

PA153 Natural Language Processing

05 – Semantics II

(logical representation, from sentence to discourse)

Karel Pala, Zuzana Nevěřilová

NLP Centre, FI MU, Brno

23th October 2019

- 1 Lexical Meaning and Context
- 2 Context
- 3 Sentential Semantics
- 4 Logical Semantics
- 5 Verbs as Predicates
- 6 Discourse Semantics

Lexical Meaning and Context

lexical meaning: isolated word meaning (as presented in the lecture Semantics I)

autosemantic (plnovýznamové, autosémantické) lexical units (LU):

- nouns
- adjectives
- verbs
- adverbs

free, to can, to leave, so, in fact

synsemantic (pomocné, sysémantické) LUs: from, why, how, this, umm

lexical meaning: isolated word meaning (as presented in the lecture

Semantics I)

autosemantic (přívýznamová, autošémantická) lexical units (LU):

- nouns
- adjectives
- verbs
- adverbs

free, to can, to leave, so, in fact

syntactic (pomocné, syntaktická) LUs: from, why, how, this, umm

Některé LU mají izolovaný význam, uvedené LU jsou do jisté míry protipříklady. Např. „černý“ znamená „neplatící“, ale jen v kolokaci „černý pasažér“, v anglickém příkladu je to “free” ve v kontextu “freerider”. Pomocná a způsobová (modální) slovesa mají význam oslabený. U některých LU má smysl mluvit o významu skutečně jen v kontextu.

Lexical Meaning and Context

Krakutel z jejich mrusy se blutkal, načež potom tračil všechny své stěvače.

Vyšetřovatel jopuz hrych vlády tre naštvál, bruvěž slekym rozzuřil kruky moré posluchače.

Vyšetřovatel z jejich vlády se naštvál, načež potom rozzuřil všechny své posluchače.

Krakulci z jejich mravy se blůtkal, načel potom tračil všechny své stěvače.

Vyletlovali jepez hrych vlády tre nářval, bruvěš slékym rozuzlí kruky moré posluchače.

Vyletloval z jejich vlády se nářval, načel potom rozuzlí všechny své posluchače.

U těchto vět zkusíme určit slovní druhy a větné členy. U té první věty by to mělo být snazší než u té druhé. V první větě jsou plnovýznamová slova nahrazena nesmyslnými slovy, ale koncovka je zachována. V druhé větě jsou nahrazena všechna neplnovýznamová slova, konec slova je opět zachován.

Jde o ukázkou důležitosti funkčních slov pro pochopení významu věty. Význam těchto slov si uvědomíme právě jen v kontextu věty.

When we replace autosemantic words in the sentence by pseudowords with the same endings, we are still able to guess POS or even syntactic function of the pseudoword. However, when we do the same with the synsemantic word, the task becomes more difficult. This toy example demonstrates what are synsemantic words good for.

Context

- **verbal context (verbální kontext): what was said, what will follow**
the surrounding of a word/sentence, not citing entities out of context
- **situational context (situační kontext): place, time, number and nature of communication partners, their mutual relations)**
indexicals (deixis)
pragmatic presupposition of speaker (presupozice mluvčího)
- **social context (sociální kontext): education, social group, experience,**
...

Context

How to explore the verbal context:

- in prehistoric pre-corpus times: introspection
- in case of small/medium size corpora: concordance
- in case of large corpora: word sketch (slovní profily)

kandidát

czTenTen12 [Majka] frekvence = 213578 (39.3 v milionu)

<u>a_modifier</u>	<u>90268</u>	<u>-1.4</u>	<u>post_na</u>	<u>45490</u>	<u>-7.3</u>	<u>gen_2</u>	<u>22232</u>	<u>-0.9</u>
prezidentský	<u>5105</u>	10.02	post	<u>2878</u>	8.8	hodnost	<u>258</u>	7.01
žhavý	<u>2404</u>	9.23	primátor	<u>1617</u>	8.14	zvolení	<u>161</u>	6.62
závislý	<u>4647</u>	8.74	prezident	<u>3946</u>	7.52	pětice	<u>139</u>	6.49
vhodný	<u>9792</u>	8.43	senátor	<u>701</u>	7.35	nominace	<u>272</u>	6.31
republikánský	<u>1055</u>	8.34	eurokomisara	<u>216</u>	7.27	slyšení	<u>109</u>	6.28
navržený	<u>1516</u>	8.28	pozice	<u>4181</u>	7.2	výběr	<u>1822</u>	6.11
horký	<u>2315</u>	8.19	děkan	<u>382</u>	7.15	představování	<u>62</u>	6.11
opoziční	<u>745</u>	7.42	rektor	<u>321</u>	7.03	navrhování	<u>91</u>	6.08

