

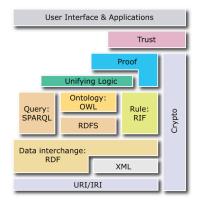
FOUNDATIONS OF SEMANTIC WEB TECHNOLOGIES

Semantics of RDF(S)

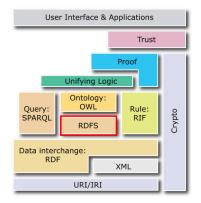
Sebastian Rudolph













Agenda



Motivation and Considerations













Agenda



Motivation and Considerations



- 3 RDF Entailment
- 4 RDFS Entailment
- 5 Downsides of RDF(S)



Why Formal Semantics?

- after introduction of RDF(S), criticism of tool developers: different tools were incompatible (despite the existing specification)
- e.g. triple stores:
 - same RDF document
 - same SPARQL query
 - different answers
- thus a model-theoretic formal semantics was defined for RDF(S)



- to start with: what are the sentences in RDF(S)?
 - basic elements (vocabulary V): IRIs, bnodes and literals (these are not sentences themselves)
 - every triple

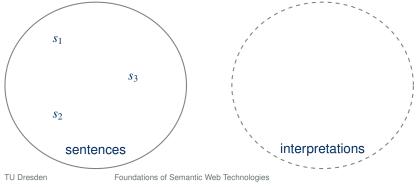
 $(s, p, o) \in (IRI \cup bnode) \times IRI \times (IRI \cup bnode \cup literal)$

is a sentence

- every finite set of triples (denoted: graph) is a sentence

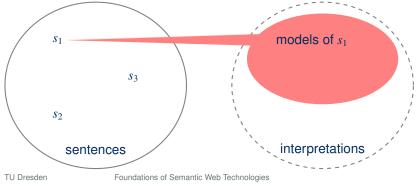


- consequence relation that defines when an RDF(S) graph G' logically follows from an RDF(S) graph G, i.e. $G \models G'$
- model-theoretic semantics: we define a set of interpretations and stipulate under which conditions an interpretation is a model of a graph



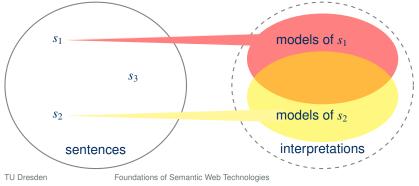


- consequence relation that defines when an RDF(S) graph G' logically follows from an RDF(S) graph G, i.e. $G \models G'$
- model-theoretic semantics: we define a set of interpretations and stipulate under which conditions an interpretation is a model of a graph



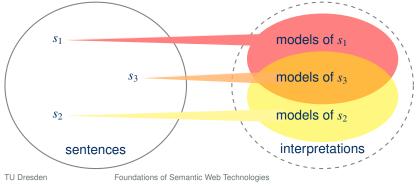


- consequence relation that defines when an RDF(S) graph G' logically follows from an RDF(S) graph G, i.e. $G \models G'$
- model-theoretic semantics: we define a set of interpretations and stipulate under which conditions an interpretation is a model of a graph



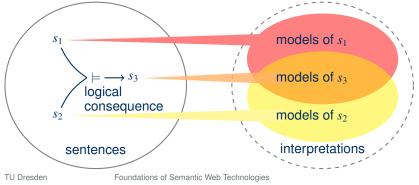


- consequence relation that defines when an RDF(S) graph G' logically follows from an RDF(S) graph G, i.e. $G \models G'$
- model-theoretic semantics: we define a set of interpretations and stipulate under which conditions an interpretation is a model of a graph





- consequence relation that defines when an RDF(S) graph G' logically follows from an RDF(S) graph G, i.e. $G \models G'$
- model-theoretic semantics: we define a set of interpretations and stipulate under which conditions an interpretation is a model of a graph





• we proceed stepwise:

simple interpretations					



• we proceed stepwise:

simple interpretations							
	RDF interpretations						



• we proceed stepwise:

si	mp	ole interpretations					
	RDF interpretations						
RDFS interpretations							



• we proceed stepwise:

simple interpretations							
	RDF interpretations						
RDFS interpretations							
]				

