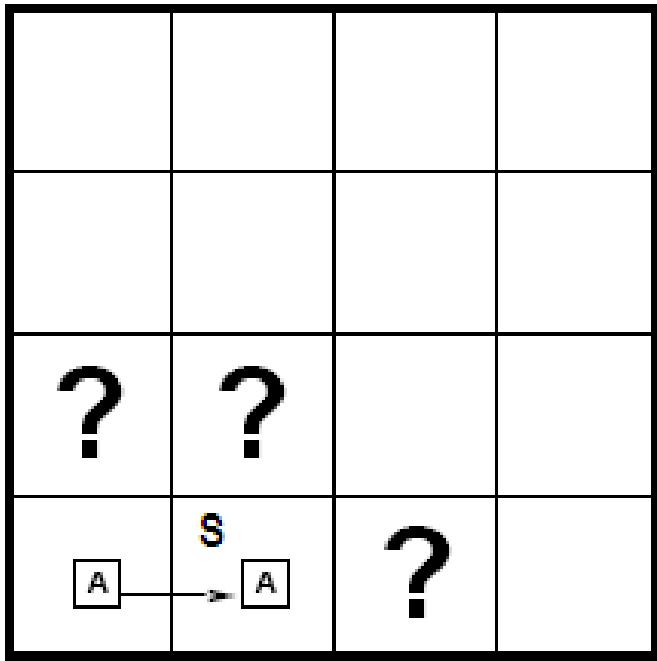
**$TB \models \alpha$  akkor és csak akkor, ha  $M(TB) \subseteq M(\alpha)$**



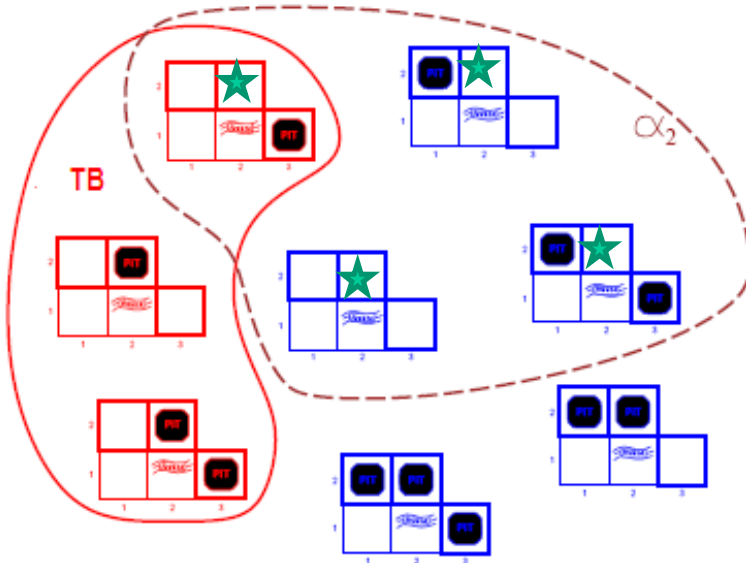
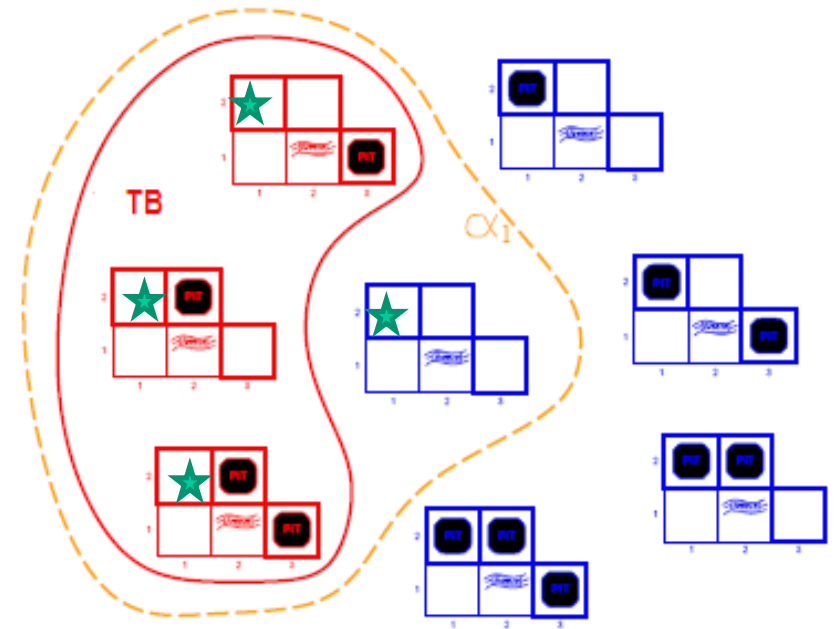
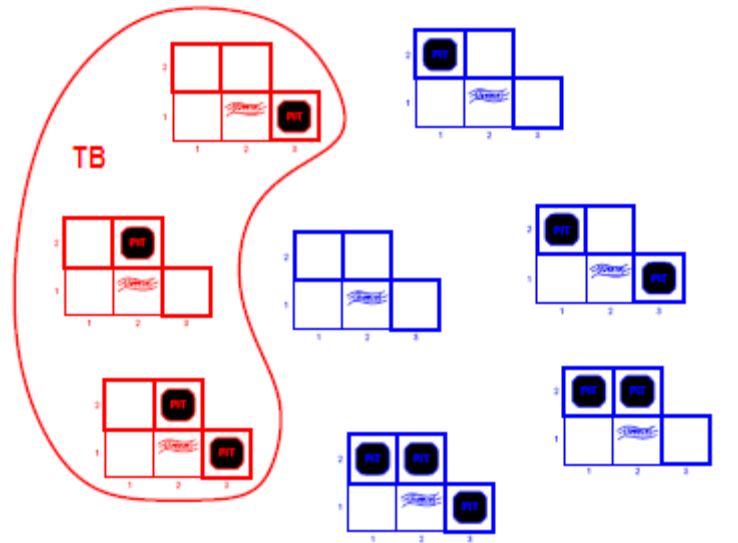
# Vonzat és modellek Wumpus világban

