# The Logic of the Semantic Web

Enrico Franconi

Free University of Bozen-Bolzano, Italy

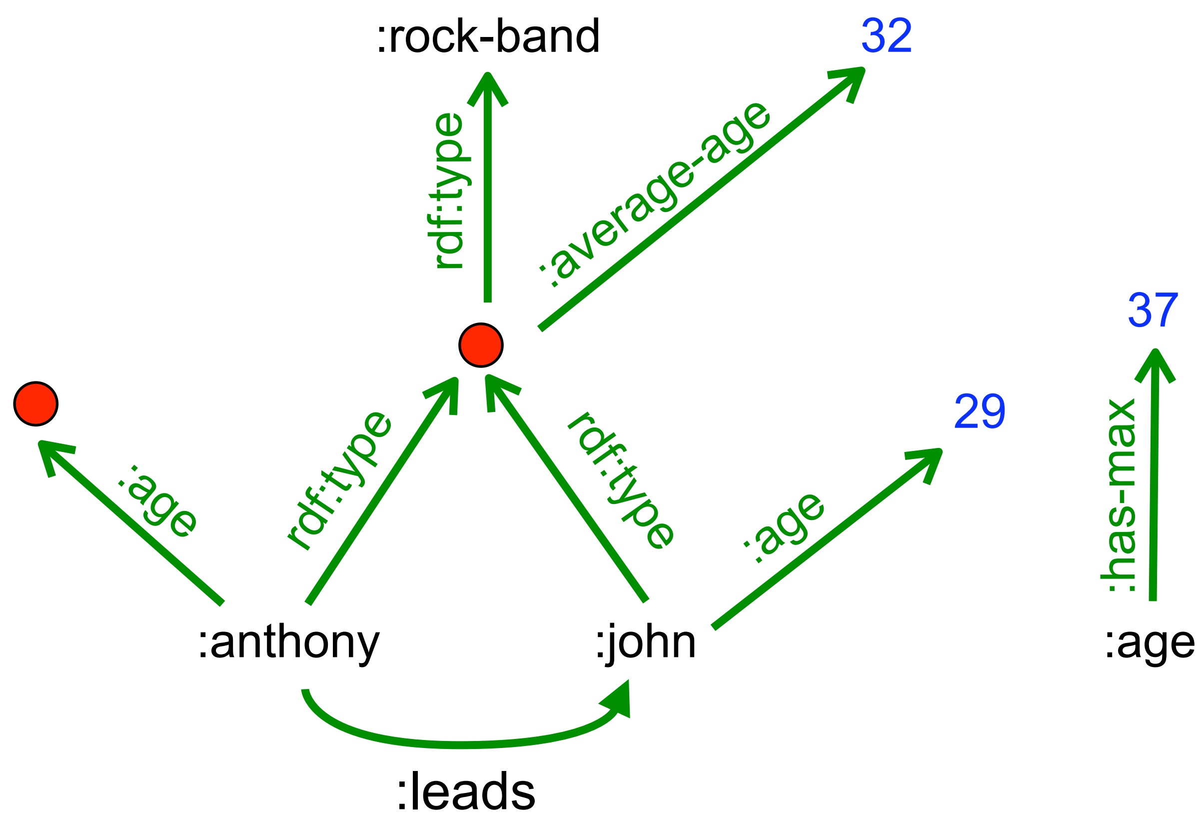
# RDF in the real world

- RDF and SPARQL are W3C standards
- Widespread use for metadata representation, e.g.
  - Apple (MCF)
  - Adobe (XMP)
  - Mozilla/Firefox
- Oracle supports RDF, and provides an extension of SQL to query RDF data
- HP has a big lab (in Bristol) developing specialised data stores for RDF (Jena)
- ...but: research is beyond practice