• the more we restrict the set of interpretations, the stronger the consequence relation becomes



Agenda



Motivation and Considerations



- 3 RDF Entailment
- 4 RDFS Entailment
- 5 Downsides of RDF(S)

TU Dresden

Foundations of Semantic Web Technologies



Definition (Simple Interpretation)

A simple Interpretation ${\mathcal I}$ for a vocabulary V consists of

- IR, a non-empty set of resources, also referred to as domain, with
- LV \subseteq IR the set of literal values, that contains (at least) all untyped literals from V, and
- IP, the set of properties of \mathcal{I} ;
- I_S, a function, mapping IRIs from V to the union of the sets IR and IP, i.e., I_S: V \rightarrow IR \cup IP,
- $I_{EXT},$ a function, mapping every property to a set of pairs from IR, i.e., $I_{EXT}\colon IP\to 2^{IR\times IR}$ and
- IL, a function mapping typed literals from V into the set IR of resources.



- IR is also called domain or universe of discourse of ${\mathcal I}$
- I_{EXT}(*p*) is also referred to as the extension of the property *p*

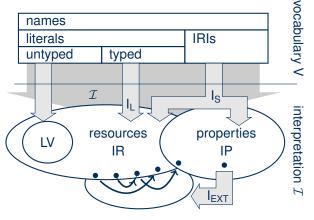
Definition (interpretation function)

based on I_L and I_S, we define $\cdot^{\mathcal{I}}$ as follows:

- every untyped literal "a" is mapped to a : $("a")^{\mathcal{I}} = a$
- every untyped literal with language information "a"@t is mapped to the pair $\langle a, t \rangle$, that is: ("a"@t)^{\mathcal{I}} = $\langle a, t \rangle$,
- every typed literal *l* is mapped to $I_L(l)$, that is: $l^{\mathcal{I}} = I_L(l)$ and
- every IRI *i* is mapped to $I_{S}(i)$, hence: $i^{\mathcal{I}} = I_{S}(i)$.



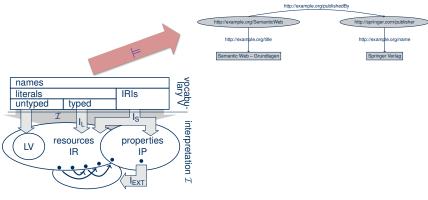
Interpretation (schematic):



Foundations of Semantic Web Technologies



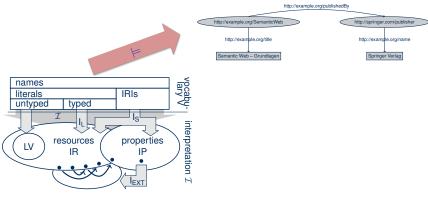
• Question: When is a given interpretation a model of a graph?



Foundations of Semantic Web Technologies

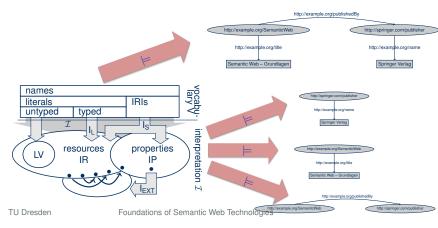


- Question: When is a given interpretation a model of a graph?
- ... if it is a model for every triple of the graph!



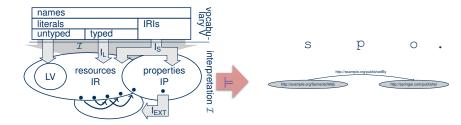


- Question: When is a given interpretation a model of a graph?
- ... if it is a model for every triple of the graph!





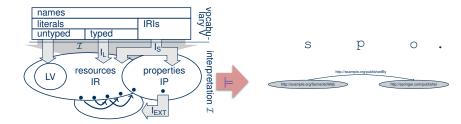
• Question: When is a given interpretation a model of a triple?



