

Jednoduchá sémantika RDF

Spracované podľa

dokumentov W3C

- <http://www.w3.org/TR/rdf11-mt/>

a prednášok

- <http://www.inf.tu-dresden.de/content/institutes/ki/cl/study/summer14/fswt/slides/FSWT2014-L4-RDSF-Semantics.pdf>

Vzťah RDF k formálnej logike

Vo formálnej logike je interpretácia zobrazenie, ktoré zobrazí vety formálneho jazyka na tvrdenia v nejakej doméne.

Keďže RDF grafy predstavujú vety jazyka RDF, môžeme používať terminológiu formálnej logiky a hovoriť o

- **interpretáciách** slov slovníka RDF grafu
- **pravdivostnej hodnote** viet - trojíc a RDF grafov
- **modeloch** RDF grafov
- **logickej ekvivalencii** RDF grafov
- **logickom dôsledku**, t.j. trojica alebo graf je logickým dôsledkom iného grafu.

K tomu je potrebné **presne - formálne definovať**, čo môže **byť interpretácia trojice** resp. **grafu**.

Skôr než si formálne definujeme čo je interpretácia RDFgrafu zopakujme si formálnu definíciu jeho syntaxe.

Formálna syntax jazyka RDF

opakovanie

Zopakujme si:

Čo sú **slová** jazyka RDF?

- **IRI** (zobecnené URI)
- **bnodes** (prázdne uzly)
- **literály**

Čo je **veta** jazyka RDF?

- každá rdf-trojica - **triple**:
 $(s, p, o) \cdot (IRI \cup \text{bnode}) \times IRI \times (IRI \cup \text{bnode} \cup \text{literal})$

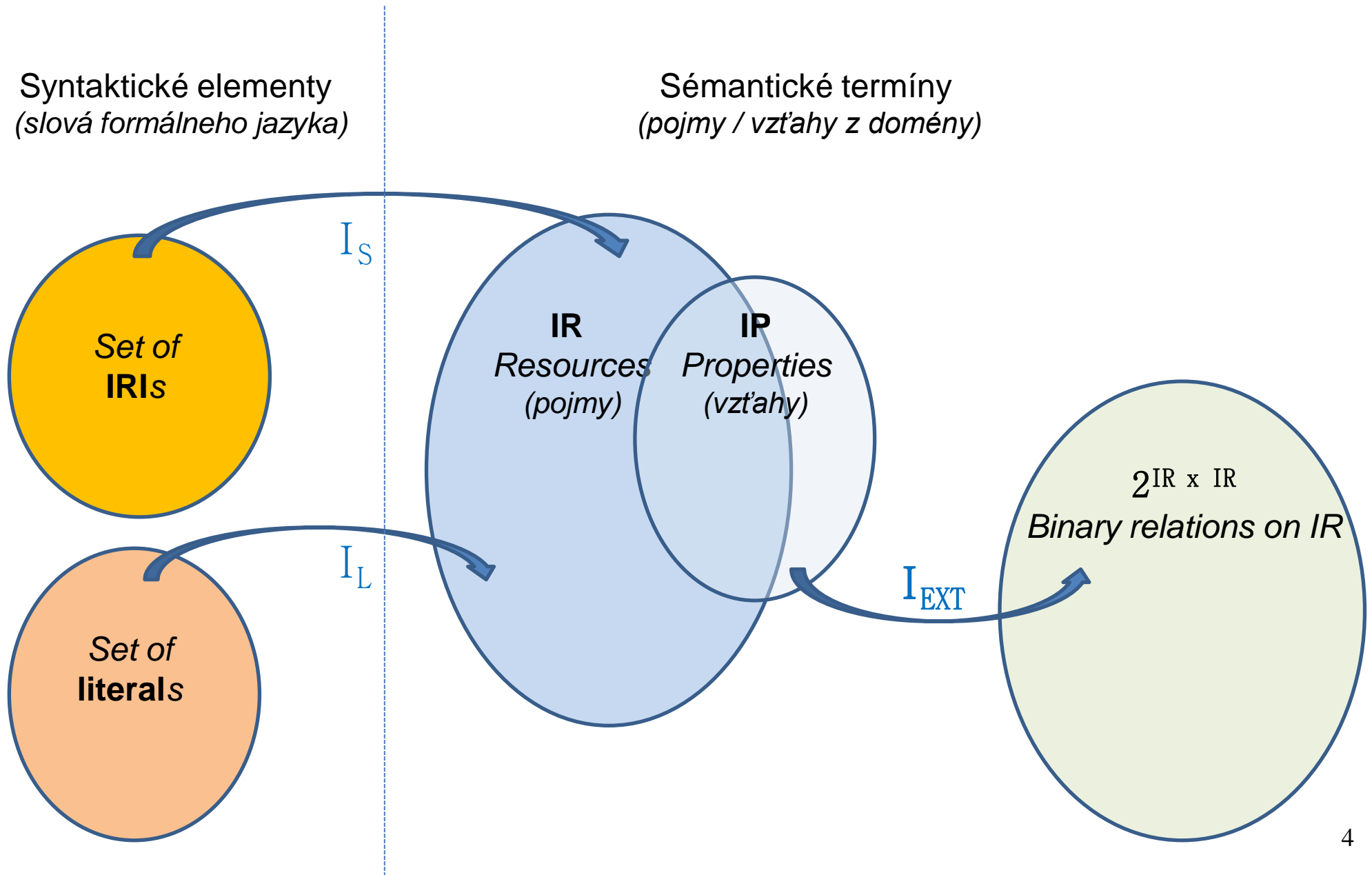
je veta.