# RDF

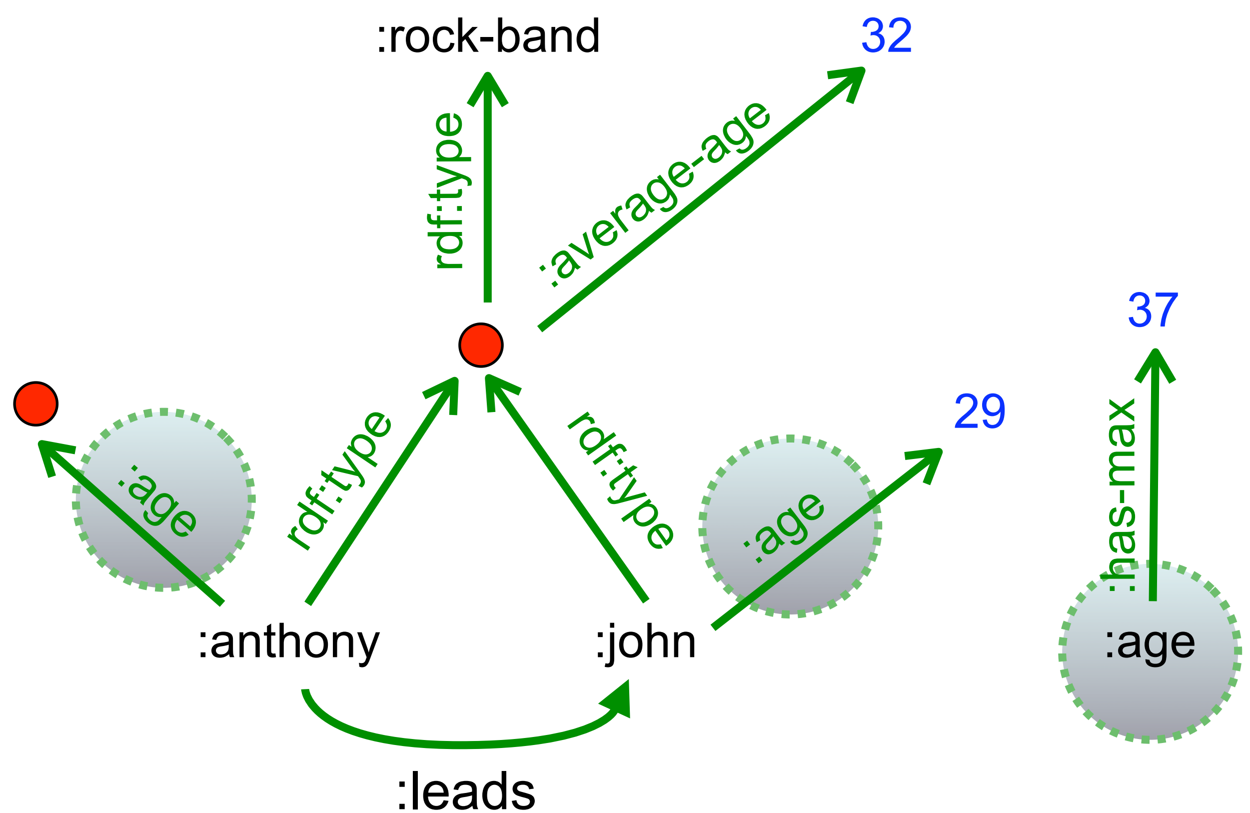
- A node- and edge-labelled directed graph
  - edges are called properties
  - the left node in a directed edge is called subject
  - the right node in a directed edge is called object
- Typical notations are:
  - **p(s,o)**
  - **Triple(s, p, o)**
  - **s p o.**
- Labels are URIs, literals, or bnodes

# Example




:antony :leads :john.
:antony :age _:a.
:antony rdf:type _:b.
_:b rdf:type :rock-band.
:age :has-max 37.
...

# Example




:antony :leads :john.
:antony :age _:a.
:antony rdf:type _:b.
_:b rdf:type :rock-band.
:age :has-max 37.
...

# RDF peculiarities

- Some labels are anonymous: bnodes
- The alphabets of labels for nodes and for properties are not disjoint: a coreference is possible between nodes and properties
- There is a special pre-defined non well-founded "**rdf:type**" property, with the intended meaning of "is-element-of"

# Meaning of RDF graphs

- We want to provide a model-theoretic semantics to RDF graphs, in order to properly define entailment and query answering
- We consider here a simplified RDF language:
  - no restrictions on literals
    - in normative RDF literals are not allowed in subject position
  - no restrictions on properties
    - in normative RDF bnodes are not allowed in property position
  - no "axiomatic" knowledge

# RDF semantics (atoms)

$$\mathcal{I} = \langle \Delta^{\mathcal{I}}, \cdot^{\mathcal{I}}, \cdot^{\mathcal{I}_p} \rangle$$

$$\begin{array}{rcl} \cdot^{\mathcal{I}} & : & \mathbb{U} \cup \mathbb{L} \mapsto \Delta^{\mathcal{I}} \\ \cdot^{\mathcal{I}_p} & : & \Delta^{\mathcal{I}} \mapsto 2^{\Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}} \\ \alpha & : & \mathbb{B} \mapsto \Delta^{\mathcal{I}} \end{array}$$

# RDF semantics (atoms)

$$\mathcal{I} = \langle \Delta^{\mathcal{I}}, \cdot^{\mathcal{I}}, \cdot^{\mathcal{I}_p} \rangle$$

$$\begin{array}{lcl} \cdot^{\mathcal{I}} & : & \mathbb{U} \cup \mathbb{L} \mapsto \Delta^{\mathcal{I}} \\ \cdot^{\mathcal{I}_p} & : & \Delta^{\mathcal{I}} \mapsto 2^{\Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}} \\ \alpha & : & \mathbb{B} \mapsto \Delta^{\mathcal{I}} \end{array}$$

$$\begin{array}{rcl} u^{\mathcal{I},\alpha} & = & u^{\mathcal{I}} \\ l^{\mathcal{I},\alpha} & = & l^{\mathcal{I}} \\ b^{\mathcal{I},\alpha} & = & \alpha(b) \end{array}$$

