

Magi

edík

31. března 2017

Obsah

I	Teoretické pojednání	2
1	Princetonský WordNet	3
1.1	Motivace vzniku	3
1.1.1	Strojově čitelné slovníky	5
1.1.2	Od slovníků k WordNetu	5
1.2	K vlivu psycholingvistiky na organizaci WordNetu	6
1.3	Organizace WordNetu	7
1.3.1	Struktura hesla	9
1.4	Sémantické vztahy WordNetu	9
1.4.1	Frekvenční distribuce sémantických vztahů ve Word- Netu	10
1.4.2	Substantiva	10
1.4.3	Antonymie	12
2	Další wordnety	13
2.1	EuroWordNet	13
2.2	BalkaNet	13
	Seznam literatury	14

Část I

Teoretické pojednání

Kapitola 1

Princetonský WordNet

Princetonský WordNet je prvním wordnet vůbec. Vznikal na univerzite v Princetону pod G. A. Millerem od poloviny 80. let 20. století. Vzhledem k tomu, že byl prvním wordnetem, bylo k němu referováno jako k WordNetu, bez přívlastku. Ačkoliv tento stav v podstatě přetrvává dodnes, oproti době jeho vzniku se situace změnila, vzniklo několik dalších wordnetů a nastala tudíž potřeba je rozlišit. V anglickém prostředí se obvykle pojmem WordNet míní ten princetonský a všechny ostatní wordnety mají přívlastek. Příkladem necht je Balkanet či Eurowordnet. Ačkoliv v mezinárodním prostředí je obvyklé přívlastek „princetonský“ používat, bude tato práce pracovat s následujícím rozlišením:

- *WordNet* (ve významu princetonský WordNet)
- *wordnet* (ve významu obecné sémantické sítě založené ideou, obsahově či strukturálně na WordNetu)
- konkrétní wordnety, např. *Balkanet*

a samozřejmě bych to actually mohl do-drzovat, tohle jsem si vymyslel až po na-psani tehle kapitoly, lol

1.1 Motivace vzniku

Od počátků snah o zpracování přirozeného jazyka (NLP, natural language processing) bylo nutné poskytnout programu data o lexiku ve zpracovávaném textu, ať už ona data byla jakákoliv. Kupříkladu pro překlad se mělo za to, že stačí ekvivalentní dvojice ve zdrojovém a cílovém jazyce, později se přidal kontext v případě statistického strojového překladu spolu s dalšími informacemi, jako je například slovní druh. Tradičně se lexikální materiál ukládá způsobem nikoliv diametrálně odlišným od papírových slovníků určených pro lidské uživatele. Ty obvykle obsahují abecedně (či podle jiného indexu) seřazené jednotlivé záznamy s potřebnými informacemi o slovech, z nichž pak program může čerpat při zpracování textu.

cit?

Jak uvádí Pala; Ševeček, (2013), uspořádání lexikálního materiálu v takovéto formě je sice vhodné pro člověka, ale nikoliv pro strojové zpracování, a to z několika důvodů. Kromě toho, že vyhledávání v abecedním seznamu je relativně pomalé, struktura tradičního slovníku kvůli onomu abecednímu řazení inherentně vzdaluje slova, jež člověk chápe jako nějakým způsobem blízká. Tato blízkost může vyplývat ze vztahu volné synonymie, antonymie, podřazenosti, nadřazenosti, etc. Pokud si tedy například uživatel výkladového slovníku nepříliš obeznámený s daným jazykem vyhledá určité heslo, dozví se sice pravděpodobně jeho význam, ale nebude schopen své znalosti prohlubovat dále zjištěním, kupříkladu jaké je slovo odpovídá opačnému významu.

nejaka citace?, dohledat neco, jak takový slovníky byly uloženy...

