

1. Bevezetés

A természetes nyelv lexikai készlete felfogható olyan rendszerként, amelynek elemei jelentésükön keresztül kölcsönös megszorításokat rónak ki egymásra. Ezeket a megszorításokat az elemek logikai kapcsolatain keresztül lehet felderíteni, és összességük körülhatárolja a nyelv „szándékolt” interpretációját, azaz azt az ontológiát, amelyhez a nyelvet implicit elkötelezettség fűzi. A MEO-projekt egyik almoduljának célkitűzése, hogy ezen implicit ontológiai elkötelezettséget empirikus, alulról felfelé építkező módon feltárja azon keresztül, hogy precíz reprezentációt rendel lexikai elemek egy jelentős méretűnek tekinthető halmazához. E munka hasznos mellékterméke egy olyan, a magyar nyelvre elkészített lexikon (szótár) létrejötte, amelyben az egyes szemémák (jelentéses egységek) jelentése precíz formalizmus segítségével van ábrázolva. Ez a lexikon azután a legkülönbözőbb természetes nyelvi mesterségesintelligencia-alkalmazások számára nyújthat felbecsülhetetlen segítséget.

2. Reprezentációs nyelv

A természetes nyelv jelentéses egységeinek leírása számos problémát felvet. Az egyik legalapvetőbb probléma az, hogy a természetes nyelv jelentéses egységei — kevés kivételtől eltekintve — nem rendelkeznek definícióval, és ilyen nem is adható rájuk. Ez azt jelenti, hogy a definíciók esetében megkövetelt „akkor és csak akkor” jellegű klauzulák a természetes nyelvi szemémák esetében többnyire nem írhatók fel. Ez nem jelenti azonban azt, hogy egyes szükséges illetve elégséges feltételek esetenként ne lennének meglehetősen bizonyossággal meghatározhatók.

Egy másik, ettől eltérő problémakör a természetes nyelv logikai komplexitásával kapcsolatos. Ezt a ténytet indirekt módon is jelzik a szemantikusok és logikusok által a modellezés céljaira kidolgozott formális nyelvek bősége,

amelyek közül talán a Richard Montague nevéhez fűződő típuselméleti intenzionális logika és annak „leszármazottai” a legismertebbek. Ez a reprezentációs nyelv a típusos λ -kalkulus lehetséges világ szemantikára épülő változata, amely alkalmas arra, hogy a természetes nyelv jelentéses egységei számára nagy kifejezőerejű keretelméletként szolgáljon. A Montague-féle keretrendszer azonban kevésbé alkalmas arra, hogy gyakorlati ontológiafeltáró megközelítés alapjául szolgáljon, alapvetően a következő — számunkra is lényeges — okoknál fogva. A Montague-szemantika a „hagyományos” logikai paradigmába illeszkedik, amelynek „ontológiai elkötelezettsége” minimális. Ennek következtében a konkrét, „életszerű” alkalmazások esetében a reprezentációk nagyon gyorsan elbonyolódnak, hiszen a gazdagabb ontológiát feltételező kifejezések jelentését is igen kevés számú rögzített alaptípus kombinációjaként kell előállítani. Ez gyorsan szükségessé teszi magasabbrendű operátorok bevezetését, amit még bátorít is az ω -rendű típuselmélet nagy kifejezőereje, amelynek segítségével könnyen előállíthatók függvényeken működő függvények, azokon működő függvények, s.í.t. Míg azonban elméleti szempontból ez ellen semmiféle kifogás nem emelhető, gyakorlati jelentésleírások céljára ez a rendszer nem alkalmas.

A reprezentációk elbonyolódásának azonban gátat lehet vetni oly módon, hogy a bonyolultság forrását, azaz a természetes nyelv ontológiai típusgazdagságát a Montague-szemantikák ontológiailag statikus megközelítésével szemben egy dinamikus típusozási eljárással ragadjuk meg. Röviden arról van szó, hogy ha a jelentésleíró nyelvet kibővítjük egy olyan eszközzel, amely lehetővé teszi a „reifikációt” bármely ponton ahol arra szükség van, akkor a jelentésleíró nyelv szintaktikai komplexitása kezelhető határokon belül marad, mert a leírandó jelentés bonyolultságát a reifikációs mechanizmus a leíró nyelv formális szemantikájába relegálja. Ennek részleteit alább tárgyaljuk. Fontos megjegyezni azonban, hogy a reifikációs mechanizmus pusztán a jelentésleíró nyelv szintaxisának egyszerűsítésére szolgáló eszköz, amely a dinamikus típusképzés lehetővé tételén keresztül elősegíti az információ kompakt formában történő leírását, de természetesen nem képes csökkenteni a szemémák inherens szemantikai bonyolultságát.