- každá konečná množina rdf-trojíc - **rdf-graf** je veta.

Jednoduchá interpretácia – diagram

Syntaktické elementy
(slová formálneho jazyka)

Sémantické termíny
(pojmy / vzťahy z domény)



Jednoduchá interpretácia

IRI a literálov

Jednoduchá interpretácia je definovaná pomocou usporiadanej 5-tice

$$I = (IR, IP, I_{EXT}, I_S, I_L)$$

kde je:

- **IR** – neprázdna množina, ktorá nazývame tiež nazývame univerzum alebo doména I .
- **IP** – množina properties (vzťahov) I .
- **I_{EXT}** – funkcia, ktorá priradí každému vzťahu z IP množinu usporiadaných dvojíc prvkov z IR (t.j. binárnu reláciu nad IR).
- **I_S** – funkcia, ktorá zobrazuje **IRI** na prvky zjednotenia $IR \cup IP$.
- **I_L** – parciálna funkcia, ktorá zobrazuje **literály** na prvky univerza IR .

Poznámka: Množiny IR a IP môžu mať prienik, to umožňuje vyslovovať aj tvrdenia aj o predikátoch.

Pravdivostná hodnota základných trojíc a grafov

Funkcie I_S , I_L , ktoré definujú interpretáciu pre IRI a literály, môžeme spojiť do jednej interpretačnej funkcie I :

$$I(x) = I_S(x) \text{ ak } x \text{ je IRI a} \\ = I_L(x) \text{ ak } x \text{ je literál}$$

Túto interpretáciu môžeme ďalej rozšíriť aj na ostatné syntaktické štruktúry základného grafu, teda **trojice a rdf-grafy**.

Pozn. zatiaľ neriešime prázdne uzly, pretože základný rdf-graf prázdne uzly nemá.

Interpretačná funkcia priradí trojiciam a grafom priradí pravdivostnú hodnotu, na základe nasledujúcich sémantických podmienok:

1. ak $E = (s \ p \ o)$ je **základná trojica**, potom $I(E) = \text{true}$ práve vtedy, keď $I(p) \in IP$ a zároveň $\langle I(s), I(o) \rangle \in I_{EXT}(I(p))$.

Teda trojica je pravdivá ak je predikát interpretovaný ako vzťah v doméne D a pojmy (resource), ktoré sú interpretáciou subjektu a objektu sú v doméne v tomto vzťahu.

