

Mesterséges intelligencia - 6. előadás

Logikai ágens, elsőrendű logika

Rezolúció

- Egy Q-ról akarjuk eldönteni, hogy igaz e vagy hamis.
- lépései:
 - TB átírása klóz formába
 - Q kérdés negálása, negált Q kérdés átírása klóz formára
 - kiterjesztés $TB' = TB \cup \neg Q$
 - rezolúciós lépés ciklikus elvégzése $\rightarrow$ üres rezolvenst kapunk
 - tehát TB' egy ellentmondást tartalmaz, és mivel TB igaz volt biztosan $\neg Q$ miatt romlott el, azaz Q igaz kell, hogy legyen!
- **klóz formára hozás**
 1. ekvivalencia elhagyása: $A \Leftrightarrow B = (A \rightarrow B) \wedge (B \rightarrow A)$
 2. implikáció elhagyása: $A \rightarrow B = \neg A \vee B$
 3. negálás atomi formulák szintjére (de Morgan): $\neg(A \vee B) = \neg A \wedge \neg B$
 4. diszjunkciók literálok szintjére (disztributivitás): $(A \wedge B) \vee C = (A \vee C) \wedge (B \vee C)$
 5. konjunkciók elhagyása (diszjunktív klózekra bontás): csak $\neg$ és $\vee$ maradhat, azaz az és-ek helyett külön mondattá szedjük őket
- probléma: túl sok ítéletszimbólumot kell kezelni, nehézkes a változások kezelése $\rightarrow$ lassú a következtetési eljárás
 - ahol jó: formális modell-ell, MI tervkészítés, auto tesztminta-generálás, szoftver-verifikáció, stb

Elsőrendű logika

- **objektumok:** más obj-tól megkülönböztethető identitású, **tulajdonságokkal** rendelkező dolog
- **relációk:** kapcsolatok, lehet n-elemű vagy egyelemű (ezek a tulajdonságok)
- **függvény jellegű relációk:** egy érték egy adott bemenetre (pl: apja, egyel több mint)
- szintaktika és szemantika
 - **konstans szimbólumok:** X, János, K21
 - **predikátum szimbólumok:** kerek(X), bátyja(József, Anna)
 - **függvény szimbólumok:** János = apja(Béla), apja(János, Béla) = Igaz
- építőelemek:
 - **term:** egy objektumra vonatkozó kifejezés (János, bal-lába(apja(János)), ...)
 - **atom:** bátyja(Géza, Árpád)
 - **atomi mondat:** predikátum szimbólum és az argumentumait jelentő termek **tényeket fejez ki** (igazságértéke van)
 - **összetett mondat:** bátyja(Misi, János) $\rightarrow$ házas(apja(Misi), anyja(János))
 - **kvantorok:**
 - **univerzális kvantor:** $\forall (\forall x. macska(x) \rightarrow emlés(x))$
 - **egzisztenciális kvantor:** $\exists (\exists x. magyka(x) \wedge úszik(x))$

- egyenlősség: (term1 = term2) adott implementáció mellett és akkor és csak akkor, ha term1 és term2 ugyanarra az objektumra vonatkozik
- $\forall$ és $\exists$ kapcsolata:
 - $\forall x \neg \text{szeret}(x, \text{Répa}) = \neg \exists x \text{ szeret}(x, \text{Répa})$
 - $\forall x \text{ szeret}(x, \text{Fagylalt}) = \neg \exists x \neg \text{szeret}(x, \text{Fagylalt})$
- fontos: mivel mi adjuk a konstansok, predikátumok és függvényszimbólumok neveit → fontos, hogy ha kettő kettő ugyan azt jelenti, azt logikailag is leírjuk! (ne csak mi tudjuk, hogy a nagytestű és dőgnagy ugyanaz!)

Modell alapú következtetés?

- probléma:
 - összes modell felsorolása nagyon sok, DE megszámlálható
 - függvények egymásba ágyazhatók végtelen mélységben! (apja-fia végtelen ismétlése)
 - megoldás1: redukció ítéletlogikai következtetésre
 - megoldás2: következtetési lépések kiterjesztése
- **redukció ítéletlogikai következtetésre** → grounding
 - változó nélküli term → **alapterm (grounding)**
 - változó nélküli predikátum kalkulus → **ítélet kalkulus**
 - új TB = minden univerzális mondat minden lehetséges példányosítása
 - egy alapmondat vonzata az új TB-nek, ha vonzata volt az eredeti TB-nek is
 - függvény = végtelen sok alapterm → határt kell húzni nekik
 - **teljes**: minden igaz állítás belátható
 - **vonzat csak félig eldönthető**: állítás hamis volta nem mutatható ki
 - bizonyítási eljárásnak igaz állítás esetén **van kilépési pontja**
 - hamis állítás esetén **nincs**, közben **nem dönthető el**
 - AZAZ gépen véges erőforrások miatt elvben nem mutatható ki egy állításhalmaz konzisztenciája
- **következtetési lépések kiterjesztése**
 - univerzális kvantor eliminálása
 - $$\frac{A \quad \frac{A \rightarrow B}{B} \quad p(A) \quad \forall x, p(x) \rightarrow q(x)}{q(A)}$$
 - illesztés = unifikálás (mit mivel) + behelyettesítés (konkrét érték)
 - egzisztenciális kvantor eliminálása → skolemizálás
 - $$\frac{\exists x q(x, A)}{q(B_s, A)}$$
 - B_s = Skolem konstans, a feladatban önálló léttel NEM rendelkező objektum
 - lehet függvény is! (ami minden x-hez egy külön skolem konstanst rendel hozzá)
 - $$\frac{\forall \text{ember} \exists \text{szív. van-szíve}(\text{ember}, \text{szív})}{\forall \text{ember van-szíve}(\text{ember}, f_{\text{szív}}(\text{ember}))}$$
 - szükség van az $f_{\text{szív}}$ Skolem függvényre, ami megadja az adott ember szívét
 - egyesítés megy, ha:

x	y	eredmény
változó-1	változó-2	✓
változó	konstans	✓
konstans	változó	✓
konstans-1	konstans-2	✗
változó	f(változó)	✗

