

Homomorphismes d'hypergraphes pour la subsomption en RDF/RDFS

Jean-François Baget

Projet EXMO

INRIA Rhône-Alpes

jean-francois.baget@inrialpes.fr

Langages et Modèles à Objets

LMO 2004

Lille – 15-17 mars 2004



La recommandation RDF



W3C Recommendation

- RDF/XML Syntax Specification (Revised)
Dave Beckett, ed.
- RDF Vocabulary Description Language 1.0: RDF Schema
Dan Brickley, R.V. Guha, eds.
- RDF Primer
Frank Manola, Eric Miller, eds.
- Resource Description Framework (RDF): Concepts and Abstract Syntax
Graham Klyne, Jeremy Carroll, eds.
- RDF Semantics
Patrick Hayes, ed.
- RDF Test Cases
Jan Grant, Dave Beckett, eds.

Plan de la présentation

- La syntaxe RDF
- La sémantique RDFL
- L'inférence en RDFL: un homomorphisme
- Extension à RDF
- Extension à RDFS
- Comparaison avec CORESE
- Bénéfices de la reformulation
- RDFL + Règles

Plan de la présentation

- La syntaxe RDF
- La sémantique RDFL
- L'inférence en RDFL: un homomorphisme
- Extension à RDF
- Extension à RDFS
- Comparaison avec CORESE
- Bénéfices
- RDFL + Règles

Le vocabulaire RDF

- **Urirefs:**
 - Désignent des ressources sur le web (ex: URL)
- **Littéraux**
 - Simples: représentation lexicale = valeur
 - Typés: valeur (VAL) déterminée par le type comme par le représentation
- **Variables**

Le triplet RDF

(SUJET

PREDICAT

OBJET)