# RDF semantics (atoms)

$$\mathcal{I} = \langle \Delta^{\mathcal{I}}, \cdot^{\mathcal{I}}, \cdot^{\mathcal{I}_p} \rangle$$

$$\begin{array}{lcl} \cdot^{\mathcal{I}} & : & \mathbb{U} \cup \mathbb{L} \mapsto \Delta^{\mathcal{I}} \\ \cdot^{\mathcal{I}_p} & : & \Delta^{\mathcal{I}} \mapsto 2^{\Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}} \\ \alpha & : & \mathbb{B} \mapsto \Delta^{\mathcal{I}} \end{array} \qquad \begin{array}{rcl} u^{\mathcal{I},\alpha} & = & u^{\mathcal{I}} \\ l^{\mathcal{I},\alpha} & = & l^{\mathcal{I}} \\ b^{\mathcal{I},\alpha} & = & \alpha(b) \end{array}$$

$$\mathcal{I}, \alpha \models p(s, o) \quad \text{iff} \quad \langle s^{\mathcal{I},\alpha}, o^{\mathcal{I},\alpha} \rangle \in (p^{\mathcal{I},\alpha})^{\mathcal{I}_p}$$

# RDF semantics (atoms)

$$\mathcal{I} = \langle \Delta^{\mathcal{I}}, \cdot^{\mathcal{I}}, \cdot^{\mathcal{I}_p} \rangle$$

$$\begin{array}{rcl} \cdot^{\mathcal{I}} & : & \mathbb{U} \cup \mathbb{L} \mapsto \Delta^{\mathcal{I}} \\ \cdot^{\mathcal{I}_p} & : & \Delta^{\mathcal{I}} \mapsto 2^{\Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}} \\ \alpha & : & \mathbb{B} \mapsto \Delta^{\mathcal{I}} \end{array} \qquad \begin{array}{rcl} u^{\mathcal{I},\alpha} & = & u^{\mathcal{I}} \\ l^{\mathcal{I},\alpha} & = & l^{\mathcal{I}} \\ b^{\mathcal{I},\alpha} & = & \alpha(b) \end{array}$$

$$\mathcal{I}, \alpha \models p(s, o) \quad \text{iff} \quad \langle s^{\mathcal{I},\alpha}, o^{\mathcal{I},\alpha} \rangle \in (p^{\mathcal{I},\alpha})^{\mathcal{I}_p}$$

$$\mathcal{I}, \alpha \models p(s, o) \quad \text{iff} \quad \langle s^{\mathcal{I},\alpha}, o^{\mathcal{I},\alpha} \rangle \in p^{\mathcal{I}_p,\alpha} \quad \text{(FOL)}$$

# RDF semantics (atoms)

$$\mathcal{I} = \langle \Delta^{\mathcal{I}}, \cdot^{\mathcal{I}}, \cdot^{\mathcal{I}_p} \rangle$$

$$\begin{array}{rcl} \cdot^{\mathcal{I}} & : & \mathbb{U} \cup \mathbb{L} \mapsto \Delta^{\mathcal{I}} \\ \cdot^{\mathcal{I}_p} & : & \Delta^{\mathcal{I}} \mapsto 2^{\Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}} \\ \alpha & : & \mathbb{B} \mapsto \Delta^{\mathcal{I}} \end{array} \qquad \begin{array}{rcl} u^{\mathcal{I},\alpha} & = & u^{\mathcal{I}} \\ l^{\mathcal{I},\alpha} & = & l^{\mathcal{I}} \\ b^{\mathcal{I},\alpha} & = & \alpha(b) \end{array}$$

$$\mathcal{I}, \alpha \models p(s, o) \quad \text{iff} \quad \langle s^{\mathcal{I},\alpha}, o^{\mathcal{I},\alpha} \rangle \in (p^{\mathcal{I},\alpha})^{\mathcal{I}_p}$$

$$\mathcal{I}, \alpha \models p(s, o) \quad \text{iff} \quad \langle s^{\mathcal{I},\alpha}, p^{\mathcal{I},\alpha}, o^{\mathcal{I},\alpha} \rangle \in T^{\mathcal{I}}$$

# RDF semantics and entailment

- Non-atomic formulas:

$$\mathcal{I},\alpha \models \{p_1(s_1,o_1), p_2(s_2,o_2), \cdots\} \quad \underline{\text{iff}}$$