Context in the Word Sketch

How the word sketch is determined

it is calculated

word sketch grammar

there are also multi word sketches



Sentential Semantics (větná sémantika)

sentence meaning: sentence of individual words + meaning of the syntactic structure

Compositionality Principle (princip kompozicionality):

- The meaning of the whole is a function of the meaning of the parts and the mode of combining them.
- The meaning of a complex expression is uniquely determined by the meaning of its constituents and the syntactic construction used to combine them.

Logical Semantics (logická sémantika)

here, **logic** is the instrument

- lexical meanings are omitted
- predicate structure of verbs or deverbatives is transformed to predicate structure in the logic
- sentences are transformed into propositions (having a truth value)

Logical Semantics and Predicate Logic

Propositional Logic (výroková logika), usually not strong enough
First Order Predicate Logic [Mendelson, 1997] (predikátová logika 1. řádu, FOPL)

- terms (termy): variables x , functions $f(x)$
- predicate symbols $P(x)$
- logical connectives (logické spojky) $\vee, \wedge, \neg, \Rightarrow, \Leftrightarrow$
- quantifiers (kvantifikátory) \forall, \exists
- equality/identity symbol (symbol rovnosti) $=$
- non-logical symbols (mimologické symboly): constants

Logical Semantics and Predicate Logic

All students who obtain less than 10 points get the F grade.

$\forall x : Obtain(x, y) \wedge (y < 10) \Rightarrow Grade(x, "F")$

(*Obtain*(x, y) – x obtained y points, *Grade*(x, y) – x got the y grade)

John obtained 3 points.

Obtain("John", 3)

Conclusion

$Obtain("John", 3) \wedge (3 < 10) \Rightarrow Grade("John", "F")$

is true

Logical Semantics and Predicate Logic: Try it yourself

Komu se nelení, tomu se zelení.

Every student who takes Analysis also takes Geometry.

$\forall x : Nelenit(x) \Rightarrow Zelenit(x)$

$\forall x : (Student(x) \wedge Takes(x, Analysis) \Rightarrow Takes(x, Geometry))$

Honza se odrazil od podlahy a vyskočil do dvou metrů.

John bounced off the floor and jumped up to 2 meters.

Odražit("Honza", "podlaha")

Vyskočit("Honza", "2 metry")

Bounce("John", "floor")

Jump("John", "2meters")

FOPL Limitations

- not all NL constructions are propositions, e.g.
 - ▶ Hello. Nice to meet you.
 - ▶ If I ever saw one.
 - ▶ May the force be with you.
- not all propositions are distinguishable in FOPL
I was there vs. I am there vs. I will be there
- not all propositions are 1st order
All people have common properties.
 $\exists Property \forall x : Property(x)$
- not all NL quantifiers exist in FOPL
 - ▶ vast majority
 - ▶ the majority
 - ▶ many
 - ▶ a few
 - ▶ few

Translation from NL to FOPL is by Conventions

Live(x, y) – x lives y

Neil lives in Brno.

Live("Neil", "Brno")

Neil lived on Saturday.

Live("Neil", "Saturday")

Live(x, y) – x lives y and y is a place

typed logics

Verbs as Predicates

verbs (and deverbatives) can be considered as predicates, other parts of sentence can be considered as arguments of the predicate

valency

Verbs as Predicates

The boy broke the window.

A stone flew into the window and broke it.

