

PLIN037 Sémantika a počítače

Zuzana Nevěřilová
xpopelk@fi.muni.cz

Centrum zpracování přirozeného jazyka, B203
Fakulta informatiky, Masarykova univerzita

7. března 2019

Logická forma

Typovaný formalismus

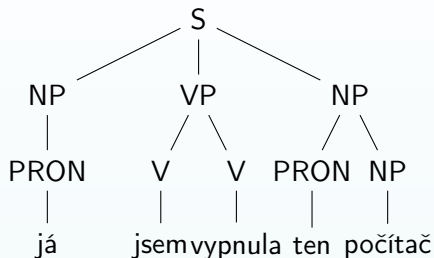
Slovesné a jiné valence

Parafráze

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

Význam, kontext a logická forma



Syntaktická analýza

Logická forma

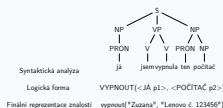
VYPNOUT(<JÁ p1>, <POČÍTAČ p2>)

Finální reprezentace znalostí

vypnout("Zuzana", "Lenovo č. 123456")

└ Logická forma

└ Význam, kontext a logická forma

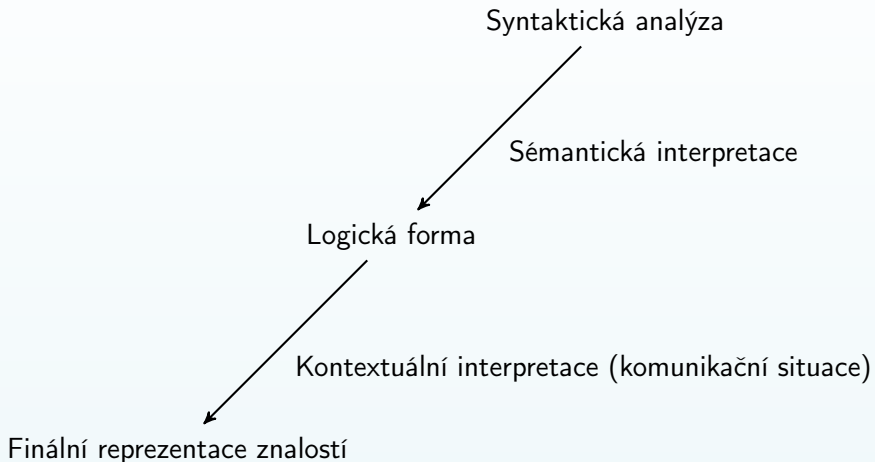


Jednou z možností, jak popsat význam věty genericky, bez ohledu na kontext, je její logická forma. Na jedné straně se zdá, že nemůžeme význam popsat bez znalosti kontextu. Často nám stačí jen „jazykový“ kontext, situační nepotřebujeme, abychom větě rozuměli.

Kdyby tato možnost vůbec neexistovala (a byli bychom odkázáni vždy jen na kontext), měly by věty jako *Já jsem vypnula ten počítač*. smysl, jen pokud bychom věděli, o jakém počítači mluvíme.

Pokud tedy vezmeme za své, že věta *Já jsem vypnula ten počítač*. má nějaký smysl bez ohledu na situaci, můžeme si dovolit mluvit o její logické formě. Větě také rozumíme proto, že víme něco o světě a máme nějakou představu, co znamená *vypnout* a *počítač*.

Logická forma



Logická forma

predikátová logika prvního řádu:

- (n-ární) predikát $P(x)$: pro objekt reprezentovaný proměnnou x platí P
 $být_doma(pepa)$: Pepa je doma.
- operátory: \vee, \wedge, \neg
 $\neg být_doma(pepa)$: Pepa není doma.
- implikace \Rightarrow
 $\neg pršet() \Rightarrow \neg zmoknout(x)$ Nebude-li pršet, nezmoknem.
- kvantifikátory: \exists, \forall
 $\exists x : být_doma(x)$: Někdo je doma.
 $\forall x : \neg být_doma(x)$: Nikdo není doma.

Převod věty na logickou formu

- sloveso je predikát (*koupit*)
- vlastnost je predikát (*koupený*)
- jedinec je proměnná (*Alík*)
- třída je proměnná (*pes*)
- operátory (*a, nebo, ale*)
- kvantifikátory (*všichni, nikdo, někdo, některý*)

Převod věty na logickou formu

Cvičení:

1. Petr i Linda bydlí v Brně. ($\text{bydlet}(x,y)$)
2. Někteří studenti na seminář nepřišli. ($\text{přijít}(x,y)$)
3. Někdo má rád holky, jiný zase vdolky. ($\text{mít_řád}(x,y)$)
4. Někdo přišel, ale někdo ne.
5. Oslovil jsem ji, protože se mi líbila. ($\text{oslovit}(x,y)$, $\text{líbit}(x,y)$)

Převod věty na logickou formu

Možná řešení:

1. Petr i Linda bydlí v Brně. ($\text{bydlet}(x,y)$)
 $\text{bydlet}(\text{Petr}, \text{Brno}) \wedge \text{bydlet}(\text{Linda}, \text{Brno})$
2. Někteří studenti na seminář nepřišli. ($\text{přijít}(x,y)$)
 $\exists \text{student} : \neg \text{přijít}(\text{student}, \text{seminář})$
3. Někdo má rád holky, jiný zase vdolky. ($\text{mít_rād}(x,y)$)
 $\exists x_1 : \text{rād}(x_1, \text{holka}) \wedge \exists x_2 : \text{rād}(x_2, \text{vdolek})$
4. Někdo přišel, ale někdo ne.
 $\exists x_1 : \text{přijít}(x_1, y) \wedge \exists x_2 : \neg \text{přijít}(x_2, y)$
5. Oslovil jsem ji, protože se mi líbila. ($\text{oslovit}(x,y)$, $\text{líbit}(x,y)$)
 $\text{líbit}(x, y) \Rightarrow \text{oslovit}(x, y)$

Převod věty na logickou formu

Problém: víceznačnost přirozeného jazyka

Každý kluk vlastní mobil.

