

A melléknevek beillesztése a Magyar WordNetbe

Gyarmati Ágnes¹, Almási Attila², Szauter Dóra²

¹ MTA Nyelvtudományi Intézet, Nyelvtechnológiai Osztály
1068 Budapest, Benczúr u. 33.
agnes@nytud.hu

² MTA-SZTE Mesterséges Intelligencia Kutatócsoport
6720 Szeged, Aradi Vértanúk tere 1.
{vizipal, szauter.dora}@freemail.hu

Kivonat: A wordnet lexikai adatbázisban a melléknevek rendezése merőben eltér a főnévi és igei struktúrától: a melléknevek közti legmeghatározóbb reláció nem a hipo-hipernímia viszony, hanem az antonímia, mely reláció mentén (kiegészítve a jelentéshasonlósági *similar_to* relációval) a deskriptív melléknevek clusteres struktúrába szervezhetőek. A Magyar WordNet melléknévi részének építéskor felmerültek nyelvspecifikus, valamint általános, a wordnet rendszerét érintő problémák. Nyelvspecifikus probléma az egyes szavak beilleszthetőségének kérdése, valamint a különböző „wordnetes” nyelvek közötti megfeleltetések megadása. Az antonímia reláció szigorúan két synset között van definiálva, mégis található a wordnetben „antonimahármasok.” Cikkünkben ezekre a kérdésekre a rendezésére kínálunk konstruktív megoldást, mellyel a szemantikai viszonyok árnyaltabbá válnak a wordnetben.

1 Bevezetés

A *wordnet* lexikai adatbázisban az egyes szavak helyét szemantikai relációk határozzák meg. A melléknevek rendezése merőben eltér a főnévi és igei struktúrától. A melléknevek közti legmeghatározóbb, a melléknévi struktúra jellegét kialakító reláció nem a főneveknél és az igéknél megismert *hipo-hipernímia* viszony, hanem az *antonímia*. Ennek eredményeként a melléknevek nagy része ún. *clusteres* rendszerbe szerveződik.

A Magyar WordNet (a továbbiakban HuWN)⁵¹ melléknévi részének készítésénél több különböző szempontot kell figyelembe venni. Jellegéből adódóan a Princeton WordNet (PWN) mintájára készül, és a magyar nyelv keretein belül igyekszik megtartani a PWN struktúráját, mind a literálok, mind a relációk felvételénél. A két nyelv között azonban komoly lexikai, valamint asszociációs különbségek vannak, ennek

⁵¹ Az adatbázis a 2005 tavaszán indult Magyar Ontológia Építése projekt keretén belül készül, mely a Szegedi Tudományegyetem, a MorphoLogic Kft. és a Nyelvtudományi Intézet közös projektuma (GVOP – 2004 – 3.1.1.). A szerzők köszönetüket fejezik ki a projektben dolgozó kollégáiknak, elsősorban Kuti Juditnak és Varasdi Károlynak a HuWN melléknévi rendszerének kialakításában, ezáltal – közvetve vagy közvetlenül – a jelen tanulmány megírásában nyújtott segítségükért.

következtében a PWN melléknévi részének pusztá lefordításával nem kapjuk meg a HuWN megfelelő részét.

Tanulmányunk a HuWN melléknévi részének építésekor felmerült általános és nyelvspecifikus problémákról ad számot, melyek nyomán szükségessé vált a szófaji kategorizáció vizsgálata, valamint néhány újabb reláció felvétele a wordnetbe.

2 A Princeton WordNet melléknévrendszere

A PWN melléknévi synsetjei többségükben mellékneveket tartalmaznak, de módosítóként szereplő főnevek, egyes igei származékok (folyamatos és befejezett melléknévi igenevek), valamint előljárós szószerkezetek is kerültek közéjük. A PWN jelenleg 19500 melléknevet tartalmaz, melyeket 10000 szinonimacsoportba (a wordnet terminológiájával synsetbe) rendeztek. Ezek nagy részét deskriptív és relációs melléknevek alkotják, referenciamódosító melléknevek kisebb számban jelennek meg köztük [1].

