
Ontológiák — egy matematikus-informatikus szemével

Szeredi Péter
szeredi@cs.bme.hu

Budapesti Műszaki és Gazdaságtudományi Egyetem
Számítástudományi és Információelméleti Tanszék

-
- ① Mi az ontológia, mire jó, hogyan csináljuk?
 - ② Matematikai alapok: leíró logikák
 - ③ OWL – Web Ontology Language

MI AZ ONTOLÓGIA?

Fogalmi rendszer, azaz
egy adott (szak)területen ismert

- fogalmak, pl. **SokgyermekesSzülő**, **SzívbillentyűZavar**
- kapcsolatok (relációk, szerepek), pl. **gyereke**, **helye**
- terminológiai axiómák (háttértudás):
 - hierarchiák, pl. a **SzívbillentyűZavar** az_egy **Szívbetegség**
(az_egy = is_a)
 - definíciós axiómák, pl. **SokgyermekesSzülő** \equiv
olyan **Ember**, akinek van legalább 3 **gyereke**.
 - általános axiómák, pl.
ha valaki **Boldog Ember**, akkor minden **gyereke** is **Boldog**
- adataxiómák (metainformációk):
 - fogalmi állítások, pl. **Elemér Boldog**
 - szerepállítások, pl. **Elemér gyereke Eleonóra**

MIRE JÓK AZ ONTOLÓGIÁK

Tisztázzuk az alapvető fogalmakat és relációkat

Idevonatkozó tudásunkat *formálisan* is meg tudjuk fogalmazni

→ adatszinten (adatdoboz, A-doboz), pl.

gyereke(**Elemér**, **Eleonóra**)

→ terminológiai szinten (terminológiai doboz, T-doboz), pl.

SzívbillentyűZavar \sqsubseteq **Szívbetegség**

Boldog \sqcap **Ember** \sqsubseteq \forall gyereke.**Boldog**

Következtetéseket tudunk végezni

→ adatszinten, pl. ha tudjuk, hogy **Elemér** **Ember** és **Elemér**nek

gyereke **Eleonóra**, **Eulália** és **Erika**, akkor besorolhatjuk **Elemért** a

SokgyermekesSzülő fogalomba.

→ terminológiai szinten, pl. kikövetkeztethetjük, hogy minden

SokgyermekesSzülő egyben **Szülő** is.

MI KELL AHHOZ, HOGY ONTOLÓGIÁT CSINÁLJUNK

Nyelvi formalizmus

- matematikai: ún. leíró logika, elsőrendű logika, ...
- informatikai: (szemantikus világháló:) RDF (Resource Description Framework), RDF séma, OWL (Web Ontology Language), ...

Informatikai eszközök

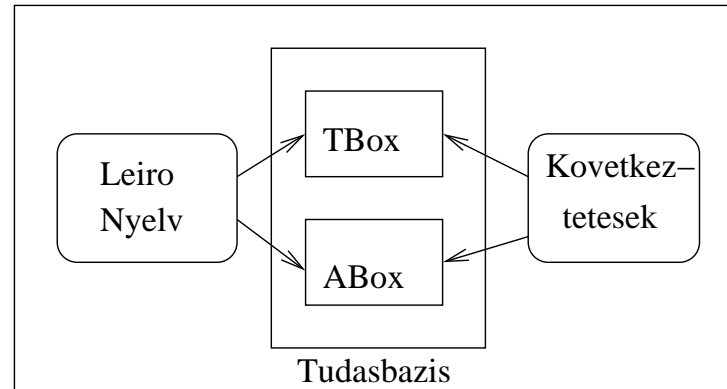
- Ontológiaszerkesztő, pl. Protégé
- Következtető, pl. Racer, KAON2

Már létező és elérhető ontológiák

- általános ontológiák, pl. Dublin Core, Magyar Egységes Ontológia
- szakterületi ontológiák, pl. Galen (orvosi szakterületen) ...