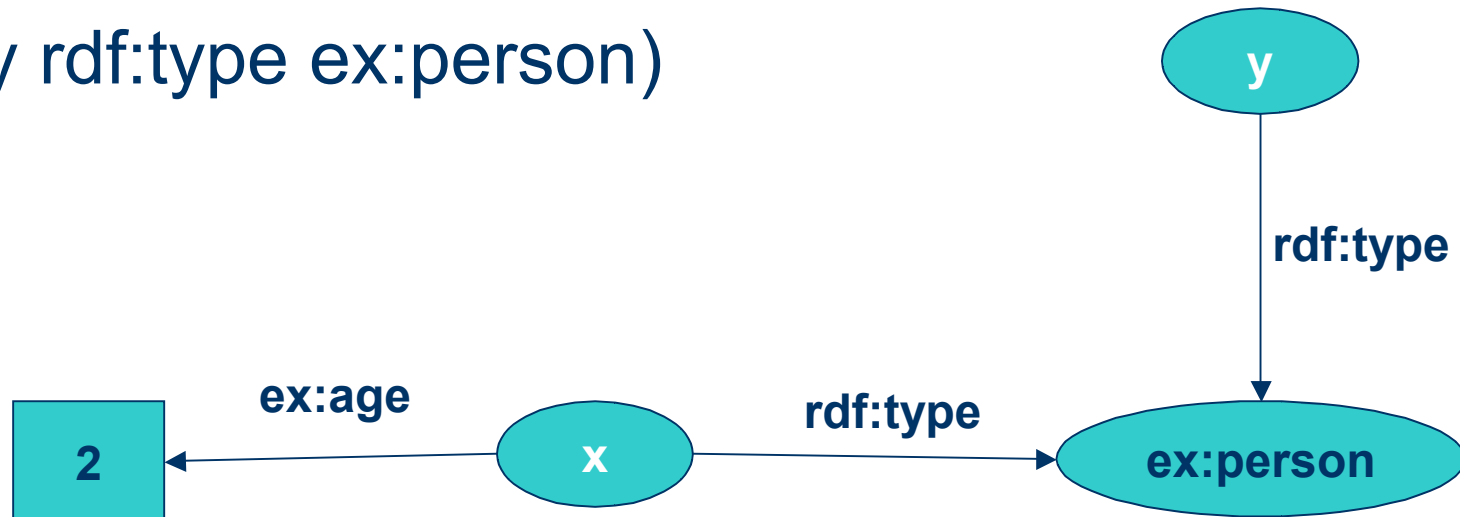
- uriref
- variable

- uriref

- uriref
- variable
- littéral

Le document/graphe RDF

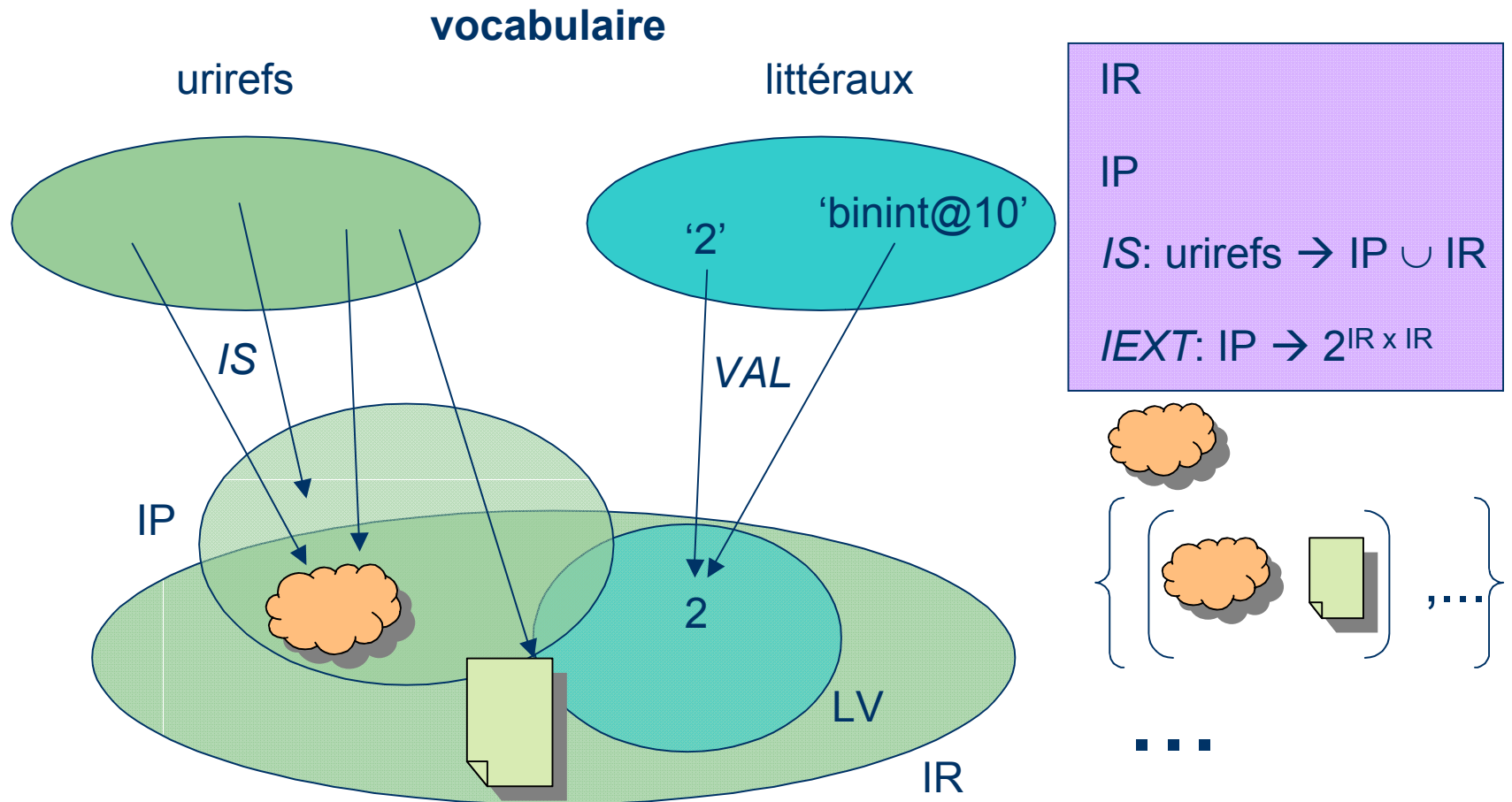
```
{  
  (x rdf:type ex:person)  
  (x ex:age "2")  
  (y rdf:type ex:person)  
}
```



Plan de la présentation

- La syntaxe RDF
- La sémantique RDFL [Pat. Hayes, 2004]
- L'inférence en RDFL: un homomorphisme
- Extension à RDF
- Extension à RDFS
- Comparaison avec CORESE
- Bénéfices
- RDFL + Règles

Interprétation du vocabulaire



Modèle d'un document

L'interprétation du vocabulaire est un modèle pour un document G ssi il existe $IV: \text{VAR}(G) \rightarrow \text{IR}$ telle que pour tout triplet:



$$(IS/IV(S) , IS/IV/VAL(O)) \in IEXT(IS(P))$$

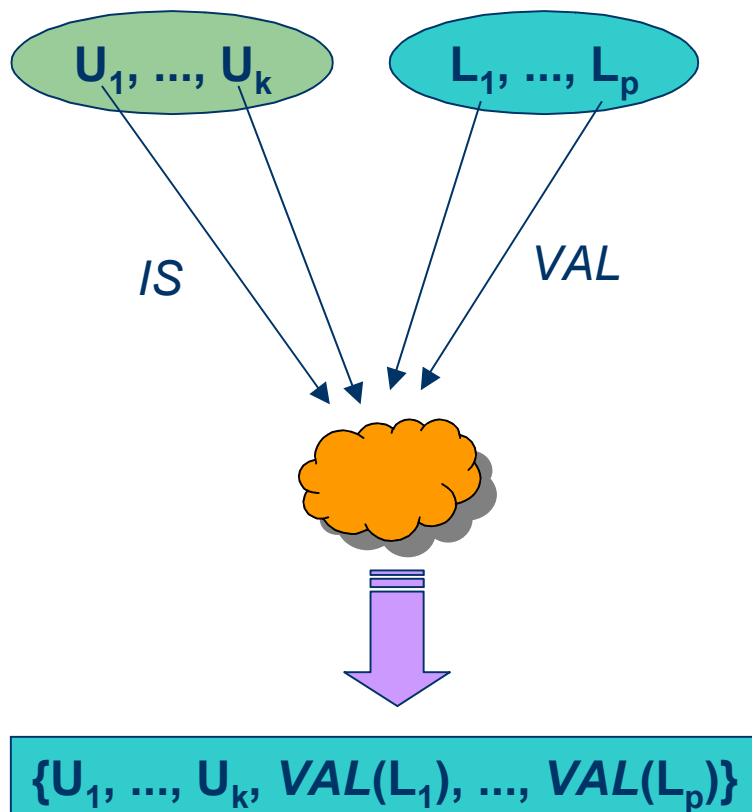
IS/IV/VAL: Preuve

Un document H est conséquence sémantique d'un document G ssi tout modèle de G est modèle de H .

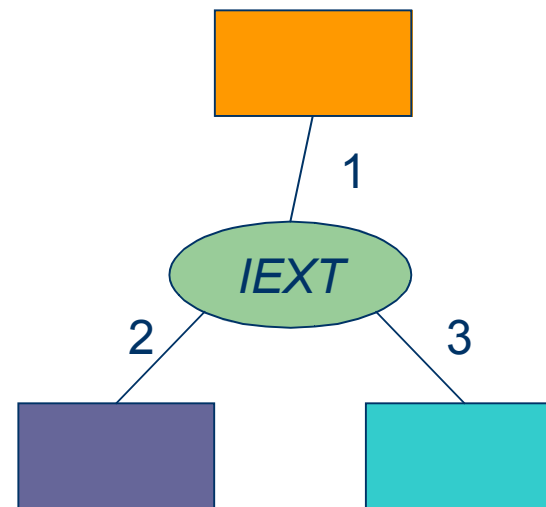
Plan de la présentation

- La syntaxe RDF
- La sémantique RDFL
- L'inférence en RDFL: un homomorphisme
- Extension à RDF
- Extension à RDFS
- Comparaison avec CORESE
- Bénéfices
- RDFL + Règles

