

PLIN037 Sémantika a počítače

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

24. září 2013

PLIN037 Sémantika a počítače

Předmět PLIN037 Sémantika a počítače je podpořen projektem OP VK Mezi bohemistikou a informatikou. Inovace vysokoškolské výuky češtiny v kontextu počítačového zpracování přirozeného jazyka (INOVA.CZ).

www.projekt-inova.cz



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Organizace předmětu v podzimním semestru 2013

Význam a počítače

Co je význam

Co studuje sémantika?

Reprezentace znalostí

Co se bude dít v průběhu semestru?

- přednášky
- cvičení
 - práce s algoritmy NLP (za body)
 - vytváření datových zdrojů (za body)
- prezentace (za body)

Co musíte umět?

- pracovat s textovým korpusem
- porozumět kódu programovacího jazyka
- porozumět anglickým (nebo francouzským) odborným textům

Co se bude dít na konci semestru?

písemný test, součet bodů posbíraných v průběhu semestru

Proč je nauka o významu obtížná

- nedokážeme formálně přesně popsat význam
- jeden a týž znak se používá pro označení více významů (víceznačnost)
- věty v přirozeném jazyce se nevyjadřují úplně (implicitní vs. explicitní informace)
- k významovým posunům dochází častěji než by se zdálo

Sémantika

- nauka o významu **znaků**
- znak = něco, co něco zastupuje (reprezentant)
- v našem případě nauka o významu morfémů, slov, slovních spojení, vět, promluv

Co není význam

- reference: jednorožec? nikdo? já? vynálezce hračky jo-jo
- (vědecká) znalost: sůl je chlorid sodný
- užití: jaký by pak byl význam celé promluvy?
- pravdivostní hodnota: $2 + 2 = 4$

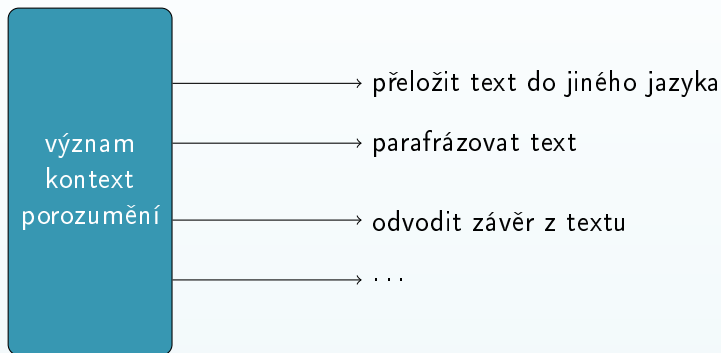
Co je význam

- klasické konceptualistické pojetí: význam je strukturovaná myšlenka (koncept) v mysli
John Locke [in [Goddard, 2011]]:

[W]ords, in their primary or immediate signification, stand for nothing but the ideas in the mind of him that uses them . . .

- strukturalistické pojetí: význam se skládá ze souhrnu vazeb, které má toto slovo s ostatními
- extensionalismus: rozdíly ve významu slov lze nejlépe zkoumat ve vztahu k extenzím (referencím)
extensionalistické pojetí významu bývá prezentováno „mapami užití“
- moderní konceptualismus: méně filozofie, více praktického pojetí

Jak se projevuje porozumění?



Co studuje sémantika?

- sémantické relace: synonymie, antonymie, hyperonymie, meronymie . . .
- taxonomické a kolektivní nadřazené koncepty
- vyplývání, kontradikce, parafráze
- vágnost – slovo může být užito v různých kontextech
- víceznačnost – slovo má víc významů (a různé gramatické vlastnosti)
- prototypy
- konotace – „subtilní aspekty významu“
- kolokace, ustálená slovní spojení

Reprezentace znalostí

- jazyková znalost
- znalost „světa“

Jazykové znalosti

znaky jazyka, segmentace na věty, segmentace na slova, určení slovních druhů, syntax jazyka, seznam slov jazyka („slovník“)

Znalosti o světě

- je možné provádět jazykovou analýzu bez znalosti významů jednotlivých slov
- na určité úrovni tuto znalost potřebujeme alespoň na úrovni vazby mezi slovy

sémantické relace: synonymie, antonymie, hyperonymie, meronymie



Goddard, C. (2011).

Semantic Analysis: A Practical Introduction.

Oxford Textbooks in Linguistics. Oxford University Press.