Foundations of Semantic Web Technologies

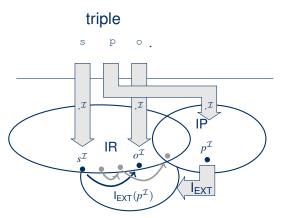


- Question: When is a given interpretation a model of a triple?
- ... if all subject, predicate, and object are contained in V and additionally $\langle s^{\mathcal{I}}, o^{\mathcal{I}} \rangle \in I_{\mathsf{EXT}}(p^{\mathcal{I}})$ holds





schematically:



Foundations of Semantic Web Technologies

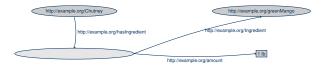


- ... oops, we forgot the bnodes!
- let A be a function mapping all bnodes to elements of IR
- given an interpretation I, let I + A behave just like I on the vocabulary, and additionally for every bnode _:label let (_:label)^{I+A} = A(_:label)
- now, an interpretation I is a model of an RDF graph G, if there exists an A such that all triples are satisfied w.r.t. I + A



Simple Interpretations: Example

given graph G:



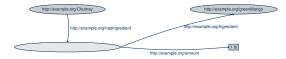
and interpretation \mathcal{I} :

IR	=	$\{c,g,h,z,l,m,1 \text{ lb}\}$	I_S	=	ex:Chutney	$\mapsto c$
IP	=	$\{h, z, m\}$			ex:greenMango	$\mapsto g$
		{1 lb}			ex:hasIngredient	$\mapsto h$
I _{EXT}	=	$h \mapsto \{\langle c, l \rangle\}$			ex:ingredient	$\mapsto z$
		$z \mapsto \{ \langle l, g \rangle \}$			ex:amount	$\mapsto m$
		$m\mapsto\{\langle l,{\rm 1lb}\rangle\}$	ΙL		is the "empty function"	

Is \mathcal{I} a model of G?



Simple Interpretations: Example



• If we pick $A: _:id1 \mapsto l$, then we get

 $\begin{array}{ll} \langle \mathrm{ex:Chutney}^{\mathcal{I}+A}, _:\mathrm{idl}^{\mathcal{I}+A} \rangle &= \langle c, l \rangle & \in \mathsf{I}_{\mathsf{EXT}}(h) = \mathsf{I}_{\mathsf{EXT}}(\mathrm{ex:hasIngredient}^{\mathcal{I}+A}) \\ \langle _:\mathrm{idl}^{\mathcal{I}+A}, \mathrm{ex:greenMango}^{\mathcal{I}+A} \rangle &= \langle l, g \rangle & \in \mathsf{I}_{\mathsf{EXT}}(z) = \mathsf{I}_{\mathsf{EXT}}(\mathrm{ex:ingredient}^{\mathcal{I}+A}) \\ \langle _:\mathrm{idl}^{\mathcal{I}+A}, \texttt{"1 } \texttt{lb}^{\texttt{"I}+A} \rangle &= \langle l, 1 \ \texttt{lb} \rangle \in \mathsf{I}_{\mathsf{EXT}}(m) = \mathsf{I}_{\mathsf{EXT}}(\mathrm{ex:amount}^{\mathcal{I}+A}) \end{array}$

• Therefore, \mathcal{I} is a model of G.



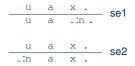
Simple Entailment

- definition of simple interpretations fixes the notion of simple entailment for RDF graphs
- question: how can this (abstractly defined) semantics be turned something computable
- answer: deduction rules



Simple Entailment

deduction rules for simple entailment:



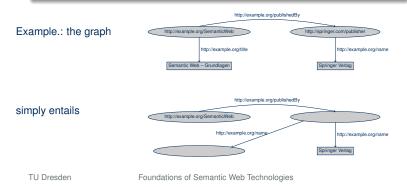
 precondition for applying this rule: the bnode has not already been associated with another IRI or literal



Simple Entailment

Theorem

A graph G_2 is simply entailed by a graph G_1 if G_1 can be extended to a graph G'_1 by applying the rules se1 and se2 such that G_2 is contained in G'_1 .