2.1. Jerry Hobbs jelentésleíró nyelve

Jerry R. Hobbs 25 éven keresztül volt az SRI International Mesterséges Intelligencia Központjának vezető kutatója, jelenleg a Dél-Kaliforniai Egye-

tem Informatikai Intézetének alkalmazottja. Fő kutatási területe a „common-sense knowledge” illetve a természetes nyelvben tárolt információ lehetséges kódolási eljárásainak vizsgálata.

Hobbs az 1980-as évek közepétől kezdve dolgozik egy olyan reprezentációs nyelven, amelynek a segítségével a természetes nyelvben lexikailag tárolódó információk relatíve egyszerűen, de mégis pontosan írhatók le. Ezt a célt egy a davidsoni megközelítést radikálisan kiterjesztő megközelítés segítségével éri el.

2.1.1. Történeti előzmények

A davidsoni alapok. Donald Davidson (Davidson 67)

- (1) Brutus stabbed Caesar.
- (2) Brutus stabbed Caesar with a dagger.
- (3) Brutus stabbed Caesar with a dagger at noon.

A hagyományos elsőrendű leírás ezeket a mondatokat rendre mint $\text{stabbed}(b, c)$, $\text{stabbed}(b, c, d)$ illetve $\text{stabbed}(b, c, d, 12)$ lenne kénytelen ábrázolni, mivel a határozói módosítást csak egy-egy újabb argumentum hozzáadásával képes kezelni. Davidson ötlete az, hogy a predikátumokat egyetlenegy eseményargumentummal kiegészítve a nehézség elháríthatóvá válik, mivel a módosítók nagy része ezután mint az eseményargumentumra vonatkozó, a „magkijelentéshez” konjunkcióval csatolt predikátum kezelhető.

$$\exists e \text{ stabbed}(e, b, c) \tag{2.1}$$

$$\exists e (\text{stabbed}(e, b, c) \wedge \text{with}(e, d)) \tag{2.2}$$

$$\exists e (\text{stabbed}(e, b, c) \wedge \text{with}(e, d) \wedge \text{at}(e, 12)) \tag{2.3}$$

Hobbs reprezentációs nyelve ennek a davidsoni ötletnek a radikális kiterjesztésére épül.

2.2. Hobbs reprezentációs nyelvének sajátosságai

Hobbs egy nagyon fontos ponton általánosítja Davidson-t. Egy „komprehenziós séma” segítségével posztulálja, hogy bármely n -argumentumú predikátumhoz tartozik egy $n + 1$ -argumentumú predikátum, amelynek van egy kitüntetett eseményargumentuma (ezt alább több példán be fogjuk mutatni). Ennek az általánosításnak azonban közvetlen következménye, hogy az eseményváltozó már nem csak a davidsoni értelemben vett eseményekre, hanem tetszőleges *tényállásra* vonatkozhat. Más szóval Hobbs tetszőleges (ténylegesen fennálló vagy pusztán lehetséges) tények fölött kvantifikál és azt, hogy egy tény *valójában* is fennáll, egy sajátos predikátummal, a $\text{Rexist}(e)$ („really exists”) predikátummal fejezi ki. Egy $p(a)$ mondat igazsága így ekvivalens módon kifejezhető lesz a $\exists e(p'(e, a) \wedge \text{Rexist}(e))$ kijelentéssel, ahol a vesszős predikátum első argumentuma egy tényekre utaló változó. Általános esetben tehát:

$$\forall x_1, x_2, \dots, x_n (p(x_1, x_2, \dots, x_n) \equiv \exists e(\text{Rexist}(e) \wedge p'(e, x_1, x_2, \dots, x_n))) \quad (\mathbf{C})$$

Hobbs magyarázata szerint a felsővessző mint operátor egy olyan predikátumot állít elő a kiinduló predikátumból, amelynek első argumentuma olyan körülményeket (*conditions*) tud denotálni, amelyek mellett a $p(x_1, x_2, \dots, x_n)$ kijelentés igaz. Egzisztenciálisan kvantifikálva ezt a változót olyan formulát kapunk, amely azt állítja, hogy *van* olyan körülmény, amelyek között a $p(x_1, x_2, \dots, x_n)$ kijelentés igaz; ezt azután kiegészítve a Rexist predikátummal a kapott formula azt állítja, hogy a $p(x_1, x_2, \dots, x_n)$ kijelentés *ténylegesen* (aktuálisan) is igaz.