MATEMATIKAI HÁTTÉR: LEÍRÓ LOGIKÁK

LEIRÓ LOGIKÁK MINT A TUDÁSREPREZENTÁCIÓ ESZKÖZEI



- Tudásbázis (KB, knowledge base) = T-doboz (TBox) + A-doboz (ABox):
- T-doboz = terminológiai doboz = terminológiai állítások, **háttértudás**: $\text{Anya} \equiv \text{Ember} \sqcap \text{Nőnemű} \sqcap \exists \text{gyereke}.\top$
- A-doboz = adatdoboz = adatállítások, konkrét **metainformációk**:
 $\text{Ember}(\text{ÉVA}) \text{ Nőnemű}(\text{ÉVA}) \text{ gyereke}(\text{ÉVA}, \text{MIKLÓS})$

PÉLDA TERMINOLÓGIAI KÖVETKEZTETÉSRE

Tudásbázis

T-doboz

anya = ember és nőnemű és van gyereke.

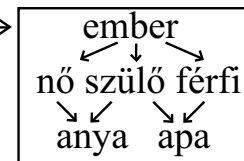
nő = ember és nőnemű

férfi = ember és nem nőnemű

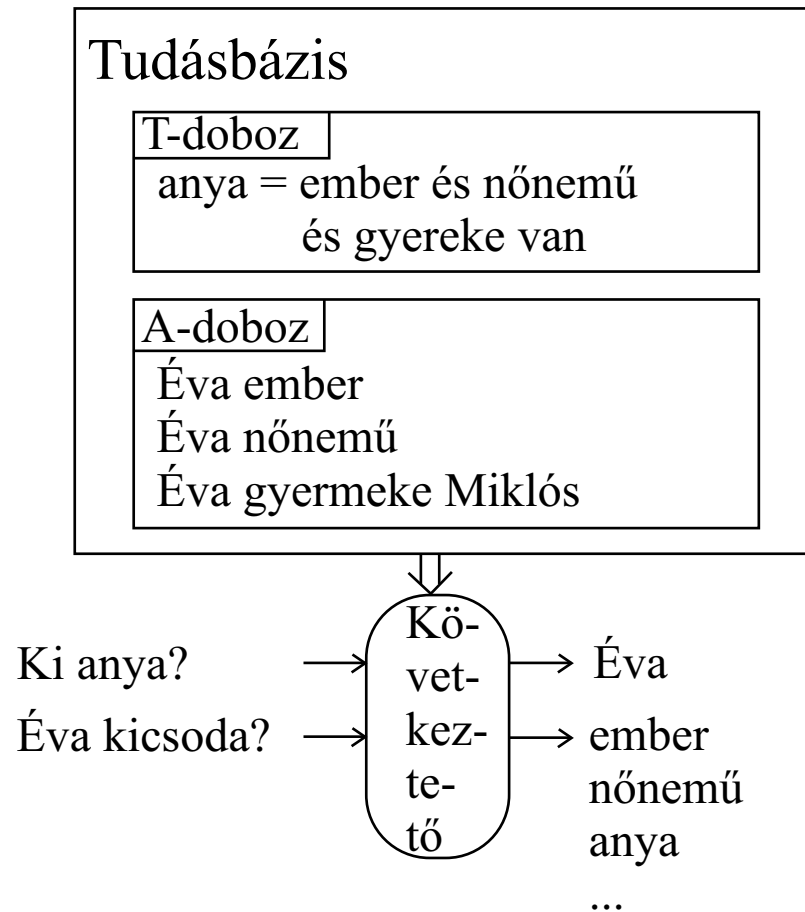
szülő = ember és van gyereke

apa = férfi és szülő

- (1) Konzisztens-e a T-doboz? → Kö- → Igen. (1)
(2) Minden anya szülő? → vet- → Igen. (2)
(3) Minden szülő férfi? → kez- → Nem. (3)
(4) Lehet-e férfi anya? → te- → Nem. (4)
(5) Mi a fogalmak hierarchiája? → tő → (5)



PÉLDA ADATKÖVETKEZTETÉSRE



A LEÍRÓ LOGIKÁK

Miről szólnak?