Agenda













RDF interpretations

RDF interpretations are specific simple interpretations, where additional conditions are imposed on the URIs of the RDF vocabulary

rdf:type rdf:Property rdf:XMLLiteral rdf:nil rdf:List rdf:Statement rdf:subject rdf:predicate rdf:object rdf:first rdf:rest rdf:Seq rdf:Bag rdf:Alt rdf:_1 rdf:_2 ...

inorder to realize their intended semantics.



Conditions for RDF Interpretations

An RDF interpretation for a vocabulary V is a simple interpretation for the vocabulary V \cup V_{RDF} that additionally satisfies the following conditions:

1. $x \in \mathsf{IP}$ exactly if $\langle x, \mathsf{rdf}: \mathsf{Property}^{\mathcal{I}} \rangle \in \mathsf{I}_{\mathsf{EXT}}(\mathsf{rdf}: \mathsf{type}^{\mathcal{I}})$.



Conditions for RDF Interpretations

An RDF interpretation for a vocabulary V is a simple interpretation for the vocabulary V \cup V_{RDF} that additionally satisfies the following conditions:

1. $x \in \mathsf{IP}$ exactly if $\langle x, \mathsf{rdf}: \mathsf{Property}^{\mathcal{I}} \rangle \in \mathsf{I}_{\mathsf{EXT}}(\mathsf{rdf}: \mathsf{type}^{\mathcal{I}})$.

"For every triple predicate we can infer that it is an member of the class of all properties."



An RDF interpretation for a vocabulary V is a simple interpretation for the vocabulary V \cup V_{RDF} that additionally satisfies the following conditions:

1. $x \in \mathsf{IP}$ exactly if $\langle x, \mathsf{rdf}: \mathsf{Property}^{\mathcal{I}} \rangle \in \mathsf{I}_{\mathsf{EXT}}(\mathsf{rdf}: \mathsf{type}^{\mathcal{I}})$.

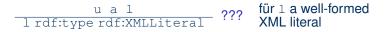
"For every triple predicate we can infer that it is an member of the class of all properties."



2. If "s"^rdf:XMLLiteral is contained in V and ${\tt s}$ is a well-formed XML literal, then

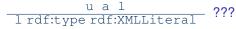
-
$$I_L("s"^rdf:XMLLiteral) \in LV;$$

- $\langle I_L("s"^{rdf:XMLLiteral}), rdf:XMLLiteral^{\mathcal{I}} \rangle \in I_{EXT}(rdf:type^{\mathcal{I}})$





2. If "s"^rdf:XMLLiteral is contained in V and ${\tt s}$ is a well-formed XML literal, then



??? für 1 a well-formed XML literal

 $-\tau$

Oops, literals must not occur in subject position!

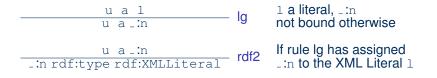


- 2. If "s"^rdf:XMLLiteral is contained in V and ${\tt s}$ is a well-formed XML literal, then
 - I_L ("s"^rdf:XMLLiteral) is the XML value of s;
 - $I_L("s"^rdf:XMLLiteral) \in LV;$
 - $\langle I_L("s"^{rdf:XMLLiteral}), rdf:XMLLiteral^{\mathcal{I}} \rangle \in I_{EXT}(rdf:type^{\mathcal{I}})$



2. If "s"^rdf:XMLLiteral is contained in V and s is a well-formed XML literal, then

$$\langle I_{L}("s"^{rdf:XMLLiteral}), rdf:XMLLiteral^{\mathcal{I}} \rangle \in I_{EXT}(rdf:type^{\mathcal{I}})$$





- 3. If "s"^rdf:XMLLiteral is contained in V and s is an ill-formed XML literal, then
 - $I_L("s"^rdf: \texttt{XMLLiteral}) \not\in \mathsf{LV}$ and
 - $\langle I_L("s"^{rdf:XMLLiteral}), rdf:XMLLiteral^{\mathcal{I}} \rangle \notin I_{EXT}(rdf:type^{\mathcal{I}}).$



- Note: *x* is a property exactly if it is linked to the resource denoted by rdf:Property via the rdf:type property (this has the direct consequence that in every RDF interpretation holds IP ⊆ IR).
- The value space of the rdf:XMLLiteral datatype contains for every well-formed XML string exactly one so-called XML value. The RDF specs only stipulate that this value is neither an XML string itself nor a data value of any XML Schema datatype nor a Unicode string.



