

# Modélisation et raisonnement sur les parcours de soins

T. Guyet  
thomas.guyet@inria.fr  
Inria/AlstroSight



# Brève présentation

- Inria researcher, AIStroSight Team, Lyon
- Projects
  - ANSM-PEPS (Pharmaco-Epidémiologie des Produits de Santé)
  - co-holder of the AI-RACLES chair (Inria/APHP/CS) with E. Audureau (APHP)
  - PEPR Santé Numérique SafePaw
- Board member of the French Association for Artificial Intelligence (AFIA)

## Personal research axes

- Temporal data modeling and analysis
- Symbolic data analysis
  - (sequential) (temporal) pattern mining techniques
  - knowledge reasoning (logic programming and a bit of Semantic Web)
- Machine learning (time series analysis)
- Care pathway analysis

# Outline

- 1 **Parcours de soins**
  - Motivation
  - Représentation du parcours de soins
  - Les tâches
- 2 **Représentation des parcours de soins**
  - Représentation de l'information temporelle
  - Temps avec logiques modales
  - Ontologies about time
  - Conclusions
- 3 **Requêtage de données temporelles**
  - Contexte
  - Querying with temporal pattern matching
  - Querying with reasoning on care pathways: a web semantic approach
  - Conclusion
- 4 **Conclusion**

# Motivations

- La **trajectoire de soins** propose une vue longitudinale d'un patient
- Elle s'oppose à une vue instantannée d'un état du patient

## Pourquoi s'intéresser aux données longitudinales

- Caractérisation de l'ensemble des soins délivrés (e.g. *errance médicale*)
- Identifier des effets à long terme (épidémiologie)
- Comprendre l'organisation des soins
- Enrichir les analyses classiques avec “plus” de données (disponibles)
- Spécial France: Passage de la T2A à la tarification de la prise en charge
  - Quels sont les parcours de soins effectifs ?
  - Comment évaluer (medico-economiquement) ces parcours de soins ?

## Opportunité

- Collecte de données longitudinales : données hospitalières, données médico-administratives, ...

# Exemple

Un homme de 35 ans a été admis à l'hôpital pour un gonflement périorbitaire, une rougeur et une douleur le 24 mai 2014. On lui a alors diagnostiqué une cellulite périorbitaire. Il a été traité avec de la clindamycine intraveineuse (IV) et de la ciprofloxacine IV, ce qui a réduit la rougeur et le gonflement de l'orbite. Cependant, le deuxième jour suivant le traitement antibiotique, il a développé des nausées et des douleurs abdominales dans le quadrant supérieur droit, et ses tests de la fonction hépatique (LFT) ont commencé à augmenter. Un diagnostic de lésion hépatique idiosyncrasique d'origine médicamenteuse a été posé.

## Quelques questions d'intérêt (1/2)

- À quelle date le patient a-t-il développé des nausées ?
- Quelle(s) relation(s) temporelle(s) y a-t-il entre l'"*admission à l'hôpital*" et l'"*augmentation des tests de la fonction hépatique*" ?
- Le traitement par "*fluoroquinolones*" commence-t-il avant le "*diagnostic d'une lésion hépatique*" ?
- Quels événements se sont produits avant le "*diagnostic de lésion hépatique idiosyncrasique d'origine médicamenteuse*" ?

# Exemple

Un homme de 35 ans a été admis à l'hôpital pour un gonflement périorbitaire, une rougeur et une douleur le 24 mai 2014. On lui a alors diagnostiqué une cellulite périorbitaire. Il a été traité avec de la clindamycine intraveineuse (IV) et de la ciprofloxacine IV, ce qui a réduit la rougeur et le gonflement de l'orbite. Cependant, le deuxième jour suivant le traitement antibiotique, il a développé des nausées et des douleurs abdominales dans le quadrant supérieur droit, et ses tests de la fonction hépatique (LFT) ont commencé à augmenter. Un diagnostic de lésion hépatique idiosyncrasique d'origine médicamenteuse a été posé.

## Quelques questions d'intérêt (2/2) $\Rightarrow$ raisonnements avancés

- Existe-t-il d'autres patients similaires ?
- Ce patient a-t-il été pris en charge selon les bonnes pratiques recommandées (*guidelines*) ?
- La survenue du problème hépatique était-il prévisible ?

# Représentation et raisonnement sur les parcours

⇒ Ces questions peuvent paraître simples à l'humain ... elles sont difficiles à automatiser

Pour répondre à ces questions nous avons besoin:

- d'un formalisme de représentation de ces informations
- d'extraire les informations pour les formaliser<sup>a</sup>
- de raisonner sur les connaissances formalisées (inférer de nouveau faits/répondre à une question/...)