- *fogalmak*: a modellezni kívánt világ egyedeiből álló *halmazok* (RDF, OO megfelelője: osztály)
- *szerepek*: a modellezni kívánt világ egyedei között fennálló kétargumentumú *relációk*, azaz egyedpárok halmazai (RDF megfelelője tulajdonság, OO megfelelője: asszociáció)

Hogyan épülnek fel?

- *atomi* fogalmak és szerepek: eleve adottak, pl. Ember, gyereke
- *összetett* fogalmak és szerepek: atomi fogalmakból és szerepekből különféle *konstruktorokkal* felépített kifejezések, pl. $\text{Ember} \sqcap \text{Nőnemű} \sqcap \exists \text{gyereke}.$
 gyereke^-

PÉLDÁK FOGALMI AXIÓMÁKRA

Az Anya nem más, mint olyan Ember aki Nőnemű és van gyereke.

$$\text{Anya} \equiv \text{Ember} \cap \text{Nőnemű} \cap \exists \text{gyereke.} \top$$

Minden Tigris Emlős.

$$\text{Tigris} \sqsubseteq \text{Emlős}$$

A boldog emberek gyerekei is boldogak.

$$\text{Boldog} \cap \text{Ember} \sqsubseteq \forall \text{gyereke.} \text{Boldog}$$

A gyermektelen emberek boldogak

$$\forall \text{gyereke.} \perp \cap \text{Ember} \sqsubseteq \text{Boldog}$$

PÉLDÁK SZEREPAXIÓMÁKRA

A gyereke viszonyban levők leszármazottja viszonyban is vannak.

gyereke \sqsubseteq leszármazottja

A szülője kapcsolat a gyereke kapcsolat megfordítottja (inverze).

szülője \equiv gyereke⁻

A leszármazottja reláció tranzitív

Trans(leszármazottja)

AZ \mathcal{ALC} NYELV SZINTAXISA

- Az \mathcal{ALC} fogalomkifejezések (röviden fogalmak) szintaxisa:

$C \rightarrow A$		(atomi fogalom)	egy halmaz, pl: Ember
\top		(tetőjel, top)	az összes objektum halmaza
\perp		(fenékJel, bottom)	az üres halmaz
$\neg C$		(negálás)	
$C \sqcap D$		(metszet)	
$C \sqcup D$		(unió)	
$\forall R.C$		(értékkorlátozás)	azon egyedek, amelyek minden R -je C -beli
$\exists R.C$		(létezési k.)	azon egyedek, amelyekhez van C -beli R

A atomi fogalom, C, D összetett fogalmak

AZ *ALC* NYELV – CSALÁDI PÉLDÁK

Nő	≡	Ember \sqcap Nőnemű
Férfi	≡	Ember \sqcap \neg Nő
Anya	≡	Nő \sqcap \exists gyereke. \top
Apa	≡	Férfi \sqcap \exists gyereke. \top
Szülő	≡	Anya \sqcup Apa
Nagyanya	≡	Anya \sqcap \exists gyereke.Szülő
Feleség	≡	Nő \sqcap \exists férje. \top
LányosSzülő	≡	Szülő \sqcap \forall gyereke.Nőnemű
Ember	\sqsubseteq	\forall gyereke.Ember
Ember	\sqsubseteq	\forall férje.Férfi
Ember \sqcap \exists férje. \top	\sqsubseteq	Nő

ADATBÁZISOK ÉS ADATDOBOZOK

Adatbázis: zárt világ

- pontosan az szerepel az adatbázisban, ami igaz
- példa-adatbázis (egyetlen tábla egy sorral):
gyereke(ÉVA,KÁIN) – ÉVANak egyetlen gyereke van