• additional requirement: every RDF interpretation must be a model of the following "axiomatic" triples:

rdf:type	rdf:type	rdf:Property .
rdf:subject	rdf:type	rdf:Property .
rdf:predicate	rdf:type	rdf:Property .
rdf:object	rdf:type	rdf:Property .
rdf:first	rdf:type	rdf:Property .
rdf:rest	rdf:type	rdf:Property .
rdf:value	rdf:type	rdf:Property .
rdf:_1	rdf:type	rdf:Property .
rdf:_2	rdf:type	rdf:Property .
	rdf:type	rdf:Property .
rdf:nil	rdf:type	rdf:List .

every axiomatic triple "u a x ." can always be derived

Foundations of Semantic Web Technologies

rdfax

u a x



- Theorem: A graph G_2 is RDF-entailed by a graph G_1 , if there is a graph G'_1 , such that
 - G'_1 can be derived from G_1 via lg, rdf1, rdf2 and rdfax and
 - $-G_2$ is simply entailed by G_1' .
- note: two-stage deduction process



Agenda









5 Downsides of RDF(S)



...RDFS interpretations are specific RDF interpretations, where additional constraints are imposed for the URIs of the RDFS vocabulary

rdfs:domain rdfs:range rdfs:Literal rdfs:Datatype rdfs:subClassOf rdfs:subPropert rdfs:member rdfs:ContainerM rdfs:comment rdfs:seeAlso rdfs:label

rdfs:range rdfs:Resource rdfs:Datatype rdfs:Class rdfs:subPropertyOf rdfs:Container rdfs:ContainerMembershipProperty rdfs:seeAlso rdfs:isDefinedBy

such that the intended semantics of these URIs is realized.



- for the sake of easier representation, we introduce given an interpretation \mathcal{I} a function I_{CEXT} that maps resources to sets of resources (thus: I_{CEXT} : $IR \rightarrow 2^{IR}$) by letting $I_{CEXT}(y)$ contain exactly those elements x, for which $\langle x, y \rangle$ is contained in $I_{EXT}(\text{rdf:type}^{\mathcal{I}})$. We call $I_{CEXT}(y)$ the (class) extension of y.
- moreover, we let IC be the extension of the specific IRI rdfs:Class, hence: IC = I_{CEXT}(rdfs:Class^I).
- note: both I_{CEXT} as well as IC are fully determined by $\cdot^{\mathcal{I}}$ and I_{EXT} .



An RDFS interpretation for a vocabulary V is an RDF interpretation for the vocabulary V \cup V_{RDFS}, that additionally satisfies the following criteria:

- IR = I_{CEXT}(rdfs:Resource^I) Every resource is of type rdfs:Resource.
- LV = I_{CEXT}(rdfs:Literal^I) Every untyped and every well-formed typed literal is of type rdfs:Literal.
- If ⟨x, y⟩ ∈ I_{EXT}(rdfs:domain^T) and ⟨u, v⟩ ∈ I_{EXT}(x), then u ∈ I_{CEXT}(y). If the property rdfs:domain connects x with y and the property x connects the resources u and v, then u is of type y.



- If ⟨x, y⟩ ∈ I_{EXT}(rdfs:range^T) and ⟨u, v⟩ ∈ I_{EXT}(x), then v ∈ I_{CEXT}(y).
 If the property rdfs:range connects x with y and the property x connects the resources u and v, then v is of type y.
- I_{EXT}(rdfs:subPropertyOf^I) is reflexive and transitive on IP. The rdfs:subPropertyOf property connects every property with itself. Moreover, if rdfs:subPropertyOf connects a property *x* with a property *y* and additionally *y* with a property *z*, then rdfs:subPropertyOf also connects *x* directly with *z*.