---

<sup>a</sup>hors champs de cette présentation! Si texte voir Review of Temporal Reasoning in the Clinical Domain for Timeline Extraction [OM21]

⇒ Choix des approches dites *formelles* (plus que statistiques)

Deux dimensions essentielles

- "ontologique" : La "ciprofloxacin" est un antibiotique de la famille des "fluoroquinolones"
- "temporel" : deux jours après le "24 mai 2023", on est le "26 mai 2023"

# Représentation ontologique



- Event : représente la nature de l'évènement médical
- Time : représente l'information temporelle associée à l'évènement

Un homme de 35 ans a été admis à l'hôpital pour un gonflement périorbitaire, une rougeur et une douleur le 24 mai 2014. On lui a alors diagnostiqué une cellulite périorbitaire. Il a été traité avec de la clindamycine en IV et de la ciprofloxacine IV, ce qui a réduit la rougeur et le gonflement de l'orbite. Cependant, le deuxième jour suivant le traitement antibiotique, il a développé des nausées et des douleurs abdominales dans le quadrant supérieur droit, et ses tests de la fonction hépatique ont commencé à augmenter. Un diagnostic de lésion hépatique idiosyncrasique d'origine médicamenteuse a été posé. Il sortira au bout de 10 jours d'hospitalisation.

- Quelles sont les évènements que vous identifiez dans ce texte ?
- Quelles sont les notions temporelles que vous identifiez ?



# Représentation ontologique



- Event : représente la nature de l'évènement médical
- Time : représente l'information temporelle associée à l'évènement

Un homme de 35 ans a été **admis à l'hôpital** pour un gonflement périorbitaire, une rougeur et une douleur le 24 mai 2014. On lui a alors diagnostiqué une cellulite périorbitaire. Il a été **traité avec de la clindamycine en IV et de la ciprofloxacine IV**, ce qui a **réduit la rougeur et le gonflement de l'orbite**. Cependant, le deuxième jour suivant le traitement antibiotique, **il a développé des nausées et des douleurs abdominales** dans le quadrant supérieur droit, et ses tests de la fonction hépatique ont commencé à augmenter. Un **diagnostic de lésion hépatique** idiosyncrasique d'origine médicamenteuse a été posé. Il **sortira** au bout de 10 jours d'hospitalisation.

- Quelles sont les évènements que vous identifiez dans ce texte ?
- Quelles sont les notions temporelles que vous identifiez ?

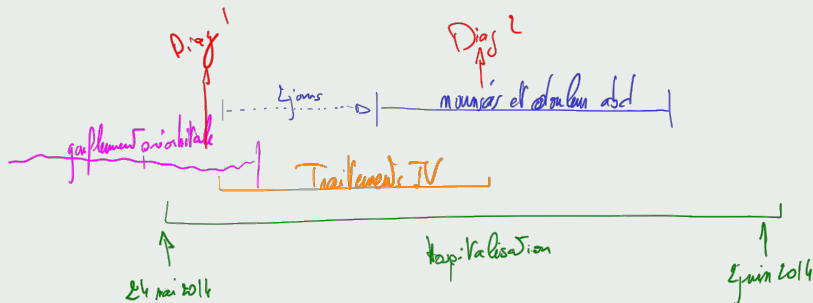
# Représentation ontologique

Un homme de 35 ans a été admis à l'hôpital pour un gonflement périorbitaire, une rougeur et une douleur le 24 mai 2014. On lui a alors diagnostiqué une cellulite périorbitaire. Il a été traité avec de la clindamycine en IV et de la ciprofloxacine IV, ce qui a réduit la rougeur et le gonflement de l'orbite. Cependant, le deuxième jour suivant le traitement antibiotique, il a développé des nausées et des douleurs abdominales dans le quadrant supérieur droit, et ses tests de la fonction hépatique ont commencé à augmenter. Un diagnostic de lésion hépatique idiosyncrasique d'origine médicamenteuse a été posé. Il sortira au bout de 10 jours d'hospitalisation.

- Hospitalisation: interval débutant le 24 mai, terminé le 2 juin
- Traitement 1 : durée au moins 2 jours, pendant l'hospitalisation
- Réduction rougeur: Entre le début du traitement et le début des nausées
- Nausées: débute 2 jours après le début de traitement 1
- Augmentation: en même temps que nausées
- Diagnostic: après l'augmentation et pendant l'hospitalisation

# Représentation ontologique

## Schéma du parcours



## Première question posée

→ comment représenter formellement ces différentes informations ?