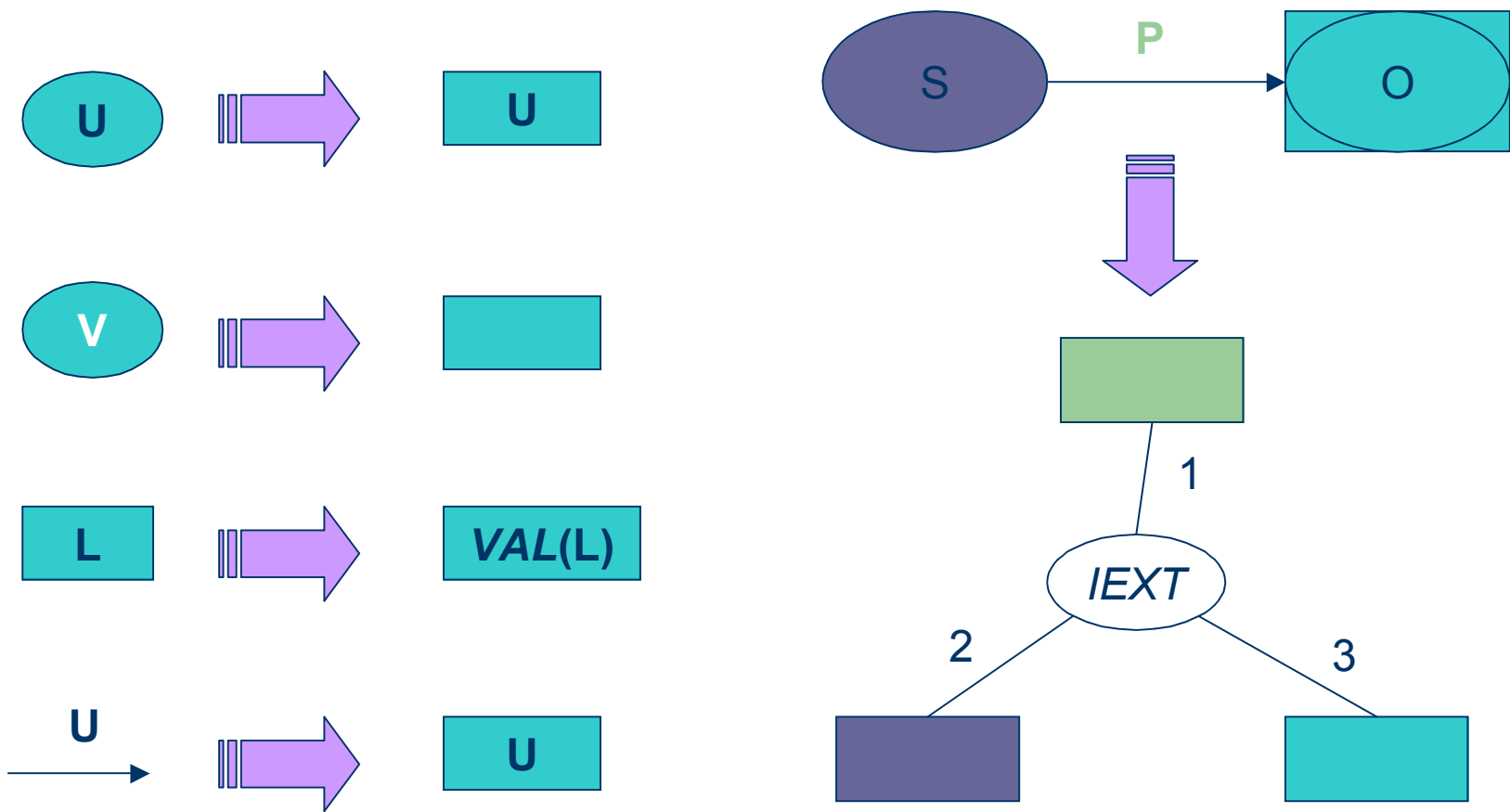
L'interprétation: un hypergraphe



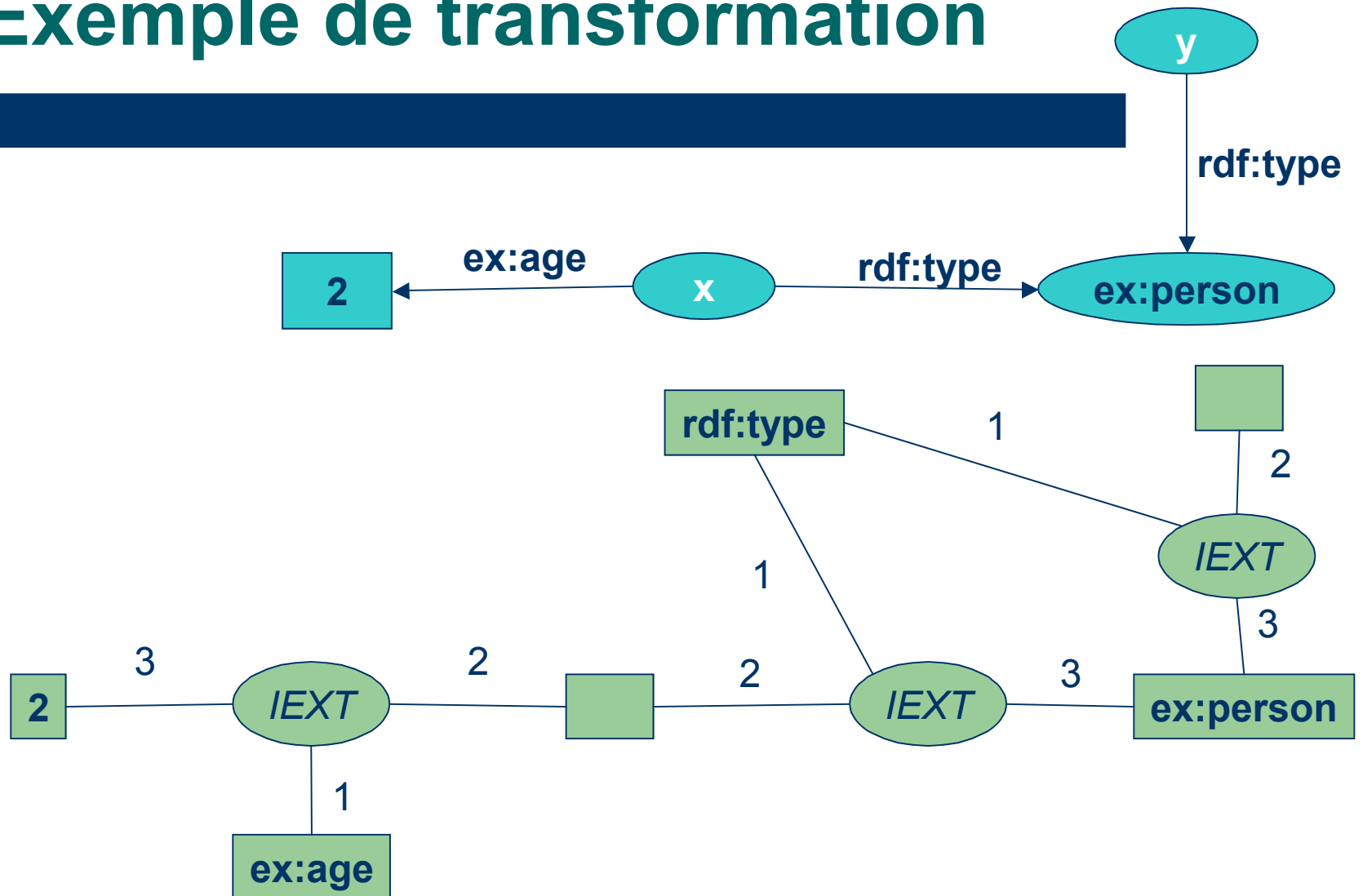
$$\left(\text{cloud}_1, \text{cloud}_2 \right) \in IEXT \left(\text{cloud}_3 \right)$$



