

PLIN021 Sémantická analýza v praxi

OP VK Mezi bohemistikou a informatikou
www.projekt-inova.cz

Zuzana Nevěřilová
xpopelk@fi.muni.cz

Centrum zpracování přirozeného jazyka, B203
Fakulta informatiky, Masarykova univerzita

29. října 2013

Vágnost a subjektivita

Víceznačnost

Je víceznačnost problém?

Zkoumání významu

Význam zkoumáme z pohledu počítačové lingvistiky. Chceme se dobrat nějakého formálního modelu významu. Co nám v tom brání?

- význam je těžké definovat
- lidé jazyk používají „nepořádně“

Zkoumání významu

Řeč je jakési bludiště cest. Přijdeš z jedné strany a vyznáš se tu; přijdeš na totéž místo z jiné strany, a už se tu nevyznáš.

[Wittgenstein, 1953]

Vágnost a subjektivita významů

muž = dospělý člověk mužského pohlaví

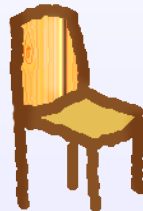
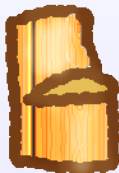
Je 18letý člověk mužského pohlaví muž?

Je 17letý člověk mužského pohlaví muž?

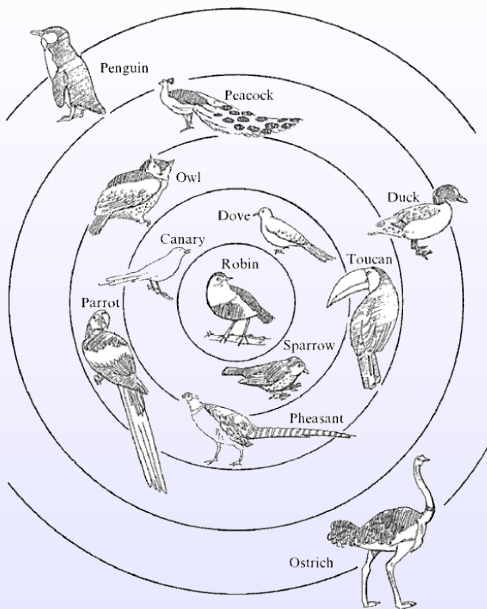
Je 16letý člověk mužského pohlaví muž?

Vágnost a subjektivita významů

špalek židle



Jsou významy diskrétní, nebo spojitý?



Aitchison,
2003 in
[Goddard, 2011]

Vágnost a subjektivita významů

Wittgensteinův příklad: hra – deskové hry, karetní hry, míčové hry, Olympijské hry, dětské hry ...

Co mají společného?

zábava, soutěž, vítězství, dovednost ...

rodinná podobnost – stavba těla, barva očí, temperament ...

E. Rosch ukázala, že kategorie jsou často organizovány okolo typických reprezentantů.[Langacker, 1987]



Víceznačnost, homonymie, polysémie

existuje na různých úrovních: přípony, koncovky, slova, slovní spojení, věty

- homonymie – náhodný jev: úplná h. (líčit, kolej) a částečná h. (stát, los ...)
- polysémie – přirozený jev: kohout, strom, kulhat ...

Víceznačnost, homonymie, polysémie

Jak rozeznat polysémii od úplné homonymie?

Nevíme.

Záleží na tom?

Je víceznačnost problém v NLP?

rychlý \Rightarrow fast

auto \Rightarrow car

rychlé auto \Rightarrow fast car

vysoký \Rightarrow high

škola \Rightarrow school

vysoká škola \Rightarrow university?

Je víceznačnost problém v NLP?

Odpověď: jak kdy, záleží na aplikaci . . .

Jak rozlišovat významy (sense)?

Kolik významů má slovo kočka?

- SSJČ: 7
- SSČ: 2
- PSJČ: 10
- Slovník českých synonym: 4
- Český WordNet: 3

Jak rozlišovat významy (sense)?

Praktické problémy: **granularita** a **užívání**

Jakou granularitu vlastně po aplikacích *požadujeme*?

Požadujeme, aby aplikace „znaly“ všechny významy, nebo jen ty, které se běžně užívají?

Granularita významu (sense): kočka

- 1. malá kočkovitá šelma, chovaná v domácnostech
- 2. malá n. středně velká šelma s hustým kožichem; zool. rod Felis
- 3. samice kočkovité šelmy vůbec
- 4. ob. kožišina na límci, kolem krku n. ramen
- 5. kocovina (Haš.)
- 6. věc připomínající někt. vlastnost kočky
- 7. druh dětek

Granularita významu (sense): kočka

- 2. malá n. středně velká šelma s hustým kožichem; zool. rod Felis
 - 1. malá kočkovitá šelma, chovaná v domácnostech
- 3. samice kočkovité šelmy vůbec
- 4. ob. kožišina na límci, kolem krku n. ramen
- 5. kocovina (Haš.)
- 6. věc připomínající někt. vlastnost kočky
- 7. druh dūtek

Granularita významu (sense): kočka





- 2. malá n. středně velká šelma s hustým kožichem; zool. rod Felis
 - 1. malá kočkovitá šelma, chovaná v domácnostech
 - 3. samice kočkovité šelmy vůbec
- 4. ob. kožišina na límci, kolem krku n. ramen
- 5. kocovina (Haš.)
- 6. věc připomínající někt. vlastnost kočky
- 7. druh dětek

Granularita významu (sense): kočka

- 2. malá n. středně velká šelma s hustým kožichem; zool. rod Felis
 - 1. malá kočkovitá šelma, chovaná v domácnostech
 - 3. samice kočkovité šelmy vůbec
- 4. ob. kožišina na límci, kolem krku n. ramen
- 6. věc připomínající někt. vlastnost kočky

- 5. kocovina (Haš.)

- 7. druh důtek

-  Erk, K., McCarthy, D., and Gaylord, N. (2009).
Investigations on Word Senses and Word Usages.
In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP*, pages 10–18, Suntec, Singapore. Association for Computational Linguistics.
-  Goddard, C. (2011).
Semantic Analysis: A Practical Introduction.
Oxford Textbooks in Linguistics. Oxford University Press.
-  Langacker, R. W. (1987).
Foundations of cognitive grammar: Theoretical Prerequisites.
Stanford University Press, Stanford, CA.
Vol 1, 1987(Hardcover), 1999(Paperback).
-  Wittgenstein, L. (1953).
Philosophical Investigations.
Basil Blackwell, Oxford.