The window broke.

to break: AG(person) ART(artifact) INS(instrument)

boy person

stone instrument

window artifact

Valency Lexicons

Czech

- Vallex: In VALLEX 2.x, there are roughly 2,730 lexeme entries containing together around 6,460 lexical units (“senses”).¹
- VerbaLex:²
 - ▶ 21,032 literals (verb + sense)
 - ▶ 10,469 verb lemmata

English

- VerbNet: 8537 total verbs represented³

¹<http://ufal.mff.cuni.cz/vallex/2.6/doc/home.html>

²<http://nlp.fi.muni.cz/cs/VerbaLex>

³<http://verbs.colorado.edu/verb-index/index.php>

Valency Lexicons: Vallex

VALLEX 2.6

alphabet

class

functors

forms

aspect

control

reflex.

recipr.

complexity

VALEVAL

- F (10)
- **G**
- H (51)
- CH (22)
- I (17)
- J (13)
- K (73)
- L (37)
- M (53)
- N (133)

- zlobit, zlobivat
- zlobit se, zlobívat se
- zlomit se, zlámat se
- zmačknout, zmačkat
- zmačknout se, zmačkat se
- zmahat, zmoci/zmoci
- zmahat se, zmoci se/zmoci se
- zmapovat

zmařit^{pf}

1 ≈ **zkazit; zničit**

-frame: **ACT**₁^{obl} **PAT**₄^{obl} **BEN**₃^{typ} **MEANS**₇^{typ}

-example: zmařil celé jednání svou nezodpovědností; zmařil mu život

-rfi: pass: jeho podvatrné plány se naštěstí dopředu zmařily

Valency Lexicons: VerbaLex

Verb classes

- admit-64 (65)
- adopt-91 (4)
- allow-63 (69)
- animal_sounds-38 (60)
- approve-75 (91)
- assessment-34 (50)
- **avoid-52 (51)**
- banish-10.2 (55)
- battle-36.3 (8)
- bodyinternalmotion-49 (131)
- build-26.1-4 (7)

Verb class

"destroy-44"

- babrat₁
- bořit₁
- bořit₂
- bourat₁
- bourat₄
- brakovat₃
- brát₂₇
- břit₁
- demolovat₁

zmařit₃^{pf} **zničit**₅^{pf} **rozbít**₃^{pf}

mařit₃^{impf} **ničit**₅^{impf} **rozbíjet**₃^{impf}

1 zmařit₃, mařit₃ ≈

2 rozbít₃, rozbíjet₃, zničit₅, ničit₅ ≈

-frame: **GROUP** <institution:1> ^{obl}_{i1} **VERB** ^{obl} **GROUP** <institution:1> ^{obl}_{i4}

-example: *policie rozbila zločinecký gang (pf)*

Valency Lexicons: VerbNet

ROLES

- AGENT [+INT_CONTROL]
- PATIENT [+CONCRETE]
- INSTRUMENT [+CONCRETE]

FRAMES

NP V NP

EXAMPLE "The Romans destroyed the city."

SYNTAX AGENT V PATIENT

SEMANTICS CAUSE(AGENT, E) DESTROYED(RERESULT(E), PATIENT)

NP V NP PP.INSTRUMENT

EXAMPLE "The builders destroyed the warehouse with explosives."

SYNTAX AGENT V PATIENT {WITH} INSTRUMENT

SEMANTICS CAUSE(AGENT, E) USE(DURING(E), AGENT, INSTRUMENT) DESTROYED(RERESULT(E), PATIENT)

NP.INSTRUMENT V NP

EXAMPLE "The explosives destroyed the warehouse."

SYNTAX INSTRUMENT V PATIENT

SEMANTICS CAUSE(?AGENT, E) USE(DURING(E), ?AGENT, INSTRUMENT) DESTROYED(RERESULT(E), PATIENT)