(1,1)-ben semmi érzet, lépés (2,1)-re, szellőt érezni, ...

Lehetséges modellek a (?) -re, csakis csapdát feltételezve:  
3 db bináris változó, 8 db lehetséges modell



# TB = Wumpus világ szabályai + megfigyelések



$\alpha1$  = „[1,2] biztonságos”, ★

**$TB \models \alpha1$** , modellellenőrzéssel

$\alpha2$  = „[2,2] biztonságos”, ★

**$TB \not\models \alpha2$** , modellellenőrzéssel

**Ha adott egy TB, a következtetés:**

- (1) létrehozhat új  $\alpha$  mondatot, amely vonzata a TB-nak
- (2) ha adott  $\alpha$  mondat, eldöntheti hogy  $\alpha$  vonzata-e a TB-nak.

**Igazságtartó** következtetés (extenzionális), 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 ami bebizonyított, az igaz is.  $A \vdash B$  -ből  $A \models B$

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

Pl. „Van bűz az  $[1,1]$ -ben, vagy nincs bűz az  $[1,1]$ -ben”

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

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

Melyik állítás érvényes, kielégíthető, ill. kielégíthetetlen?

# Ítéletkalkulus (Arisztotelész, Euklidész, i.e. 300, Leibnitz, Boole, XVII sz.)

## Szintaktika:

**igaz** és **hamis** logikai konstansok,  $P1, P2, \dots$  ítélet szimbólumok  
ítéletmondatok,  $\neg P, P1 \wedge P2, P1 \vee P2, P1 \rightarrow P2, P1 \Leftrightarrow P2$  is ítéletmondatok.

precedencia:  $\neg, \wedge, \vee, \rightarrow$  és  $\Leftrightarrow$ .

## Szemantika

Minden modell igaz/hamis értéket rendel minden ítéletszimbólumhoz,  
pl.  $P1 = \text{igaz}, P2 = \text{hamis}, P3 = \text{igaz}, \dots$

Tetszőleges mondat szemantikája - rekurzív elemzés, pl.:

$$\neg P1 \wedge (P2 \vee P3) = \text{hamis} \wedge (\text{hamis} \vee \text{igaz}) = \text{hamis} \wedge \text{igaz} = \text{hamis}$$

## Összekötő jelek igazságtáblái

| $P1$ | $P2$ | $\neg P$ | $P1 \wedge P2$ | $P1 \vee P2$ | $P1 \rightarrow P2$ | $P1 \Leftrightarrow P2$ |
|------|------|----------|----------------|--------------|---------------------|-------------------------|
| 0    | 0    | 1        | 0              | 0            | 1                   | 1                       |
| 0    | 1    | 1        | 0              | 1            | 1                   | 0                       |
| 1    | 0    | 0        | 0              | 1            | 0                   | 0                       |
| 1    | 1    | 0        | 1              | 1            | 1                   | 1                       |

N argumentumon értelmezett logikai függvény  $2^{2^N}$



# Az ítéletlogika ekvivalencia szabályai

$$(a \wedge b) \Leftrightarrow (b \wedge a)$$

$\wedge$  kommutativitása

$$(a \vee b) \Leftrightarrow (b \vee a)$$

$\vee$  kommutativitása

$$((a \wedge b) \wedge c) \Leftrightarrow (a \wedge (b \wedge c))$$

$\wedge$  asszociativitása

$$((a \vee b) \vee c) \Leftrightarrow (a \vee (b \vee c))$$

$\vee$  asszociativitása

$$\neg(\neg a) \Leftrightarrow a$$

dupla negálás eliminálása

$$(a \rightarrow b) \Leftrightarrow (\neg b \rightarrow \neg a)$$

kontrapozíció

$$(a \rightarrow b) \Leftrightarrow (\neg a \vee b)$$

**implikáció elimináció**

$$(a \Leftrightarrow b) \Leftrightarrow ((a \rightarrow b) \wedge (b \rightarrow a))$$

ekvivalencia eliminálás

$$\neg(a \wedge b) \Leftrightarrow (\neg a \vee \neg b)$$

De Morgan

$$\neg(a \vee b) \Leftrightarrow (\neg a \wedge \neg b)$$

De Morgan

$$(a \wedge (b \vee c)) \Leftrightarrow ((a \wedge b) \vee (a \wedge c))$$

$\wedge$  disztributivitása  $\vee$  felett

**DNF**

$$(a \vee (b \wedge c)) \Leftrightarrow ((a \vee b) \wedge (a \vee c))$$

$\vee$   $\wedge$  disztributivitása  $\wedge$  felett

**CNF**

két állítás logikailag ekvivalens, ha ugyanazokban a modellekben igaz:

