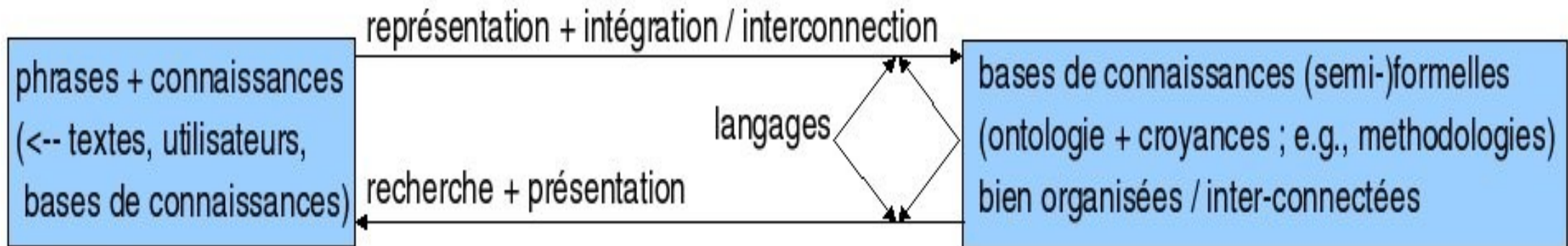


Construction collaborative et passant à l'échelle de base de connaissances

Philippe Martin

MCF à l'ESIROI - STIM



Plan

1. **C.V.**
2. **Petite ontologie des ontologies *et* solution pour une onto-terminologie**
3. **Briques de solution pour un Web de connaissances :**
protocoles / techniques, ontologie multi-sources, langages
 - 2.1. **Protocoles de collaboration** (intra-BC, inter-BCs, ...)
 - 2....
4. **Collaboration avec l'équipe ECD** -> généralisation de modèles / ontologies / algorithmes ?
 - 3.1. Exemple

1. C.V.

- 1993-1996. Doctorant à l'INRIA - **structuration de BCs + documents**
- 1997-1999. CR2 en Australie - **BCs dans des documents Web** - WebKB (www.webkb.org)
- 2000-2007. CR1 + MCF + Chef de Projet en Australie - **construction collaborative d'une grande BC**
- 2008-2009. Responsable du projet PACA-**RFID** Grande Distribution à Eurécom - **spécifications**
- A partir de sept. 2009. MCF à l'ESIROI-STIM - **HDR** (<http://www.webkb.org/doc/papers/hdr/>)

2. Petite ontologie des ontologies

```
knowledge_base
  part: 1 (ontoterminology
    part: 1 (ontology
      > top_level_ontology //">" means "subtype:"
        lexical_ontology domain_ontology,
      part: 1..* (formal_term
        definition: 0..* formal_definition)
        0..* formal_definition
      )
      0..1 (thesaurus
        part: 1..* (informal_term
          definition: 0..* informal_definition)
        )
    )
  1 (base_of_facts part: assertion);
```

```

term
  > {(formal_term informal_term)}; //partition complete de sous-types

statement
  > {(definition assertion)}
  {( (formal_statement
      part: 1 formal_grammar_with_interpretation_in_a_logic
          1..* formal_term 0 informal_term)
    (informal_statement
      part: 0 formal_grammar_with_interpretation_in_a_logic
          0..* formal_term 0..* informal_term)
    (semi-formal_statement
      part: 1 formal_grammar_with_interpretation_in_a_logic
          0..* formal_term 1..* informal_term)
  )});

```

2. Solution pour une onto-terminologie

Connection de termes formels et informel dans une même hiérarchie de spécialisation étendue

```
"right" //string ; ".>" means "extended_specialization:"
.> (en#"right" //english word "right"
  .> (wn#"right" //english word "right" in WordNet
    .> (wn#right //1 meaning of en#"right" in WordNet
      > (pm#wn#right //P.M.'s interpretation of wn#right
        .> (pm#right_of_any_animal
          .< (en#"animal right"
            .< pm#"animal related concept"),
            :<=> (wn#right owner=> any pm#animal),
            .> (pm#right_of_every_animal_in_2010
              pm#owner=> every pm#animal,
              .> (pm#Garfield_right_in_2010
                owner=> Garfield)
            ))))));
```

3. Briques de solution pour un Web de connaissances : protocoles / techniques, ontologie multi-sources, langages

"*Web de connaissances*" : Web sémantique, global, construit de manière collaborative, bien organisé