Le document: un hypergraphe



Exemple de transformation



La preuve: un homomorphisme

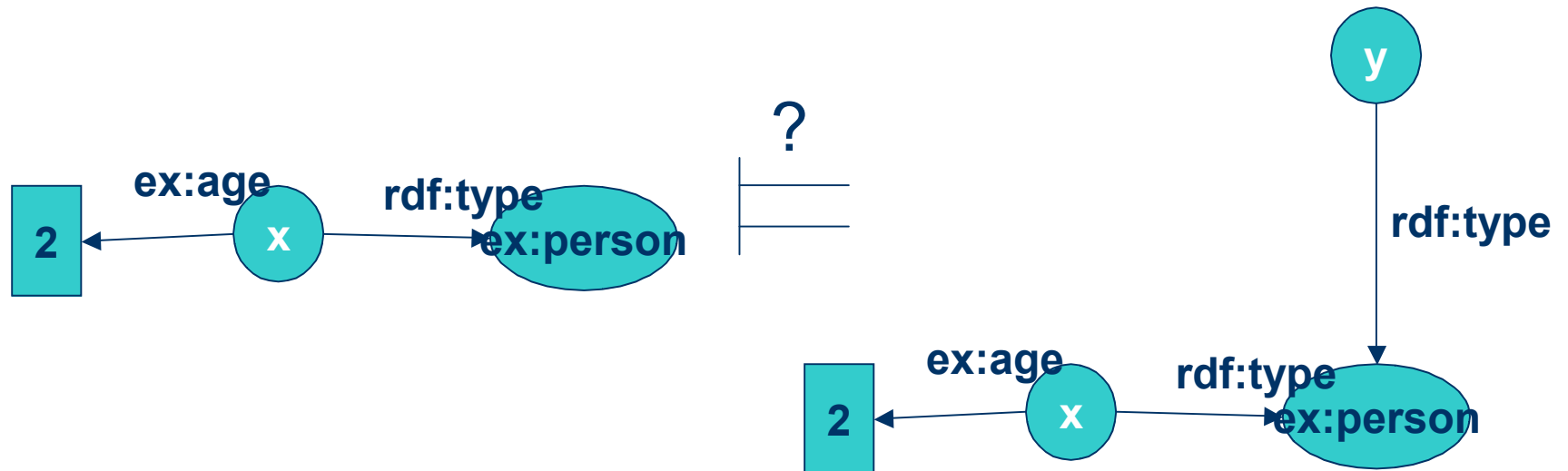
Soient G et H deux hypergraphes. Une *projection* de H dans G est une application Π de $V(H)$ dans $V(G)$ telle que:

- $\forall x \in V(H), \text{etiq}(\Pi(x)) \leq \text{etiq}(x)$
- $\forall r = (x_1, \dots, x_k) \in U(H), \exists r' = (\Pi(x_1), \dots, \Pi(x_k)) \in U(G) / \text{etiq}(r') \leq \text{etiq}(r)$