Az analitikus filozófiában az utóbbi években egyre többször kerül terítékre az ún. *truth-makers* elmélete (ezt bevett terminológia híján esetleg „az igazságalapok elméletnek” nevezhetjük). Eszerint az elmélet szerint az igazság fogalmával kapcsolatba hozható dolgok két osztályba sorolhatók: igazsághordozók (*truth-bearers*): ezek azok a dolgok, amik egyáltalán igazak vagy hamisak lehetnek (alapvetően a mondatok és a kijelentések), illetve az igazságalapok (*truth-makers*): ezek azok a dolgok, amelyek igazzá vagy hamissá teszik az igazsághordozókat. Az elfogadott definíció szerint egy igazsághordozó igazságalapja az, aminek létezése kikényszeríti („necessitates”) az igazsághordozó igazzá válását. Davidson fent említett elmélete az igazságalap-elmélet egy sajátos alkalmazásának tekinthető¹, ahol a kijelentések (Davidson cikkében a cselekvésleíró mondatok) mint igazsághordozók igazságalapjául intencionális cse-

¹Davidson extenzionalista, azaz az egzisztenciális kvantifikáció nála mindig az aktuálisan létező entitások felett történik; Hobbs elejti ezt a megszorítást a possibilista

lekvési események szolgálnak. Hobbs megközelítése ezzel szemben az igazságalap-elmélet korlátlan elfogadására épül.

Illusztráljuk most ezt a megközelítést Hobbs egy példájával. A

(4) Maybe the boy wanted to build a boat.

mondat „hagyományos” reprezentációja Hobbs szerint valahogy így nézne ki:

$$\exists x(\text{boy}(x) \wedge \diamond \text{PAST}(\text{WANT}(x, \lambda z(\exists y(\text{boat}(y) \wedge \text{Quick}(\text{build}(z, y)))))),$$

míg a reifikáció alkalmazásán keresztül a fentiek a következőkre egyszerűsödnek:

$$\exists e_0, e_1, e_2, e_3, e_4, x, y(\text{Rexist}(e_0) \wedge \text{Past}'(e_0, e_1) \wedge \text{possible}'(e_1, e_2) \wedge \text{want}'(e_2, x, e_3) \wedge \text{quick}'(e_3, e_4) \wedge \text{build}'(e_4, x, y) \wedge \text{boy}(x) \wedge \text{boat}(y))$$

2.3. Thematikus szerepek

A thematikus szerep (vagy théta-szerep) fogalma a nyelvtudományból származik, ahol abból a megfigyelésből kiindulva, hogy az egyes igék argumentumaikra tipikus, ám rendkívül általános feltételeket rónak ki, a nyelvészek körülhatároltak néhány ilyen típust. A thematikus szerep (θ -szerep) fogalma azonban igen általánosan is definiálható. Voltaképpen tetszőleges predikátum (nem csak ige) esetében is van értelme. Általános esetben egy θ thematikus szerep azonos argumentumszámú predikátumok valamelyik argumentumhelyét jellemzi. Így a thematikus szerepek két tényezőtől függenek: n -argumentumú predikátumok (pl. igék) egy \mathcal{P} halmazának megválasztásától (n rögzített), illetve attól, hogy pontosan *melyik* argumentumhelyet vizsgáljuk. A legáltalánosabb szinten tehát nem beszélhetünk thematikus szerepről *per se*, csak egy *adott predikátumhalmazban* a predikátumok i . argumentumhelyét jellemző thematikus szerepről.

Általánosan azt mondhatjuk, hogy adott \mathcal{P} n -argumentumú igékből (n -argumentumú predikátumokból) álló igealalmaz (predikátumhalmaz) elemeinek i . argumentumhelyét jellemző thematikus szerep nem más, mint azon tulajdonságok (1-argumentumú predikátumok) összessége, amelyeket a halmazban található igék (predikátumok) az i . argumentumukra vonatkozóan mind megkövetelnek:

kvantifikáció kedvéért, és az aktuális entitások feletti kvantifikációt ennek és a Rexist predikátumnak a segítségével állítja elő.

$$\theta(x_i, \mathcal{P}) \stackrel{\text{def}}{=} \bigcap_{P \in \mathcal{P}} \{ \{ \Phi \mid \forall x_1 \forall x_2 \dots \forall x_n (P(x_1, x_2, \dots, x_i, \dots, x_n) \rightarrow \Phi(x_i)) \} \}$$

Ebből a szokásos thétaszerep-fogalomhoz úgy jutunk, hogy a $\theta(x_i, \mathcal{P})$ halmaz elemei közül csak a legspecifikusabbakat tartjuk meg (a „legkisebbeket” a tulajdonságok (predikátumok) implikációs (generikus) rendezése szerint).