- Protocoles / techniques pour intégrer, rechercher ou évaluer des connaissances, collaborativement, dans une ou plusieurs bases
- Ontologie multi-sources intégrant "sans perte" des ontologies lexicales, de haut-niveau, de langages, de méthodologies, ...
- Des langages (*1 modèle, plusieurs syntaxes*) concis, expressifs et paramétrables pour représenter, rechercher, comparer et combiner des connaissances

Hypothèses :

- 1. ces briques sont toutes nécessaires pour passer à l'échelle dans le "partage" de connaissances
- 2. un nombre suffisant de fournisseurs d'informations pourront et accepteront d'utiliser ces briques dans leurs représentations afin que leurs informations soient plus facilement trouvées, comparées, (ré-)utilisées, ...
-> chercheurs, enseignants, étudiants, ...
- 3. tous les conflits sémantiques peuvent être résolus en ajoutant plus de précision jusqu'à aboutir à exprimer des observations ou des préférences
-> résoudre des conflits augmente l'organisation de la BC

3.1. Protocole pour l'édition collaborative d'une BC

Problématique : maintien de l'organisation ;

collaboration sans discussion, compromis, comité de sélection

Approche :

1. Tout objet (catégorie, relation, phrase, ...) a un créateur -> pas de conflit lexical + filtrage
2. Tout utilisateur peut ajouter un objet, ou détruire ses objets, sauf si *conflit-intra*
3. *Conflit-inter* lors de l'ajout / destruction d'une catégorie / définition -> clonage
4. *Conflit-inter* lors de l'ajout / destruction d'une croyance -> rejet mais "corrections" possibles

Exemple : ` `every bird is agent of a flight'_[John] has for **corrective_restriction**

`most healthy French birds are able to be agent of a flight' '_[Joe]

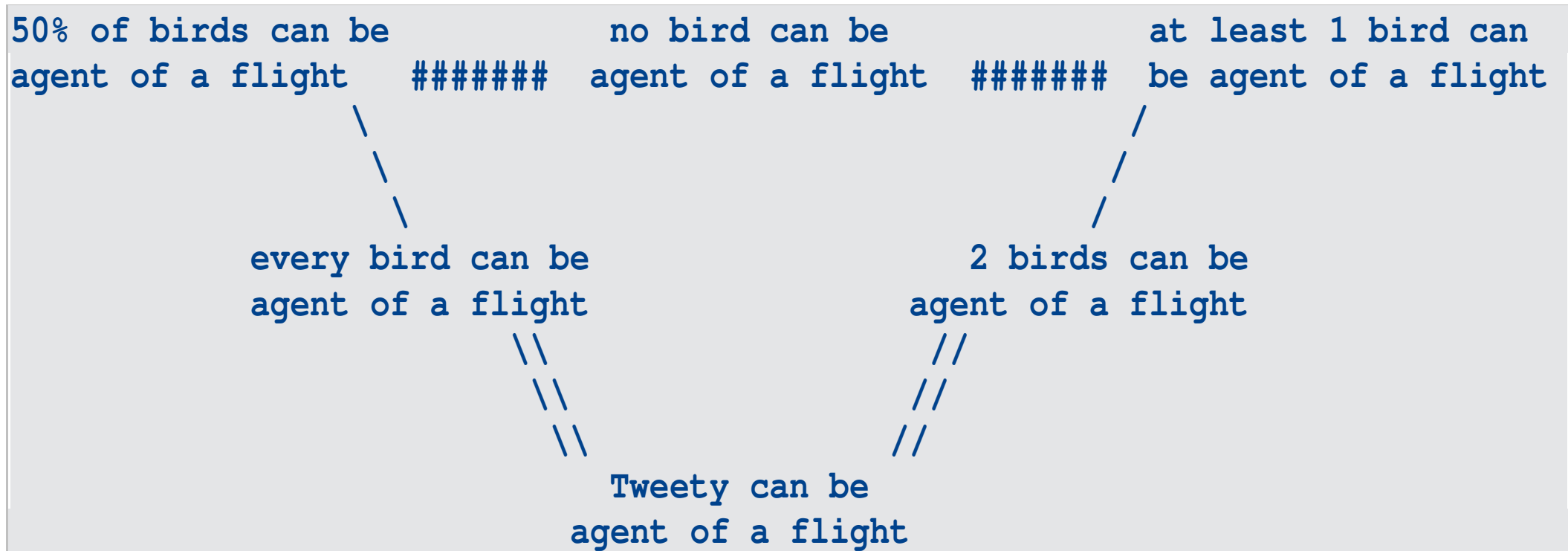
Complément de : approches basées sur des modules (-> versions)

Intérêts : wikis, états de l'art, mémoires d'entreprise, catalogues, ressources d'apprentissage, ...