Dalším všeobecným problémem při využití tradičních slovníků k počítačovému zpracování jazyka je fakt, že lexikografové předpokládají u uživatele slovníku značné encyklopedické znalosti. Zařazují tak do slovníku jen informace dle jejich názoru důležité pro rozlišení (*differentia specifica*) a zařazující do kontextu či přiřazující k určité nadřazené třídě objektů (*genus proximum*). Vyhledá-li si tedy člověk ve Slovníku spisovného jazyka českého heslo *vlk*, zjistí následující:

citace

vlk: psovitá šelma šedě (n. šedožlutě) zbarvená, žijící v Evropě, Asii a v Sev. Americe

Definice a priori předpokládá, že uživatel je obeznámen s tím, co je šelma a co je pes. Pokud takovou znalostí neslyne (což je vcelku představitelné například u cizince), je nucen si tato slova ve slovníku najít a podívat se na jejich definice (pomiňme nyní netriviální úkol převést slovo *psovitá* na základní tvar *pes*). Pokud nerozumí definicím ani nadřazených slov, musí pokračovat v hierarchii dále a dále.

Z uvedeného případu plyne, že jakkoliv je možné správným vyhledáváním hyperonym¹ dospět k tomu, že *vlk* je konkrétní entita našeho vesmíru, živá bytost o čtyřech končetinách, savec nějakým způsobem příbuzný se psovi, má šedou srst etc., je takový proces dosti komplikovaný. Příklad s cizincem se sice nemusí zdát zcela relevantní, protože se dá předpokládat, že daný člověk má, byť v jiném jazyce, stejné základní znalosti předpokládané lexikografy jako člověk český. Situace je však dramaticky jiná u počítače. Na rozdíl od člověka totiž počítač nemá žádné předchozí znalosti, tudíž musí projít celým procesem popsaným výše, aby byl schopen kupříkladu určit, že *vlk* může umřít (ježto je živá bytost). Protože však tradiční slovníky typu SSJČ byly vytvářené pro papírové médium, neobsahují žádné propojení ve stylu *toto je odkaz na hyperonym*, a počítač tudíž jen těžko může zjišťovat, na které vlastně slovo se to má podívat, aby se dobral podstaty pojmu *vlk*.

¹nadřazené slovo

1.1.1 Strojově čitelné slovníky

V zájmu automatizace vyhledávání ve slovníku vznikaly tzv. strojově čitelné slovníky², což je pojem souhrnně označující lexikální databáze. Podle množství informací, které taková databáze obsahuje, pak lze tyto dělit na slovníky, taxonomie a ontologie. Je evidentní, že obyčejný slovník neobsahuje oproti tradičnímu papírovému slovníku navíc žádné metainformace, takže je počítač při jeho užívání v podstatě omezen na elektronický listovač (Miller et al., 1990).

Míra, jakou se strojově čitelný slovník odlišuje od pouhé zdigitalizované formy papírového slovníku a přiblíží se k pokročilé lexikální databázi, lze vyjádřit v několika stupních. V případě, že slovník má jednotlivé významy uspořádány v hierarchii dle nadřazenosti–podřazenosti, lze jej označit za taxonomii, tedy systém s hlubší strukturou než pouze abecedním řazením hesel.

nejakej
link, kde
budou vý-
znamy/sen-
ses
vysvetleny

Dalším stupněm už je skutečná lexikální databáze, která má jednotlivé významy propojeny rozličnými vztahy, počínaje onou základní hyperonymií a hyponymií a pokračuje kupříkladu vztahy meronymie³ či antonymie⁴. Kromě vztahů mezi významy bude taková lexikální databáze obsahovat zřejmě i další informace, tedy nějaké kategorie slov, jejich popis, etc. Databáze tak popsaných významů propojených sémantickými vztahy může být nazývána ontologií. (Garshol, n.d.)