Az alábbiakban röviden bemutatjuk a PWN melléknévi struktúrájának azt a részét, melynek ismerete a HuWN 3. fejezetbeli tárgyalásához elengedhetetlenül szükséges. Ennek értelmében a következőkben *deskriptív* melléknevekkel foglalkozunk, és csak megjegyezzük, hogy a *referenciamódosító* melléknevek egy része szintén beilleszthető a wordnetbe a *deskriptív* melléknevek közé, valamint hogy a *relációs* melléknevek kezelése merőben eltér a deskriptívékétől (ez utóbbi két melléknévosztályról, és a wordnetbe való integrálásukról ld. [1]).

2.1 Deskriptív melléknevek

A 'melléknév' szó hallatára rendszerint a deskriptív melléknevekre gondolunk. A deskriptív melléknév egy attribútum értékét adja a főnévnek. Pl. az *alacsony* és *magas* melléknevek a *MAGASSÁG* attribútumának (ellentétes pólusán levő) értékeit képviselik, és adják át azután a vonatkozó főnévnek, főneveknek. [1]

A deskriptív melléknevek szemantikai szerveződése teljességgel eltér a főnevek vagy az igék szerveződésétől, ugyanis melléknevek esetében nem beszélhetünk hiponímia relációról, ami hierarchiába rendezné őket. Deskriptív melléknevek közti alapvető szemantikai kapcsolat az *antonímia*, melynek jelentőségére először szóaszszociációs kísérletek során derült fény. Felnőtt beszélők esetén pl. a *jó* melléknévvel kapcsolatban elhangzott leggyakoribb asszociáció a *rossz* melléknév, a *rossz* esetén pedig a *jó* volt. Az asszociáció kölcsönössége a melléknevekkel kapcsolatos adatok egy szemebetűnő tulajdonsága, érdemes ezt a lexikai információt a wordnetben is jelölni.

Antonim kapcsolat csak szavak, *literálok* között állhat fenn, erre utalnak a szóaszszociációs kísérletek eredményei is. Tegyük fel, hogy a főnevekhez és igékhez hasonlóan a (deskriptív) mellékneveket is szinonimacsoportokba, azaz synsetekbe osztjuk jelentésük alapján. Mint láttuk, ezek a synsetek nem szervezhetők hierarchiába *hipohipernímia* relációk mentén, erre sokkal alkalmasabb az *antonímia*.

Tekintsük a következő két lehetséges synsetet: {*heavy, weighty*} ('nehéz', 'súlyos') és {*light, weightless*} ('könnyű', 'súlytalan'), jelentésük alapján ez a két synset oppozi-

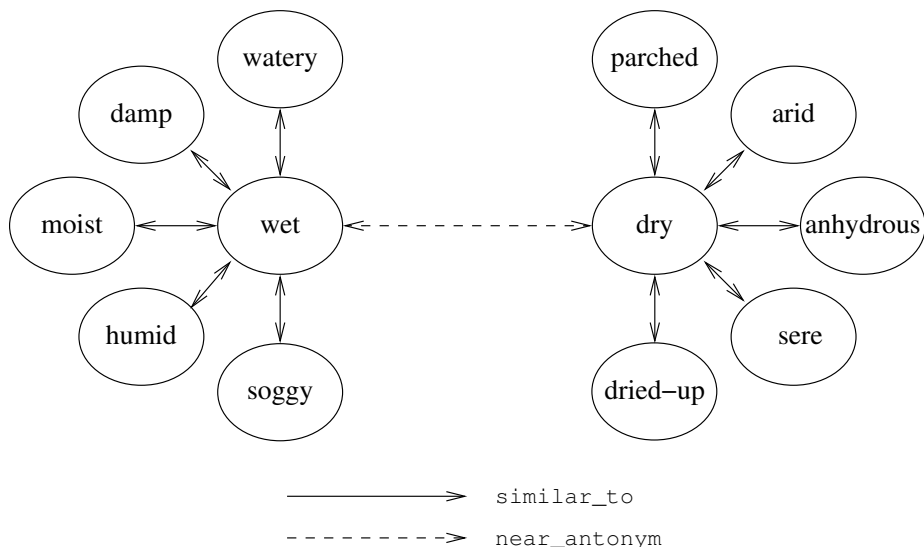
cióban áll egymással. Ennek ellenére nem mondhatjuk, hogy egy antonim párral állunk szemben, hiszen a synset elemei egyenjogú tagjai az őket tartalmazó synsetnek, így ha egy reláció a synset egyik literáljára fennáll, akkor definíció szerint minden további literál esetében fenn kell állnia a relációnak. Tehát ha a *heavy-light* antonim párt elfogadjuk, akkor létezőnek kell feltételeznünk a *heavy-weightless*, valamint a *weighty-light* párok antonimaságát is, márpedig ezek a viszonyok nem valóságok.