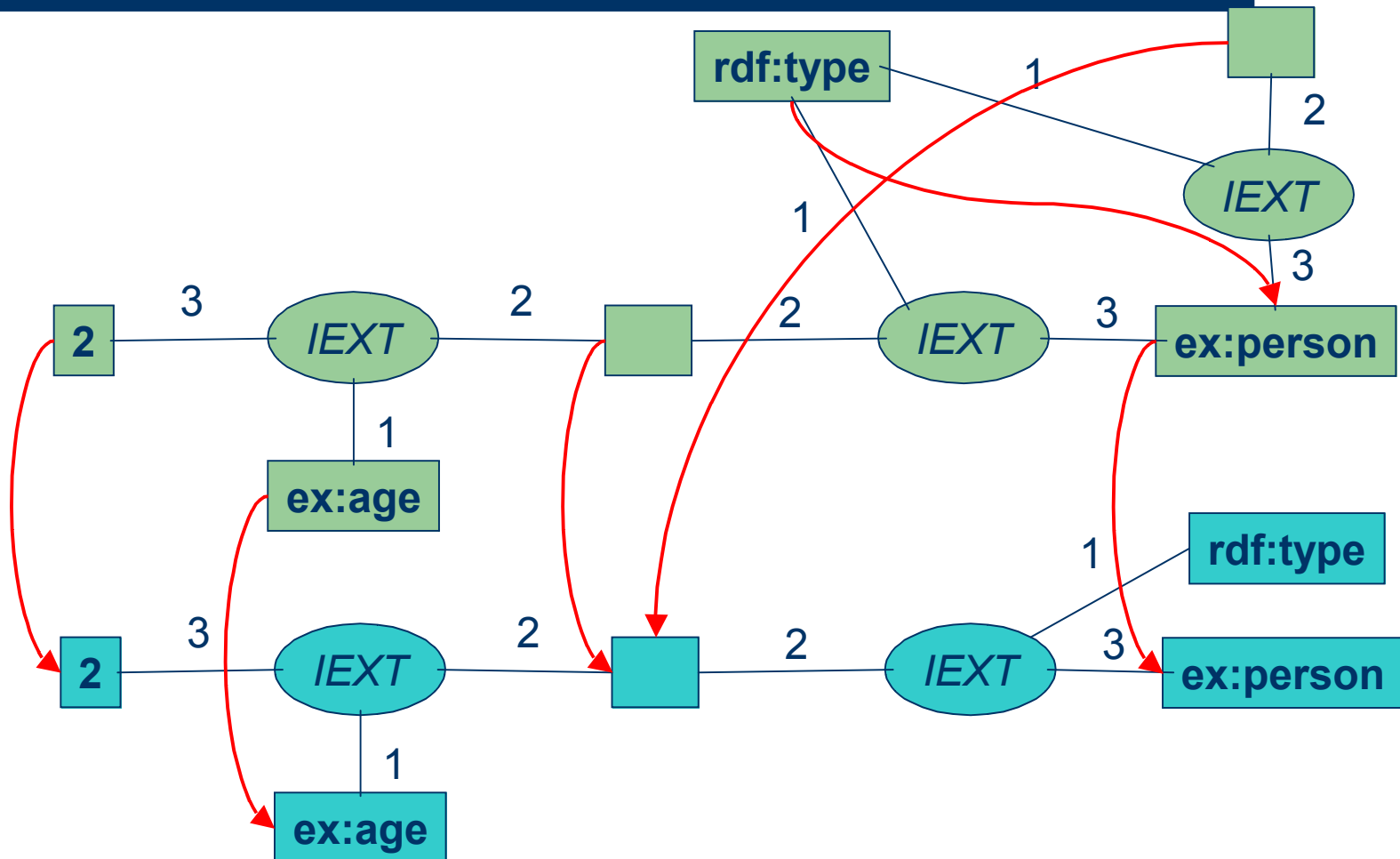
Lemme: La preuve de satisfiabilité est une projection du transformé du document dans le transformé de l'interprétation.

Conséquence et homomorphisme

Théorème: H est conséquence sémantique de G si et seulement si le transformé de H se projette dans le transformé de G.



Exemple



Plan de la présentation

- La syntaxe RDF
- La sémantique RDFL
- L'inférence en RDFL: un homomorphisme
- Extension à RDF
- Extension à RDFS
- Comparaison avec CORESE
- Bénéfices
- RDFL + Règles

Le vocabulaire réservé RDF

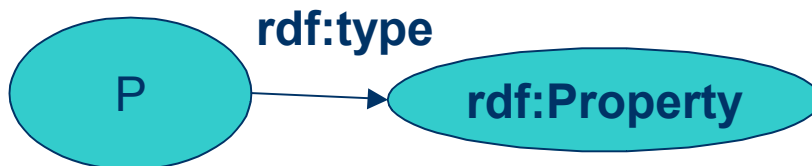
- `rdf:type`
- `rdf:Property`
- `rdf:XMLLiteral`
- `rdf:nil`
- `rdf:List`
- `rdf:Statement`
- `rdf:subject`
- `rdf:value`
- `rdf:Alt`
- `rdf:predicate`
- `rdf:object`
- `rdf:first`
- `rdf:rest`
- `rdf:Seq`
- `rdf:Bag`
- `rdf:_1`
- `rdf:_2`
- `rdf:_...`

Une interprétation est un modèle...

