

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

5. března 2015

Logická forma

Typovaný formalismus

Slovesné a jiné valence

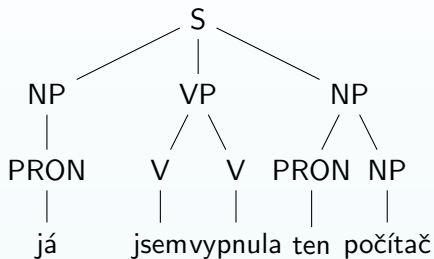
Parafráze

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

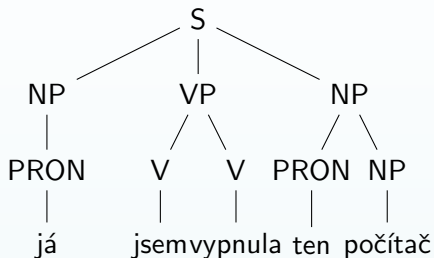
Význam, kontext a logická forma

Syntaktická analýza



Význam, kontext a logická forma

Syntaktická analýza

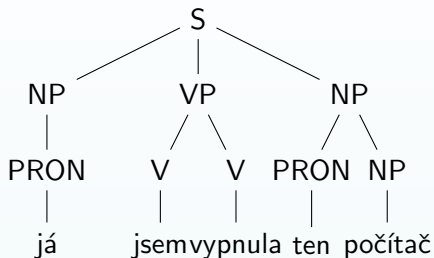


Logická forma

VYPNOUT(<JÁ p1>, <POČÍTAČ p2>)

Význam, kontext a logická forma

Syntaktická analýza



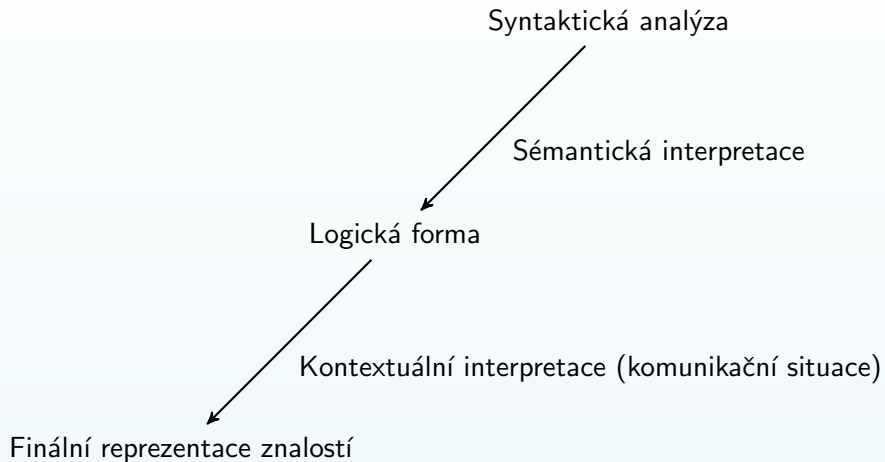
Logická forma

VYPNOUT(<JÁ p1>, <POČÍTAČ p2>)

Finální reprezentace znalostí

vypnout("Zuzana", "Lenovo č. 123456")

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
- vlastnost je predikát
- jedinec je proměnná
- třída je proměnná
- 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. ($\text{mít rád}(x,y)$)
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

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. ($\text{mít} \text{ rád}(x,y)$)
 $\exists x_1 : \text{rád}(x_1, \text{holka}) \wedge \exists x_2 : \text{rád}(x_2, \text{vdolek})$
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.

$$\forall k : \text{brečet}(k)$$

Někteří kluci brečí.

$$\exists k : \text{brečet}(k)$$

Žádní kluci nebrečí.

$$\neg \exists k : \text{brečet}(k)$$

Ne všichni kluci brečí.

$$\neg \forall k : \text{brečet}(k)$$

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

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.

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 řády

P1 = Sam pracuje.

Logické formule a řady

$P1 = \text{Sam pracuje.}$

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

Logické formule a jejich relativizace

P1 = Sam pozoruje kočku.

Logické formule a jejich relativizace

P1 = Sam pozoruje kočku.

- Sam pozoruje kočku, která honí myš.
- Sam pozoruje kočku, která honí myš, která snědla chleba.
- Sam pozoruje kočku, která honí myš, která snědla chleba, který zbyl od včerejška.
- ...

Logická forma: typovaný formalismus

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

bydlet(x (jednotlivec), y (místo))

bydlet(Praha, teta)

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]

Cvičení

najděte 20 dvojic parafrází:

- ve vlastních textech
- v novinových člancích
- v překladech
- v testech čtenářských dovedností
- v titulkovaných pořadech



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.