1.1.2 Od slovníků k WordNetu

Výše uvedená opozice papírového slovníku a ontologie ilustruje rozdíly tradičního slovníku a počítačově zpracovatelné lexikální databáze. Už ze samotného významu takové databáze je evidentní jeden klíčový rozdíl – tradiční slovníky, jsoucne řazené abecedně, od sebe oddalují některá hesla, jež by bylo vhodné mít pohromadě (Pala; Ševeček, 2013). Příkladem budiž *kostra* a její části, např. *lebka*. V SSČ⁵ i SSJČ se u *lebky* uvádí, že jde o *kostru hlavy*. Lze tedy s jistou rezervou tvrdit, že heslo obsahuje své holonymum⁶, opačně to však již nefunguje. Z celkem evidentních důvodů nejsou u hesla *kostra* uvedeny všechny její části. Tento příklad příhodně ukazuje i jistou nesystémovost tradičních slovníků, která je pro počítačové zpracování fatální.

Naznačeny tedy byly vlastnosti, jež by databáze významů měla oproti tradičnímu slovníku mít, aby byla použitelná pro počítačové zpracování přirozeného jazyka. Především jde o systémovost vztahů. Hypero/hyponymie je vztah oboustranný, tudíž by mělo být možné se stejnou cestou dostat od

²machine readable dictionary

³vztah je částí, tedy např. *dveře* je meronymem *trolejbusu*

⁴protikladu

⁵Slovník spisovné češtiny

⁶vztah opačný k meronymii; tedy např. *dům* je holonymem pro *okno*, *dveře*, *práh* etc.

nadřazeného slova k podřazenému a naopak. Dále jsou podstatné pojmenované sémantické vztahy mezi slovy. Díky nim je totiž možno jednoznačně určit, které slovo (či slova) je v takové databázi konkrétnímu slovu nadřazené, které je jeho specifikací, označením jeho částí, etc.

S touto myšlenkou vznikl WordNet - lexikální síť provázaná sémantickými vztahy, která by dle poznatků psycholingvistiky odrážela uspořádání lexikálního materiálu v lidském mozku (o tom v dalších kapitolách). (Pala; Ševeček, 2013) Zde by bylo na místě poznamenat, že ačkoliv se tak z odstavců výše může čtenáři jevit a i všeobecně je to často tvrzeno, WordNet není ontologií v pravém slova smyslu, protože něco něco.. https://en.wikipedia.org/wiki/WordNet#WordNet_as_a_lexical_ontology

ref

a tady tomu vůbec nerozumím, ale přijde mi to relevantní

1.2 K vlivu psycholingvistiky na organizaci WordNetu

G. A. Miller, který je tvůrcem WordNetu, se po spolupráci s Chomským na základních kapitolách jeho *Handbook of Mathematical Psychology*, která se věnuje spíše syntaktickému popisu jazyka, se zaměřil společně s Johnsonem-Lairdem na výzkum, jakým způsobem je lexikální materiál uložen v lidském mozku. Tento přístup je označován jako psycholingvistika a jeho počátky jsou spojeny s průzkumem asociací a budování modelu mentálního slovníku člověka. Výchozí myšlenka, jež se odráží i ve způsobu organizace WordNetu, spočívá v tom, že slovní zásoba není organizována abecedně, jako tomu je v tradičních slovnících, ale spíše konceptuálně.

cit

Jednou z otázek tohoto směru bylo, jakým způsobem je organizována paměť. Aby člověk byl schopen určit pravdivostní hodnotu výroku **Kanárek může létat**, musí použít svou dlouhodobou paměť. Její organizace je pak možná (minimálně) dvěma způsoby. První, redundantní, by vypadal tak, že by u každého ptáka bylo uloženo, že je schopen létat. Druhý, již na první pohled výrazně méně náročný na úložný prostor, by příznak schopnosti létat měl uložený pouze u kategorie pták. V případě, že by bylo třeba zjistit, zda kanárek léta, by bylo nutno pak zapojit inferenční proces ve stylu *kanárek je pták, tudíž může létat*. (Collins; Quillian, 1969)

Jak Collins; Quillian, (1969) dále uvádí, lze předpokládat, že v případě prvního způsobu organizace paměti by člověk mohl kteroukoliv informaci o příznacích (vlastnostech) z paměti vyvolat za konstantní čas. Naproti tomu v případě způsobu druhého by extrakce příznaku z významu v hierarchii položeného výše měla trvat delší čas než extrakce příznaku přítomného přímo u významu, jenž je objektem věty. Důvodem by měla být nutnost zapojení inferenčního procesu.