Az antonímia relációnak a literálok közötti definiálásához érv továbbá az is, hogy az angolban a deskriptív melléknevek többségének ellentéte morfológiai szabály alkalmazásával jön létre. A jelentés polaritásának megváltozását egy negatív prefixumnak a szóhoz való hozzákapcsolása váltja ki (pl- *un-*, *in-*, és allomorfiái az *il-*, *im-*, *ir-*). A morfológiai szabályok szóalakokra vonatkoznak, nem pedig szójelentésekre. Ennek következményeként is el kell vetnünk, hogy az antonímia viszony a jelentések között húzódik.

Az antonímia a wordnet terminológiája szerint nem szemantikai, hanem lexikális reláció. A wordnet tehát nem konceptuális oppozícióként értelmezi az antonimaságot, hanem asszociatív relációként, amely (a prototipikalitás-elmélet szerint) központi helyet elfoglaló lexikalizálódott fogalmak között áll fenn. A konceptuális oppozíció *indirekt antonímia* viszonyként jeleníti meg, az alábbiakban ismertetett módon.

A deskriptív melléknevek tipikusan kétpólusú attribútumok értékeit rendelik hozzá a főnévi fejekhez, ellentétes értelmű melléknevek egy adott attribútum ellentétes értékeit fejezik ki. Ezek az antonímia viszonytal rendelkező melléknevek kitüntetett szereppel rendelkeznek az attribútumértékek halmazában, azaz abban a dimenzióban, melyre vonatkozóan értékeket jelölnek. Ezek a speciális helyzetű, a tartománynak mintegy a két pólusát lexikalizáló melléknevek lesznek az adott dimenzió *központi* vagy *fokális synsetei*,⁵² melyek *near_antonym* relációval vannak összekötve. Köréjük csoportosulnak az adott dimenzió egyéb melléknevei, melyeknek nincsen antonim párjuk, azaz nincs *közvetlen antonimájuk*. Ezeket az ún. *szatelit synseteket* jelentéshasonlóság alapján, a hozzájuk közelebb álló fokális synsethez kapcsoljuk *similar_to* relációval, így a szatelit synseteknek is lesz – a hozzájuk kapcsolódó fokális synseten keresztül – közvetett antonimájuk. Ezt a clusteres szerkezetet szemlélteti a következő ábra (1).

⁵² A wordnet terminológiájában a synset szinonimacsoportot jelent, itt csak az egyszerűség és a hagyomány kedvéért tartjuk meg az elnevezést, bár a szinonimák, mint a következőkben látni fogjuk, nem a synsetben, hanem körülötte helyezkednek el.



1. ábra: A wordnet kétpólusú melléknévi szerkezete

Vannak olyan clusterok, melyeket mintegy mesterségesen egészítettek ki egy fokális synsettel, miután a két pólus közül az egyikhez nem találtak sem közvetlen, sem közvetett antonimát, de magának a dimenzióknak a létjogosultsága nem volt megkérdőjelezhető. Ilyen például az *angry* melléknév esete, mely ezáltal az *angry-unangry* dimenzióban az egyetlen valóban lexikalizált fokális synset.

Miután bemutattuk a wordnet melléknévi struktúrájának alapjait, rátérhetünk a HuWN melléknévi rendszerének tárgyalására.