$A \Leftrightarrow B$  a.cs.a, ha  $A \models B$  és  $B \models A$

$TB \models \alpha$  a.cs.a., ha  $(TB \rightarrow \alpha)$  érvényes

$TB \models \alpha$  a.cs.a., ha  $(TB \wedge \neg \alpha)$  kielégíthetetlen (reductio ad absurdum)

# Következtetés (bizonyítások)

## **Modell-ellenőrzés – mondat kielégíthető-e (SAT satisfiability)**

Igazságtábla listázás (exponenciális)

Visszalépéses keresés

Heurisztikus keresés modellek terében (helyes, de nem teljes)  
(lokális kereséssel)

## **Következtetési szabályok** alkalmazása

Új mondatok legális generálása régebbi mondatokból.

## **Bizonyítás:** következtetési szabályok sorozata.

speciális operátorok **megszokott** keresési algoritmusokban.

- keresés általános operátorokkal (természetes dedukció)
- teljes keresés teljes operátorral általános logikában (rezolúció, exp.!) )
- teljes keresés teljes operátorral redukált logikában  
(Horn-klózek, Modus Ponens, redukált komplexitás!)

(más kombinációk nem garantálják a következtetés teljességét)

**Egy történet:** „Ádám, Béla és Csaba közül valaki egy üvegház ablakát törte ki. 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'."

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 a mondanivalójuk:

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

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

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

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

| A | B | C | SA | SB | SC | TB = SA $\wedge$ SB $\wedge$ 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                               |

felsorolós bizonyítás,  
 $2^N$  sor,  $O(2^N)$  komplexitás (NP)

Keressük  $TB = SA \wedge SB \wedge SC =$   
 $\neg B \wedge C \wedge (C \wedge (\neg B \vee \neg A)) \wedge (\neg A \rightarrow \neg C)$   
 modelljét (ahol a TB kielégíthető,  
 azaz igaz lesz)

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

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

$\neg A \rightarrow$  Hamis igazához szükséges,  
 hogy  $\neg A$  legyen hamis, azaz  
**A** legyen **igaz**

(Vajon más kombinációkra is?)

Visszalépéses keresés modellek  
 Terében - Davis-Putnam, stb.