Pokus, kterým podpořili Collins; Quillian, (1969) druhý, neredundantní, způsob organizace paměti, spočíval v tom, že testovací subjekty, dobrovolníci z řad zaměstnanců společnosti Bolt Beranek and Newman, měly určovat,

zda je výrok pravdivý, či nepravdivý. Měli tak činit co nejpřesněji a v co nejkratším čase, přičemž byla měřena rychlost jejich reakce. Ukázalo se, reakční doba při určování pravdivosti výroku *Kanárek umí létat*⁷ je delší než při určování pravdivosti výroku *Kanárek umí zpívat*⁸ a ještě delší při určování výroku *Kanárek má kůži*⁹. Důvodem pro tyto progresivní prodlevy podle nich právě byla zvětšující se vzdálenost od významu *kanárka* ke významu, u něhož byl uložen příslušný příznak, tedy *umí zpívat*, *umí létat*, resp. *má kůži*. Příznak *umí zpívat* totiž je pravděpodobně uložen přímo u *kanárka*, jelikož jej odlišuje od ostatních ptáků, zatímco příznak *umí létat* je obecným znakem ptáků, tudíž je uložen u významu *pták*. V poslední řadě pak příznak *má kůži* bude patrně uložen u významu *zvíře*, který je oněch tří v hierarchii nejvýše, a ze všech tudíž od *kanárka* nejdále.

Jelikož WordNet G. A. Millera je založen na podobném principu „lze cit
říci jistým způsobem odráží organizaci lexika v lidském mozku.“

1.3 Organizace WordNetu

Ve WordNetu lze nalézt informace autosémantikách, tedy substantivech, adjektivech, slovesech a příslovcích (Vossen, 1998). Synsémantika (např. předložky, spojky etc.) nebyla zahrnuta, jelikož se zdá, že jsou uložena odděleně od slov plnovýznamových.

Teorii, že jsou funkční slova uchovávána jako součást syntaktikonu, podpořil kupříkladu Garrett, (1982) při svém pozorování afatických pacientů. Stejně jako je výběr slovních druhů zastoupených ve WordNetu podložen předpokládanou organizací lexikálního materiálu v lidském mozku, je i samotné seskupování lexika podle slovních druhů (na rozdíl od abecedního řazení v tradičních slovnících) podloženo výsledky průzkumu dalšího aspektu jeho organizace. Fillenbaum; Jones, (1965) provedli asociační test na anglicky mluvících subjektech, kteří měli za úkol uvést první slovo, které je napadne v kontextu dobře známých slov, která jim byla předkládána. Tato slova náležela k různým slovním druhům a ukázalo se, že ve většině případů náleží asociované slovo ke stejnému slovnímu druhu jako slovo, které asociaci vyvolalo. Substantiva vyvolala asociaci na substantivum v 79 % případů, adjektiva v 65 % případů a slovesa v 43 % případů. Ačkoliv není zřejmé, jak je znalost o slovním druhu určitého slova získávána, lze z uvedených dat předpokládat, že slovní druh je vskutku primární organizační vlastností lexikálního materiálu v lidském mozku a informace o něm je snadno dostupná (alespoň intuitivně).