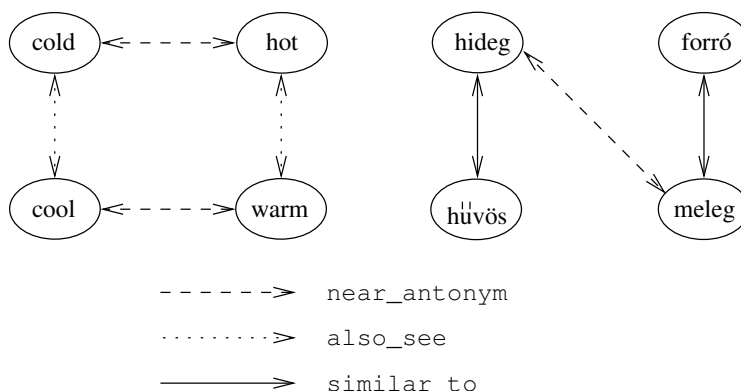
3 Melléknevek a Magyar WordNetben

A Magyar WordNet melléknévi része nem a PWN-nek egyszerűen magyarra fordított változata, elkészítése körültekintő munkát igényel. Erre elsősorban a két nyelv közötti lexikai, ill. asszociációs különbségek miatt van szükség, emellett igyekszünk a PWN-ben előforduló következetlenségeket kiküszöbölni. Ebben a fejezetben csak deskriptív melléknevekkel foglalkozunk.

3.1 Nyelvi sajátosságok

Mint azt a 2. fejezetben, a PWN melléknévi rendszerének ismertetésekor láttuk, az antonímia reláció konkrét szavak, nem pedig az általuk jelölt konceptuális tartalmak között állhat fenn. Ezt szem előtt tartva nem meglepő, hogy a HuWN struktúrája néhány kisebb eltérést mutat a PWN struktúrájához képest. Az alábbi (2) ábra egy ilyen esetet szemléltet.⁵³

⁵³ Az ún. *also_see* reláció kapcsolja össze azokat az egymáshoz nagyon hasonló dimenziókat meghatározó fokális synseteket, melyek a hagyományos wordneti értelemben egyetlen közös



2. ábra: A hideg-meleg tartomány nyelvi különbözősége

Konceptuálisan csak egy ellentétpár létezik ebben a dimenzióban: a *hideg* és a *meleg* tartomány. Lexikális szinten azonban az angolban két oppozíció is jelen van: *cold-hot*, ill. *cool-warm*. Mind a négy melléknévnek találunk lexikalizálódott magyar megfelelőt (*hideg, forró, hűvös*, ill. *meleg*), de a közöttük lévő relációkat nem vehetjük át, ugyanis a magyarban csak egyetlen ellentét realizálódott. Az angol példa alapján feltételezhető *hideg-forró* pár csak a *forró* szó irányából létezhetne, hiszen a *forró* szóról egy magyar nyelvi beszélő a *hideg* (esetleg a *jéghideg*) szóra asszociál, míg a *hideg* esetében a *melegre*.

Magában a szókészletben is mutatkoznak eltérések. Vannak az angolban (így a PWN-ben is) olyan melléknévek, melyekhez nincs megfelelő magyar melléknév. Ez az eset többféleképpen állhat elő. Egyfelől eredhet ez a lexikalizálódott magyar megfelelő teljes hiányából (pl. *unattractive* – 'nem vonzó'). Másfelől egy, az angolban melléknévvel kifejezett tartalom a magyarban akár más szófajú szóként is lexikalizálódhat, pl. *afraid* (mn) – *fél* (ige). Az ilyen típusú megfelelések jelölésére vezettük be újonnan az *eq_xpos_synonym* relációt, mely a szófaji határokon átvélő szinonimitást hivatott jelezni.