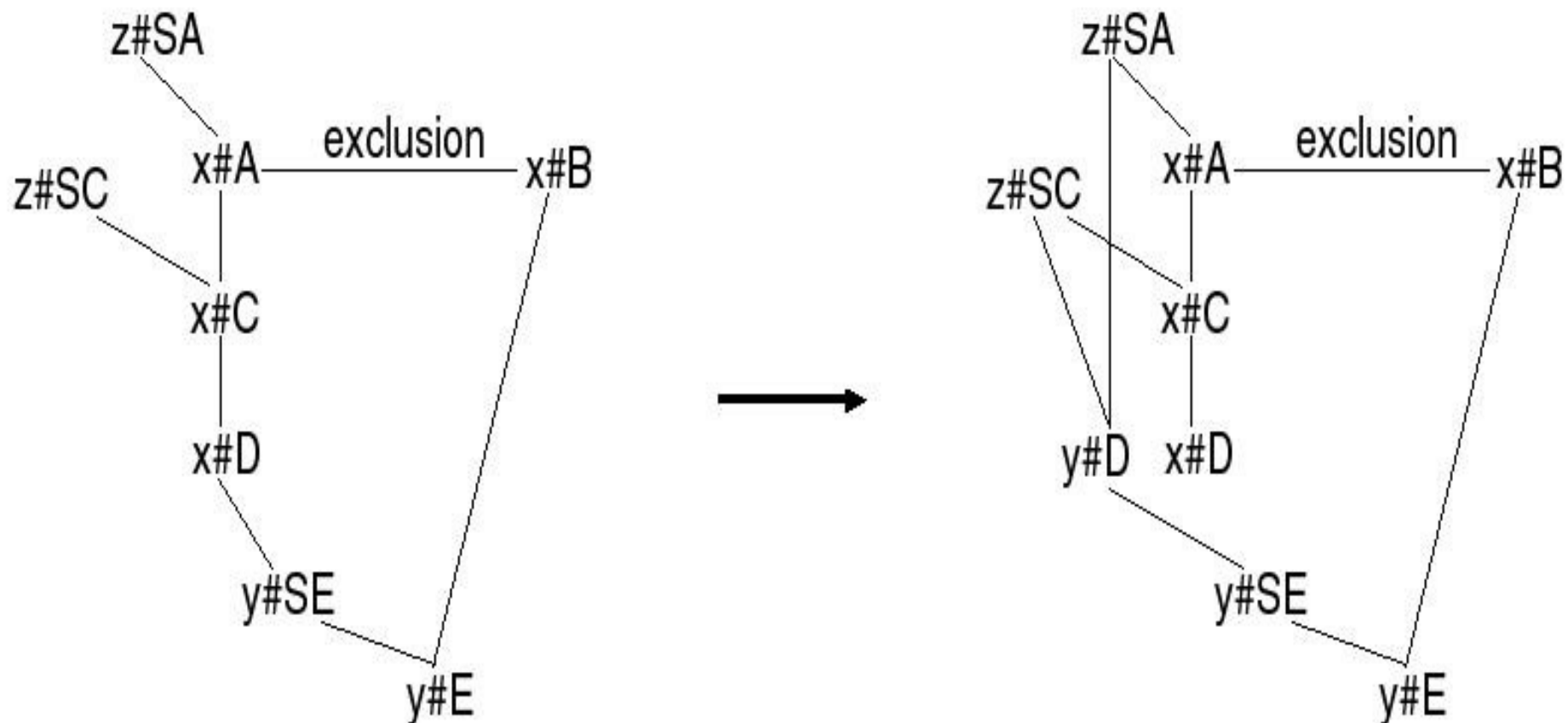
Conflit entre 2 phrases X et Y <- lorsque l'on ignore les quantificateurs dans X et Y,
X spécialise Y ou Y spécialise X
(sauf si "simple instantiation")
(contextes et collections comparés niveau par niveau)

Avec quantificateurs -> contradiction ou généralisation partielle / totale entre X et Y

Exemples de relations / conflits détectés par WebKB-2 :



Exemple de résolution automatique de conflits par clonage de catégories après ajout d'une relation d'exclusion :



3.2. Protocole pour l'évaluation collaborative de connaissances et de leurs auteurs

Problématique : précision, flexibilité

Approche : représentation des évaluations individuelles, des fonctions de combinaison et de leur résultats. Exemple pour **l'utilité d'une croyance** :

- **Son intérêt moyen** <- intérêts individuels pour cette croyance
<- utilité de chacun de ces utilisateurs
<- l'utilité de ses croyances
<- nombre de ses évaluations de croyances d'autres auteurs
- **Son état de confirmation** <- nombre de (contre-)arguments sur (contre-)arguments

Intérêts pour : enseignants, chercheurs, ...