- If ⟨x, y⟩ ∈ I_{EXT}(rdfs:subPropertyOf^I), then x, y ∈ IP and I_{EXT}(x) ⊆ I_{EXT}(y). If rdfs:subPropertyOf connects x with y, then both x and y are properties every pair of resources contained in the extension of x is also contained in the extension of y.
- If x ∈ IC, then ⟨x, rdfs:Resource^T⟩ ∈ I_{EXT}(rdfs:subClassOf^T). If x represents a class, then it has to be a subclass of the class of all resources, i.e., the pair containing x and rdfs:Resource is in the extension of rdfs:subClassOf.



- If ⟨x, y⟩ ∈ I_{EXT}(rdfs:subClassOf^I), then x, y ∈ IC and I_{CEXT}(x) ⊆ I_{CEXT}(y).
 If x and y are connected via the rdfs:subClassOf property, then both x and y are classes and the (class) extension of x is a subset of the (class) extension of y.
- I_{EXT}(rdfs:subClassOf^T) is reflexive and transitive on IC. The rdfs:subClassOf property connects every class to itself. Moreover, whenever this property connects a class *x* with a class *y* and a class *y* with a class *z*, then it also directly connects *x* with *z*.



- If $x \in I_{CEXT}$ (rdfs:ContainerMembershipProperty^{*I*}), then $\langle x, rdfs:member^{$ *I* $} \rangle \in I_{EXT}$ (rdfs:subPropertyOf^{*I*}). If x is a property of the type rdfs:ContainerMembershipProperty, then it is rdfs:subPropertyOf-connected with the property rdfs:member.
- If $x \in I_{CEXT}(rdfs:Datatype^{\mathcal{I}})$, then $\langle x, rdfs:Literal^{\mathcal{I}} \rangle \in I_{EXT}(rdfs:subClassOf^{\mathcal{I}})$. If some x is typed as element of the class rdfs:Datatype, then it must be a subclass of the class of all literal values (denoted by rdfs:Literal).
- ... additionally we require satisfaction of the following axiomatic triples:



rdf:tvpe rdfs:domain rdfs:range rdfs:subPropertyOf rdfs:subClassOf rdf:subject rdf:predicate rdf:object rdfs:member rdf:first rdf:rest rdfs:seeAlso rdfs:isDefinedBy rdfs:comment rdfs:label rdf:value

rdfs:domain rdfs:domain

- rdfs:Resource .
- rdf:Property .
- rdf:Property .
- rdf:Property . rdfs:Class .
- rais:class .
- rdf:Statement .
- rdf:Statement .
- rdf:Statement .
- rdfs:Resource .
- rdf:List .
- rdf:List .
- rdfs:Resource .



rdf:type	rdf
rdfs:domain	rdf
rdfs:range	rdf
rdfs:subPropertyOf	rdf
rdfs:subClassOf	rdf
rdf:subject	rdf
rdf:predicate	rdf
rdf:object	rdf
rdfs:member	rdf
rdf:first	rdf
rdf:rest	rdf
rdfs:seeAlso	rdf
rdfs:isDefinedBy	rdf
rdfs:comment	rdf
rdfs:label	rdf
rdf:value	rdf

s:range s:range

- rdfs:Class .
- rdfs:Class .
- rdfs:Class .
- rdf:Property .
- rdfs:Class .
- rdfs:Resource .
 rdf:List .
- rdfs:Resource .
- rdfs:Resource .
- rdfs:Literal .
- rdfs:Literal .
- rdfs:Resource .