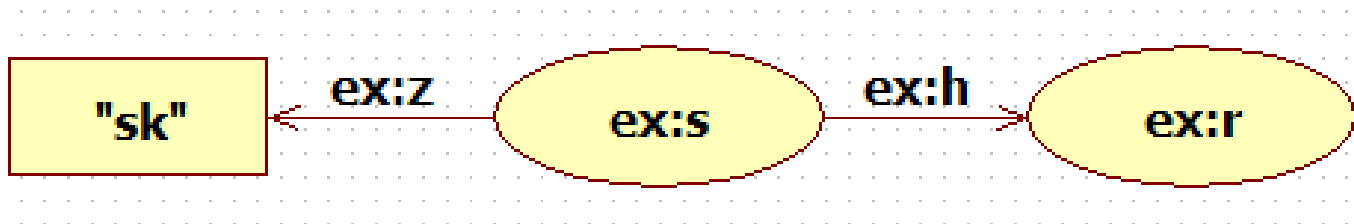
2. ak je E **základný graf**, potom $I(E) = \text{true}$ práve vtedy, keď pre každú trojicu $t \in E$ platí $I(t) = \text{true}$

Teda graf je pravdivý ak sú pravdivé všetky trojice.

Príklad

rdf-graf **G**:

ex:s ex:h ex:r.
ex:s ex:z "sk".



Intrepretácia I1.

IR = {Slovensko, Rakusko, Rumunsko, hranici, maSkratku, "SK", "A"}

IP = {hranici, maSkratku, hlavneMesto}

IEXT(hranici) = {<Slovensko,Rakusko>}

IEXT(maSkratku) = {<Slovensko,"SK">;<Rakusko,"A">}

IS(ex:z) = maSkratku

IS(ex:h) = hranici

IS(ex:s) = Slovensko

IS(ex:r) = Rakusko

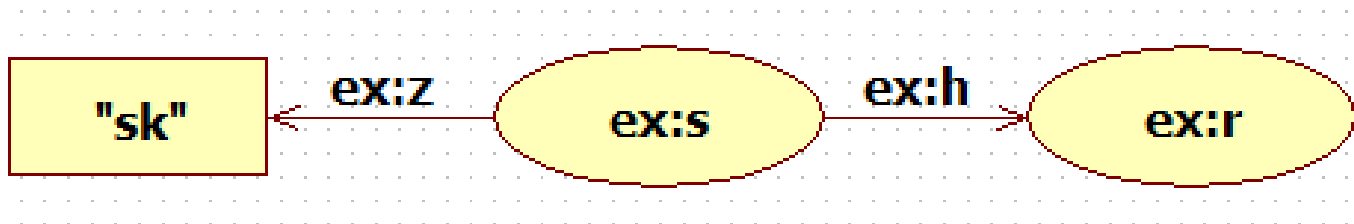
IL("sk") = "SK"

Graf G je v interpretácii I1 pravdivý

Príklad

rdf-graf G:

ex:s ex:h ex:r.
ex:s ex:z "sk".



Intrepretácia I2.

IR = {Slovensko, Rakusko, Rumunsko, hranici, maSkratku, "SK", "A"}

IP = {hranici, maSkratku, hlavneMesto}

IEXT(hranici) = {<Slovensko,Rakusko>}

IEXT(maSkratku) = {<Slovensko,"SK">;<Rakusko,"A">}

IS(ex:z) = maSkratku

IS(ex:h) = hranici

IS(ex:s) = Slovensko

IS(ex:r) = Rumunsko

IL("sk") = "SK"

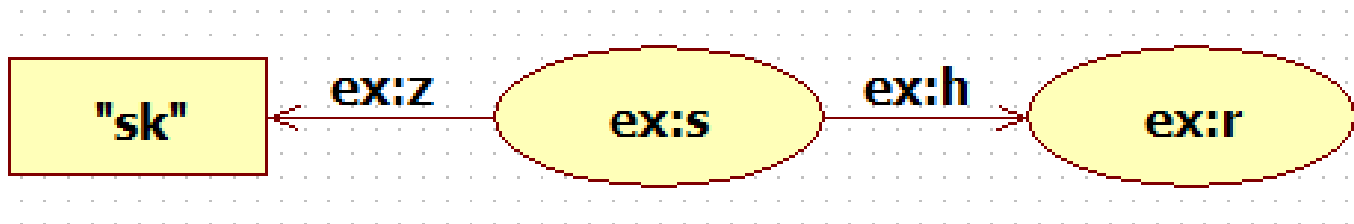
← zmena oproti I1

Graf G nie je v interpretácii I2 pravdivý, prečo?

Príklad

rdf-graf G:

ex:s ex:h ex:r.
ex:s ex:z "sk".