Complément de : Co4, CHAO, DILIGENT, CoAKTinG, SomeWhere, Knowledge Zone, ...

Publié dans : ICCS 2005, E-learn 2006

3.3. Protocole pour le partage de connaissances entre BCs

Problématique : BC virtuelle ; pas de redondances / contradictions *implicites* entre BCs
(combinaison des avantages de la distribution physique et de la centralisation des informations)

Approche : le choix de la BC pour une requête ne doit pas avoir d'importance

- > réplication de connaissances (et transmission de requêtes) entre serveurs de BCs
- > chaque serveur s'engage à être un *nexus* pour 1 ou plusieurs objets (type, phrase, ...)
 - > il s'engage à stocker n'importe quelle connaissance directe sur cet objet
 - > pour tout autre objet, il stocke des pointeurs sur au moins un de ses nexus
 - > il vérifie périodiquement d'autres nexus pour intégrer les connaissances dont il est un nexus

Complément de : SomeWhere, KB-Grid, S-OGSA, DOGMA, HCOME, ...

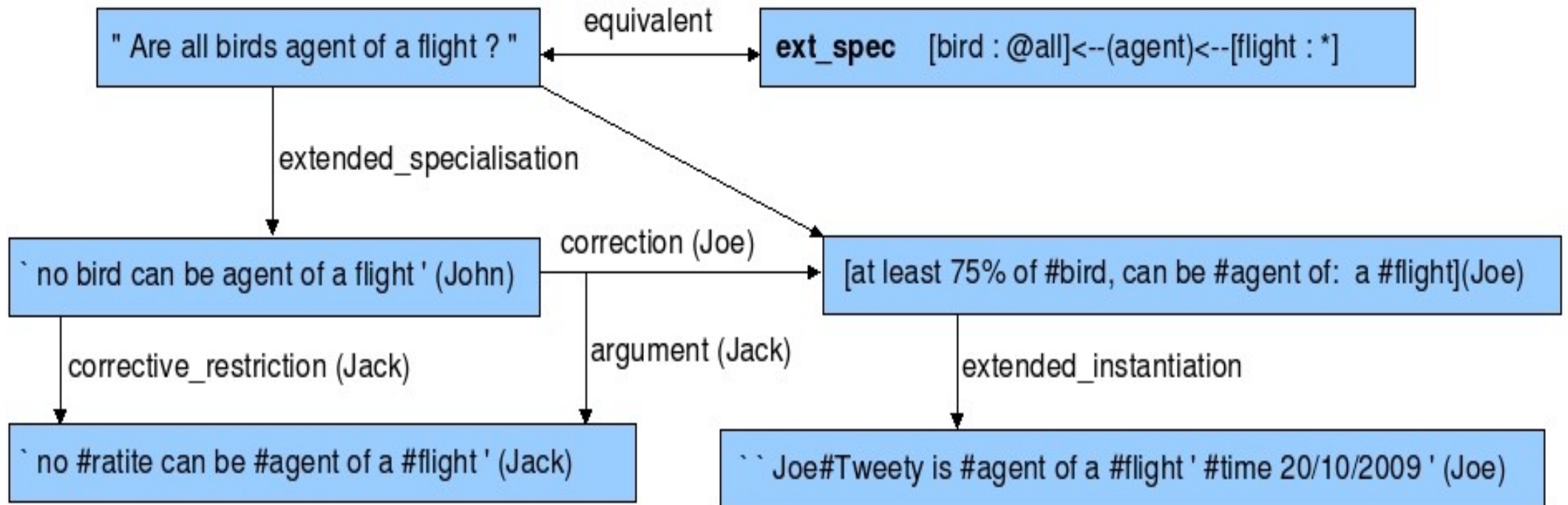
Intérêts : compétition + centralisation + distribution