| 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  |

## Visszalépéses keresés

a,  $a = 0$

$C1 : a \vee b$

$C2 : c \vee d$

$C3 : a \vee e \vee f$

$C4 : \neg b \vee \neg f \vee g$

$C5 : \neg f \vee h$

$C6 : \neg b \vee \neg h \vee i$

$C7 : \neg g \vee \neg i$

$C1 : a \vee b$

$C2 : c \vee d$

$C3 : a \vee e \vee f$

$C4 : \neg b \vee \neg f \vee g$

$C5 : \neg f \vee h$

$C6 : \neg b \vee \neg h \vee i$

$C7 : \neg g \vee \neg i$

$b, b = 1$

$C1 : a \vee b$

$C2 : c \vee d$

$C3 : a \vee e \vee f$

$C4 : \neg b \vee \neg f \vee g$

$C5 : \neg f \vee h$

$C6 : \neg b \vee \neg h \vee i$

$C7 : \neg g \vee \neg i$

c,  $c = 0$

$C1 : a \vee b$

$C2 : c \vee d$

$C3 : a \vee e \vee f$

$C4 : \neg b \vee \neg f \vee g$

$C5 : \neg f \vee h$

$C6 : \neg b \vee \neg h \vee i$

$C7 : \neg g \vee \neg i$

$d, d = 1$

$C1 : a \vee b$

$C2 : c \vee d$

$C3 : a \vee e \vee f$

$C4 : \neg b \vee \neg f \vee g$

$C5 : \neg f \vee h$

$C6 : \neg b \vee \neg h \vee i$

$C7 : \neg g \vee \neg i$

e,  $e = 0$

$C1 : a \vee b$

$C2 : c \vee d$

$C3 : a \vee e \vee f$

$C4 : \neg b \vee \neg f \vee g$

$C5 : \neg f \vee h$

$C6 : \neg b \vee \neg h \vee i$

$C7 : \neg g \vee \neg i$

$f, f = 1$

$C1 : a \vee b$

$C2 : c \vee d$

$C3 : a \vee e \vee f$

$C4 : \neg b \vee \neg f \vee g$

$C5 : \neg f \vee h$

$C6 : \neg b \vee \neg h \vee i$

$C7 : \neg g \vee \neg i$

$g, g = 1$

$C1 : a \vee b$

$C2 : c \vee d$

$C3 : a \vee e \vee f$

$C4 : \neg b \vee \neg f \vee g$

$C5 : \neg f \vee h$

$C6 : \neg b \vee \neg h \vee i$

$C7 : \neg g \vee \neg i$

TB = **klózik halmaza**, klóz egy CNF

$i, i = 0$

$h, h = 1$

ellentmondás

$C1 : a \vee b$

$C1 : a \vee b$

$C2 : c \vee d$

$C2 : c \vee d$

$C3 : a \vee e \vee f$

$C3 : a \vee e \vee f$

$C4 : \neg b \vee \neg f \vee g$

$C4 : \neg b \vee \neg f \vee g$

$C5 : \neg f \vee h$

$C5 : \neg f \vee h$

$C6 : \neg b \vee \neg h \vee i$

**$C6$**  :  $\neg b \vee \neg h \vee i$

$C7 : \neg g \vee \neg i$

$C7 : \neg g \vee \neg i$

Vissza,  $e = 1$ , legközelebbi  
szabad elágazás

# Az ítéletlogika következtetési mintái (operátorok)

## Modus Ponens

(Implikáció eliminálása)

néha elég is

$$A \rightarrow B$$

$$\frac{A}{B}$$

## AND eliminálása

$$\frac{A_1 \wedge A_2 \wedge \dots \wedge A_n}{A_1, A_2, \dots, A_n}$$

gépi világhoz

## AND bevezetése

$$\frac{A_1, A_2, \dots, A_n}{A_1 \wedge A_2 \wedge \dots \wedge A_n}$$

TB = 1 állítás!

## OR bevezetése

$$\frac{A_i}{A_1 \vee A_2 \vee \dots \vee A_n}$$

## Dupla negálás eliminálása

$$\frac{\neg \neg A}{A}$$

## Elemi (egység)rezolúció

$$\frac{A \vee B \quad \neg B}{A}$$

## Rezolúció

$$\frac{B \vee A \quad \neg B \vee G}{A \vee G}$$

# Logikai következtetések fajtái

**Tudás átalakítás  
(manipulálás) modellje**

## **Dedukció** (már az ógörögök)

formálisan érvényes következtetés  
olyan tények származtatása, amelyek a premisszákból mindenképpen  
következnek (un. 'igazságtartó' eljárás)

|                                  |                          |
|----------------------------------|--------------------------|
| Ha kutya nagy, akkor sokat eszik | (feltétel)               |
| <u>kutya nagy</u>                | (tapasztalat, evidencia) |
| sokat eszik                      | (következtetés)          |

**Tömörítés, általánosítás:  
példákból tanulás modellje!**

## **Indukció** (középkori arabok, skolasztikusok)

|                                               |                          |
|-----------------------------------------------|--------------------------|
| Kutya1 nagy, Kutya2 nagy, ..., Kutya1000 nagy | (tapasztalat)            |
| Minden kutya nagy                             | (induktív általánosítás) |

miért fontos? sok tény helyett egyetlen egy (új) tény  
probléma mérete csökken,  
exponenciális jelleg kevésbé zavaró,  
de formálisan nem igaz! Miért?



# Abdukció - 'belátás' folyamata (Charles Pierce, 1839–1914 )

ha kutya nagy, akkor sokat eszik (feltétel)  
sokat eszik (tapasztalat)  
kutya nagy (hipotézis)

**Diagnózis,  
magyarázatadás  
modellje!**

ha kutya nagy, akkor sokat eszik

ha kutya éhes, akkor sokat eszik

sokat eszik (tapasztalat)

éhes, vagy nagy?

(melyik hipotézis esélyesebb)

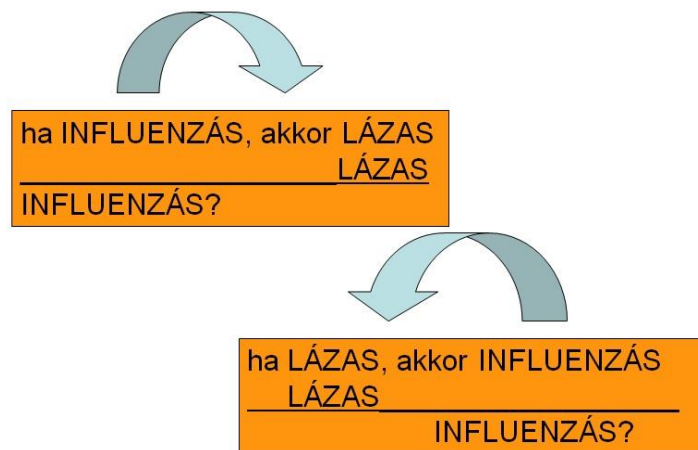
formálisan miért nem igaz? de miért nagyon fontos?

Természetes rendszermodell  
(a feladat megfogalmazása)

ha 'egy rendszer X állapotban van',  
akkor 'Y a rendszer megfigyelhető viselkedése'

Természetes kérdés (feladat megoldása)  
milyen állapotban van a rendszer, ha konkrét  
információt (Y) kapok a viselkedéséről?

Tudjuk, hogy a beteg lázas.  
Dedukció, vagy abdukció?