A nyelvészek pragmatikus alapon csoportosították az igéket, miközben párhuzamosan kialakították a csoportokhoz tartozó thétaszerepeket is. Kialakították pl. az ún. *ágentív igék* osztályát, amiben olyan igéket szerepeltetnek, amelyek (felszínen alanyként megjelenő) argumentumára *nagyjából* a következő érvényes: {aktív-az-eseményben, felelős-kezdeményezője-az-eseménynek}. Ha tehát egy ige teljesíti azt a feltételt, hogy bármely vele képzett igaz mondatban az alanyi argumentumára igazak e halmaz elemei, akkor a szóbanforgó ige eleme az ágentív igék osztályának.

A nyelvészek által használt thematikus szerepek főbb típusai a következők.

- Ágens (AG): tudatos, a cselekvésben aktív, felelős érte, és ő kezdeményezi (esetleg folyamatosan kontrollálja) azt (*Brutus a Brutus megölte Cézárt* mondatban)
- Páciens (PAT): egy ágens által kifejtett hatás, változás elszenvedője; az esemény során tulajdonságai az ige jelentéséből megjósolható módon megváltoznak (*Cézár a Brutus megölte Cézárt* mondatban)
- Eszköz (INSTR): a cselekvés közvetlen eszköze; jelenléte feltételezi egy ágens és egy páciens jelenlétét is; az ágens az eszköz közvetlen manipulálásával gyakorol hatást a cselekvés páciensére (a *tőr* a *Brutus egy tőrrel ledöfte Cézárt* mondatban)
- Experiens (EXP): tudatos szereplő; az ige az ő tudatállapotáról illetve annak valamilyen megváltozásáról tesz állítást (*János fél a pókoktól, Mari megijedt a kutyától*)
- Természeti erő: önmagától, és emberi eszközökkel kontrollálhatatlanul működő tényező (*A vihar tönkretette a termést, A szél betörte az ablakot*)

A tematikus szerep fogalma a mi számunkra két szempontból is lényeges. Egyrészt azért, mert a többargumentumú predikátumkonstansok argumentumai között nemigen lehet máshogy különbséget tenni (a sorrend nyilván konvencionális, ám a konvenció alapjául többnyire éppen a tematikus szerepek szolgálnak). Másrészt azért, mert a reifikációval együttesen használva lehetővé teszik a többargumentumú relációk kétargumentumúra redukálását. Terence Parsons (Parsons 90)-ben ugyanis módszeresebb továbbfejleszti Davidson eredeti elméletét egy olyan irányban, amelyben az eseményben résztvevő szereplőket egy-egy tematikus szerep köti magához az eseményhez. Például a

(5) Brutus egy tőrrel ledöfte Cézárt.

mondat parsonsi ábrázolása így néz ki:

$$\exists e(\text{ledöfés}(e) \wedge \text{AG}(\text{brutus}, e) \wedge \text{PAT}(\text{cézár}, e) \wedge \text{INSTR}(\text{tőr}, e))$$