Publié dans : ICCS 2005, ICUT 2007

3.4. Normalisation

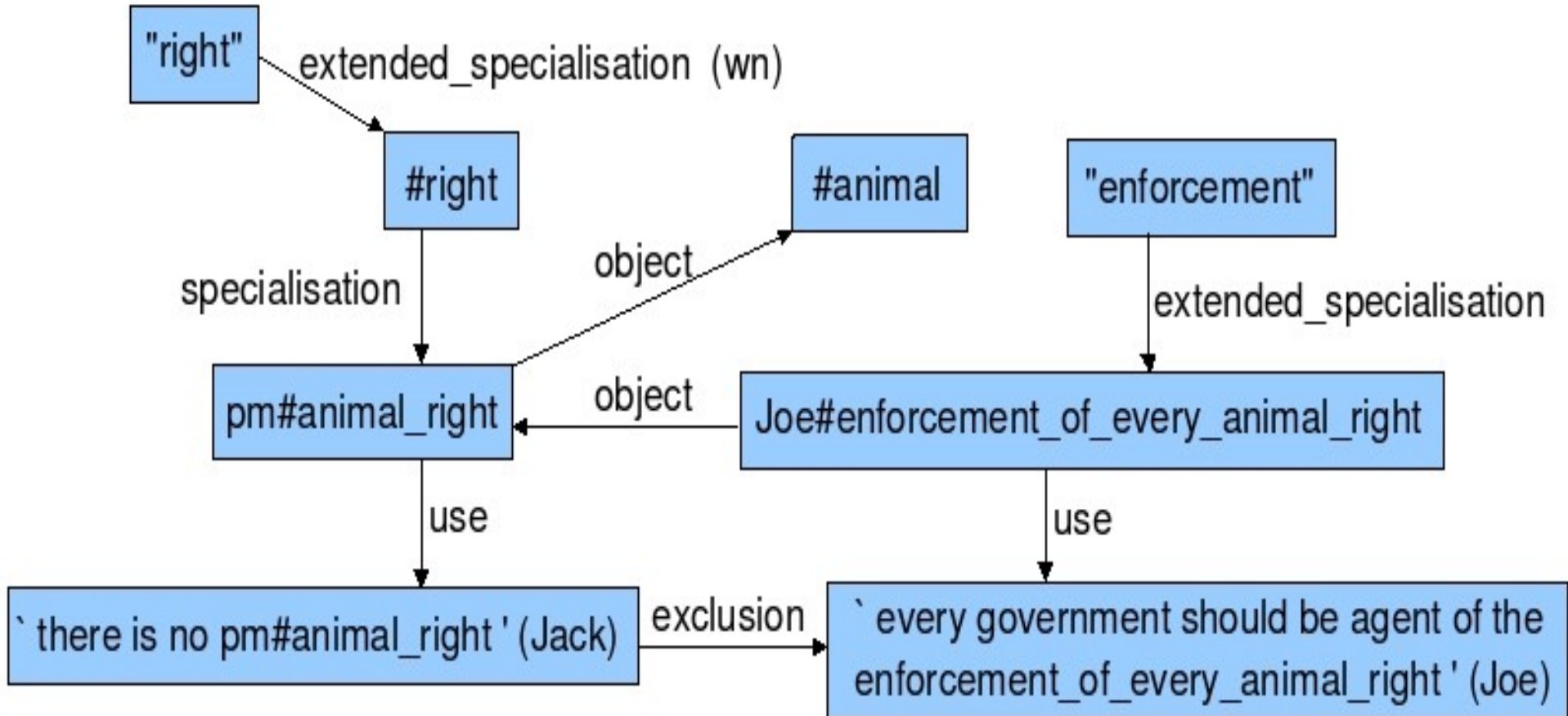
Problématique : maximiser la précision, organisation, comparabilité, ré-utilisabilité des connaissances

- **Dans la BC.** Exemples de règles encouragées par l'ontologie et les langages proposés :
 - *Règles lexicales* : utilisation de noms communs au singulier, pas de style Intercap
 - *Règles structurales / sémantiques* : utilisation de
 - termes précis, de supertypes et types du 1er ordre
 - relations simples, binaires, transitives (e.g., extended_generalization)
 - arguments sur relations de correction (plutôt que des relations d'objection)
 - contextes (temps, espace, source, ...), cardinalités, exclusions / corrections
- **Dans les fichiers d'entrée** -> les sections suivent l'ontologie de haut niveau



Exemple d'intégration / organisation / normalisation :

Sources : "governments should enforce all animal rights"(Joe), "animals have no rights to be enforced"(Jack)



Normalisation - un contre-exemple?

