

Mesterséges intelligencia - 5. előadás

Logikai ágens, szabályalapú rendszerek

Történelem

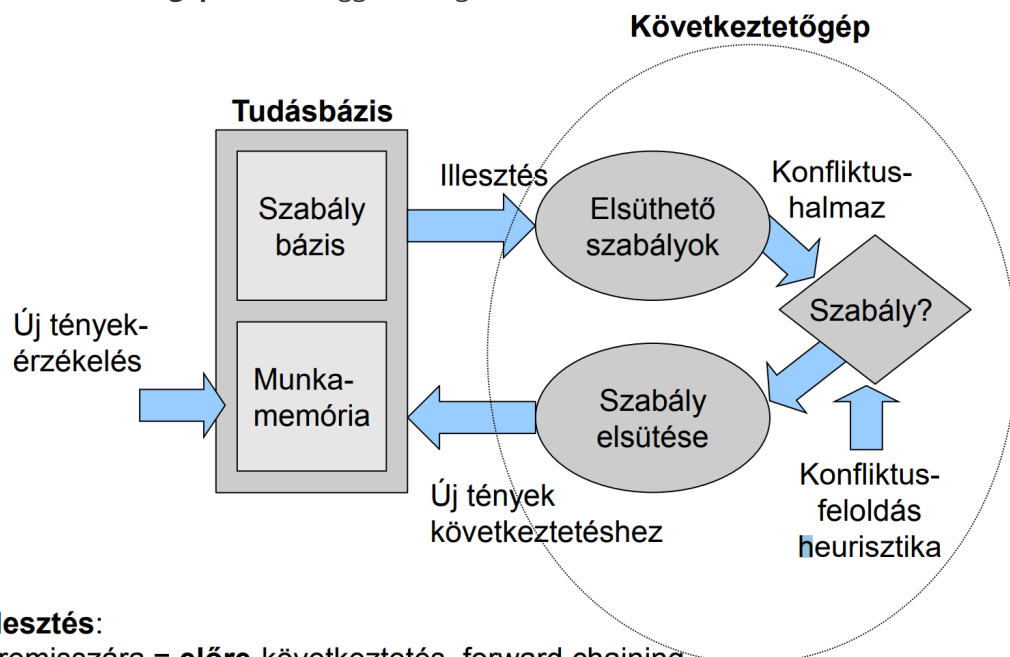
- Mycin → szakorvosi szintű első szabályalapú rendszer → tünetek alapján betegség és kezelés javaslat

Logika szerepe a komplex MI alkalmazásokban

- logikai szabályalapú következtetés és beavatkozás
 - egy vagy több feltétel logikai állításként történő megfogalmazása
 - reakciók, beavatkozások megadása logikai szabályok formájában
- IBM Watson (2011): Jeopardy kvízzjáték → beszédértés, következtetés, játék

Logika

- 2 része van
 - **tudás bázis:** feladatspecifikus információ (elméletek + észrevételek)
 - **következtető gép:** feladatfüggetlen algoritmusok



Illesztés:

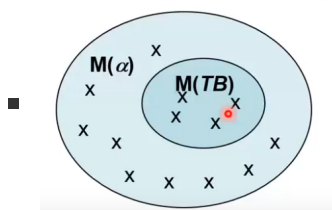
Premisszára = **előre**-következtetés, forward chaining

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

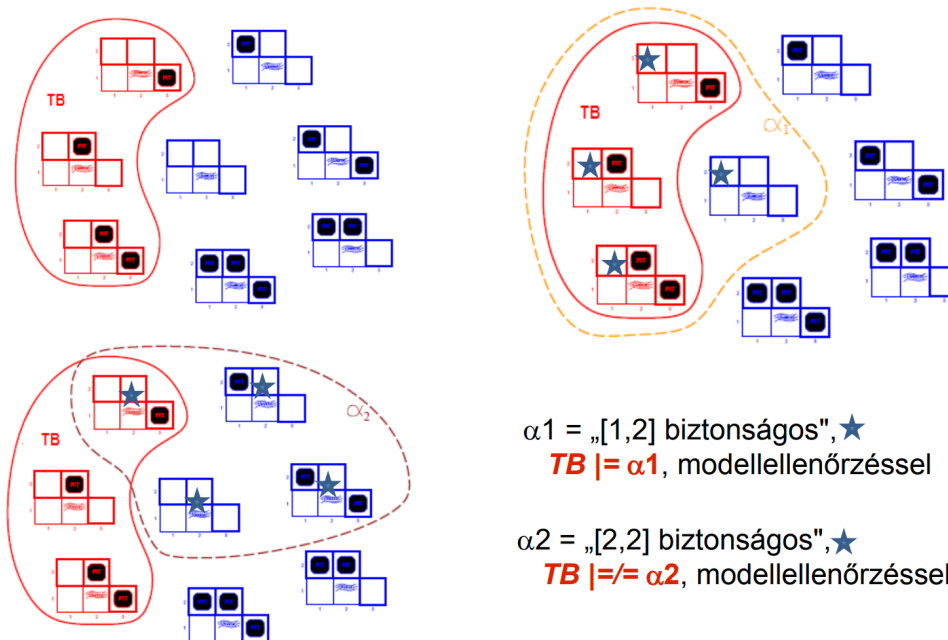
- példák ilyen rendszerekre: jogi szakértői rendszerek, ügyfélszolgálat, csetbot, stb
- előnyei:
 - látjuk a levezetést is! (hogy jutottunk pl a diagnózisig?)