$\exists m \forall k : \textit{vlastnit}(k, m)$

$\forall k \exists m : \textit{vlastnit}(k, m)$

Každou hodinu zemře jeden Američan na rakovinu.

$\forall h \exists a : \textit{zemřít}(a, h)$

$\exists a \forall h : \textit{zemřít}(a, h)$

Převod věty na logickou formu

Problém: spojky a logické spojky

- „a“ není vždy \wedge (Honza kopl a balón přistál v brance.)
- „ale“ je \wedge (Někdo přišel na seminář, ale ne všichni.)
- „nebo“ není vždy \vee : Bydlím buď v Praze, nebo v Brně.

Převod věty na logickou formu

Problém: nedostatek kvantifikátorů

všichni, nikdo, každý: \forall

někdo, aspoň jeden, některý: \exists

pár, málo, hodně, skoro všichni, většina, . . .

neuvedené kvantifikátory: lidé = všichni lidé?

Někdo má rád všechny holky, jiný má rád všechny vdolky.

Převod věty na logickou formu

Problém: negace a kvantifikátory

Všichni kluci brečí.

$\forall k : brečet(k)$

Někteří kluci brečí.

$\exists k : brečet(k)$

Žádní kluci nebrečí.

$\neg \exists k : brečet(k)$

Ne všichni kluci brečí.

$\neg \forall k : brečet(k)$

Převod věty na logickou formu

Problém: ne všechny části věty mají ekvivalent v PL1

nejen, v podstatě, mimochodem . . .

přací věty, rozkazovací věty

Převod věty na logickou formu

Problém: logická forma neodráží realitu (není potřeba znalost světa)

K obědu bude jízdní kolo s kusem betonu.

Převod věty na logickou formu

Problém: ne všechny věty jsou prvního řádu

Lhát se nemá.

Slovo myš je jednoslabičné.

Logické formule a spojky

P1 = Sam pracuje.

P2 = Sam vydělává peníze.

- Sam pracuje a vydělává peníze.
- Sam pracuje, aby vydělal peníze.
- Sam pracuje, a tím vydělává peníze.

Logické formule a řady

P1 = Sam pracuje.

- Max řekl, že Sam pracuje.
- Ida slyšela, jak Max řekl, že Sam pracuje.
- Majka chtěla, aby Ida slyšela, jak Max řekl, že Sam pracuje.
- ...

Logická forma: typovaný formalismus

milovat(x (jednotlivec), y (jednotlivec))

bydlet(x (jednotlivec), y (místo))

bydlet(Praha, teta)

2019-03-07

PLIN037 Sémantika a počítače

└ Typovaný formalismus

└ Logická forma: typovaný formalismus

mřovat(x (jednotlivec), y (jednotlivec))
bydlit(x (jednotlivec), y (misto))
bydlit(Praha, teta)

Jeden z typovaných formalismů je transparentní intenzionální logika (TIL).

Logická forma: sloveso jako predikát

na sloveso je možné pohlížet jako na n-ární predikát (milovat(x,y),
bydlet(x,y), smát_se(x), ...)

také na vlastnosti je možné pohlížet jako na predikáty
(být_modrý(x), ...)

valenční slovníky: Vallex, VerbaLex

Parafráze

- (1) *Wonderworks Ltd. constructed the new bridge.*
- (2) *The new bridge was constructed by Wonderworks Ltd.*
- (3) *Wonderworks Ltd. is the constructor of the new bridge.*

Příklad z [Androutsopoulos and Malakasiotis, 2009]

Parafráze

“approximate conceptual equivalence among outwardly different material.”

Beaugrande and Dressler (1981, page 50) in
[Bhagat and Hovy, 2013]

sémantická ekvivalence, ale pragmatické rozdíly

Hranice parafráze

- (1) *Wonderworks Ltd. constructed the new bridge.*
- (2) *The new bridge was constructed by Wonderworks Ltd.*
- (3) *Wonderworks Ltd. is the constructor of the new bridge.*

Příklad z [Androutsopoulos and Malakasiotis, 2009]

- (1) Wondersworks Ltd. constructed the new bridge.
(2) The new bridge was constructed by Wondersworks Ltd.
(3) Wondersworks Ltd. is the constructor of the new bridge.

Příklad z [Androustopoulos and Malakasioti, 2009]

Pohlížet na parafráze optikou logické formy (tj. parafráze jsou věty, jejichž logická forma je ekvivalentní) je příliš přísné. Srovnej s definicí plné synonymie, o které řada odborníků tvrdí, že neexistuje (kdyby existovala, bylo by jedno ze slov zbytečné a přestalo by se používat). Lidé vyhodnotí jako parafráze věty, které tvrdí podobnou věc, nevíme ale, jak moc tvrzení mohou být rozdílná, aby se o nich stále dalo mluvit jako o parafrázích.

Cvičení

najděte 20 dvojic parafrází:

- ve vlastních textech
- v novinových článcích
- v překladech
- v testech čtenářských dovedností



Allen, J. (1995).

Natural Language Understanding (2nd ed.).

Benjamin-Cummings Publishing Co., Inc., Redwood City, CA, USA.



Androutsopoulos, I. and Malakasiotis, P. (2009).

A survey of paraphrasing and textual entailment methods.

CoRR, abs/0912.3747.



Bhagat, R. and Hovy, E. (2013).

What is a paraphrase?

Computational Linguistics, 39(3):463–472.