# Différentes notions : clarification

- Représenter ?? mais pour en faire quoi ??

## Trois notions distinctes

- 1 **Trajectoire de soins**
- 2 **Trajectoire observée/décrite**
- 3 **Parcours de soins**

# Différentes notions : clarification

## Trois notions distinctes

### 1 Trajectoire de soins

- ensemble des évènements de soins d'une personne
- il s'agit de la donnée "réelle" : uniquement théorique

### 2 Trajectoire observée/décrite

- *description* d'une trajectoire
  - les informations qu'elle contient sont partielles
  - offre un point de vue particulier sur la trajectoire
- *données* longitudinales d'un **patient particulier**

### 3 Parcours de soins

- description d'une abstraction d'un enchaînement des évènements de soins
- correspond à un **type de patients**
- par exemple un *guidelines*

# Différentes notions : clarification

## Trois notions distinctes

### ① Trajectoire de soins

- ensemble des évènements de soins d'une personne
- il s'agit de la donnée "réelle" : uniquement théorique

### ② Trajectoire observée/décrite

- *description* d'une trajectoire
  - les informations qu'elle contient sont partielles
  - offre un point de vue particulier sur la trajectoire
- *données* longitudinales d'un **patient particulier**

### ③ Parcours de soins

- description d'une abstraction d'un enchaînement des évènements de soins
- correspond à un **type de patients**
- par exemple un *guidelines*

**Trajectoire = Trajectoire observée**

Si dans la suite on parle de trajectoire, il ne faut pas perdre de vue qu'il s'agit systématiquement d'une trajectoire observée

## Différentes notions : exemple

Un homme de 35 ans a été admis à l'hôpital pour un gonflement périorbitaire, une rougeur et une douleur le 24 mai 2014. On lui a alors diagnostiqué une cellulite périorbitaire. Il a été traité avec de la clindamycine intraveineuse et de la ciprofloxacine, ce qui a réduit la rougeur et le gonflement de l'orbite. Cependant, le deuxième jour suivant le traitement antibiotique, il a développé des nausées et des douleurs abdominales dans le quadrant supérieur droit, et ses tests de la fonction hépatique ont commencé à augmenter. Un diagnostic de lésion hépatique idiosyncrasique d'origine médicamenteuse a été posé.

### Trajectoire observée (point de vue medico-administratif)

- Hospitalisation avec le code CIM H050 au 24 mai (Inflammation aiguë de l'orbite [Cellulite orbite])
- Délivrance de J01FF01 et J01MA02 (codes ATC) du 24 mai au 26 mai
- Test NABM 0516 (ALAT,TGP) le 24 mai
- Test NABM 0516 (ALAT,TGP) le 25 mai
- Test NABM 0516 (ALAT,TGP) le 26 mai
- CIM S361 (Lésion traumatique du foie et de la vésicule biliaire) le 26 mai

## Différentes notions : exemple

Un homme de 35 ans a été admis à l'hôpital pour un gonflement périorbitaire, une rougeur et une douleur le 24 mai 2014. On lui a alors diagnostiqué une cellulite périorbitaire. Il a été traité avec de la clindamycine intraveineuse et de la ciprofloxacine, ce qui a réduit la rougeur et le gonflement de l'orbite. Cependant, le deuxième jour suivant le traitement antibiotique, il a développé des nausées et des douleurs abdominales dans le quadrant supérieur droit, et ses tests de la fonction hépatique ont commencé à augmenter. Un diagnostic de lésion hépatique idiosyncrasique d'origine médicamenteuse a été posé.

### Parcours de traitement de l'inflammation aiguë de l'orbite<sup>a</sup>

<sup>a</sup>fictif

- Diagnostique H050 à  $J_0$
- Démarrage immédiat d'un traitement antibiotique par Fluoroquinolones (J01MA) de  $J_0$  à  $J + 7$ , dosage entre 20 et 30 mg/j IV
- Suivi de la fonction hépatique pendant toute la durée du traitement jusqu'à 1 semaine après



# Différentes sources de données

## Pour les trajectoires

- Toutes données médicales à laquelle on peut associer du temps (et un patient)
- Usuellement:
  - Données intrahospitalière
    - Données structurées
    - Comptes rendus d'hospitalisation (textes)
  - Données de remboursement de soins
  - Données de cohorte / de patientelle