Intrepretácia I3.

IR = {Slovensko, Rakusko, Rumunsko, hranici, maSkratku, "SK", "A"}

IP = {hranici, maSkratku, hlavneMesto}

IEXT(hranici) = {<Rakusko, Slovensko>}

← zmena oproti I1

IEXT(maSkratku) = {<Slovensko, "SK">; <Rakusko, "A">}

IS(ex:z) = maSkratku

IS(ex:h) = hranici

IS(ex:s) = Slovensko

IS(ex:r) = Rakusko

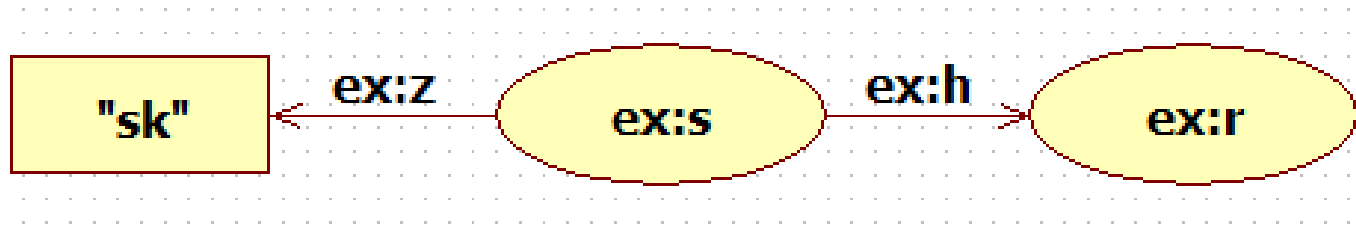
IL("sk") = "SK"

Graf G nie je v interpretácii I3 pravdivý, prečo?

Príklad

rdf-graf G:

ex:s ex:h ex:r.
ex:s ex:z "sk".



Intrepretácia I4.

IR = {Slovensko, Rakusko, Rumunsko, "SK", "A"}

IP = {hranici, maSkratku, hlavneMesto}

IEXT(hranici) = {<Slovensko,Rakusko>}

IEXT(maSkratku) = {<Slovensko,"A">;<Rakusko,"SK">} ← zmena oproti I1

IS(ex:z) = maSkratku

IS(ex:h) = hranici

IS(ex:s) = Slovensko

IS(ex:r) = Rakusko

•IL("sk") = "A"

← zmena oproti I1

Graf G je v interpretácii I4 opat pravdivý, prečo?

Jednoduchá interpretácia grafov s prázdnyimi uzlami

Nech G je graf a I jednoduchá interpretácia. Nech A je nejaká funkcia zobrazujúca všetky prázdne uzly grafu G na prvky z IR .

Funkcie I a A môžeme opäť spojiť do jednej funkcie IA :

$$IA(x) = I(x) \text{ ak } x \text{ je } IRI \text{ alebo literál} \\ = A(x) \text{ ak } x \text{ prázdny uzol}$$

Túto funkciu môžeme ďalej rozšíriť, tak že trojiciam resp. grafom priradí logickú hodnotu true alebo false:

1. ak $E = (s \ p \ o)$ je **trojica**, potom $IA(E) = true$ práve vtedy, keď $I(p) \in IP$ a zároveň $\langle IA(s), IA(o) \rangle \in I_{EXT}(I(p))$.
2. ak je G **graf**, potom $IA(G) = true$ práve vtedy, keď pre každú trojicu $t \in G$ platí $IA(t) = true$
3. Nech G je graf a I je jednoduchá interpretácia, potom **G je pravdivý v interpretácii I** t.j. $I(G) = true$, ak existuje taká funkcia A zobrazujúca všetky prázdne uzly grafu G na prvky univerza I , pre ktorú $IA(G) = true$.