2.4. Az „ontológiai létra”

3. Mikroelméletek

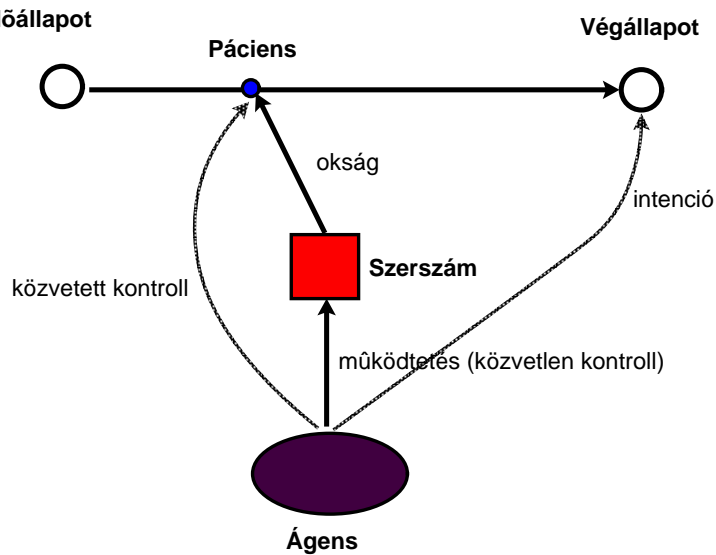
A mikroelmélet fogalma a CYC ontológiában széleskörűen használatos fogalom. A mikroelméletek olyan tudásterületek (egy csoportba tartozó állítások), amelyek osztoznak bizonyos általánosnak tekinthető axiómákban. A mikroelmélet fogalma a jelentésleírások esetében is alapvetőnek tekinthető, mert ezen keresztül ragadható meg az azonos „jelentésmezőkhöz” tartozó kifejezésekben lévő közös tartalom. Az egy szemantikai mezőhöz tartozó mikroelméletet jellemző axiómák meghatározása azért fontos, mert segítségükkel a mezőhöz tartozó egyes szemémák jelentésének leírása „parametrizálhatóvá válik”. Ezt az eljárást részletesen az eszközfogalmak példáján mutatjuk be.

3.1. Az eszközfogalmak mikroelmélete

3.1.1. Informális leírás

Minden eszköz feltételez egy Ágenst, aki használja, és egy Pácienst, amin az Ágens a eszköz segítségével valamilyen változtatást hajt végre. Ezt a vi-

szonyrendszert a következő ábra szemlélteti:



Az Ágens a Páciens egy s_1 kezdőállapotból egy szándékolt s_2 végállapotba szeretné eljuttatni. Az ehhez szükséges állapotváltozást az Ágens a Munkaeszköz működtetésének útján közvetett módon kontrollálja. Az Eszköz működtetése a Páciens állapotára oksági úton hatást fejt ki, aminek következtében a Páciens tulajdonságai a kívánt módon változnak.

Minden általunk ismert eszköz esetében birtokában vagyunk annak a tudásnak, hogy azt mikor használnánk. Az eszközhasználat általános feltétele egy *problémaszituáció* beállása, amit a következő jellemez:

A z objektum s_1 állapotban van és x Ágens azt akarja, hogy z az s_2 állapotban legyen:²

$$s'_1(e_1, z) \wedge s'_2(e_2, z) \wedge \text{akar}'(e_3, x, e_2) \wedge \text{és}(e_1, e_3)$$

Az Eszköz működtetése okozza az átmeneti folyamat lezajlását. Az átmeneti folyamat fogalmát ezen a szinten primitívumnak tekintjük.

²A felső vessző Hobbs reifikációs operátora.

A z entitás átmeneti folyamatát (e) miközben z s_1 -ből s_2 -be megy át, egy 4-argumentumú funktor jelöli, amelyben két eseményszerűség is meg van említve, a Páciens kezdő- és végállapota:³

$$\text{átmenetifolyamat}(e, z, e_1, e_2) \wedge s'_1(e_1, z) \wedge s'_2(e_2, z)$$

Azt, hogy az Ágens az eszközt működteti, önálló eseményszerűségnek vettük. Ennek többek között az az oka, hogy a nyelvek gyakran lexikalizált módon is jelölik ezt a folyamatot (pl. *kalapál*), továbbá, az eszköz jellemzéséhez hozzátartozik működtetési módjának a leírása. A működtetést a következő funktor³ jelöli:

Az x Ágens működteti az y Szerszámot az e eseményszerűségben (folyamatban):

$$\text{működtetés}(e, x, y)$$

Az eszköz működtetése és a Páciens átmeneti folyamata között oksági kapcsolat áll fenn. Ezt azért érdemes külön is felvenni, mert hasznos lehet a nyelvi kifejezések elemzésénél. Például, az a mondat, hogy

(6) Brutus egy törrel megölte Cézárt

elemezhető úgy, hogy Brutus tett valamit, ami Cézár halálát okozta, és amit Brutus tett, az egy tör működtetése volt.

Összefoglalva: az y eszköz x Ágens általi működtetése okozza z s_1 -ből s_2 -be vezető átmeneti folyamatát:

$$\text{működtetés}(e_0, x, y) \Rightarrow \exists e_1, e_2, e_3 (\text{átmenetifolyamat}(e_3, z, e_1, e_2) \wedge s'_1(e_1, z) \wedge s'_2(e_2, z) \wedge \text{okoz}(e_0, e_3)) \quad (3.1)$$

Ez a formula az eszközök tartományára vonatkozó mikroelmélet egyik (alapvető) axiómája.