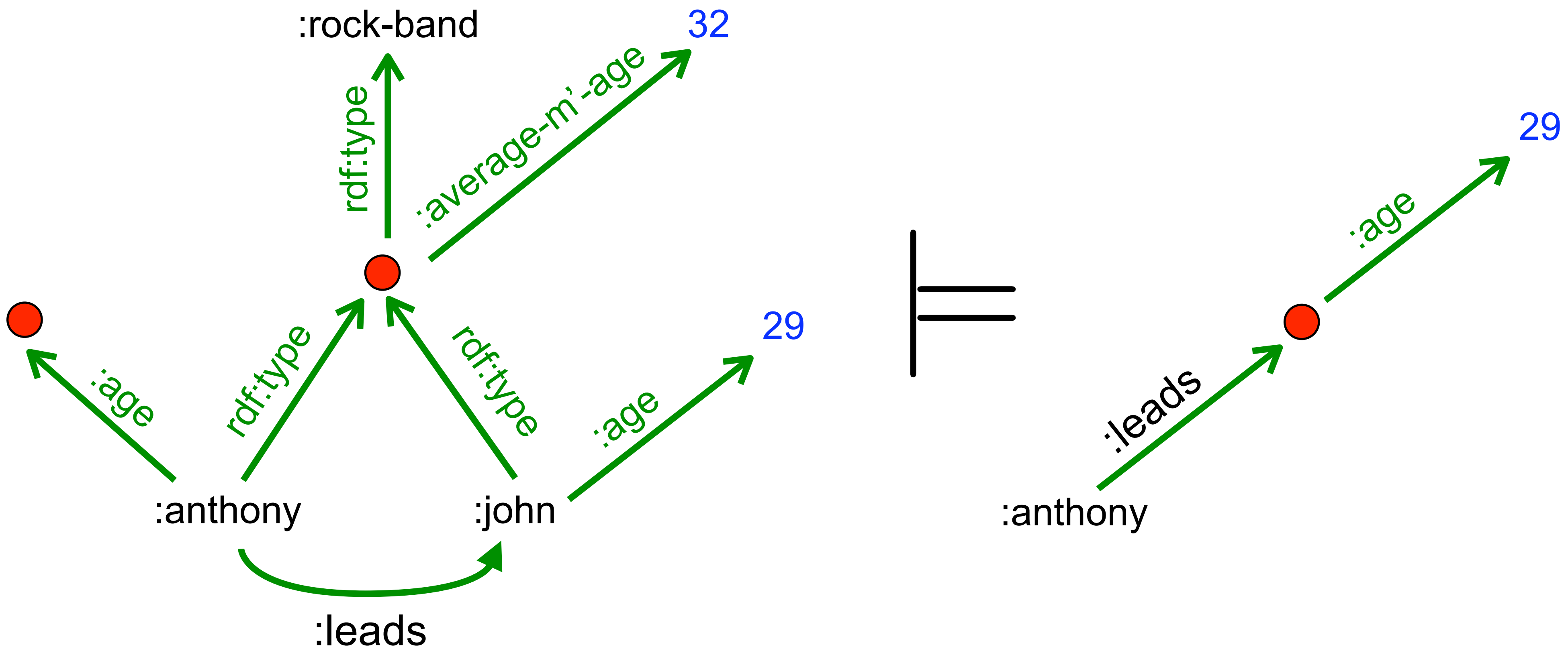
$$\mathcal{I},\alpha \models p_1(s_1,o_1) \quad \text{and}$$

$$\mathcal{I},\alpha \models \{p_2(s_2,o_2), \cdots\}$$

- $\mathcal{I}$ is a model of an RDF graph G, written $\mathcal{I} \models G$, if there exists an $\alpha$ such that $\mathcal{I},\alpha \models G$

- An RDF graph G entails an RDF graph H ( $G \models H$ ) iff for any $\mathcal{I}$ such that $\mathcal{I} \models G$ then $\mathcal{I} \models H$

# Example



:rock-band
32
rdf:type
:average-m'-age
rdf:type
rdf:type
:age
:age
29
:anthony
:john
:leads
29
:age
:leads
:anthony

# RDF and FOL

[_, Tessaris, 2004] The models of an RDF graph

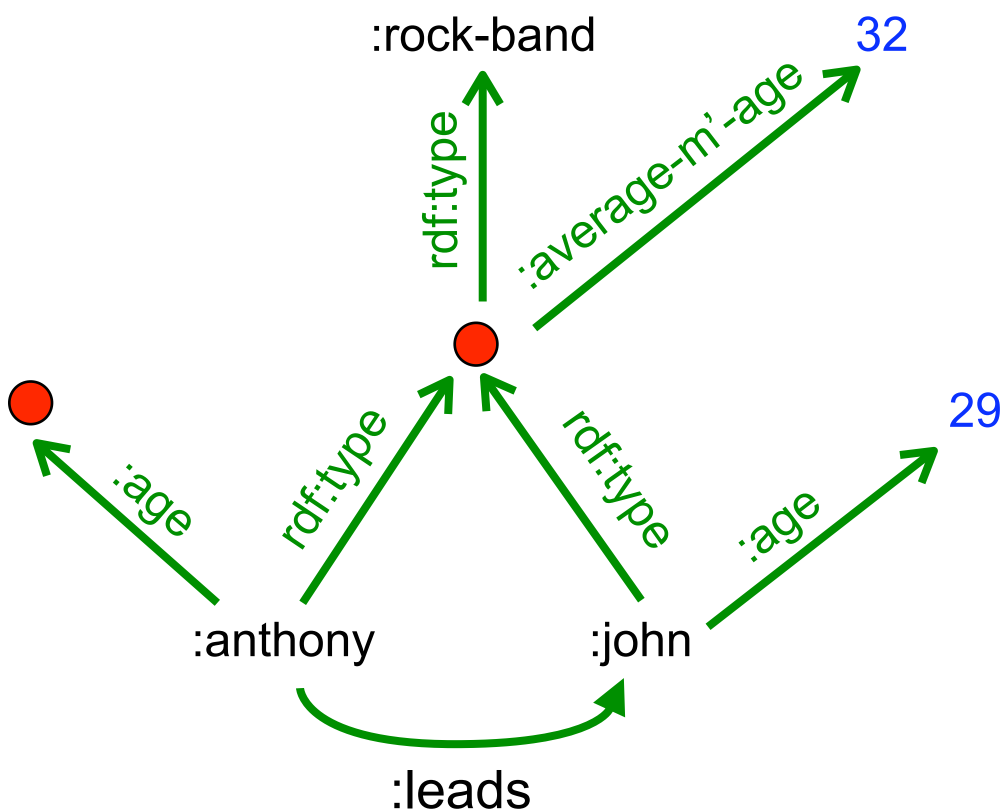
$$\mathcal{I} \models \{p_1(s_1, o_1), p_2(s_2, o_2), \cdots\}$$