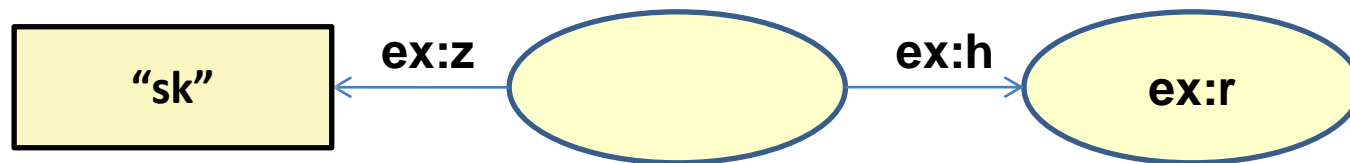
Jednoducho povedané: Graf, ktorý obsahuje prázdne uzly je pravdivý v I , ak existuje také zobrazenie jeho prázdnych uzlov na prvky univerza I , v ktorom budú všetky trojice pravdivé.

Inak: Graf je pravdivý v I , ak existuje jeho základná inštancia, ktorá je pravdivá v I .

Príklad

rdf-graf G:

`_:b ex:h ex:r.`
`_:b ex:z "sk".`



Intrepretácia I1.

$IR = \{\text{Slovensko, Rakusko, Rumunsko, hranici, maSkratku, "SK", "A"}\}$

$IP = \{\text{hranici, maSkratku, hlavneMesto}\}$

$IEXT(\text{hranici}) = \{\langle \text{Slovensko, Rakusko} \rangle\}$

$IEXT(\text{maSkratku}) = \{\langle \text{Slovensko, "SK"} \rangle; \langle \text{Rakusko, "A"} \rangle\}$

$IS(\text{ex:z}) = \text{maSkratku}$

$IS(\text{ex:h}) = \text{hranici}$

$IS(\text{ex:r}) = \text{Rakusko}$

$IL(\text{"sk"}) = \text{"SK"}$

Graf G je v interpretácii I1 pravdivý, pretože za prázdny uzol môžeme dosadiť:

$IS(_:b) = \text{Slovensko}$

Jednoduchý dôsledok (simple entailment)

Rovnako ako vo formálnej logike hovoríme o splniteľnosti, logickom dôsledku a modeli formule alebo teórie môžeme hovoriť o splniteľnosti ... trojice resp. grafu.

- Interpretácia I je modelom trojice t keď $I(t) = \text{true}$ teda, ak sa trojica zobrazí na pravdivé tvrdenie.

Hovoríme tiež, že t je (jednoducho) pravdivá v I resp. I (jednoducho) spĺňa t , a značíme to:

$$I \models t$$

- Interpretácia I je modelom grafu G keď $I(G) = \text{true}$ teda, ak sa všetky trojice zobrazia na pravdivé tvrdenia.

Hovoríme tiež, že G je (jednoducho) pravdivý v I resp. I (jednoducho) spĺňa G , a značíme to:

$$I \models G$$

- Graf H je jednoduchým logickým dôsledkom grafu G , keď je H je (jednoducho) pravdivý v každej interpretácii, v ktorej je G je (jednoducho) pravdivý. Hovoríme tiež, že H jednoducho vyplýva z G a značíme to:

$$G \models H$$

- Grafy G a H sú (jednoducho) logicky ekvivalentné, keď sú si navzájom jednoduchým dôsledkom.

Interpolačná lemma

Vlastnosti jednoduchého dôsledku:

- Každý podgraf grafu je jeho jednoduchým dôsledkom.
- Graf je jednoduchým dôsledkom každej svojej inštancie.

Interpolation lemma:

Graf E je jednoduchým dôsledkom grafu G práve vtedy, keď existuje inštancia grafu E , ktorá je podgrafom G .

Dôkaz: nájdete v <http://www.w3.org/TR/rdf11-mt/#proofs-of-some-results-informative>

Táto lemma kompletne charakterizuje jednoduchý dôsledok len pomocou syntaktických pojmov. Zároveň dáva návod ako zistiť či jeden graf vyplýva z druhého.

Niektoré ďalšie vlastnosti:

- Ak je S podgraf S' a E je jednoduchým dôsledkom S , potom je E aj jednoduchým dôsledkom S' .
- Ak E obsahuje IRI, ktoré sa nevyskytuje v S , potom E nemôže byť jednoduchým dôsledkom S .
- Prázdny graf je jednoduchým dôsledkom každého grafu, naopak, žiadny neprázdny graf nie je jednoduchým dôsledkom prázdneho grafu.

Dôkaz: *téma na cvičenie a DÚ* - tvrdenia vyplývajú viacmenej priamo z definícií.