Seskupování konceptů podle slovního druhu a zřejmě navzdory snaze o ekonomii ukládání informací, kterou se lidský mozek vyznačuje, zprostředko-

⁷ angl. A canary can sing

⁸ angl. A canary can sing

⁹ angl. A canary has skin

cit:
https://word-
net.prince-
ton.edu/word-
net/man/wn-
g-
loss.7WN.html

```

graph LR
    entita --> struktura
    entita --> bytost
    struktura --> konzejni
    struktura --> bazmek
    konzejni --> dopr_prostredky[dopr. prostředek]
    konzejni --> bazmek
    bazmek --> sub_bazmek[sub bazmek]
    bazmek --> dopr_prostredky
    bytost --> zviru
    bytost --> olouk
    dopr_prostredky --> bicykl
    dopr_prostredky --> auto
    dopr_prostredky --> el
    sub_bazmek --> dopr_prostredky
    
```

Hierarchická organizace významů substantiv¹¹ ve WordNetu tak, jak byla naznačena na grafu 1.1 na straně 8, je podpořena i poznatkem, že lidé

¹⁰derivationally related form

¹¹ a možná i dalších slovních druhů

jsou schopni velice rychle zpracovávat anaforické a kataforické výrazy a komparativní konstrukce. Například ve větě *Vlastnil pušku, ale z té zbraně se nikdy nevystřelilo*. je každému čtenáři zřejmé, že výraz *ta zbraň* odkazuje k výrazu *puška*.¹² Co se zmíněného zpracovávání komparace týče, lze říci, že nelze porovnávat dvě substantiva, která jsou provázána sémantickým vztahem hyperonymie-hyponymie. Výrok *Puška je bezpečnější než zbraň* je zcela nesmyslný. (Pala; Ševeček, 2013)

1.3.1 Struktura hesla

1.4 Sémantické vztahy WordNetu

Jak bylo naznačeno výše, koncept WordNetu je založen na lexikální sémantice, tedy představě, že slovo je kombinací slovní formy a významu, nebo slovního významu. Slovní forma je projevem „fyzickým“, tedy je to vyřčená či napsaná instance významu. Jak je zjevné z přirozeného jazyka, nelze počítat s tím, že by zobrazení významu na formu bylo bijektivní, tedy každý význam byl namapován jedna ku jedné na slovní formu. Mapování je v přirozeném jazyce tzv. více ku více, tedy jedna forma může zastupovat více významů a jeden význam může být vyjádřen více formami. Je velmi časté, že jedna slovní forma zastupuje více významů. Kupříkladu slovní forma *koruna* může zastupovat význam měny, vrcholku stromu, vladařského odznaku, etc. Toto zobrazení se nazývá homonymií¹³. Možný je samozřejmě i opačný případ, kdy více významům slouží k vyjádření jedna slovní forma, což je nazýváno polysémií¹⁴. Příkladem může být forma *kolej*, již lze interpretovat jako referenci k stopě po voze, případně dvojici kolejnic jako vodící dráze pro dopravní prostředky a zároveň jako zařízení vysoké školy pro ubytování studentů.

cit. SSJC

cit SSJC

Miller et al., (1990) popisují výše naznačené vztahy pomocí takzvané lexikální matice. Ta názorně zobrazuje formy synonymní (F_1 a F_2) a formy polysémií (F_2):

tabulka z miller1990introduction pg. 4: <http://i.imgur.com/sohtwe5.png>

V dalších podkapitolách budou rozebrány podrobněji, nikoliv však vyčerpávajícím způsobem, sémantické vztahy konstituující Wordnet. Jelikož se sémantické vztahy pro jednotlivé slovní druhy liší, bude tato kapitola strukturována primárně právě podle slovních druhů a až sekundárně podle sémantických vztahů.

tady by
jeste slo
pokracovat
opisova-
nim dalsi
casti toho
clanku
(miller1990in-
troduction
pg 5 »>)

¹²Zajímavé je uvažovat o významu této věty při vypuštění deiktického *ta* – zdá se, že pak už anaforický odkaz k *pušce* nefunguje a z věty se stává jakési nesmyslné spojení dvou výroků – konkrétního (*Vlastnil pušku* a zcela obecného (a evidentně nepravdivého) *ze zbraně se nikdy nevystřelilo*.