Valency Frame and the Meaning

Verb classes	Verb class "destroy-44"	
<ul style="list-style-type: none">• admit-64 (65)• adopt-91 (4)• allow-63 (69)• animal_sounds-38 (60)• approve-75 (91)• assessment-34 (50)• avoid-52 (51)• banish-10.2 (55)• battle-36.3 (8)• bodyinternalmotion-49 (131)• build-26.1-4 (7)	<ul style="list-style-type: none">• babrat₁• bořit₁• bořit₂• bourat₁• bourat₄• brakovat₃• brát₂₇• břitdit₁• demolovat₁	<p>zmařit₃^{pf} zničit₅^{pf} rozbít₃^{pf}</p> <p>mařit₃^{impf} ničit₅^{impf} rozbíjet₃^{impf}</p> <p>1 zmařit₃, mařit₃ ≈</p> <p>2 rozbít₃, rozbíjet₃, zničit₅, ničit₅ ≈</p> <p>-frame: GROUP <institution:1>_{i1} VERB^{obl} GROUP <institution:1>_{i4}</p> <p>-example: <i>policie <u>rozbila</u> zločinecký gang (pf)</i></p>

- verb synset (see WordNet synset)
- translation
- verb class (verbs of communication, abolition ... [Wu and Palmer, 1994])
- verb frames (slovesné rámce) – usage
- number of slots
- syntactic information (generic order, grammatical case, preposition)
- semantic roles
- selectional restrictions (výběrová omezení) – typical representatives

Connecting Verb Frames: Building a Network

description of stereotypical situations: scripts (scénáře): to buy, to mine, to die

- 1970's proposals: Schank, Abelson, Minsky
- applications: from 1990's FrameNet ⁴

⁴<https://framenet.icsi.berkeley.edu/fndrupal/>

FrameNet

Definition:

This transparent noun frame is concerned with **Units** for measuring the **Area** of regions.
Hillary lives in the middle of **10 ACRES** of wheat.

Paul owns **20 HECTARES** of land.

Semantic Type: Transparent Noun

FEs:

Core:

Area □

Excludes: Occupant

The region whose surface is being measured.

Klaas has a 10 **ACRE** peach **orchard**

Count □

The number of **Units**.

Smiley owns **fifteen** **HECTARES** of prime real estate.

Discourse Semantics (sémantika diskurzu, utterance meaning)

- subword
- word
- word expression (slovní spojení)
- clause (holá věta)
- coherent (souvislý) text

Discourse Semantics (sémantika diskurzu, utterance meaning)

Two coffees and an apple pie.

- Would you like to order something?
- What would you like to pay for?
- What did you have for breakfast?
- What were your dreams about?
- What did the guy ask for?
- What did the guy attack you with?
- What is the presentation about?
- ...

Discourse Semantics: Anaphora Resolution

exophora (extralinguistic, often deictic)

What is *this*?

endophora (intralinguistic)

- **anaphora** (zpětný odkaz) – Susan dropped the plate. **It** shattered loudly.
- **cataphora** (dopředný odkaz) – When **he** arrived home, John went to sleep.
- **self-reference (koreference)**: I

Anaphora Types

English

- pronominal: John found the love of his life
- definite nominal phrase: The relationship did not last long.
- quantifier/ordinal: He started a new one.
- adverb: Fred was angry, and so was I.
- verb phrase: If Fred buys a new bike, I will do it as well.

Czech

- pronominal: Petr si ukrojil chleba a snědl *ho*.
- synonym: Petr si ukrojil chleba a pak *krajíc* snědl.
- hyperonym: Petr si ukrojil chleba a pak *jídlo* snědl.
- ellipsis: [Petr] Snědl chleba. Marie taky.

Functional Sentence Perspective

aktuální členění větné, funkční větná perspektiva, topic-focus articulation, functional sentence perspective, FSP

distinguishes known and new parts of the sentence

work of Brno linguist J. Firbas

considerable work by Prague Linguistic Circle (V. Mathesius, currently E. Hajicova)

topic, theme (téma) – comment, rheme, focus (réma)

The dog bit the little girl.

The little girl was bitten by the dog.

References I



Mendelson, E. (1997).

Introduction to Mathematical Logic.

Discrete Mathematics and Its Applications Series. Chapman & Hall.



Wu, Z. and Palmer, M. (1994).

Verbs semantics and lexical selection.

In *Proceedings of the 32nd annual meeting on Association for Computational Linguistics*, ACL '94, pages 133–138, Stroudsburg, PA, USA. Association for Computational Linguistics.