## 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**.

Cook (1971): **egy mondathalmaz kielégíthetőségi vizsgálata NP-teljes** (de nem minden ítéletlogikai következtetés  $O(2^n)$  időt igényel! Ld. előbb)

# 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.

ha  $TB_1 \models a$  , 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.

$\Sigma$ : amíg a probléma (a szükséges absztrakció szintjén) tökéletesen statikus, a monotonitással nyerünk, különben ráfizethetünk.

## Horn klózok:

$$P_1 \wedge P_2 \dots \wedge \dots P_n \rightarrow Q$$
$$\neg P_1 \vee \neg P_2 \dots \vee \dots \neg P_n \vee Q$$

- Legfeljebb egy pozitív literál, pontosan egy pozitív literál = határozott klóz
- Egy hasznos mondat típus, amelyre létezik a TB méretében **lineáris idejű** következtetési eljárás.

Két fontos speciális eset:

$$P_1 \wedge P_2 \wedge \dots P_n \rightarrow \text{Hamis, ua., mint } \neg P_i \vee \dots \neg P_n$$
$$\text{Igaz} \rightarrow Q \quad \text{ua., mint } Q.$$

Nem minden TB-beli állítás írható fel Horn klózként, de azoknál, amelyeknél ez megtehető: alkalmazzuk Modus Ponens-t, ameddig marad alkalmazható következtetés.

Modus Ponens **teljes** bizonyítási lépés a **Horn klózok** tudásbázisában!