are the same as the models of the FOL formula

$$\mathcal{I} \models_{\text{FOL}} \exists \bar{b}.\ T(s_1, p_1, o_1) \wedge T(s_2, p_2, o_2) \wedge \cdots$$

where $\bar{b}$ is the set of bnode names appearing in the graph

# Example




∃x,y. T(:antony, :leads, :john) ∧ T(:antony, :age, x) ∧ T(:antony, rdf:type, y) ∧
T(y, rdf:type :rock-band) ∧ ...

# Complexity of RDF entailment

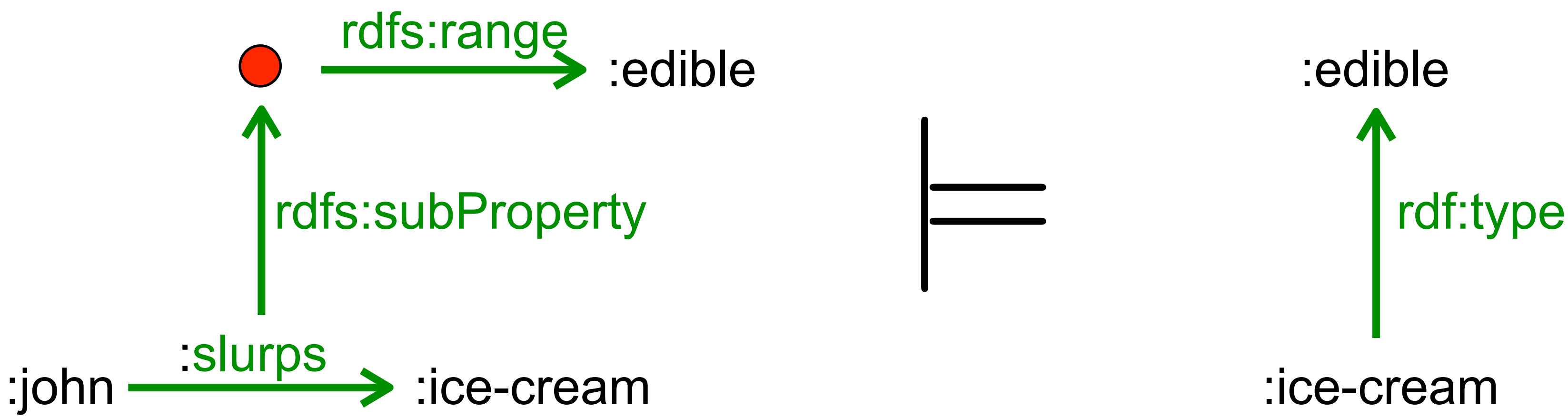
$$G \models H$$

- NP-complete in the size of the graphs
- Polynomial in the size of the entailing graph G
- Algorithm: reduction to conjunctive query containment
  - Typically implemention: graph homomorphism

# RDFS

- RDFS adds to the signature properties with a fixed semantics
  - **rdf:type** (= is-element-of)
  - **rdfs:subclass**
  - **rdfs:subproperty**
  - **rdfs:domain**
  - **rdfs:range**
- Note that the (above) properties are also elements of the domain

# Example



rdfs:range
:edible
rdfs:subProperty
:slurps
:john
:ice-cream
⊨
:edible
rdf:type
:ice-cream

# Normative semantics of RDFS

$$\text{rdfs:subclass}^{\mathcal{I}_p,\alpha} \subseteq$$
$$\{\langle u,v\rangle \in \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}} \mid$$
$$\forall x.\ (x,u) \in \text{rdf:type}^{\mathcal{I}_p,\alpha} \rightarrow (x,v) \in \text{rdf:type}^{\mathcal{I}_p,\alpha}\}$$

$$\text{rdfs:domain}^{\mathcal{I}_p,\alpha} \subseteq$$
$$\{\langle u,v\rangle \in \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}} \mid$$
$$\forall x,y.\ (x,y) \in u^{\mathcal{I}_p,\alpha} \rightarrow (x,v) \in \text{rdf:type}^{\mathcal{I}_p,\alpha}\}$$

# Normative semantics of RDFS (in FOL)

$$\forall u, v.$$
$$T(u, \text{rdfs:subclass}, v) \rightarrow$$
$$\forall x.\ T(x, \text{rdf:type}, u) \rightarrow T(x, \text{rdf:type}, v)$$