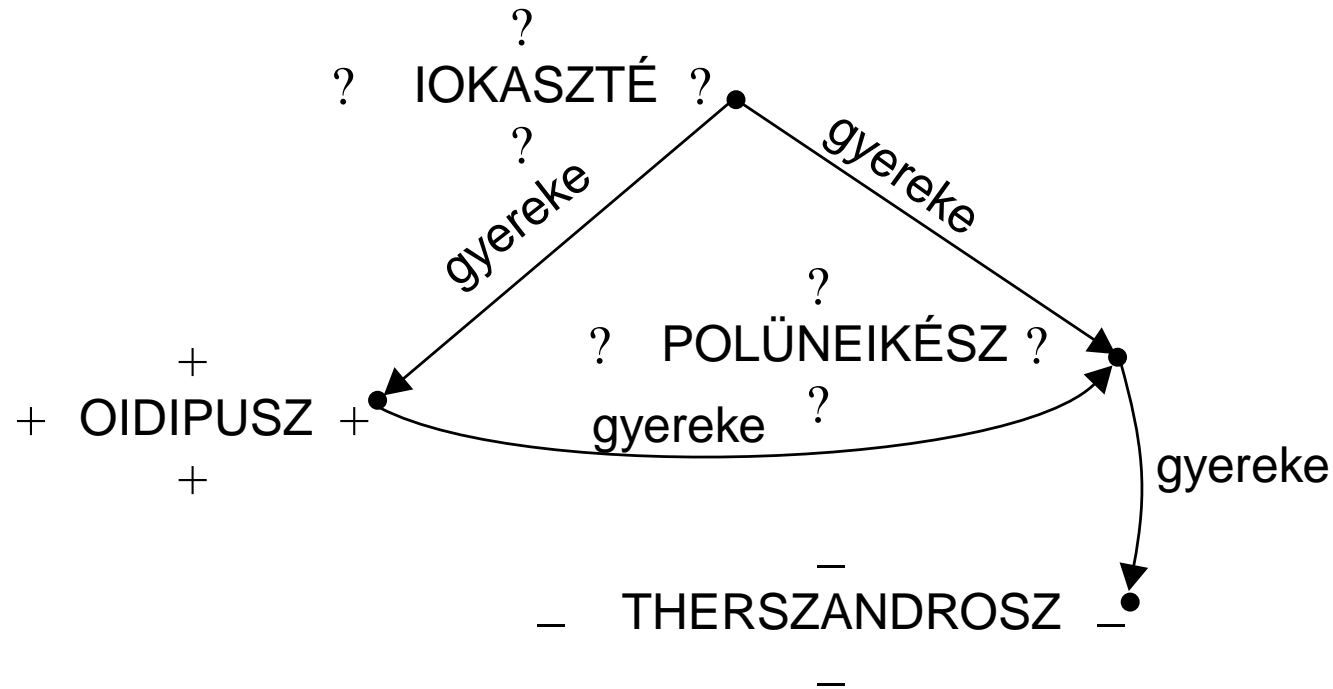
Leíró logikák: nyílt világ

- egyszerre több lehetséges világot kell lefedni
- ami szerepel a tudásbázisunkban az igaz, de más is lehet igaz
- negatív információt is megadhatunk, pl. \neg Ember(ZEUSZ)
- lehet, hogy $\neg C(X)$ és $C(X)$ egyike sem ismert
- példa-tudásbázis (egyetlen adatállítás):
gyereke(ÉVA,KÁIN) – ÉVANak lehet több gyereke is

A szemantikus technológiák körében (pl. a világhálón)

- fontos a nyílt világ szemlélet,
- nem-teljes információ birtokában kell következtetni.

PÉLDA NYÍLT VILÁGBAN VALÓ KÖVETKEZTETÉSRE



→ Az adatdoboz:

gyereke(IOKASZTÉ,OIDIPUSZ)

gyereke(IOKASZTÉ,POLÜNEIKÉSZ)

gyereke(OIDIPUSZ,POLÜNEIKÉSZ)

gyereke(POLÜNEIKÉSZ,THERSZAN.)