## Pour les parcours

- Description de guidelines
- Données descriptives
  - souvent issues d'études épidémiologiques

→ données quantitatives (datées), précises

→ données qualitatives, *incomplètes*

# Différents traitements informatiques des trajectoires

- Représentation/visualisation
- Enrichir une trajectoire (inference)
- Adhérence d'une trajectoire à un parcours (recherche)
- Conformance d'une trajectoire à un parcours (guidelines)
- Clustering de trajectoires
- Généralisation d'un parcours à partir d'un ensemble de trajectoires
- Prédiction/classification de trajectoires

# Différents traitements informatiques des trajectoires

## Sur des trajectoires ou parcours individuels

- Représentation/visualisation
- Enrichir une trajectoire (inference): e.g. abstraction temporelle, reconstruction d'évènements
- Vérification de la consistance d'un parcours

## Confrontation de trajectoires avec un parcours

- Adhérence d'une trajectoire à un parcours (recherche)
- Conformance d'une trajectoire à un parcours (guidelines)

## Sur des ensembles de trajectoires

- Représentation/visualisation
- Clustering de trajectoires
- Généralisation d'un ensemble de trajectoires en parcours
- Préviation/classification de trajectoires

# Outline

- 1 Parcours de soins
  - Motivation
  - Représentation du parcours de soins
  - Les tâches
- 2 Représentation des parcours de soins
  - Représentation de l'information temporelle
  - Temps avec logiques modales
  - Ontologies about time
  - Conclusions
- 3 Requêtage de données temporelles
  - Contexte
  - Querying with temporal pattern matching
  - Querying with reasoning on care pathways: a web semantic approach
  - Conclusion
- 4 Conclusion

# De l'importance des représentations

La représentation d'une trajectoire/parcours est inhérente à l'informatique

## Compromis expressivité / calculabilité

- Expressivité

- On cherche une représentation pour exprimer le plus d'information possible de la trajectoire/parcours

- Calculabilité

- On cherche des outils informatiques qui résolvent en temps raisonnable les questions

⇒ plus le formalisme est expressif et plus les calculs pour exploiter les données représentées seront lourds

Quelles représentations pour les trajectoires/parcours de soins ?

# Représentation ontologique

Un homme de 35 ans a été admis à l'hôpital pour un gonflement périorbitaire, une rougeur et une douleur le 24 mai 2014. On lui a alors diagnostiqué une cellulite périorbitaire. Il a été traité avec de la clindamycine en IV et de la ciprofloxacine IV, ce qui a réduit la rougeur et le gonflement de l'orbite. Cependant, le deuxième jour suivant le traitement antibiotique, il a développé des nausées et des douleurs abdominales dans le quadrant supérieur droit, et ses tests de la fonction hépatique ont commencé à augmenter. Un diagnostic de lésion hépatique idiosyncrasique d'origine médicamenteuse a été posé. Il sortira au bout de 10 jours d'hospitalisation.

- Hospitalisation: interval débutant le 24 mai, terminé le 2 juin
- Traitement 1 : durée au moins 2 jours, pendant l'hospitalisation
- Réduction rougeur: Entre le début du traitement et le début des nausées
- Nausées: débute 2 jours après le début de traitement 1
- Augmentation: en même temps que nausées
- Diagnostic: après l'augmentation et pendant l'hospitalisation

# Représentation formelle des informations temporelles

## Représentations

- numériques / qualitatives (symboliques)
- absolu / relative au temps courant

## Différents formalismes dont

- Représentation et gestion des contraintes temporelles
  - Algèbre des instants / algèbres des intervalles
  - CSPs
- Extension de la logique classique : logiques temporelles
  - logiques modales
  - réifiées : calcul des situations (*situation calculus*), calcul des événements (*event calculus*)

⇒ On laisse de côté le volet ontologique quelques instant ...

# Algèbre des instants [VK86, McD82]

- Objets temporels : instant, point de  $T$
- Relations binaires de base:  $\mathcal{B} = \{<, >, =\}$
- 8 relations possibles entre instants:

$$2^{\mathcal{B}} = \{\emptyset, <, >, =, \{<, >\}, \{<, =\}, \dots, \{<, >, =\}\}$$

## Opération sur les relations

- Union, intersection, complément ( $\sim$ )
  - Inverse ( $\neg$ )
  - Composition ( $\circ$ )
- ⇒ la structure  $(2^{\mathcal{B}}, \circ, =)$  est une algèbre de relations