Egy-egy új szónak a lexikai adatbázisba való felvételekor több szempontot is meg kell vizsgálni. Egyfelől meg kell győződni arról, hogy valóban létezik ilyen szó, és ezáltal a szótárba való beillesztése indokolt. Másfelől meg kell határozni az adott szó szófaját. Ezekben a vizsgálatokban támpontot nyújthatnak a hagyományos szótárak (pl. a Magyar Értelmező Kéziszótár (ÉKSZ)). Azonban egy itt elért negatív eredmény még nem jelenti automatikusan egy szó felvételének kizárását, ebben az esetben segítségünkre lehetnek korpuszok is (pl. a Magyar Nemzeti Szövegtár [4]).

A szófaj meghatározása még akkor is lehet problémás, ha az adott szó szerepel egy hagyományos szótárban. Az *alvó* szó az ÉKSZ szerint főnév vagy melléknévi igenév lehet, 'aki éppen (javában) alszik', valamint 'hálószoba' jelentéssel. Érvelésünk szerint vulkánokra vonatkoztatva ez egy lexikalizált melléknév. A melléknéveknek más szófajoktól való elkülönítésére tesztek adhatók, melyek a szótárszerkesztők segítségére lehetnek [2], [3].

synsetbe kerülnének, de egy antonímia reláció révén saját „fokális synset”-ségük kétjogosultságot nyert.

A teljesség igénye nélkül bemutatunk néhány tesztet, melyek a melléknevek és a melléknévi igenevek megkülönböztetésére szolgálnak:

-- Állítmányi szerepben csak melléknevek állhatnak, pl. *ez a hír megdöbbsentő* szemben azzal, hogy **ez a hír Pétert megdöbbsentő*.⁵⁴

-- Az alapige vonzatait a melléknevek nem tarthatják meg (a szabad határozó ez alól kivétel), pl. a *Pétert megdöbbsentő hír*, illetve **ez a hír Pétert megdöbbsentő*.

-- Csak melléknevek fokozhatóak, melléknévi igenevek nem, nem mondhatjuk, hogy *Pétert megdöbbsentőbb hír*. (További tesztekhez ld. [3])

Látható, hogy a tesztek nem minden esetben megbízhatóak, csupán támpontot adnak. Ezek alapján az *alvó* szót még nem tekinthetnénk melléknévnek. A következő teszt indokolja döntésünket, melyek a lexikalizált jelzős szerkezetek (pl. *bontott tégla*, *vágott virág*) felismerését segítik.

-- A szerkezet jelentése sajátos (a *bontott tégla* kifejezés nem egyenértékű azzal, hogy 'tégla, melyet (le/szét-stb) bontanak/bontottak').

-- A jelzett szó elveszti élhangsúlyát ('*bontott tégla*, nem pedig '*bontott 'tégla*).

-- További jelzők csak a szerkezet egészére vonatkozhatnak, ezért egyrészt nem szakíthatják meg a szerkezetet, másrészt nem értelmezhető kizárólag a jelzett szóra vonatkoztatva (**bontott piszkos tégla*, **piszkos*, *bontott tégla*).

-- A jelzős szerkezetnek megfelelő névszói állítmányú mondat agrammatikus (**Az udvar közepén levő tégla bontott*).

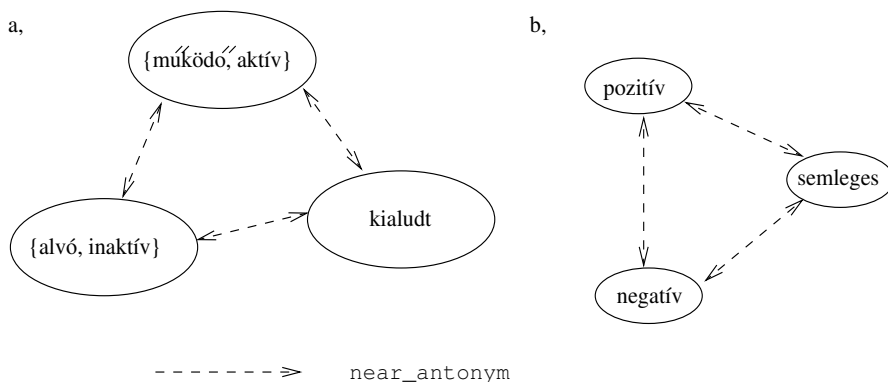
Ezek alapján az *alvó* szó melléknévként való felvétele megalapozott, hiszen érvelésünk szerint az *alvó vulkán* lexikalizálódott jelzős szerkezet. Nem mondhatjuk, hogy **a szigeten álló vulkán alvó*, sem azt, hogy **alvó nagy vulkán*, és a szerkezet egyetlen egységet alkot a hangsúly szempontjából.