- gépi bizonyítás
 - Modus Ponens alapú biz. $\rightarrow$ NEM teljes
 - elsőrendű logikai biz. $\rightarrow$ teljes
 - algorithmizált "hogyan" $\rightarrow$ **rezolúció** $\rightarrow$ DE a félig eldönthetőség megmarad
 - transzformáció klóz formára**
 - ekvivalencia elhagyása: $A \Leftrightarrow B = (A \rightarrow B) \wedge (B \rightarrow A)$
 - implikáció elhagyása: $A \rightarrow B = \neg A \vee B$
 - negálás atomi formulák szintjére (de Morgan): $\neg(A \vee B) = \neg A \wedge \neg B$, $\neg \forall x p(x) = \exists x \neg p(x)$
 - Egzisztenciális kvantorokat eltüntetni: Skolemizálás**
 - Ha szükséges, a változókat átnevezni: $\forall x p(x) \vee \forall x q(x)$ helyett $\forall x p(x) \vee \forall y q(y)$**
 - Univerzális kvantorokat balra kihejezni: $\dots \forall x \dots \forall y \dots = \forall x \forall y \dots x \dots y$**
 - diszjunkciók literálok szintjére (disztributivitás): $(A \wedge B) \vee C = (A \vee C) \wedge (B \vee C)$
 - konjunkciók elhagyása (diszjunktív klózokra bontás): csak $\neg$ és $\vee$ maradhat, plusz a kvantorok
 - ha szükséges, a változókat átnevezni**
 - univerzális kvantorokat elhagyni $\rightarrow$ csak $\neg$ és $\vee$ maradhat, azaz az és-ek helyett külön mondattá szedjük őket**
 - Rezolúció bizonyítás procedúrája $\rightarrow$ adott állítás halmaza F, a bizonyítandó állítás Q
 - Az F halmaz összes állítását konvertáljuk klóz formába $\rightarrow F'$
 - Negáljuk a Q-t, és konvertáljuk klóz formába, adjuk hozzá F'-hez
 - Ismételjük a ciklust (3-4-5 pont):
 - ellentmondásra rá nem futunk
 - előrehaladást nem tapasztaljuk
 - az erőforrások előre meghatározott mennyiségét ki nem használjuk
 - Válasszunk meg két klózt és alkalmazzunk rezolúciós lépést
 - Ha a rezolvens üres, megvan az ellentmondás. Ha nem, adjuk hozzá a többi klózhoz és folytatjuk a 3. pontot.
 - Rezolúciós stratégiák = klózek kiválasztási heurisztikái
 - Egység rezolúció $\rightarrow$ Horn-klóz alakú TB-ban az eljárás teljes, különben nem!
 - Set of Support:
 - MI A LÓFASZ EZ A 31. DIA? KONKRÉTAN EGY SZÓT NEM MOND EL BELŐLE ÉRTELMESEN... 1:01:00 előtt FUCKING TODO
 - logikai prorgamozás - Prolog
 - nem teljes** $\rightarrow$ időkorlát van!

- **Ontológia**

- cél: a tudás megosztása (érthetően). Adott tárgyszerület részletes leírását valósítja meg.
- összetevők: fogalmak (osztály), tulajdonságok és relációk a fogalmak közt (taxonómia (isA, partOf), reláció, attribútum, szerep)
- **tárgyszerület részletes leírása:**
 - megkötések → típus megkötés, számosság (1 alatti), értéktartomány (0 és 50 közti), X nagyobb mint Y, X és Y különböző
 - konkrét érték megadása
- ontológiák típusai
 - szemantika szerinti csoportosítás
 - taxonomy - relációs modell szerinti
 - thesaurus - ER
 - conceptual model - RDF/S, XTM, UML
 - logical theory - OWL, ...
 - Elsőrendű logika
- Ontológia vs Relációs séma
 - ontológia → formálisan definiált kapcsolatok az entitások közt, ember és gép is érti
 - relációs séma → implicit kapcsolatok az entitások között, interpretáció szükséges, adatbázis szemantikájának ismerete nélkül nem használható

- **festő ágens problémája**

- amiket nem kezelünk: idő múlása, ágens cselekvéseinek leírása, környezet változása
- idő bevezetése → **szituáció kalkulus**
 - idő múlása = szituációk sorozata
 - egy szituációban 1 tény igaz vagy hamis, de a következő szituációban lehet más
 - fluent = folyékony esemény
 - **hatás axiómák:** egy adott cselekvésnek mi lesz az eredménye
 - **keret axiómák:** a környezet miatt mi marad változatlan

- TODO itt még volt valami a diasorban, de arról nem beszélt (?)