

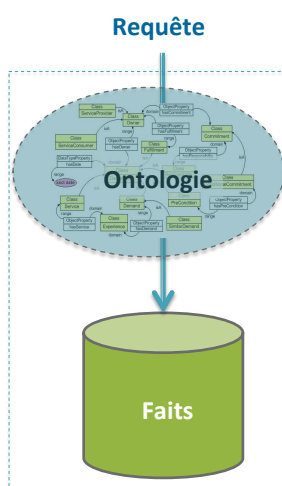


## BASES DE CONNAISSANCES

- Règles positives « à la Datalog »
- Contraintes négatives
- Règles existentielles

THÉORIE DES BASES DE CONNAISSANCES  
HMIN312M

## INTERROGATION DE BASES DE CONNAISSANCES



Les réponses se définissent par rapport à la base de connaissances

The answer to a BCQ  $Q$  in  $K$  is yes if  $K \models Q$

A tuple  $(a_1, \dots, a_k)$  of constants is a (certain) answer to  $Q(x_1, \dots, x_k)$  with respect to  $K$  if  $K \models Q[a_1, \dots, a_k]$ ,

where  $Q[a_1, \dots, a_k]$  is obtained from  $Q(x_1, \dots, x_k)$  by replacing each  $x_i$  by  $a_i$ .

Base de connaissances  $K = \text{Faits} \cup \text{Ontologie}$

## ONTOLOGIE (1)

**Concepts** Film, Personne, Acteur, ...

**Relations** joue/2, intervient/2, réalise/2, connaît/2...

**organisés par spécialisation** (hiérarchies, ordres, préordres)

Acteur *spécialisation de* Personne  
joue *spécialisation de* intervient

Ontologie  
de base

+ **autres connaissances** sur les concepts et relations

La signature de réalise est (Personne, Film)

Les concepts Film et Personne sont disjoints

Ceux qui jouent dans un même film se connaissent

Tout acteur joue dans un film

.....

## ONTOLOGIE (2)

- **Spécialisation de concepts et relations**

$\forall x (\text{Acteur}(x) \rightarrow \text{Personne}(x)) \quad \forall x \forall y (\text{joue}(x,y) \rightarrow \text{intervient}(x,y))$

- « La *signature* de réalise est (Personne, Film) »

$\forall x \forall y (\text{réalise}(x,y) \rightarrow \text{Personne}(x) \wedge \text{Film}(y))$

- « Ceux qui jouent dans un même film se connaissent »

$\forall x \forall y \forall z (\text{joue}(x,z) \wedge \text{joue}(y,z) \rightarrow \text{connaît}(x,y))$

Datalog

- « Film et Personnes sont *disjoints* »

$\forall x (\text{Film}(x) \wedge \text{Personne}(x) \rightarrow \perp)$

Contrainte négative

- « Tout acteur joue dans *un* film »

$\forall x (\text{Acteur}(x) \rightarrow \exists z (\text{Film}(z) \wedge \text{joue}(x,z)))$

Règle existentielle

## RÈGLES POSITIVES « À LA DATALOG »

- A **datalog rule** (or range-restricted rule) is of the form

$$\forall X \forall Y ( \text{Body } [X,Y] \rightarrow \text{Head } [X] )$$

where *Body* and *Head* are conjunctions of atoms (seens as sets)

Without loss of generality *Head* can be restricted to a single atom

- A rule  $R: B \rightarrow H$  is **applicable** to a factbase  $F$  if there is a homomorphism  $h$  from  $B$  to  $F$
- Applying  $R$  to  $F$  according to  $h$  consists of adding  $h(H)$  to  $F$

$R: \forall x \forall y \forall z (\text{joue}(x,z) \wedge \text{joue}(y,z) \rightarrow \text{connait}(x,y))$   
 $F = \{ \text{joue}(A,F1), \text{joue}(B,F1) \}$   
 L'application de  $R$  à  $F$  ajoute  $\text{connait}(A,B)$

## PROPERTIES OF DATALOG RULES

- $\mathcal{K} = (F, \mathcal{R})$  where  $F$  is a set of (ground) facts  
 $\mathcal{R}$  is a set of range-restricted rules

By applying rules from  $\mathcal{R}$  starting from  $F$ , a unique result is obtained:  
 the **saturation** of  $F$  (denoted here by  $F^*$ )

$F^*$  is **finite** since no new variable is created

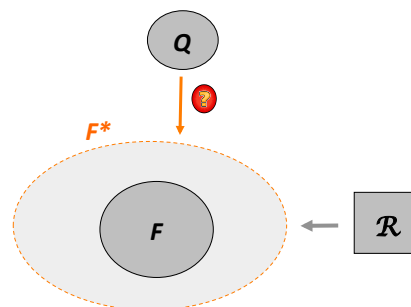
$F^*$  is a **core** (since there are no variables)

The nice properties of  $\text{FOL}(\exists, \wedge)$  are kept:

(the isomorphic model of)  $F^*$   
 is a **universal model** of  $\mathcal{K}$

Hence, for BCQs :

$\mathcal{K} \models q$  iff  $q \geq F^*$   
 (there is a homomorphism from  $q$  to  $F^*$ )



## NEGATIVE CONSTRAINTS

- A **negative constraint** is of the form

$$\forall x ( \text{Body} [x] \rightarrow \perp )$$

where *Body* is a conjunction of atoms and  $\perp$  is the absurd symbol (false)

$$\forall x (\text{Film}(x) \wedge \text{Personne}(x) \rightarrow \perp)$$

- A factbase *F* satisfies a negative constraint *C* if there is **no** homomorphism from the body of *C* to *F* (i.e. *C* seen as a rule is not applicable)

Note that  $F \cup \{C\}$  is **consistent (satisfiable)** iff *F* satisfies *C*

- A knowledge base  $K = (F, \mathcal{R}, C)$  where *C* is a set of negative constraints is **consistent** iff  $F^*$  (i.e. the saturation of *F* by  $\mathcal{R}$ ) satisfies all constraints from *C*

## EXISTENTIAL RULES

$$\forall x \forall y ( \text{Body} [x,y] \rightarrow \exists z \text{Head} [x,z] )$$

*x, y, z :  
sets of variables*

any **positive conjunction** (without functional symbols)

$$\forall x ( \text{actor}(x) \rightarrow \exists z \text{play}(x,z) )$$

$$\forall x \forall y ( \text{siblingOf}(x,y) \rightarrow \exists z (\text{parentOf}(z,x) \wedge \text{parentOf}(z,y)) )$$

Key point: ability to assert **the existence of unknown entities**

Crucial for representing ontological knowledge in **open domains**