¹³totožnost formy

¹⁴víceznačnost

1.4.1 Frekvenční distribuce sémantických vztahů ve Word-Netu

1.4.2 Substantiva

Synonymie

Synonymie je centrálním organizačním vztahem pro substantiva ve Word-Netu. Na praktických aplikacích je tento jev nejlépe pozorovatelný, jelikož při vyhledání konkrétní formy je uživateli obvykle nabídnut výběr z jednotlivých významů dané formy. Aby byly od sebe významy oné formy odlišitelné, nabídka běžně sestává z výpisu skupin forem (tzv. synsetů, o tom později), přičemž každá skupina náleží k jednomu významu a obsahuje formy danému významu přiřazené. Kupříkladu při vyhledání slova *kolo* tak je uživatel konfrontován s několika skupinami, které obsahují zhruba následující:

- *kolo* (1),
- *jízdní kolo* (1), *bicykl* (1), *kolo* (2),
- *kružnice* (1), *cívka* (1), *kolo* (3),

přičemž čísla (zde) v závorce značí index významu dané formy v daném synsetu. Reprezentace v různých aplikacích a různých wordnetech se liší (standardem bývá číslo významu psát za dvojtečku), koncept však zůstává neměnný.

Navzdory zdánlivé jednoduchosti výše uvedeného konceptu je všeobecnou otázkou, jak synonymii pojímat. Striktní teorie (obvykle připisovaná Leibnizovi) praví, že dvě slova jsou synonymní, pokud se jejich záměnou nikdy nezmění pravdivostní hodnota výroku. Lingvistickou interpretací tohoto poněkud matematicko-logického výroku může pak být, že synonymní dvě slova jsou v případě, že se jejich záměnou nikdy neporuší význam (zhruba ona pravdivostní hodnota) a gramatičnost výroku. Je nasnadě, že takto striktně synonymní slova budou pospolu v jazyce těžko přežívat, jelikož je dokázáno, že jazyk tíhne k ekonomičnosti, která by takovým soužitím dvou slov byla hrubě porušena. Pravděpodobně jedinými obecně uznávanými synonymy jsou obvykle dvojice cizího slova a domácího slova, například *internacionální* a *mezinárodní*. S relativně vysokou jistotou lze tvrdit, že jejich záměnou se nikdy pravdivostní hodnota výroku nezmění, stejně tak jako jeho gramatičnost. Stále však zůstává ve hře stylistika, která může být podobnou náhradou narušena (např. z důvodu cílové skupiny čtenářů či stylistické příznakovosti jednoho ze slov (srov. *zajímavý* a *interesantní*)). Co se tendence k ekonomičnosti jazyka týče, lze předpokládat, že v těchto případech převládá potřeba synonym k eliminaci opakování určitých slov v textu a tím zajištění jeho stylistické uhlazenosti.

Volnější interpretace synonymie také počítá s kontextem. Dvě slova jsou synonymní, jsou-li bez způsobení škod nahraditelná alespoň ve stejném kontextu. Jako příklad mohou posloužit formy **board** a **plank**. V kontextu dřevařství mohou tyto dvě formy pravděpodobně bez problému být nahrazeny jedna za druhou, ovšem v případě, že je forma **board** použita ve významu **committee**, těžko ji lze nahradit za formu **plank**, neboť by se věta obsahující takové nahrazení stala zcela nesmyslnou.

najít české příklady

Bylo by nanejvýš logické považovat synonymii za vztah diskrétní, tedy že dvě formy buďto synonymní jsou, či nejsou. Z logického hlediska to nepochybně z již uvedeného vyplývá, ovšem lingvisticko-filosofický náhled vycházející z poznatků reálného jazyka se na tuto problematiku poněkud liší. Jak bylo dokázáno, synonymie v striktním slova smyslu je velice vzácná. Její volnější interpretace je značně častější, ale také výrazně vágnější – kontext, v němž dvě formy synonymní jsou může být velmi široký, či naopak velice úzký. Záměna některých dvojic (obecně *n-tic*, lze ale předpokládat, že mnoho forem nebude mít dvě další synonymní) může měnit stylistiku a význam výpovědi více či méně, přičemž ony dvě formy stále dle daných kritérií lze považovat za synonymní. Nelze tedy než vyvodit, že synonymie, minimálně z pohledu přirozeného jazyka, je jevem graduálním, a některé formy jsou tak *synonymnější* než jiné.