# Előrecsatolt, hátracsatolt (Modus Ponens) következtetés

$$P \Rightarrow Q$$

$$L \wedge M \Rightarrow P$$

$$B \wedge L \Rightarrow M$$

$$A \wedge P \Rightarrow L$$

$$A \wedge B \Rightarrow L$$

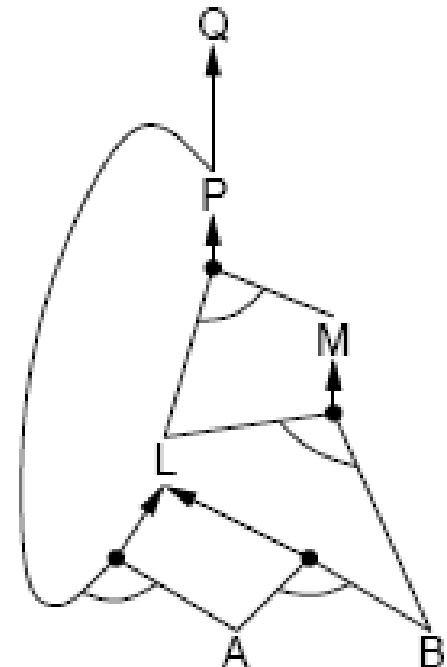
$A$

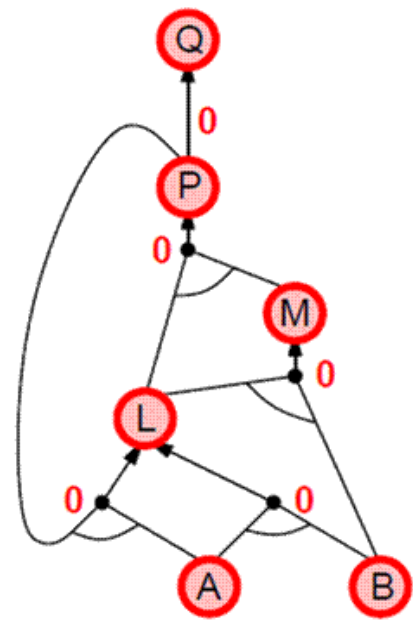
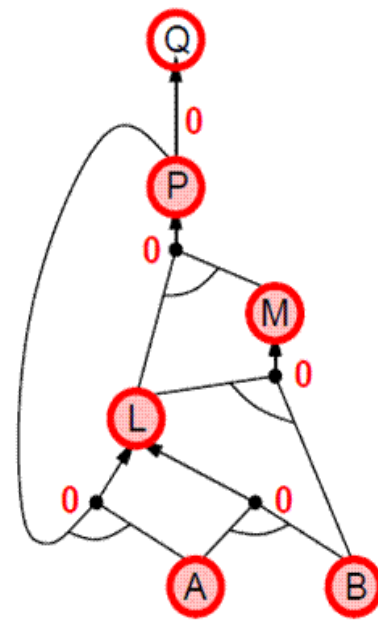
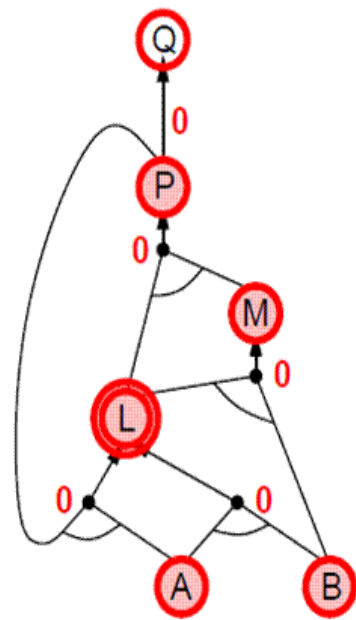
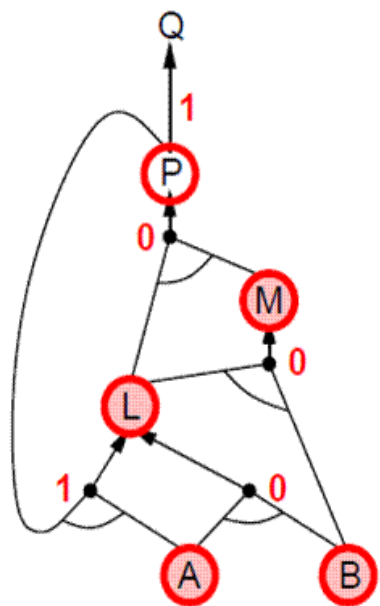
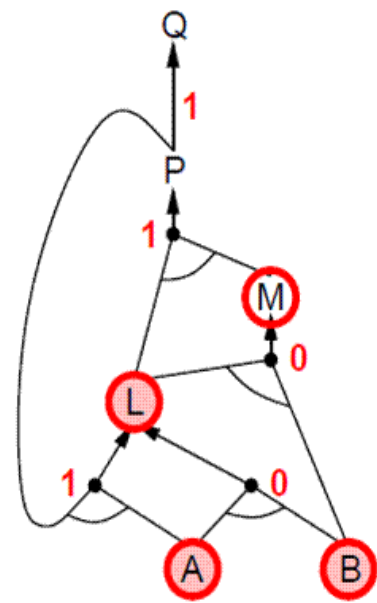
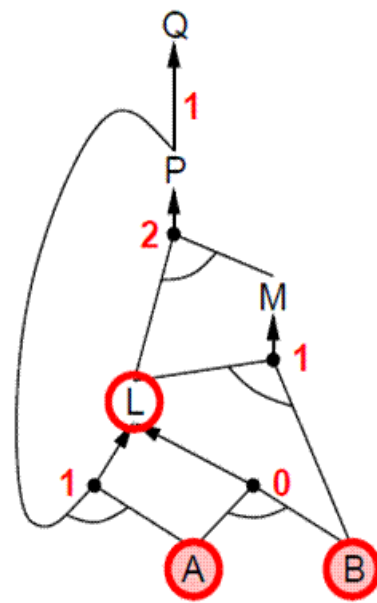
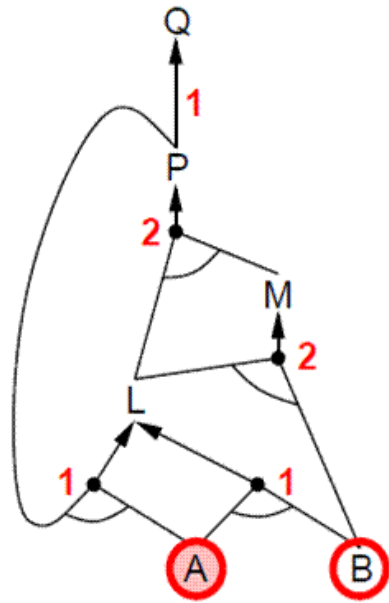
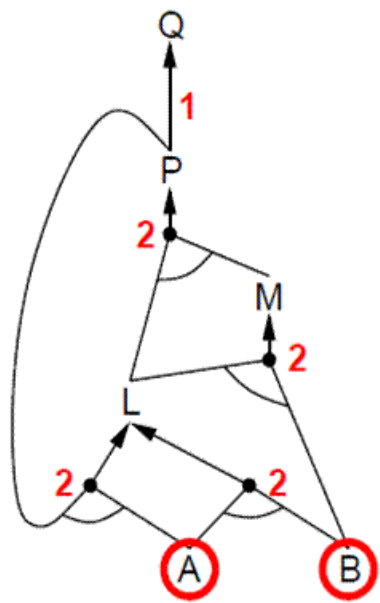
$B$