Hasonló gondolatmenet alapján beillesztésre kerültek a wordnetbe pl. az agyhalott, ázott melléknevek is, melyek egyébként sem az ÉKSZ-ben, sem A magyar nyelv nagyszótárában nem szerepelnek.

3.2 A tipikustól eltérő dimenziók

A Princeton WordNetben vannak olyan *near_antonym* relációval rendelkező melléknevek is, melyek nem egy bizonyos kétpólusú dimenzió valamely végpontját jelölik. Az alábbi ábra (3) két ilyen esetet szemléltet.

⁵⁴ A * jelentése itt: rosszulformált.



3. ábra: A tipikustól eltérő melléknévi dimenziók

Ezekben a dimenziókban a központi synsetek mintegy háromszöget alkotva állnak egymással ellentétben. A PWN alapján azt feltételezhetnénk, hogy itt többdimenziós tartományokról van szó, melyeknek egyaránt három kitüntetett „végpontjuk” van. Amellett érvelünk, hogy a wordnetben nem helyes ez az ábrázolási mód.

A wordnet melléknévrendszere nagyrészt a lexikalizált oppozíciós párokra épül, ezek kétpólusú struktúrákat határoznak meg. Felmerülhet az az igény, hogy ez a kétpólusú rendszert kiterjeszthető legyen úgy, hogy a fenti példák beilleszthetők legyenek a wordnetbe anélkül, hogy a „kivételes eset” bélyeget kellene magukra venniük.

Mint ismeretes, két adott szó esetén az antonímia reláció felvételének szükséges feltétele a kettejük közti asszociáció megléte. Vizsgáljuk meg ebből a szempontból a (3.b) ábrán látható példát.

A *pozitív-negatív* pár egyértelmű ellentétpár, mely a szóban forgó dimenzió⁵⁵ két végpontját realizálja. De mi a helyzet a *semleges* melléknévvvel? A *pozitív-semleges* pár oppozíciós viszonya hasonló módon zárható ki, mint az előző pontban a *hideg-forró* esetében: habár a *semleges* szó asszociálhatja a *pozitív* szót, fordítva ez nem áll fenn. A *negatív-semleges* pár helyzete még bizonytalanabb. Hiába érezhetünk egyfajta ellentétes viszonyt bármely kettő között, a lehetséges három antonímia reláció közül csak egyetlenegy vehető fel a wordnetben.

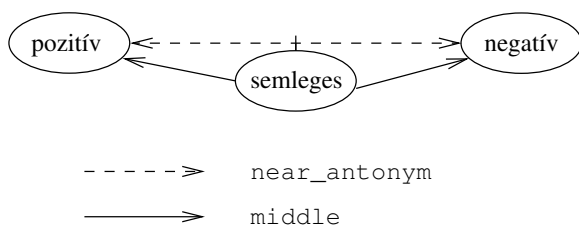
Ebből a konkrét példából is jól látható, de elvi megfontolások alapján is elvethető, hogy a wordnet filozófiájához híven ki lehessen terjeszteni a *near_antonym* reláció használhatósági körét. Egy adott szó egy adott jelentésben csak egyetlen ellentéte vonatkozó asszociációt idézhet fel, ellenkező esetben ezt az asszociációs módszert eleve nem lehetett volna alkalmazni a wordnet melléknévi struktúrájának kialakítására.⁵⁶ Ekkor ugyanis nem lehetett volna a dimenziók végpontjait egyértelműen és megbízhatóan meghatározni. Nézetünk szerint nem a wordnet alapjainak megváltoztatásában kell keresni a kiutat, hanem egyfajta továbbfejlesztésében. A következőkben erre kínálunk megoldást.