"Semantic Web Topics Ontology" de ISWC 2006 :
une ontologie de "domaines", "techniques" et "projects",
construite avec Protégé et un wiki, pour indexer des documents

Extrait (traduit de RDF+OWL vers FL) :

Knowledge_Representation

```
topic_usesTechnique: neural_networks heuristic_question_answering,  
topic_usedIn: "library classification" "information processing",  
topic_relatedTo: Knowledge_Discovery  
    (Artificial_Intelligence  
        topic_supports: Web_Services_Composition,  
        topic_usedIn: "Pattern Recognition"  
            "Computer Vision",  
        topic_subtopic: Machine_Learning);
```

3.5. Organisation des catégories de haut niveau

Problématique : - hiérarchisation et partitionnement intuitifs et systématiques des catégories
- support pour la création de représentations simples

Approche pour les concepts

- *distinctions* : situation / entity, spatial / non-spatial, roles
- *non distinction* : attribute_or_quality_or_measure, description_content/medium/container, catégories pour nouns, verbes, adverbes, adjectifs
- catégorisation par relations associées

Approche pour les relations

- catégorisation par signature (avec usage d'expressions régulières lorsque nécessaire)
- dérivation à partir de concepts avec signatures associées

1. *Things (categorization of uppermost concept types or 2nd order types)*

1.1. *Situations (either processes or states)*

1.1.1. *Genuine processes (actions, tasks, ...)*

1.1.2. *States*

1.1.3. *Situations w.r.t. to their roles: achievement, ...*

1.1.4. *Other categorizations for situations: perdurants, ...*

- 1.2. *Entities (things that are not situations)*
 - 1.2.1. *Spatial objects*
 - 1.2.1.1. *Space areas/regions*
 - 1.2.1.2. *Physical endurants: agentive entities, substances, ...*
 - 1.2.2. *Non-spatial objects*
 - 1.2.2.1. *Attributes and measures (temporal, spatial, ...)*
 - 1.2.2.2. *Non-spatial objects that are not attributes/measures*
 - 1.2.2.2.1. *Description content/mediums/containers*
 - 1.2.2.2.2. *True collections: bags, sets, types, ...*
 - 1.2.2.3. *Non-spatial objects w.r.t. their roles*
 - 1.2.3. *Entities w.r.t. to their roles*
 - 1.2.4. *Other categorizations for entities*
 - 1.2.4.1. *Endurants*
 - 1.2.4.2. *Entities w.r.t. to their (un-)divisibility: collections, ...*
- 1.3. *Things w.r.t. to their roles*
 - 1.3.1. *Concept types usable for generating relation types*
 - 1.3.2. *Things w.r.t. to other roles: mediating things, created things, ...*
- 1.4. *Other categorizations for things: continuants/occurrents, divisible/indivisible, ...*

2. Relations

2.1. Categorization of relations w.r.t. their source/destination arguments

2.1.1. Relations from situations: to time measures, to situations

2.1.2. Spatial relations from entities with spatial features

2.1.3. Relations from collections (lists, types, ontologies, ...)

2.1.4. Relations (logical/rhetorical/...) from descriptions

2.1.5. Relations from attributes or measures

2.1.6. Relations to situations

2.1.7. Relations to entities with spatial features

2.1.8. Relations to time measures

2.1.9. Relations to collections (lists, types, ontologies, strings, ...)

2.1.10. Relations to attributes or measures

2.2. Categorization of relations w.r.t. their roles

2.2.1. Attributive relations

2.2.2. Mereological relations

2.2.3. Intentional relations

2.2.4. Temporal relations

2.2.5. Object relations

2.2.6. Conceptual relations

2.2.7. Relations for particular applications

2.3. Categorization of relations w.r.t. what/who/why/.../how questions

2.4. Categorization of relations w.r.t. particular properties

2.4.1. Relations with particular mathematical properties (transitivity, ...)

2.4.2. Relations categorized w.r.t. their fixed or variable arities

3.6. Langages

3.6.1. Des langages concis et expressifs pour différentes familles de notations

3.6.2. Une seule grammaire LR(1) et un seul modèle de (re-)présentation

Complément de : modèles (CG, RDF+OWL, CL, GRIWES, Ontology Definition Meta-model, ADL, SIMILE, ...),
notations, GRDDL

Intérêts : génération / personnalisation de notations / (re-)présentations ; 1 seul parseur

3.6.3. - combinaison via un langage de script (FS)

- insertion dans des documents et indexation de n'importe quelle partie de document
- recherche et comparaison de catégories ou de représentations via une famille d'opérateurs
- génération de menus d'entrée ou de recherche à partir de l'ontologie

3.6.1. Des langages complémentaires, concis et expressifs

Problématique : intégrations, extensions et simplifications de CGLF, CGIF, KIF, N3, ...

-> réunion+amélioration des langages de graphes, frames, avec/sans quantificateurs, S-expressions, langages contrôlés

Approche : expressivité (quantificateurs numériques, collections, contextes, ...)

valeur par défaut configurables

non-répétition de relations <- sucre syntaxique pour position variable des contextes

-> aussi simples pour les cas simples que des langages peu expressifs

-> plus "normalisants" que RDF+OWL/XML, N3 et KIF

Résultats : FE (Formalized-English), FCG (Frame-CGs), FL (For-Links)

Illustration/comparaison sur divers types de

quantificateurs, définitions, collections, méta-assertions, ...