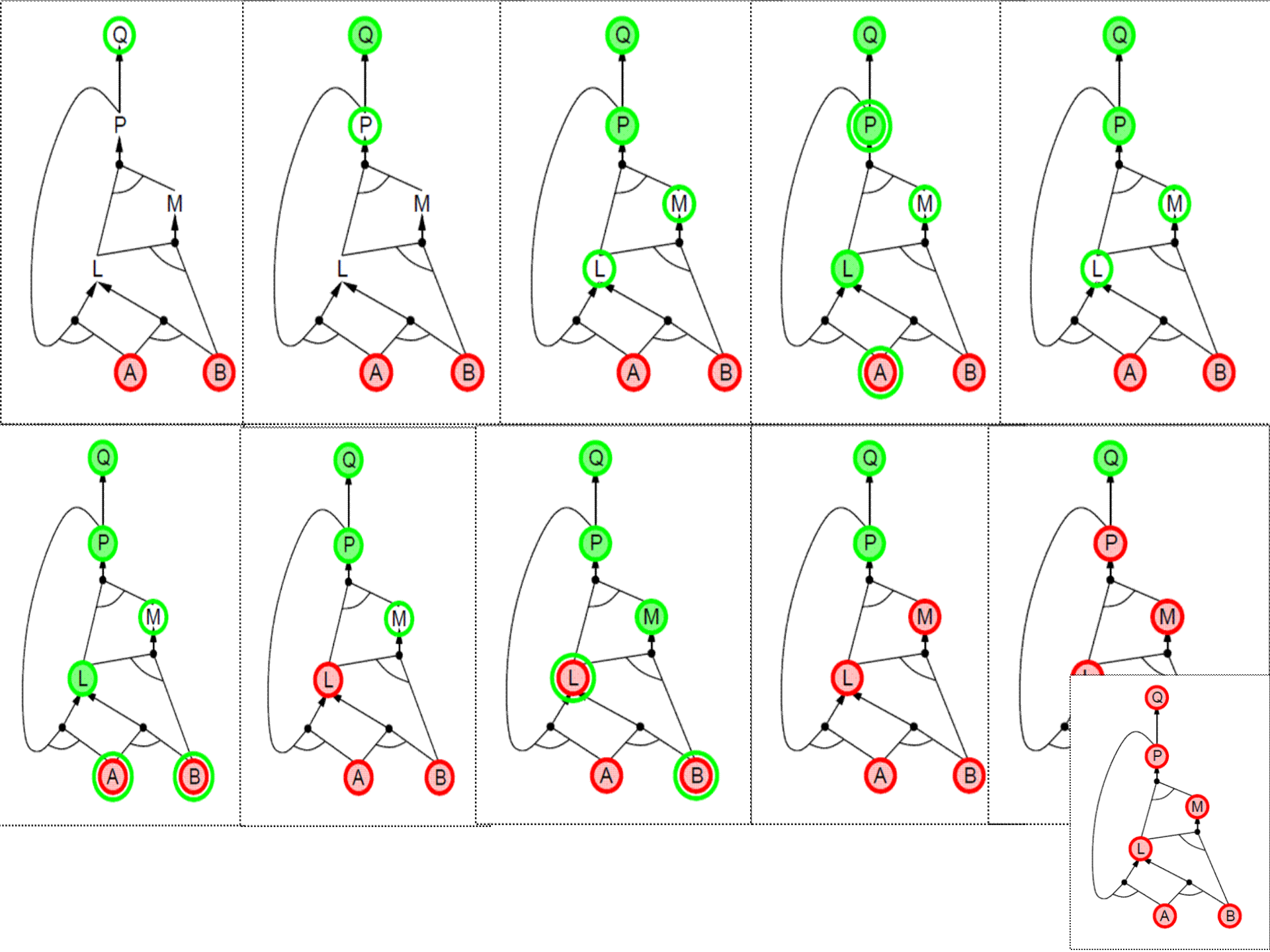
**ECs:** Minden szabályt elsütni, melynek premisszája teljesül.  
Következmény hozzáadása a TB-hoz, amíg a lekérdezendő változó értéket nem kap (fixpont).

**HCS:** A lekérdezéstől visszafelé:

- ellenőrzés, netán a lekérdezés már igaz,
- elővenni egy szabályt, melynek következménye a lekérdezés és a premisszáit bizonyítani rekurzív Hcs-sal.







# Egy ágens a Wumpus világ számára

**Tudásbázis ciklus:** ágens érzetei  $\rightarrow$  mondatok  $\rightarrow$  tudásbázis  
további érvényes mondatok érzet mondatok vonzatai.

tegyük fel:

$B_{1,2}$  = "bűz van az [1,2]-ben" (de lehetne B12, vagy XYZ)

$S_{1,2}$  = „szellő van az [1,2]-ben”, stb.

**Tudásbázis:**  $\square S_{1,1}$  ,  $\square B_{1,1}$  ,  $S_{2,1}$  ,  $\square B_{2,1}$  ,  $\square S_{1,2}$  ,  $B_{1,2}$  , ... stb.

amit az ágens tud a környezetéről általánosságban, pl.:

ha valahol nincs bűz, akkor sem ott, sem szomszédban nincs Wumpus.

**R<sub>1</sub>:**  $\neg B_{1,1} \rightarrow \neg W_{1,1} \wedge \neg W_{1,2} \wedge \neg W_{2,1}$

**R<sub>2</sub>:**  $\neg B_{2,1} \rightarrow \neg W_{1,1} \wedge \neg W_{2,1} \wedge \neg W_{2,2} \wedge \neg W_{3,1}$

**R<sub>3</sub>:**  $\neg B_{1,2} \rightarrow \neg W_{1,1} \wedge \neg W_{1,2} \wedge \neg W_{2,2} \wedge \neg W_{1,3}$

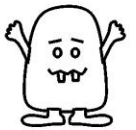
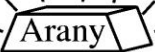
ha bűz van az [1,2]-ben, akkor egy Wumpus van ott, vagy egy vagy több szomszédos négyzetben:

**R<sub>4</sub>:**  $B_{1,2} \rightarrow W_{1,3} \vee W_{1,2} \vee W_{2,2} \vee W_{1,1}$

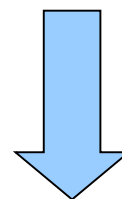


**A Wumpus megtalálása** - az ágensnek el kellene készítenie a  $TB \rightarrow W_{1,3}$  igazságtábláját, hogy megmutassa, hogy ez a mondat **érvényes** (mert akkor a Wumpus(1,3) jelenléte a TB (= ágens tudása) vonzata!).