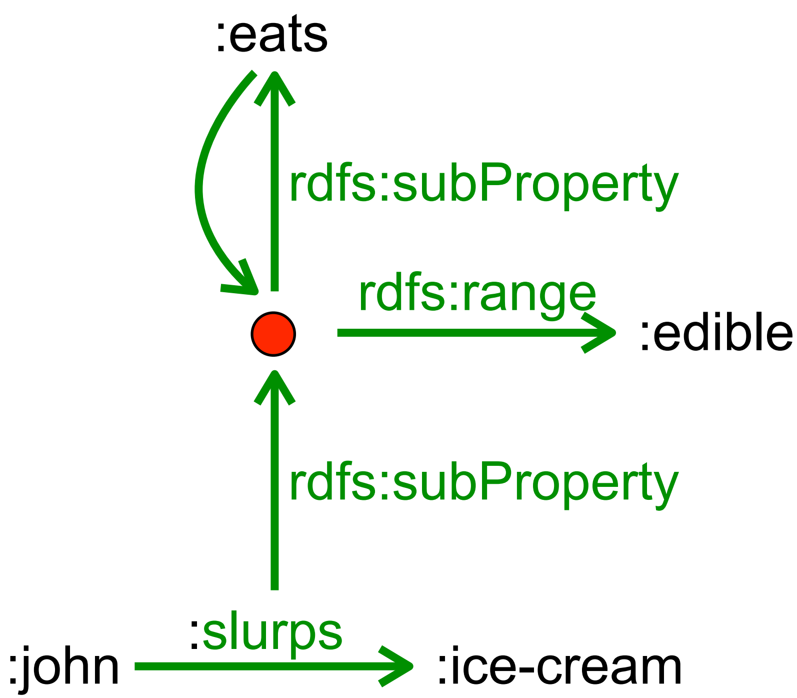
$$\forall u, v.$$
$$T(u, \text{rdfs:domain}, v) \rightarrow$$
$$\forall x, y.\ T(x, u, y) \rightarrow T(x, \text{rdf:type}, v)$$