- Si pour tout triplet



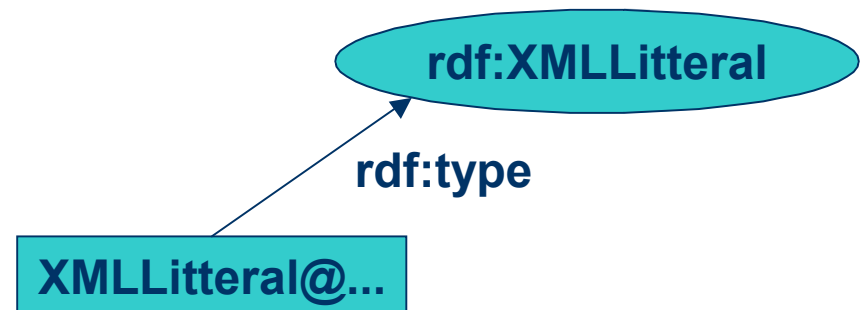
on peut déduire le triplet



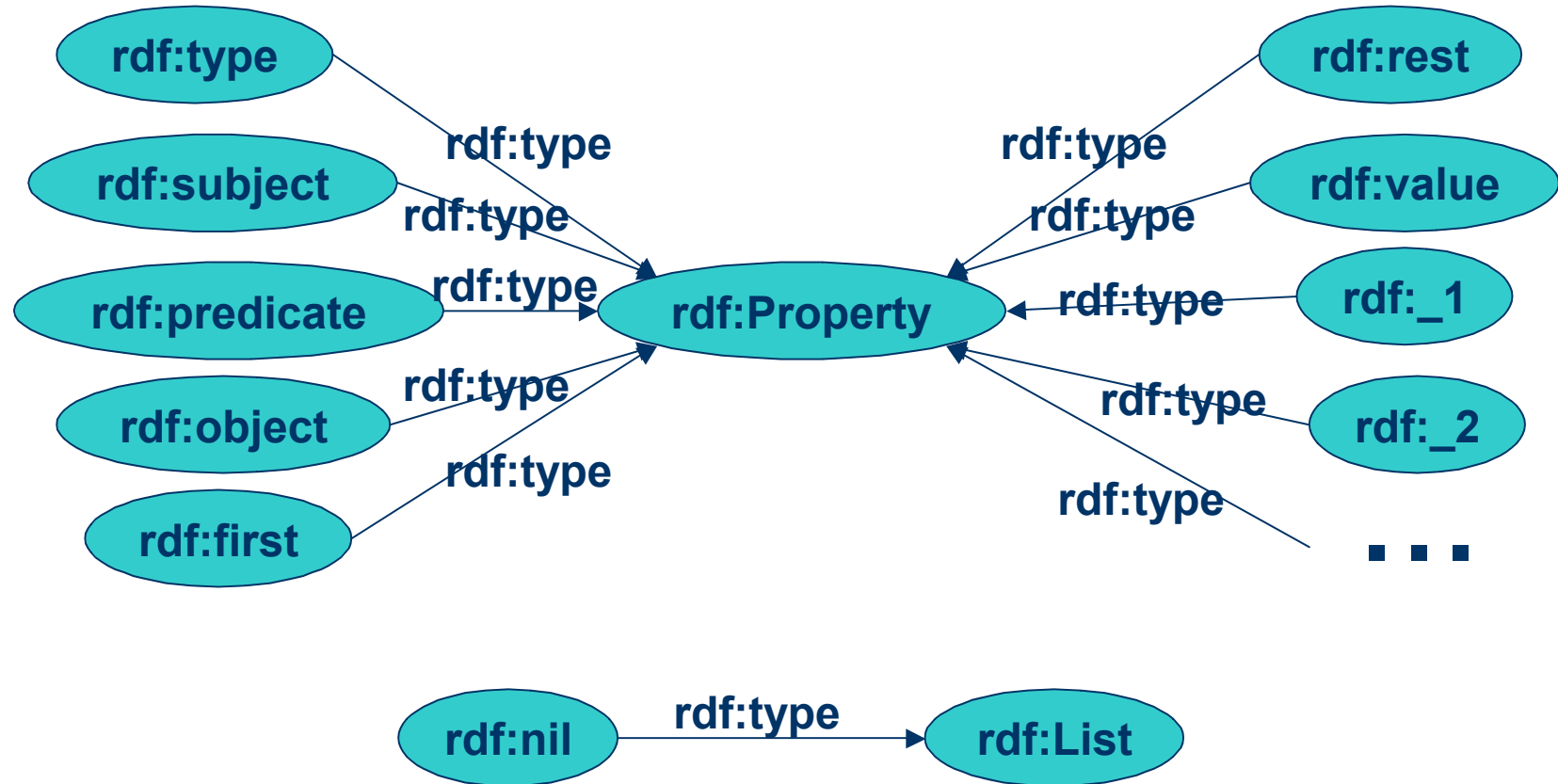
- Ssi pour tout littéral typé syntaxiquement consistant

`XMLLiteral@...`

on peut « déduire le triplet »



Les axiomes RDF



Remarque: insatisfiabilité



Le sommet est *localement inconsistent*.

Lemme: Un document RDF est satisfiable ssi tous ses sommets sont localement consistants.

Extension du résultat principal

Théorème: Un document RDF H est conséquence sémantique de G ssi G est insatisfiable ou si le transformé de H se projette dans le transformé de G .


Plan de la présentation

- La syntaxe RDF
- La sémantique RDFL
- L'inférence en RDFL: un homomorphisme
- Extension à RDF
- Extension à RDFS
- Comparaison avec CORESE
- Bénéfices
- RDFL + Règles

Le vocabulaire réservé RDFS

- rdfs:domain
- rdfs:range
- rdfs:Resource
- rdfs:Literal
- rdfs:Datatype
- rdfs:Class
- rdfs:subClassOf
- rdf:subPropertyOf
- rdfs:member
- rdfs:Container—
Membership—
Property
- rdfs:comment
- rdfs:seeAlso
- rdfs:isDefinedBy
- rdfs:label

Une interprétation est un modèle...

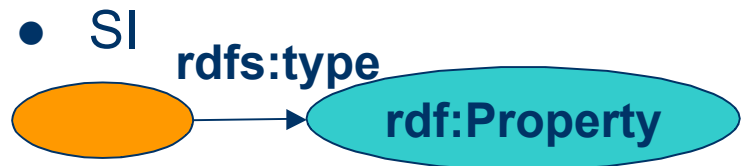
- Pour chaque sommet  (interprété par un élément de IR)