$\neg$	
$<$	$>$
$>$	$<$
$=$	$=$

$$\neg(l_1 < l_2) \wedge l_3 = l_2 \implies l_2 < l_3$$



# Algèbre des instants [VK86, McD82]

- Objets temporels : instant, point de  $T$
- Relations binaires de base:  $\mathcal{B} = \{<, >, =\}$
- 8 relations possibles entre instants:

$$2^{\mathcal{B}} = \{\emptyset, <, >, =, \{<, >\}, \{<, =\}, \dots, \{<, >, =\}\}$$

## Opération sur les relations

- Union, intersection, complément ( $\sim$ )
  - Inverse ( $\neg$ )
  - Composition ( $\circ$ )
- ⇒ la structure  $(2^{\mathcal{B}}, \circ, =)$  est une algèbre de relations

$\circ$	$<$	$>$	$=$
$<$	$<$	$?$	$<$
$>$	$?$	$>$	$>$
$=$	$<$	$>$	$=$

$$\neg(l_1 < l_2) \wedge l_3 = l_2 \implies l_2 < l_3$$

# Appliquons cela à notre exemple

On ne dispose que d'instants ... les intervalles sont représentés avec des instants de début et des instants de fin!

## Ensemble des faits (conjonctifs)

- $T_{Hd} < T_{Hf}, T_{Td} < T_{Tf}$
- $T_{Hd} \leq T_{Td} \wedge T_{Tf} \leq T_{Hf}$
- $T_R \geq T_{Td}, T_R \leq T_{Tf}$
- ...

## Nouveaux faits pouvant être inférés

$\Rightarrow T_{Hd} \leq T_R \wedge T_R \leq T_{Hf}$ : la rougeur apparaît pendant l'hospitalisation

# À vous de jouer

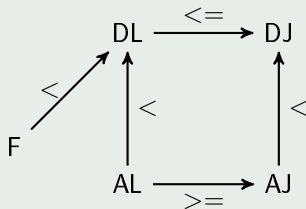
## Utilisation du reasoner de Protégé

- Créer un nouveau projet
- Définir un concept d'Instant
- Définir les notions de relation entre les Instants
  - commencer avec =, < et > seulement
  - compléter avec les <= et >= dans un second temps
- Décrire les faits représentant la situation du slide précédent
- Implémenter les règles de composition de l'algèbre des instants
  - possible avec les propriétés des relations ou avec SWRL
- Tester l'inférence (avec HermiT)
  - Regarder les explications
- charger le fichier `base_instantalgebra.ro`

# Graphe de contraintes temporelles (GCT)

## Graphe de contraintes temporelles: $(V, C)$

- $V$ : ensemble des instants
  - $C : V \times V \mapsto 2^{\mathcal{B}}$ : relations temporelles entre les instants
- Lucie n'est pas arrivée avant Jean et elle n'est pas partie après lui, mais après l'arrivée du facteur

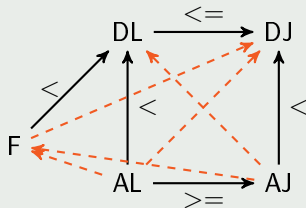


# Graphe de contraintes temporelles (GCT)

## Graphe de contraintes temporelles: $(V, C)$

- $V$ : ensemble des instants
- $C : V \times V \mapsto 2^{\mathcal{B}}$ : relations temporelles entre les instants

- Lucie n'est pas arrivée avant Jean et elle n'est pas partie après lui, mais après l'arrivée du facteur
- Jean est-il arrivé avant le facteur ?



# Consistance des GCT

## Instanciation consistante

- Consiste à déterminer une valeur (date) pour chaque instant qui satisfasse l'ensemble des contraintes temporelles.
- Problème similaire avec un assignement incomplet prédéfini

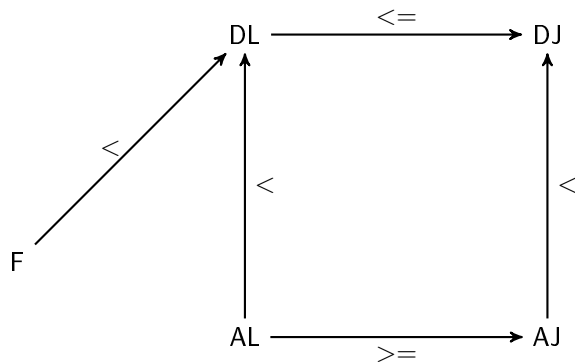
## Scénario consistant

- Consiste à compléter les relations entre les instants de manière consistante
- Si il existe un scénario consistant, il existe une instanciation consistante