Reprezentáció és manipuláció eszköze

- **szintaktika:** legálisan létrehozható szimbolikus mondatok
- **szemantika:** a világ tényei, amikre a szimbolikus mondatok vonatkoznak
- **tények:** valós dolgok a világban, amik egymással kapcsolatban vannak
 - **vonzat:** logikai konzekvencia dolgok között
 - vonzatreláció $\rightarrow \alpha$ vonzata TB-nek: **$TB \models \alpha$** , ha α minden olyan világban igaz, ahol TB is igaz.
- **következtetés:** a vonzat kiszámítása a mondatok formális manipulálásával
- **bizonyítás:** a következtetési algoritmus lépéssorozata
 - α bizonyítható TB-ből: **$TB \vdash \alpha$**
- szintaktika és szemantika összekapcsolódása $\rightarrow$ **interpretáció** kell hozzá
 - az igazság függ mind a mondat interpretációjától, mind a világ aktuális állapotától
- **modell**
 - bármely világ, ahol a mondat igaz egy bizonyos interpretációban
 - minél több információt tudunk $\rightarrow$ annál kevesebb modellünk lesz
 - **egy α mondat vonzata a TB tudásnak, ha a TB modelljei mind modelljei az α -nak is**
 - ekkor ha TB igaz, akkor α is igaz, azaz
 - **$TB \models \alpha$, akkor és csak akkor, ha $M(TB) \subseteq M(\alpha)$**



- példa:



- Ha adott egy TB, egy a következtetés
 1. létrehozhat új α mondatot, amely vonzata a TB-nak (azaz ha kikövetkeztetem, hogy [1;2] egy biztonságos mező, akkor abból létrejön egy új α mondat, ami az adott tudásbázisból következik, tehát annak vonzata)
 2. ha adott α mondat, eldöntheti hogy α vonzata-e a TB-nak (azaz ha kikövetkeztetem, hogy [1;2] egy biztonságos mező, akkor abból egy adott α mondatról eltudom

dönteni, hogy vonzata-e a tudásbázisomnak)

- **igazságtartó következtetés** (*extenzionális*), ha csak igazságfüggvényekkel dolgozik

- (igazságfüggvény logikai értéke → csak az igazságfüggvény részeitől és logikai műveletek definíciójától függ, az interpretációtól nem!)

- **formális bizonyítás** → igazságtartó következtetési eljárás lépései

- **következtetési eljárás**

- **teljes**: ha minden vonzatmondathoz talál egy bizonyítást (azaz ami igaz az bebizonyítható)

- $A \models B$ -ből $A \models \neg B$

- **helyes**: ha minden bizonyított mondat vonzatrelációban áll a felhasznált tényekkel (azaz ami bebizonyított, az igaz is)

- $A \models \neg B$ -ből $A \models B$

- **értvényes** (analitikus mondat, tautológia): ha minden világban minden lehetséges interpretációja igaz, függetlenül attól, hogy mit akar jelenteni és mi a világ állapota (pl van fal vagy nincs fal)

- **kiegészíthető**: ha létezik olyan interpretáció, hogy valamely világban igaz. Különben **kiegészíthetetlen**.

- példa:

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

- itt az első értvényes (minden világban 1), a második értvénytelen (B és Bnegált miatt), a harmadik kielégíthető (néhány világban teljesül, pl ahol $A=1$ és $B=1$)

- **Ítéletkalkulus**

- igaz és hamis logikai konstansok, precedencia: $\neg, \wedge, \vee, \rightarrow, \leftrightarrow$

- szemantika: minden modell i/h értéket rendel minden ítélet szimbólumhoz

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 a 2^N -ediken → igazságtáblás kiértékelés nem nagyon lesz használható

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
$(a \vee (b \wedge c)) \Leftrightarrow ((a \vee b) \wedge (a \vee c))$	$\vee$ $\wedge$ disztributivitása $\wedge$ felett
DNF	
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)