En: According to John, any human is a mammal
and has at most 1 head and 2 arms.
According to Jack, every human happens to have
exactly 1 head and conversely.
According to Jo, man and woman form a complete subpartition of human.

FE: `any human is a mammal
and has for part {at most 1 head, at most 2 arms}'_[John].
`any human has for part 1 head'_[Jack].
`any head is part of 1 human'_[Jack].
`human has for subtype {(man woman)}'_[Jo].

FL: human supertype: mammal __[believer: John],
part : 0..2 arm __[John]
head __[any->0..1 __[John], every->1 __[Jack],
1<-every __[Jack]],
subtype : {(man woman)} __[Jo];

Un langage textuel/graphique concis et structuré est nécessaire pour visualiser, naviguer et éditer une quantité réaliste de connaissances

```
N3:  :armPart  a rdf:Property;  rdfs:subPropertyOf :part;  rdfs:range :arm;
                                           owl:inverseOf :armPartOf.

    :headPart a rdf:Property;  rdfs:subPropertyOf :part;  rdfs:range :head;
                                           owl:inverseOf :headPartOf.

    :stuffWith1Head
      rdfs:subClassOf [a owl:Restriction;  owl:onProperty :headPart;
                                                              owl:cardinality 1].

    :stuffOf1Human
      rdfs:subClassOf [a owl:Restriction;  owl:onProperty :headPartOf;
                                                              owl:cardinality 1].
```

```
{:human a owl:Class; rdfs:subClassOf :mammal;} :believer :John.  
{:human a owl:Class; rdfs:subClassOf :mammal;  
  rdfs:subClassOf [a owl:Restriction; owl:onProperty :headPart  
                  owl:maxCardinality 1]  
  rdfs:subClassOf [a owl:Restriction; owl:onProperty :armPart  
                  owl:maxCardinality 2]  
} :believer :John.  
{@forall :b . {:a rdf:type :human; rdf:type :stuffWith1Head;}  
}:believer :Jack.  
{@forall :h . {:a rdf:type :head; rdf:type :stuffOf1Human;}  
} :believer :Jack.  
{:man rdfs:subClassOf :mammal; owl:disjointWith :woman  
} :believer :Jo.  
{:woman rdfs:subClassOf :mammal;} :believer :Jo.
```

En KIF :

```
(believer ' (defconcept human (?b) (mammal ?b)) John)
(believer ' (defconcept human (?b) (atMostN 1 '?a head (part ?b '?a))) John)
(believer ' (defconcept human (?b) (atMostN 2 '?a arm (part ?b '?a))) John)
(believer ' (forall ((?b human)) (exactlyN 1 '?a head (part ?b '?a))) Jack)
(believer ' (forall ((?a head)) (atMostN 1 '?b human (part '?b ?a))) John)
(believer ' (forall ((?a head)) (exactlyN 1 '?b human (part '?b ?a))) Jack)
(believer ' (defconcept man (?b) (and (human ?b) (not (woman ?b)))) Jo)
(believer ' (defconcept woman (?b) (and (human ?b) (not (man ?b)))) Jo)
```