A fenti sémákban az x , y , z valamint az s_1 és s_2 szabad változók szerepelnek. Az x a Szerszámot használó Ágenst, a z pedig a hozzárendelt folyamat

³Itt nincs felső vessző, mert *átmenetifolyamat* és a *működtetés* funktorok szándékolt jelentésüknél fogva csakis reifikációs jellegűek lehetnek.

Páciensét képviseli. Ezek előre nyilván nem határozhatók meg. A (3.1) formulában szereplő s_1 és s_2 szabad változók olyan paraméterek, amelyeket egy konkrét eszközfogalom esetén specializálni kell. Minden tényleges eszköz esetén „tipikus” módon megadható állapotátmenetek olyan halmaza, amit az adott eszköz elvben lehetővé tesz. Például egy fűrészsel elérhető, hogy egy eredetileg egységes relatíve puha anyagú (fa, vas, stb.) tárgy két részre váljon. Egy kalapáccsal elérhető, hogy két különálló fadarabot egy szöggel egyetlen egységgé összeerősítsünk, stb. Így például a *kalapácshoz* a következő $\langle e_1, e_2 \rangle$ pár rendelhető (és persze még más párok is):

$$\begin{aligned} \exists x_1 \exists x_2 \exists x_3 (\text{Tárgy}(x_1) \wedge \text{Tárgy}(x_2) \wedge \text{különáll}'(e_1, x_1, x_2) \wedge \\ \wedge \text{Kötőelem}(x_3) \wedge \text{összeköti}'(e_2, x_1, x_2, x_3)) \end{aligned}$$

4. Metodika

A jelentésleírások kapcsán az igéket tekintjük elsődlegesnek, a nominalizációt pedig Hobbs módján levezetettnek. Ennek oka egyszerűen az, hogy az ige szemantikailag kitüntetett: az igének vannak argumentumhelyei, amiket kitöltve állításhoz juthatunk; a többi szintaktikai kategória alapvetően csak az ige argumentumhelyeinek kitöltésekor jut szóhoz; az ige az, ami szemantikai megszorításokat (pl. szelekciós megszorításokat) ró ki az argumentumaira (v.ö. még: thematikus szerepek); az ige nominalizációja során tipikusan „argumentumelnyomás” lép fel, stb.

A thematikus szerepek meghatározását fontos, de másodlagos kérdésnek tartjuk. A „davidsonizálás” a többargumentumú predikátumok kétargumentumúvá alakításának eszköze. Azonban a predikátumok szükséges és elégséges feltételekkel történő jellemzése az elsődleges fontosságú, és ha ez sikerült valamilyen szinten, a thematikus szerepek meghatározása már nem okozhat túl nagy nehézséget.

A következő vázlatos eljárást fogadjuk el (függetlenül attól, hogy esszenciális vagy tipikus feltételek leírásáról van szó).

1. Természetes nyelvi jellemzés
 - (a) szükséges feltételek leírása

- (b) elégséges feltételek leírása
- 2. A jellemzéshez kapcsolható formális predikátumok jellemzése
 - (a) a funktornév megválasztása
 - (b) az argumentumok számának és (ontológiai) típusának meghatározása
 - (c) esetleg a predikátum argumentumhelyeinek jellemzése egzisztenciális transzparencia illetve homályosság szempontjából
 - (d) a predikátumokhoz rövid természetes nyelvi glossza készítése
- 3. A természetes nyelvi jellemzés formalizálása
 - (a) a logikai szerkezet felírása
 - (b) a predikátumok bizonyos mértékű jellemzése további („beszélő nevű”) predikátumokkal (ez a többiek számára segítség)
- 4. „Davidsonizálás:” a kettőnél több argumentumú funktorok kétargumentumúvá tétele tematikus relációk segítségével

4.1. Példa: Foglalkozások

4.2. Példa: Pszichikai fogalmak

4.3. Példa: Intézmények

4.4. Példa: Természeti fajták

Hivatkozások

[Davidson 67] Donald Davidson: The Logical Form of Action Sentences, In: *The Logic of Decision and Action*, pp. 81–95, N. Rescher (ed.), The University Press, 1967, Pittsburgh.

[Parsons 90] Terence Parsons: *Events in the Semantics of English: A Study in Subatomic Semantics*, MIT Press, 1990, Cambridge, Massachusetts.