- Két állítás logikailag ekvivalens, ha ugyanazokban a modellekben igaz, azaz $A \Leftrightarrow B$
akkor és csak akkor, ha $A \models B$ és $B \models A$
- TB $\models \alpha$ akkor és csak akkor, ha $(TB \rightarrow \alpha)$ érvényes
- TB $\models \alpha$ akkor és csak akkor, ha $(TB \wedge \neg \alpha)$ kielégíthetetlen
- **következtetés módjai**
 - modell-ellenőrzés (a mondat kielégíthető-e)
 - igazságtábla listázás
 - visszalépéses keresés
 - heurisztikus keresés modellek terében (helyes de nem teljes)
 - következtetési szabályok alkalmazása $\rightarrow$ új mondatok legális generálása régebbi mondatokból
 - bizonyításkor **speciális operátorok**
 - keresés általános operátorokkal (természetes dedukció)
 - teljes keresés teljes operátorral általános logikában (rezolúció, exponenciális!)
 - teljes keresés teljes operátorral redukált logikában (Horn-klózok, Modus Ponens, redukált komplexitás!)
- **ítéletlogika következtetési mintái**
 - **modus ponens** (implikáció eliminálása): $A \rightarrow B$ helyett B
 - AND eliminálása: $A_1 \wedge A_2 \wedge A_3 \wedge \dots$ helyett $A_1, A_2, A_3, \dots$
 - AND bevezetése: $A_1, A_2, A_3, \dots$ helyett $A_1 \wedge A_2 \wedge A_3 \wedge \dots$ (TB = 1 állítás)
 - OR bevezetése: $A_1, A_2, A_3, \dots$ helyett $A_1 \vee A_2 \vee A_3 \vee \dots$
 - Dupla negálás eliminálása: $\neg\neg A$ helyett A
 - elemi rezolúció: $A \vee B$ helyett ha Bről tudjuk, hogy negált, akkor A biztosan ponált
 - rezolúció: Ha szerepel B $\vee$ A és $\neg B \vee G$, akkor a kettő helyett $A \vee B$ használható (ha van 2 diszjunktív állításunk, és pontosan az egyik elemre igaz, hogy az egyikben ponáltan a másokban negáltan szerepel, akkor kikövetkeztethetjük, hogy a két állításban a rezolúcióban nem érintett részek maradnak csak meg)
- **logikai következtetések fajtái**
 - **dedukció**: formálisan érvényes következtetés, tudás átalakítás modellje

- pl feltétel: ha a kutya nagy, akkor sokat eszik. tapasztalat: a kutya nagy.
következtetés: sokat eszik.
- **indukció:** ezek nem formálisan igazak, tömörítés, általánosítás: példákból tanulás modellje
 - pl tapasztalat: kutya1 nagy, kutya2 nagy, ..., kutya1000 nagy. induktív általánosítás: minden kutya nagy
 - formálisan NEM igaz!
- **addukció:** belátás folyamata, diagnózis, magyarázatadás modellje
 - pl feltétel: ha a kutya nagy, sokat eszik. tapasztalat: a kutya sokat eszik. hipotézis: a kutya nagy
 - formálisan NEM igaz!
- **ítéletlogika eldönthető és teljes**
 - beláthatóak véges algoritmussal → következtetés igazságtábla módszere teljes, DE a számítási idő exponenciális
 - egy mondathalmaz kielégíthetőségi vizsgálata NP-teljes
- **ítéletlogika monoton**
 - új mondat TB-hoz hozzáadásakor az eredeti TB mondatok továbbra is vonzatai maradnak az új nagyobb TB-nak
 - ha $TB_1 \models a$, akkor $(TB_1 \cup TB_2) \models a$
 - statikus világ → akkor ez nagyon jó
 - dinamikus világ → valahogy frissíteni kell, hogy továbbra is igazak e a megszerzett tudások
- **Horn klózek**
 - ha ilyen formájú: $P_1 \wedge P_2 \wedge P_3 \wedge \dots \wedge P_n \rightarrow Q$ vagy $\neg P_1 \vee \neg P_2 \vee \neg P_3 \vee \dots \vee \neg P_n \vee Q$
 - ekkor létezik a TB méretében lineáris idejű következtetési eljárás
 - nem minden TB-beli állítás írható fel így, de ami igen → Modus Ponens-t tudunk alkalmazni
 - Modus Ponens **teljes** bizonyítási lépés a **Horn klózek** tudásbázisában
- **előrecsatolt következtetés**
 - minden olyan szabályt elsütünk, aminek premisszája teljesül, a következményt a TB-hez adjuk, amíg a kérdéses változó értéket nem kap
- **hátracsatolt következtetés**
 - kérdéses változótól visszafele lépkedünk, amíg nem találunk olyat amik igazak, és onnan mehetünk felfele.
- **Wumpus következtetés**
 - TODO feladta házinak a jófej bácsi :) 36. diától