avec

```
(defrelation atMostN (?num ?var ?type ?predicate) :=
  (exists ((?s set) (?n)) (and (size ?s ?n) (<= ?n ?num)
    (truth ^ (forall (, ?var) (=> (member , ?var , ?s)
      (and (, ?type , ?var) , ?predicate))))))
```

```
<rdf:Property rdf:ID="armPart">
  <rdfs:subPropertyOf rdf:resource="part"/>
  <owl:inverseOf rdf:ID="armPartOf"/>
  <rdfs:range rdf:resource="arm"/></rdf:Property>
<rdf:Property rdf:ID="headPart">
  <rdfs:subPropertyOf rdf:resource="part"/>
  <owl:inverseOf rdf:ID="headPartOf"/>
  <rdfs:range rdf:resource="head"/></rdf:Property>

<owl:Class rdf:about="head">
  <rdfs:subClassOf>
    <owl:Restriction><owl:onProperty rdf:resource="#headPartOf"/>
      <owl:maxCardinality rdf:datatype="&xsd;nonNegativeInteger"
        dc:creator="Jack">1
      </owl:maxCardinality></owl:Restriction> </rdfs:subClassOf>
</owl:Class>

<owl:Class rdf:about="man">
  <rdfs:subClassOf rdf:resource="mammal" dc:creator="Jo"/>
  <owl:disjointWith rdf:resource="woman" dc:creator="Jo"/>
</owl:Class>

<owl:Class rdf:about="woman">
  <rdfs:subClassOf rdf:resource="mammal" dc:creator="Jo"/>
</owl:Class>
```

```
<owl:Class rdf:about="human">
  <rdfs:subClassOf rdf:resource="mammal" dc:creator="John"/>
  <rdfs:subClassOf>
    <owl:Restriction><owl:onProperty rdf:resource="#headPart"/>
      <owl:cardinality rdf:datatype="&xsd;nonNegativeInteger"
        dc:creator="Jack">1</owl:cardinality>
    </owl:Restriction> </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Restriction><owl:onProperty rdf:resource="#headPart"/>
      <owl:maxCardinality rdf:datatype="&xsd;nonNegativeInteger"
        dc:creator="John">1</owl:maxCardinality>
    </owl:Restriction> </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Restriction><owl:onProperty rdf:resource="#armPart"/>
      <owl:maxCardinality rdf:datatype="&xsd;nonNegativeInteger"
        dc:creator="John">2
      </owl:maxCardinality></owl:Restriction> </rdfs:subClassOf>
</owl:Class>
```

FL peut aussi être utilisé pour des discussions structurées

```
"knowledge_sharing_with_an_XML-based_language_is_advantageous"  
  generalization: "knowledge_sharing_with_an_XML-based_language_is_possible"  
    __[pm],  
  argument: - "XML is a standard" __[pm]  
            - ("knowledge_management_with_classic_XML_tools_is_possible"  
              corrective_restriction:  
                "syntactic_knowledge_management_with_classic_XML_tools  
                 is_possible"__[pm]  
              )__[pm],  
  argument: "the use of URIs and Unicode is possible in XML"  
            (fg, objection: "the use of URIs and Unicode can easily be made  
                           possible in most syntaxes" __[tbl _[pm]]  
            ),
```

```

objection: - ("the use_of_XML_by_KBSs implies several tasks to manage"
              argument: "the internal_model_of_KBSs is rarely XML" __[pm]
            )__[pm],
- ` "an increase of the number of tasks *t to_manage"
    has for consequence
      "an increase of the difficulty to develop a software
      to manage *t"
  ' __[pm],
objection: - "knowledge_sharing_with_an_XML-based_language will force
              many persons (developers, specialists, etc.) to understand
              complex_XML-based_knowledge_representations" __[pm]
- ("understanding_complex_XML-based_knowledge_representations
    is difficult"
    argument: "XML is verbose" __[pm]
  )__[pm];

```


Comparaison de connaissances

Une manière "passant à l'échelle" d'afficher une comparaison d'un ensemble d'objets.

```
> compare pm#WebKB-2 km#Ontolingua on
  (support of: a is#IR_task, output_language: a km#KR_notation,
   part: a is#user_interface), maxdepth 5
```

	WebKB-2	Ontolingua
support of:		
is#IR_task	+	+
is#lexical_search	+	+
is#regular_expression_based_search	+	.
km#knowledge_retrieval_task	+	.
km#generalization_structural_retrieval	+	.
...		
output_language:		
km#KR_notation	+	+
(expressivity: km#FOL)	+	+
km#FCG	+	.
km#KIF	.	+
km#XML-based notation	+	.
km#RDF	+	-
...		

4. Collaboration avec l'équipe ECD

Possibilités :

- 1 langage (modèle) commun
- 1 ontologie commune
- 1 base de données/connaissances virtuelle commune
- 1 hiérarchie d'algorithmes commune
- ...

4.1. Exemple

Généralisation+formalisation en FL (-> OWL) des contraintes méronymiques utilisées dans IKBS

*Si une partie d'un objet n'est pas (observée comme) présente,
ses sous-parties ne peuvent être (observées comme) présentes.*

1er essai (incorrect) : $[?x \text{ part: } 0..* P \ 0..* (p \text{ part of } \Rightarrow 1 P)] \Rightarrow [?x \text{ part: } 0 p]$

2ème essai (incorrect) : $[?x \text{ part: } 0..* P \ 0..* (p \text{ part of } \Rightarrow 1 P)] \Rightarrow ![?x \text{ part: } 1..* p]$

3ème essai :
 $! [\text{an observation } ?o \text{ object: } [?x \text{ part: } 1..* P]]$
 $\Rightarrow [?o \text{ object: } [?x \text{ part: } 1..* (p \text{ part of } \Rightarrow 1P)]]$