rdfs:ContainerMembershipProperty			
	rdfs:subClassOf	rdf:Property .	
rdf:Alt	rdfs:subClassOf	rdfs:Container .	
rdf:Bag	rdfs:subClassOf	rdfs:Container .	
rdf:Seq	rdfs:subClassOf	rdfs:Container .	
rdfs:isDefinedBy	rdfs:subPropertyOf	rdfs:seeAlso .	
rdf:XMLLiteral	rdf:type	rdfs:Datatype .	
rdf:XMLLiteral	rdfs:subClassOf	rdfs:Literal .	
rdfs:Datatype	rdfs:subClassOf	rdfs:Class .	
rdf:_1	rdf:type		
	rdfs:ContainerMembershipProperty .		
rdf:_1	rdfs:domain	rdfs:Resource .	
rdf:_1	rdfs:range	rdfs:Resource .	
rdf:_2	rdf:type		
	rdfs:ContainerMe	mbershipProperty .	
TU Dresden	Foundations of Semantic Web Technologies		



Automatic inference is again realized via deduction rules:

every axiomatic triple "u a x ." rdfsax rdfsax can always be derived $\label{eq:linear} \begin{array}{c} _ u \ a \ _ : n \ . \end{array} \ gl \quad the \ converse \ of \ Rule \ lg: \ _:n \ has \ been \ assigned \ (via \ Rule \ lg) \ to \ the \ untyped \ literal \ l$...:n has been assigned (via Rule lg) to the u a l . _:n rdf:type rdfs:Literal rdfs1 untyped literal 1 a rdfs:domain x . u a y . rdfs2 implements the semantics of u rdf:type x . property domains implementis the semantics of a rdfs:range x . u a v . rdfs3 v rdf:tvpe x . property ranges x, y IRI, blank node or literal a, b IRIs u, v IRI or blank node 1 literal in blank nodes



u a x . u rdf:type rdfs:Resource . rdfs4a	the subject of every triple is a resource			
u a v . v rdf:type rdfs:Resource . rdfs4b	objects that are not literals are resources as well			
urdfs:subPropertyOfv.vrdfs:subPropertyOfx. urdfs:subPropertyOfx.				
u rdf:type rdf:Property . u rdfs:subPropertyOf u . rdfs6 reflexivity				
a rdfs:subPropertyOf b . u a y . u b y .	 rdfs7 subproperty inferences for instances 			
u rdf:type rdfs:Class . u rdf:subClassOf rdfs:Resource .	rdfs8 classes contain only resources			



u rdfs:subClassOf x . v rdf:type	v rdf:type u . rdfc0	subclassen inferences
v rdf:type	x .	for instances

<u>u rdf:type rdfs:Class .</u> rdfs10 reflexivity <u>u rdfs:subClassOf u .</u>

u rdfs:subClassOfv.vrdfs:subClassOfx. u rdfs:subClassOfx.

u rdf:type rdfs:ContainerMembershipProperty . u rdfs:subPropertyOf rdfs:member .

u rdf:type rdfs:Datatype . u rdfs:subClassOf rdfs:Literal . rdfs10 subclass of rdfs:Literal



• important definition: XML clash

ex:hasSmiley rdfs:range rdfs:Literal.

ex:evilRemark ex:hasSmiley ">:->"^rdf:XMLLiteral.

occurs if a node of type rdfs:Literal gets assigned an ill-formed literal value



Theorem:

A graph G_2 is RDFS entailed by G_1 , if there is a graph G'_1 obtained by applying the rules lg, gl, rdfax, rdf1, rdf2, rdfs1 – rdfs13 and rdfsax to G_1 , such that

- G_2 is simply entailed by G'_1 or
- G'_1 contains an XML clash.



Agenda





- 3 RDF Entailment
- 4 RDFS Entailment





What RDF(S) Cannot Do

• Certain seemingly sensible consequences are not RDFS-entailed, e.g.

ex:talksTo ex:Homo	rdfs:domain rdfs:subClas:	sOf	ex:Homo. ex:Primates.
should imply			
ex:talksTo	rdfs:domain	ex:P	rimates.

- possible solution: use a stronger, so-called "extensional" semantics (but this would be outside the standard)
- no possibility to express negation