Zaměnitelnost forem podporuje ještě jeden koncept, na němž je WordNet postaven, a to fakt, že jednotlivé významy jsou seskupovány podle slovních druhů. Jak již bylo řečeno, tento systém vede k jisté redundantnosti, jelikož zvláště v syntetických jazycích, jako je kupříkladu angličtina, lze nalézt mnoho případů, kdy identická slovní forma zastupuje více slovních druhů. Významy, které taková slovní forma zastupuje (napříč slovními druhy), mohou být velice blízké, nikdy však nebudou stejné (nelze říci, že význam slovesa *run*¹⁵ a substantiva *run*¹⁶ je identický). Jejich záměnou by se sice nestalo vůbec nic, jelikož čtenář či posluchač textu, v němž taková záměna nastala, by automaticky formu interpretoval ve prospěch správného slovního druhu, avšak pokud by slovní druh byl nějakým způsobem „vynucen“ (necht nyní čtenář pomine úvahy, jakým způsobem lze „vynutit“ slovní druh formy), stala by se výpověď zcela negramatickou a nesmyslnou.

check, nekecam?

Seskupování významů podle slovních druhů a seskupování forem dle vztahu synonymie se tedy zdá v případě lexikální databáze určené pro strojové zpracování jako vhodným konceptem. Oproti tradičním slovníkům se totiž počítačově zpracovávaná lexikální databáze nemusí potýkat s problémem lidského faktoru – jednotlivé synonymické řady je stroj schopen prohledávat, na rozdíl od člověka, velice účinně, a nahradí tak v případě, že WordNet používá člověk, neúčinné lidské procházení restříkovaného obsahu.

psáno v chvatu, mohlo by se to možná trochu uhladit...

¹⁵běžet

¹⁶běh

1.4.3 Antonymie

Kapitola 2

Další wordnety

Podle vzoru princetonského WordNetu začaly postupně vznikat i další sémantické sítě založené na stejném konceptu. Tyto sémantické sítě se samozřejmě svou strukturou do větší či menší míry liší, hlavním kritériem pro to, aby mohly být považovány za wordnet, je to, aby obsahovaly synsety a hyponyma .

nakou
kurva ci-
taci

<http://globalwordnet.org/wordnets-in-the-world/>

2.1 EuroWordNet

2.2 BalkaNet

Seznam literatury

- Collins, Allan M; Quillian, M Ross, 1969. Retrieval time from semantic memory. *Journal of verbal learning and verbal behavior*, 8(2): s. 240–247.
- Fillenbaum, Samuel; Jones, Lyle V, 1965. Grammatical contingencies in word association. *Journal of Verbal Learning and Verbal Behavior*, 4(3): s. 248–255.
- Garrett, Merrill F, 1982. Production of speech: Observations from normal and pathological language use. *Normality and pathology in cognitive functions*, s. 19–76.
- Garshol, Lars Marius. Metadata? Thesauri? Taxonomies? Topic Maps! 2004—10—26)[2010-6—19] <http://www.ontopia.net/topicmaps/materials/tm-VS—thesauri.htm1>. Dostupné také z: <http://www.ontopia.net/topicmaps/materials/tm-vs-thesauri.html#N773>.
- Miller, George A et al., 1990. Introduction to WordNet: An on-line lexical database. *International journal of lexicography*, 3(4): s. 235–244.
- Pala, Karel; Ševeček, Pavel, 2013. Česká lexikální databáze typu WordNet. *Feedback*, 48(A47).
- Vossen, Piek, 1998. Introduction to eurowordnet. *Computers and the Humanities*, 32(2-3): s. 73–89.