ALORS



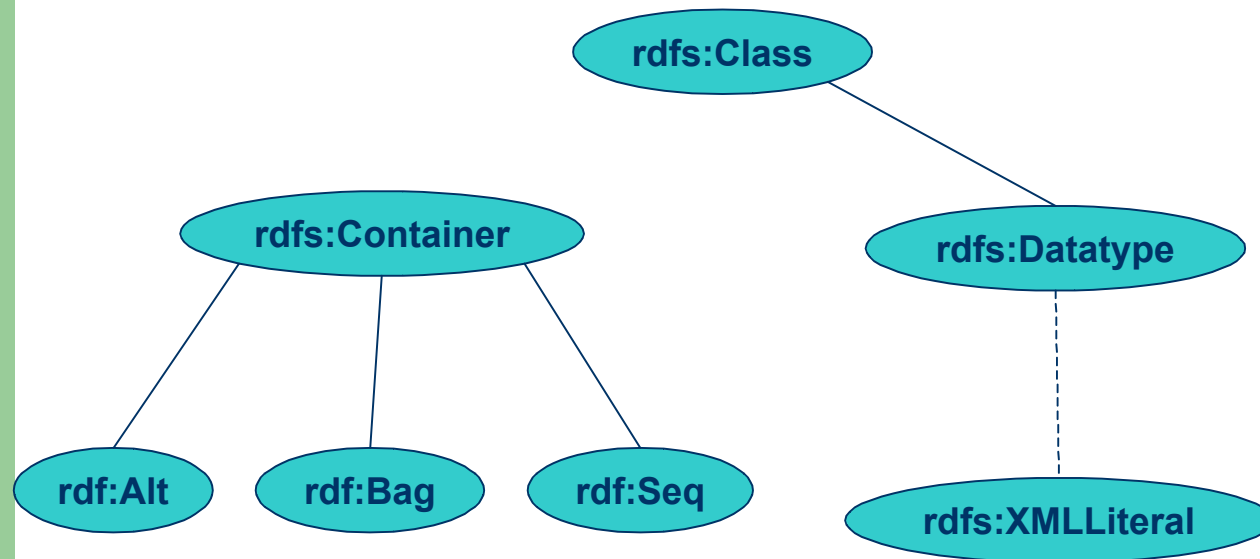
ALORS



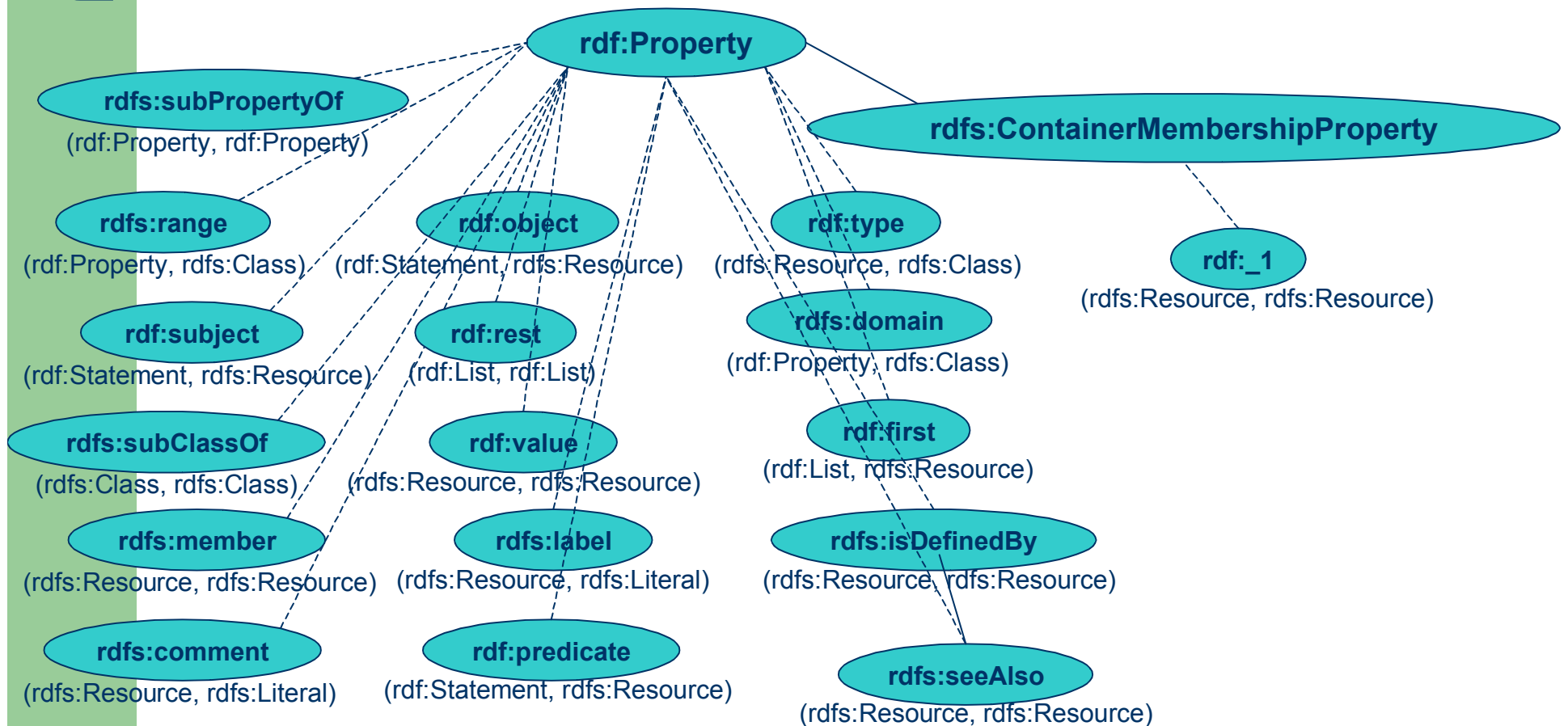
ALORS **rdfs:subPropertyOf**



Les axiomes RDFS (1)



Les axiomes RDFS (2)



Extension du résultat principal

Théorème: Un document RDFS H est conséquence sémantique de G ssi G est insatisfiable ou si le transformé de H se projette dans le transformé de G .

Plan de la présentation

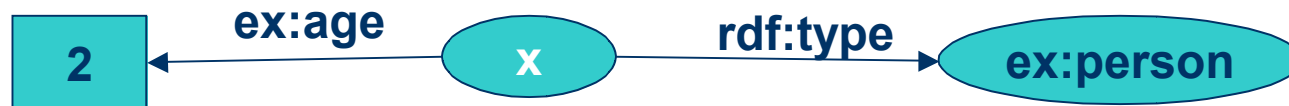
- La syntaxe RDF
- La sémantique RDFL
- L'inférence en RDFL: un homomorphisme
- Extension à RDF
- Extension à RDFS
- Comparaison avec CORESE
- Bénéfices
- RDFL + Règles

Des graphes conceptuels...

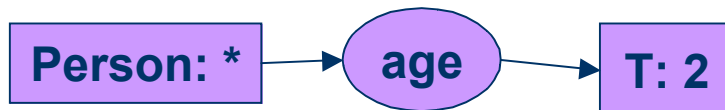
- Les hypergraphes obtenus sont des graphes conceptuels (Sowa, 84)
- « RDFS » \approx GC (Tim Berners-Lee, 2001)
- Autre transformation utilisée dans CORESE (Corby, Dieng, Hebert, ICCS 2000)