kb. 12 ítélet szimbólum van: az igazságtáblának  $2^{12} = 4096$  sora lesz, minden sorban, amelyben a TB mondat igaz, a  $W_{1,3}$ -nak is igaznak kellene lennie. Alkalmazzunk inkább a következtetési mintákat! (de természetesen dedukcióval)

|   |                                                                                              |                                                                                                                                                                                                                                                                                       |                                                                                               |                                                                                               |
|---|----------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------|
| 4 | <br>Bűz     |                                                                                                                                                                                                                                                                                       | <br>Szellő   | <br>CSAPDA   |
| 3 |             | <br>Szellő<br><br>Bűz<br><br>Arany | <br>CSAPDA   | <br>Szellő   |
| 2 | <br>Bűz    |                                                                                                                                                                                                                                                                                       | <br>Szellő |                                                                                               |
| 1 | <br>START | <br>Szellő                                                                                                                                                                                         | <br>CSAPDA | <br>Szellő |
|   | 1                                                                                            | 2                                                                                                                                                                                                                                                                                     | 3                                                                                             | 4                                                                                             |

$\square S_{1,1} \quad \square B_{1,1} \quad S_{2,1} \quad \square B_{2,1} \quad \square S_{1,2} \quad B_{1,2}, \dots$



$W_{1,3}$

1. MP:  $\neg B_{1,1}$  + R1-re:  $\neg W_{1,1} \wedge \neg W_{1,2} \wedge \neg W_{2,1}$

$R_1$ :  $\neg B_{1,1} \rightarrow \neg W_{1,1} \wedge \neg W_{1,2} \wedge \neg W_{2,1}$

**Modus Ponens**

(Implikáció eliminálása)

$A \rightarrow B$

$\frac{A}{B}$

$\frac{\neg B_{1,1} \rightarrow \neg W_{1,1} \wedge \neg W_{1,2} \wedge \neg W_{2,1} \quad \neg B_{1,1}}{\neg W_{1,1} \wedge \neg W_{1,2} \wedge \neg W_{2,1}}$

2. És-eliminálás:  $\neg W_{1,1}$ ,  $\neg W_{1,2}$ ,  $\neg W_{2,1}$

3. MP:  $\neg B_{2,1}$  + R2-re, utána és-elimináció:  $\neg W_{2,2}$ ,  $\neg W_{2,1}$ ,  $\neg W_{3,1}$

**R<sub>2</sub>:**  $\neg B_{2,1} \rightarrow \neg W_{1,1} \wedge \neg W_{2,1} \wedge \neg W_{2,2} \wedge \neg W_{3,1}$

## Modus Ponens

(Implikáció eliminálása)

$$\frac{A \rightarrow B \quad A}{B}$$

$$\frac{\neg B_{2,1} \rightarrow \neg W_{1,1} \wedge \neg W_{2,1} \wedge \neg W_{2,2} \wedge \neg W_{3,1} \quad \neg B_{2,1}}{\neg W_{1,1} \wedge \neg W_{2,1} \wedge \neg W_{2,2} \wedge \neg W_{3,1}}$$

4. MP:  $B_{1,2}$ -re és R4-re:  $W_{1,3} \vee W_{1,2} \vee W_{2,2} \vee W_{1,1}$

**R<sub>4</sub>:**  $B_{1,2} \rightarrow W_{1,3} \vee W_{1,2} \vee W_{2,2} \vee W_{1,1}$

**Modus Ponens**

(Implikáció eliminálása)

$A \rightarrow B$

$\frac{A}{B}$

$B_{1,2} \rightarrow W_{1,3} \vee W_{1,2} \vee W_{2,2} \vee W_{1,1}$

$\frac{B_{1,2}}{W_{1,3} \vee W_{1,2} \vee W_{2,2} \vee W_{1,1}}$

5. Egység rezolúció:  $W_{1,3} \vee W_{1,2} \vee W_{2,2} \vee W_{1,1}$  és  $\neg W_{1,1}$

**Elemi (egység)rezolúció**

$$\frac{A \vee B \quad \neg B}{A}$$

$$\frac{W_{1,3} \vee W_{1,2} \vee W_{2,2} \vee W_{1,1} \quad \neg W_{1,1}}{W_{1,3} \vee W_{1,2} \vee W_{2,2}}$$

az eredmény:  $W_{1,3} \vee W_{1,2} \vee W_{2,2}$

6. Egység rezolúció:  $W_{1,3} \vee W_{1,2} \vee W_{2,2}$  és  $\neg W_{2,2}$   
és az eredmény:  $W_{1,3} \vee W_{1,2}$

7. Végül, még egy rezolúció:  $W_{1,3} \vee W_{1,2}$  és  $\neg W_{1,2}$   
és az eredmény:  $W_{1,3}$   
azaz a Wumpus tényleg az [1,3]-ban van.

# Bizonyítás = Keresés

## Keresés kezdeti TB állapota