A *pozitív* és *negatív* melléknévek közötti antonímia reláció létjogosultsága vitathatatlan, ez a két melléknév alkotja egy dimenzió két végpontját. Ez a dimenzió azon-

⁵⁵ Valójában jelentéstől függően több dimenzió is szóba jöhet, de ez a gondolatmenetünk szempontjából nem jelent problémát.

⁵⁶ Fontos kiemelni, hogy egy adott jelentésről van szó. A többszörös asszociációnak éppen a több jelentés megkülönböztetésében lehet szerepe (ld. a 2. fejezetben).

ban eltér a (metaszinten) prototipikusnak mondható tartományoktól, hiszen nem jelennek meg benne fokozatok, sem a két végpont között, sem körülöttük nem helyezkednek el jelentésükben csupán graduális különbségeket mutató szavak. Ezzel szemben a van lexikalizált kifejezés, mely ennek a durva fokozatú skálának pontosan a közepét nevezi meg, ez a *semleges* melléknév (nem kapcsolható *similar_to* relációval egyik végponthoz sem). A „szabályos” dimenziókban is léteznek olyan szavak, melyek a tartományuk közepére utalnak (a *hideg-meleg* dimenzióban pl. a *langyos*), de ezek vonatkozása bizonytalan, míg a *semleges* határozott. Ezt a kitüntetett szerepet javaslatunk szerint nem az elméleti síkon is megtámadható *near_antonym* reláció felvételével kell a tartomány két központi végpontjához kapcsolni, hanem egy új, a dimenzió határozott középpontját jelző *middle* relációval. Ezáltal az ellentétviszonyok nem alkotnának egy háromszöghöz hasonló szerkezetet, mint ahogy az a PWN rendszerében leolvasható, hanem a többi tartományhoz hasonló szerkezetet hoznának létre. A különbség csak annyi, hogy ebben a dimenzióban három kitüntetett szerepű synset is van, egyfajta elfajuló háromszöget alkotva (4. ábra).



4. ábra: A middle reláció

A *middle* reláció definiálásával nemcsak a „mindenki mindenkinek ellentéte” anomáliás helyzetet lehet kiküszöbölni, hanem a szemantikai viszonyokat is jobban tükrözi a rendszer. Szükséges még megjegyezni, hogy ezzel nem csak egy egyedi, elszigetelt esetet oldottunk meg. A *middle* reláció minden olyan dimenzióban felvehető (és felveendő), melyek középső pontjára (esetleg pontszerűnek tekintett belső intervallumára) lexikalizálódott megnevezés létezik az adott nyelvben, mint pl. *alsó-felső-középső*. (Azonban ha egy melléknév nem pontosan a középpontot fejezi ki, akkor az eddigiekben megszokott *similar_to* relációval kapcsolódik az egyik végponthoz.)

Első közelítésben a (3.a)-ban bemutatott példa esetében is alkalmazható lenne a *middle* reláció. A dimenzió két ellentétes végpontja a {*működő, aktív*}, ill. a *kialudt*, az {*alvó, inaktív*} pedig a középső tartományt jelöli. Itt azonban a középpont nem pontszerű, és nem is tekinthető annak. Sőt, ami ennél meglepőbb, az {*alvó, inaktív*} synset akár *similar_to* viszonyban lehetne, hiszen az *alvó* jelző 'most éppen nem működő' vulkánra utal, tehát jelentése közelebb áll a {*működő, aktív*} synsethez. Azonban nem hagyhatjuk figyelmen kívül az *aktív* és *inaktív* szavak között feszülő ellentétet.