Une transformation moins lisible

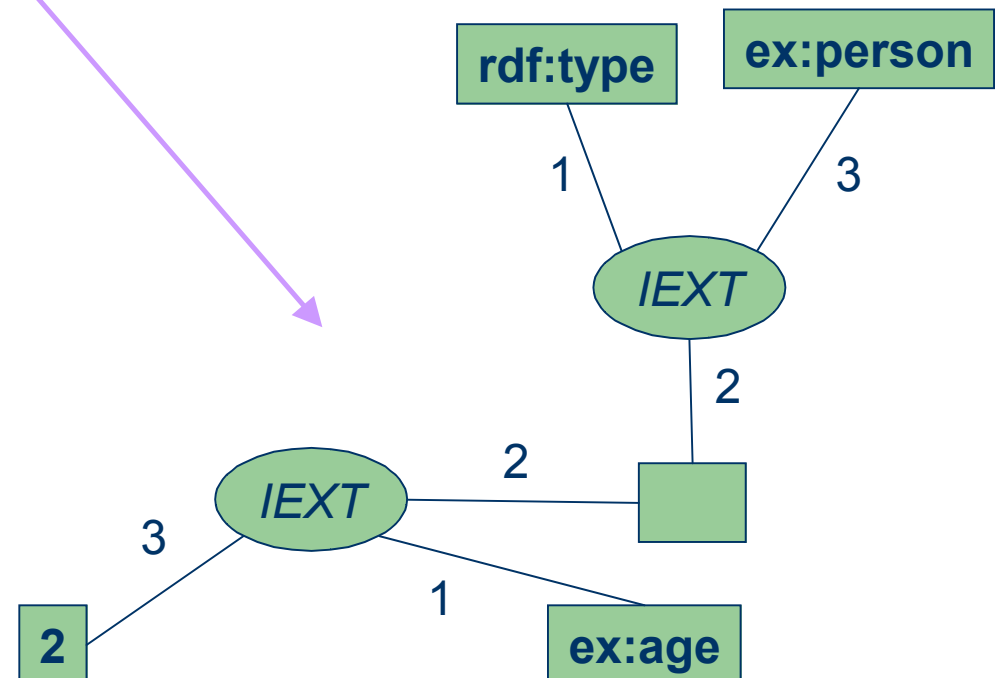
Graphe RDF/S



CORESE



Hiérarchie de types dans une autre structure: le support



Préservation de la sémantique

- CORESE: Préservation des conséquences
- ICI: Préservation de la sémantique
- Plus résistant à des extensions du modèle

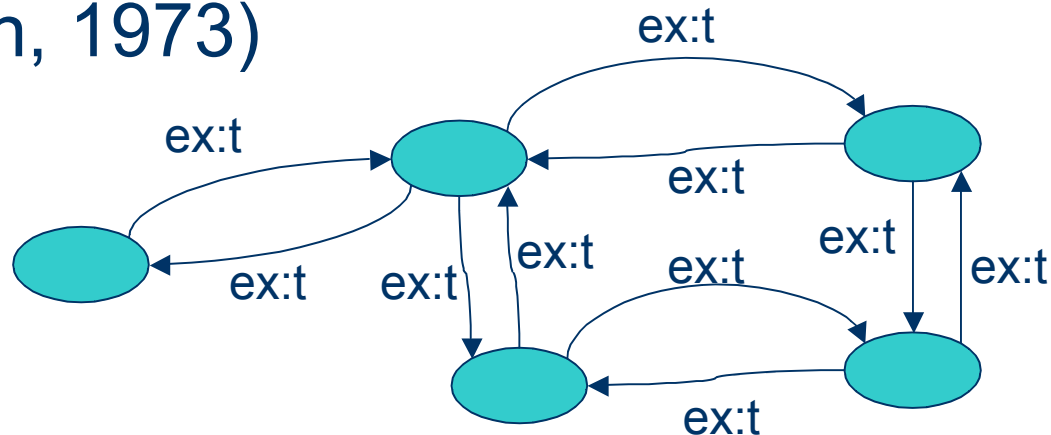
Plan de la présentation

- La syntaxe RDF
- La sémantique RDFL
- L'inférence en RDFL: un homomorphisme
- Extension à RDF
- Extension à RDFS
- Comparaison avec CORESE
- **Bénéfices**
- RDFL + Règles

Complexité

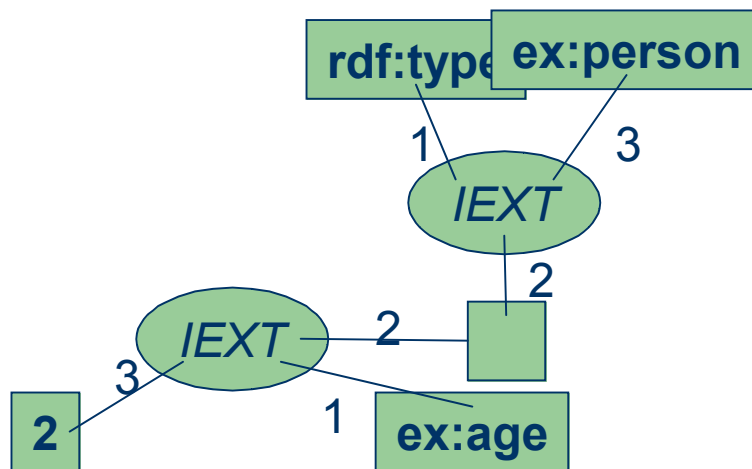
- RDF-Déduction?
→ Un problème NP-complet

(Réduction à GRAPH-HOMOMORPHISM,
Levin, 1973)



Identification au fragment FOL(\wedge, \exists)

- Utilisation de la sémantique logique des graphes conceptuels (Sowa, 84, Mugnier & Chein, RIA 96)

$$\exists X (\text{IEXT}(\text{rdf:type}, x, \text{ex:person}) \wedge \text{IEXT}(\text{ex:age}, x, 2))$$


Identification de la conséquence sémantique à la conséquence logique des formules associés.

Réseaux de contraintes

- Utilisation de l'équivalence GC / CSP (Mugnier, ICCS 2000)
 - Conservation de la structure!
- Optimisation de l'algorithme de BackTrack (Baget, ICCS 2003)
- Traduction directe des cas polynomiaux (Gottlob, Leone, Scarcello, IJCAI 1999)

Plan de la présentation

- La syntaxe RDF
- La sémantique RDFL
- L'inférence en RDFL: un homomorphisme
- Extension à RDF
- Extension à RDFS
- Comparaison avec CORESE
- Bénéfices
- RDFL + Règles

Utilisation de règles de GC (Salvat, 96)

- Règles: la prochaine étape du WS
- Un langage RDFL+Règle?
 - Simple
 - Bien étudié (GC)
 - Permet d'exprimer RDF/RDFS
 - Permet d'exprimer une (petite) partie de OWL
 - Permet de donner une sémantique satisfaisante à certains éléments du vocabulaire réservé
- Un langage, différents packages \neq plusieurs langages