Apagyilkos(OIDIPUSZ)

¬ Apagyilkos(THERSZANDROSZ)

→ Egy igaz állítás:

$(\exists \text{gyereke.} (\text{Apagyilkos} \sqcap \exists \text{gyereke.} \neg \text{Apagyilkos}))(\text{IOKASZTÉ})$

OWL – WEB ONTOLOGY LANGUAGE

AZ OWL (WEB ONTOLOGY LANGUAGE) NYELV

Az OWL résznyelvek

- OWL Light
- OWL DL
- OWL Full

Példa: a lányos szülő definíciója OWL-ben

```
<owl:Class rdf:ID="LányosSzülő">
  <owl:intersectionOf rdf:parseType="Collection">
    <owl:Class rdf:about="#Szülő" />
    <owl:Restriction>
      <owl:onProperty rdf:resource="#gyereke" />
      <owl:allValuesFrom rdf:resource="#Nőnemű" />
    </owl:Restriction>
  </owl:intersectionOf>
</owl:Class>
```

AZ OWL ÉS A LEÍRÓ LOGIKÁK

Az OWL 1.0 DL alapja a $\mathcal{SHOIN}(\mathbf{D})$ leíró logikai nyelv

- $\mathcal{S} \equiv \mathcal{ALC}_{\mathcal{R}+}$ (a \mathcal{ALC} nyelv kiegészítve tranzitív szerepekkel), azaz egyes szerepekről (pl. őse) kijelenthetjük, hogy tranzitívak.
- $\mathcal{H} \equiv$ szerephierarchiák. Egy szerephierarchia $R \sqsubseteq S$ alakú állítások halmaza, pl. minden barátja kapcsolat egyben ismerőse kapcsolat is: barátja \sqsubseteq ismerőse.
- $\mathcal{O} \equiv$ egyedfogalmak (objektumok). Olyan fogalmak, amelyeknek csak egy példánya lehet.
- $\mathcal{I} \equiv$ inverz szerepek: egy R szerep mellett annak R^{-} inverzét is használhatjuk, pl. gyereke $^{-} \equiv$ szülője.
- $\mathcal{N} \equiv$ számosság-korlátozások, azaz $\leq nR$ és $\geq nR$ alakú fogalomkifejezések pl. azon emberek akiknek legalább 3 gyereke van: (≥ 3 gyereke)
- $(\mathbf{D}) \equiv$ konkrét adattartományok: egy tulajdonság értékészlete lehet pl. egész, fűzér stb.

ÖSSZEFOGLALÁS

- Az ontológiák lehetővé teszik, hogy tudásunkat
 - formálisan
 - gépi következtetésre alkalmasanfogalmazzuk meg.
- A gépi tudásreprezentációhoz egyszerű logikai nyelvekre van szükség
 - például a leíró logikákra
- A szilárd matematikai alapokra megfelelő mérnöki módszerek és eszközök építhetők
 - például RDF, OWL

SZEMANTIKUS TECHNOLOGIÁK

Személyes tapasztalatok

- SINTAGMA (Szemantikai INtegrációs Technológia Alkalmazása Grid-alapú, Modellvezérelt Architektúrákban), NKFP projekt, 2005–2007, IQSYS, BME, SZTAKI, ...
- FUSION (Business process fusion based on Semantically-enabled Service-oriented Business Applications) EU FP6 STREP projekt, 2006-2008, SAP, ... , BME, ...
- „A szemantikus világháló és az ontológiakezelés alapjai” – BME VIK választható tárgy, heti 4 óra, 2004. óta.
- „Szeredi P., Lukácsy G., Benkő T. A szemantikus világháló elmélete és gyakorlata”, tankönyv, Typotex 2005 (angol kiadása várhatóan: Cambridge University Press, 2007 vége)
Jelenleg 50% engedménnyel (2200 Ft) kapható az Olvasók Boltjában, Budapest V.,Pesti Barnabás u. 4.