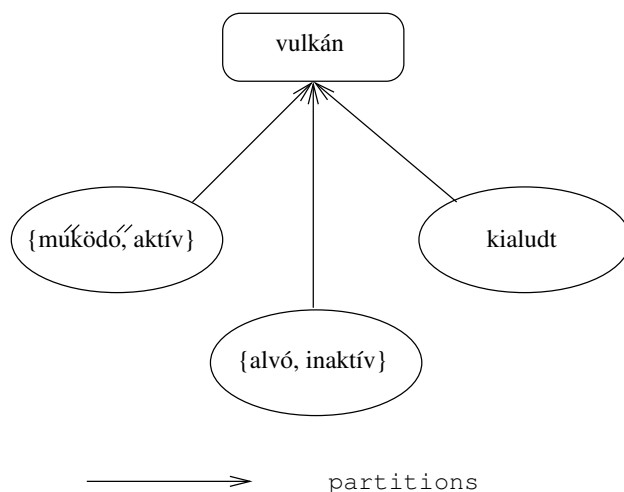
A {*működő, aktív*}, *kialudt* és {*alvó, inaktív*} synsetek által meghatározott tartomány nemcsak az egyes elemek között húzódó hasonlóság és ellentét kettőssége miatt tér el az eddig tárgyalt dimenzióktól. Ezek a melléknevek szemantikailag megszorítják, mire vonatkozhatnak: konkrét esetünkben *vulkánokra*. A wordnetnek számot kell tudnia adni erről a szemantikai kapcsolatról is. A PWN és a BalkaNet ezeket a mel-

lékneveket az antonímia relációval kapcsolja össze, és sokszor az általuk kizárólagosan módosított főnévvel nem is jelez kapcsolatot.

A megoldás ismertetése előtt megjegyezzük, hogy most sem csupán egyetlen tartományban lép fel ez a szituáció. Gondoljunk például a kizárólag *növényekre* jellemző *egynyári-kétnyári-évelő* tulajdonságokra, melyek a PWN-ben szintén háromszöget alkotó *near_antonym* relációval szerepelnek.

A szóban forgó melléknevek *particionálják* a főnév terjedelmét, azaz a főnév alá eső tárgyak halmazát az általuk jelölt tulajdonságok mentén diszjunkt részhalmazokra osztják. Az általunk felvett új reláció neve éppen ezért *partitions* (ld. 5. ábra).

A *partitions* hasonlít a wordnet *category_domain* relációjára, de nem azonos vele. A *category_domain* reláció azt a célt szolgálja, hogy megadja, egy bizonyos jelentés milyen témakörhöz kapcsolódik (pl. {*monovalent*:2 'egyvegyértékű, monovalens'} – {*chemistry*:1 'kémia, vegyészet'}), de nem mond semmit arról, hogy az adott melléknev mely főnevet módosíthatja, akkor sem, ha az kizárólagos. A *partitions* reláció definiálásával tehát teljesebbé és pontosabbá tehetjük a szemantikai kapcsolatok ábrázolását a wordnetben.



5. ábra: A partitions reláció

4 Értékelés

A HuWN melléknévi részének építése közben elért eredményeink elsősorban nem mennyiségi, hanem minőségi jellegűek. Habár a 3.1 fejezetben említettük, hogy a wordnet szókészlete bővebb (vagy bővebbé tehető), mint a hagyományos szótáraké, fontosabbnak tartjuk a minőség javítását. A 3.2 fejezetben megmutattuk, hogy hogyan lehet a wordnetben meglévő, a wordnetben elvileg tarthatatlan hármas melléknévi ellentéteket kiküszöbölni. A rendszerben lévő következetlenség felszámolásán túl a látszólag azonos, de két különböző probléma megoldásával, azaz két új reláció

definiálásával egy eszközt adtunk, mellyel a wordnetben a szemantikai viszonyok árnyaltabban ábrázolhatók.

Bibliográfia

1. Fellbaum, C.: WordNet: An Electronic Lexical Database. MIT Press (1998)
2. Kiefer, F.: Jelentéelmélet. Corvina, Budapest (2000)
3. Komlósy, A.: Régensek és vonzatok. In: Kiefer, F. (ed.): Strukturális magyar nyelvtan I. Mondattan. Akadémiai Kiadó, Budapest (1992) 299—527
4. Magyar Nemzeti Szövegtár. <http://corpus.nytud.hu/mnsz>