## Consistance du GCT

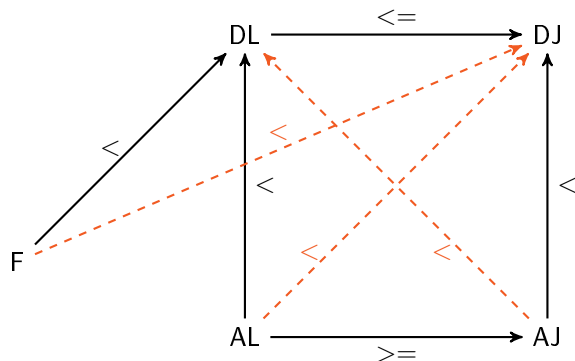
- Le GCT est consistant si il existe une instanciation consistante
  - Sinon, les contraintes temporelles ne sont simplement pas satisfiables
- méthode de résolution de ces problèmes assez classique par résolution de contraintes

## Construction d'un scénario



## Construction d'un scénario

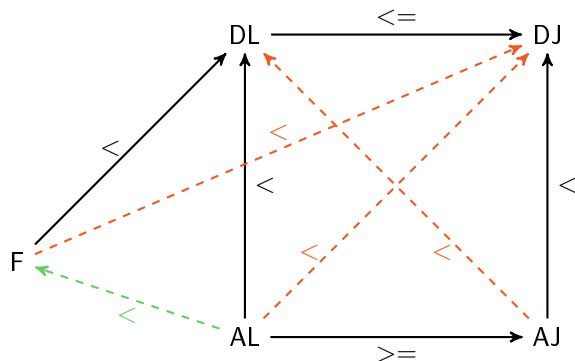
- utilisation de l'algèbre pour inférer des relations





## Construction d'un scénario

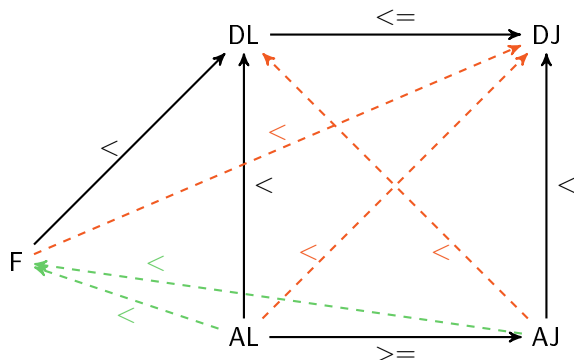
- utilisation de l'algèbre pour inférer des relations
- ajout hypothèse supplémentaire  $AL < F$



## Construction d'un scénario

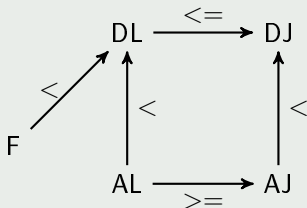
- utilisation de l'algèbre pour inférer des relations
- ajout hypothèse additionnelle  $AL < F$
- nouvelle inférence

- $P = 6$
- $AJ = 12$
- $AL = 14$
- $DL = 18$
- $DJ = 28$



# Utilisation des graphes de contraintes pour le requêtage

## Exemple de requête



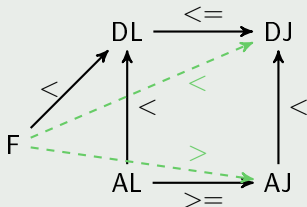
Est-ce possible que le facteur soit venu alors que Jean était présent ?

## Résolution d'une requête : raisonnement par défaut

- ❶ on ajoute la requête comme nouvelle contrainte
- ❷ on vérifie la consistance du GCT
  - si consistant alors la réponse est "oui"
  - sinon, la réponse est "non"

# Utilisation des graphes de contraintes pour le requêtage

## Exemple de requête



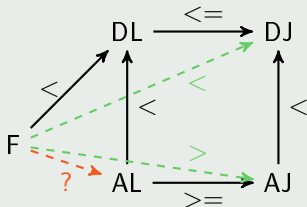
Est-ce possible que le facteur soit venu alors que Jean était présent ?