# Entailment in normative RDFS

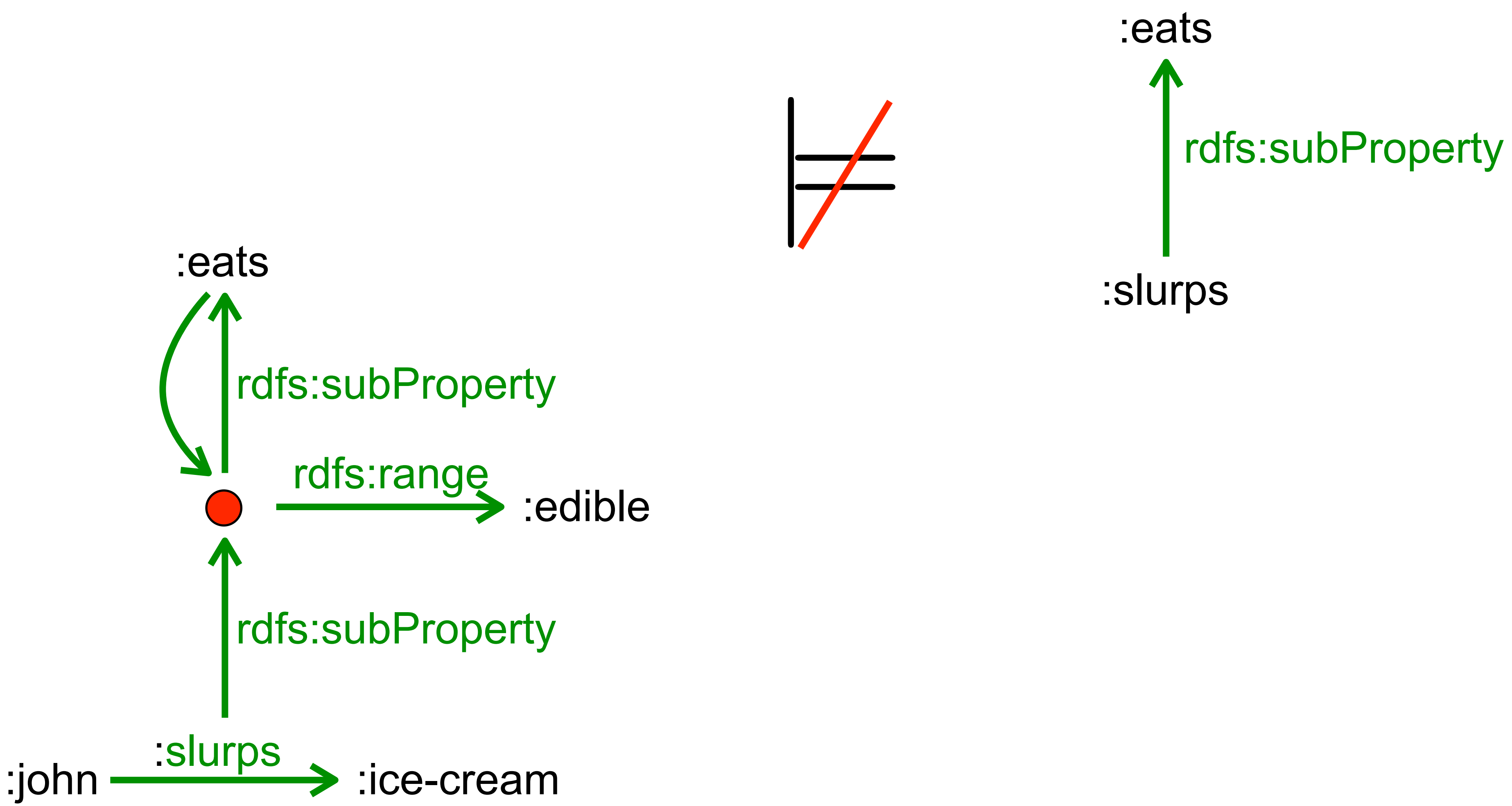
$$G \models_{\mathrm{RDFS}} H$$

- Entailment under constraints
- [ter Horst, 2005] NP-complete in the size of the graphs
  - Polynomial if H does not contain bnodes
- Algorithm: reduction to RDF entailment through a completion of graph G
  - Warning: W3C standard algorithm (by P. Hayes) is incomplete [ter Horst, 2005; Gutiérrez et al, 2004] (e.g., the previous example does not work)

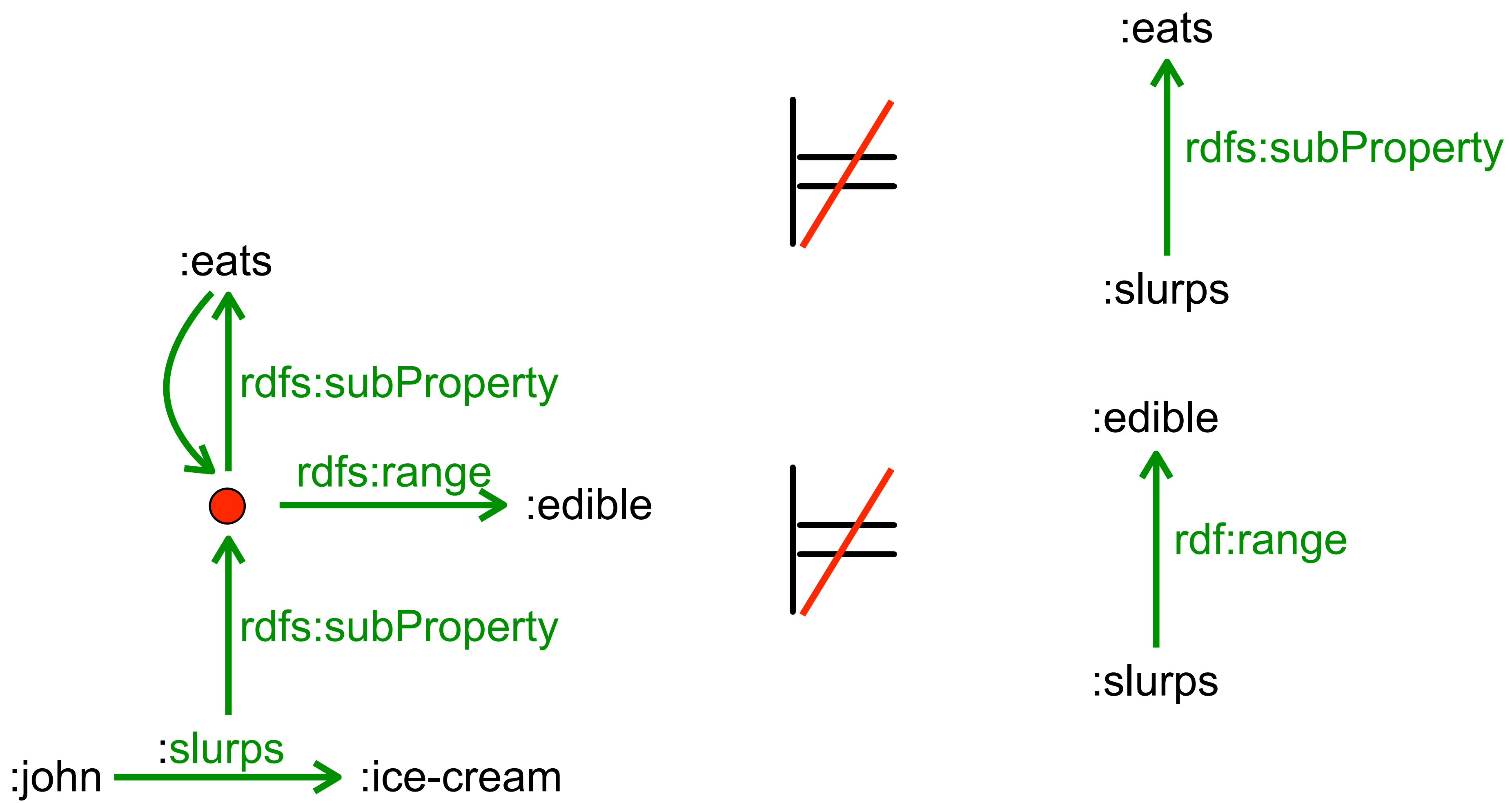
# Example revisited ☹



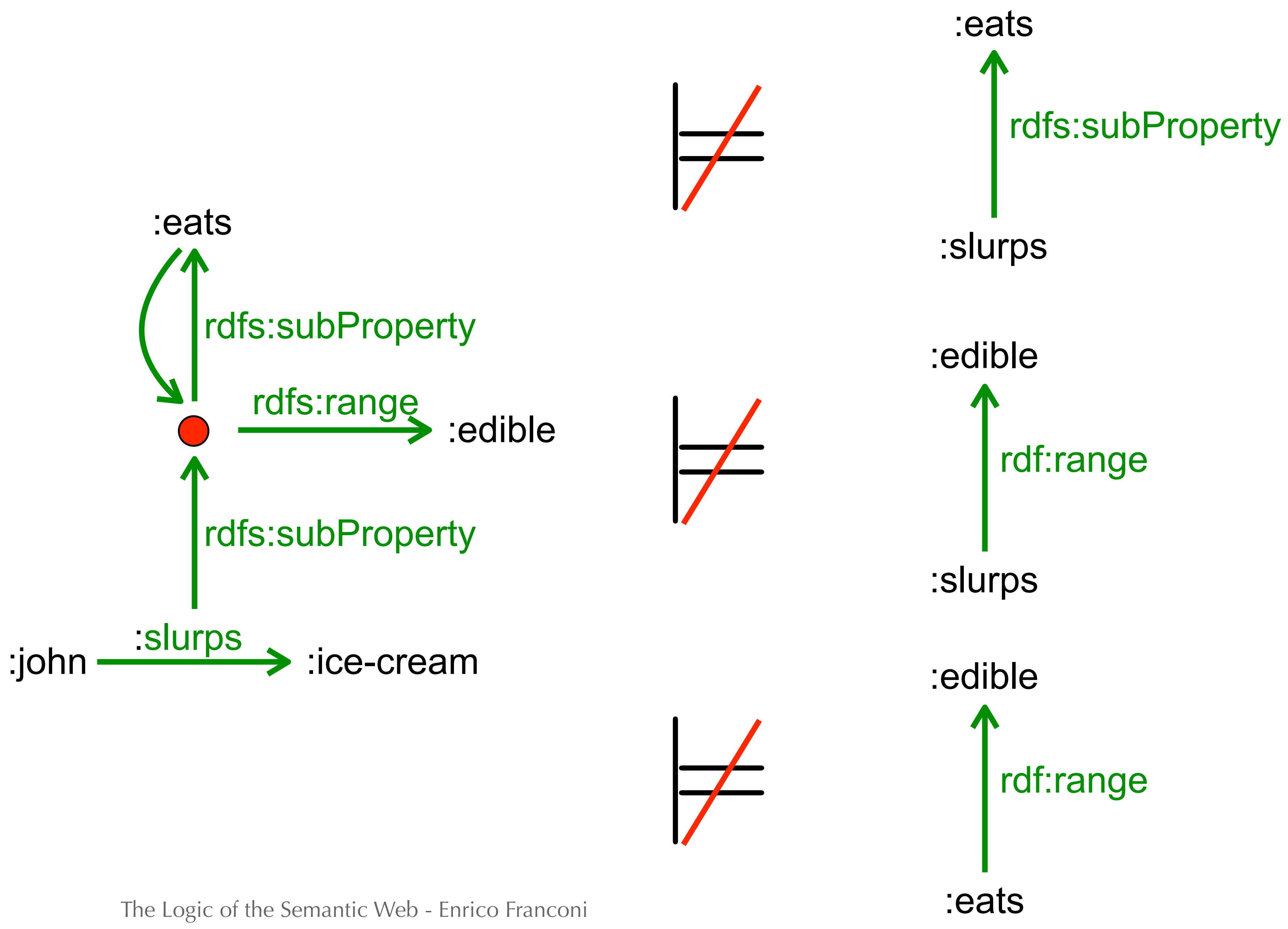
# Example revisited ☹



# Example revisited ☹



# Example revisited ☹



# Extensional semantics of RDFS

$$\forall u, v.$$
$$T(u, \text{rdfs:subclass}, v) \leftrightarrow$$
$$(\forall x.\ T(x, \text{rdf:type}, u) \rightarrow T(x, \text{rdf:type}, v))$$

$$\forall u, v.$$
$$T(u, \text{rdfs:domain}, v) \leftrightarrow$$
$$(\forall x, y.\ T(x, u, y) \rightarrow T(x, \text{rdf:type}, v))$$

# Entailment in extensional RDFS

$$G \models_{\mathrm{RDFS}^e} H$$

- Entailment under constraints
- General algorithm not known
- Complexity known only if H does not contain bnodes:
  - Theorem [_,Rosati, 2006 - unpublished]: polynomial in the size of the graphs