## Résolution d'une requête : raisonnement par défaut

- 1 on ajoute la requête comme nouvelle contrainte
- 2 on vérifie la consistance du GCT
  - si consistant alors la réponse est "oui"
  - sinon, la réponse est "non"

# Utilisation des graphes de contraintes pour le requêtage

## Exemple de requête



Est-ce possible que le facteur soit venu alors que Jean était présent ?  
"oui"

## Résolution d'une requête : raisonnement par défaut

- 1 on ajoute la requête comme nouvelle contrainte
- 2 on vérifie la consistance du GCT
  - si consistant alors la réponse est "oui"
  - sinon, la réponse est "non"

# Algèbre des instants

## Quelques conclusions

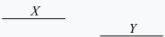

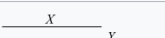
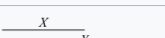
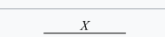
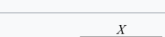
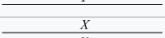
- bonne illustration de la notion de raisonnement temporel
  - objets temporels: des instants
  - raisonnement sur l'ordre d'apparition des évènements
- expressivité limitée
- inférence de nouvelles connaissances également limitée
- gestion des données sous forme d'intervalles un peu lourde
- + mécanismes de résolution simples

# Algèbre des intervalles [All83]

- Objets temporels : intervalles (couple d'instant)
  - Instant représentable comme intervalle dégénéré
- Relations binaires de base : 13 relations d'Allen

$$\mathcal{B} = \{<, m, o, s, d, f, >, mi, oi, si, di, fi, =\}$$

- $2^{\mathcal{B}} = 8292$  relations possibles

Relation	Illustration	Interprétation
$X < Y$ $Y > X$		X se déroule avant Y ( <i>before</i> et <i>after</i> )
$X m Y$ $Y mi X$		X rencontre Y ( <i>meets</i> ) ( <i>i</i> signifie <i>inverse</i> )
$X o Y$ $Y oi X$		X chevauche Y ( <i>overlaps</i> )
$X s Y$ $Y si X$		X démarre Y ( <i>starts</i> )
$X d Y$ $Y di X$		X se déroule pendant Y ( <i>during</i> )
$X f Y$ $Y fi X$		X termine Y ( <i>finishes</i> )
$X = Y$		X est égal à Y

# Exemples de représentation

- Je lis le journal **avant** le petit déjeuner:  
 $Lecture\{\prec, m\}PetitDej$
- Je lis le journal **pendant** le petit déjeuner:  
 $Lecture\{s, d, f, =\}PetitDej$
- Je lis le journal **avant ou pendant** le petit déjeuner:  
 $Lecture\{\prec, m, s, d, f, =, o\}PetitDej$

Instants :

- $Deb(lect) < Fin(Lect); Deb(PetitDej) < Fin(PetitDej);$
- $Fin(Lect)\{\prec, =\}Fin(PetitDej)$
- Je lis le journal **avant ou après** le petit déjeuner:  
 $Lecture\{\prec, m, a, mi\}PetitDej$

Instants???



# Raisonnement sur les intervalles

## Algèbre de relations temporelles

- Opérateurs sur les relations
  - Union, intersection, complément de relation temporelles
  - Inverse d'une relation:  $mi \rightarrow \neg mi$
  - Composition de relations

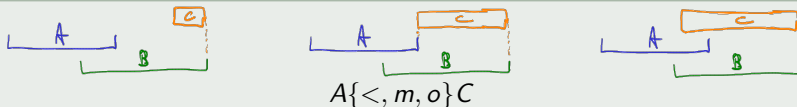
Exemple composition  $AoB$  et  $BfC$  ??

# Raisonnement sur les intervalles

## Algèbre de relations temporelles

- Opérateurs sur les relations
  - Union, intersection, complément de relation temporelles
  - Inverse d'une relation:  $mi \rightarrow \neg mi$
  - Composition de relations

## Exemple composition $AoB$ et $BfC$ ??



# Compositions dans la logique d'intervalles

$$R = \{m, s, si\} \text{ et } S = \{f, o\}$$

Les compositions entre relations non-élémentaires sont obtenues comme l'union des compositions des relations élémentaires.

$$\bullet R \circ S = \{d, o, s, <, m, oi, di, fi\}$$

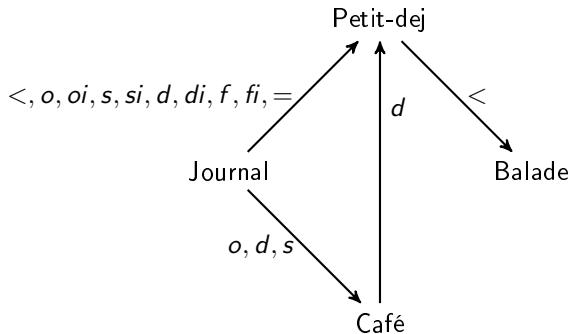
	<	m	o	f	s	d	di	si	fi	oi	mi	>
<	<	<	<	<	<	< mosd	<	< mosd	<	< mosd	< mosd	tout >
m		<	<	<	m	osd	<	osd	m	osd	ffi =	
o			< mo	< mo	o	osd	< mofidi	osd	ofidi	oioffiddissi =		
s				< mo	s	d	< mofidi	d	ssi =			
fi					o	osd	di	fifi =				
d						d	tout					
di						oioffiddissi =						

→ Raisonnement qui est plus difficile (impossible?) à mettre en place que sur les instants

→ Utilisation de règles SWRL comme alternative

# Graphe de contraintes temporelles

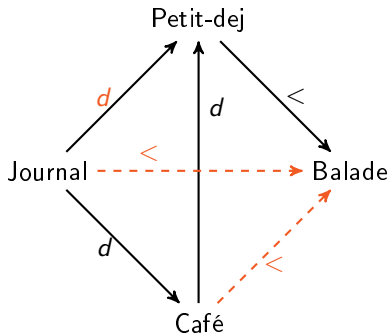
Fred lit le journal pendant qu'il prend son petit-déjeuner. Il pose son journal et boit la fin de son café. Après le petit-déjeuner, il part se balader



- Vérifier la consistance d'un graphe de contraintes temporelles d'intervalles est difficile!

# Graphe de contraintes temporelles

Fred lit le journal pendant qu'il prend son petit-déjeuner. Il pose son journal et boit la fin de son café. Après le petit-déjeuner, il part se balader



- Vérifier la consistance d'un graphe de contraintes temporelles d'intervalles est difficile!
- Exemple de scénario consistant (instanciation facilement déductible)

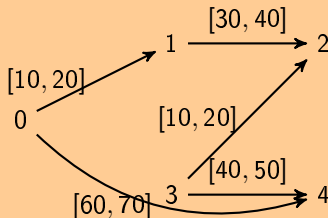
# CSP temporels quantitatif [DMP91]

## Représentation quantitative du temps

Objets temporels:

- Des instants datés (numériquement)
- Contraintes sur les délais entre les instants : interval numériques
- Domaine du temps : discret ou dense (réel ou rationnel)

Représentation sous la forme de graphe de contraintes temporelles (numériques)



# CSP temporels quantitatif [DMP91]

## Ce qu'on peut faire avec:

- Vérification si une instantiation temporelle correspond au graphe
- Vérification de la consistance du graphe: est-ce qu'une instantiation des évènements respectant ses contraintes est possible?
- Compléter les contraintes non-connues (si elles existent)
- Minimisation des intervalles temporels possibles (trouver les instantiations)

## pros/cons

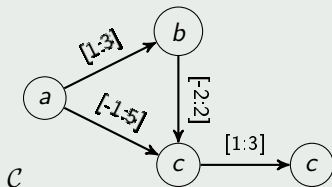
- + résolution de tâches complexes (décidable et efficace)
- + facilement interprétable
- nécessite des informations assez complètes
- raisonnement quantitatif au sein des raisonneurs du web sémantique ?

# Chroniques [DGG93]

## Chronicle: Example and graphical representation

Chronicle  $(\mathcal{E}, \mathcal{T})$  where

- $\mathcal{E} = \{a, b, c, c\}$
- $\mathcal{T} = \{(a, 1)[1 : 3](b, 2),$   
 $(a, 1)[-1 : 5](c, 3),$   
 $(b, 2)[-2 : 2](c, 3),$   
 $(c, 3)[1 : 3](c, 4)\}$



## Différences avec les CSP temporels

- Gestion explicite des évènements répétés
- Introduction de délais négatifs



# À vous de jouer

## Expérimentation du raisonnement sur graphes

⇒ Utilisation du notebook : `Temporal_network.ipynb`

1. Exemple sur le raisonnement sur graphes de contraintes temporelles
2. Exemple d'utilisation sur les parcours des exérèses pulmonaires

# Limites et extensions des algèbres d'instants et d'intervalles

## Limites de ces représentations

- Mélange des objets temporels: instants, intervalles, etc.
- Mélange de raisonnement quali et quanti [Mei96]
- Représentation d'une information incertaine ou imprécise
- Représentation explicite de la granularité temporelle

→ il existe des extensions des algèbres permettant d'intégrer ces dimensions

# Autres formalismes classiques de représentation de l'information temporelles

## Logiques modales

- LTL / CTL : linear timed logic / computation tree logic
- MTL/TPTL [AH93, BCM10]: version métriques des logiques temporelles
- MITL [BRSP16]: version sur intervalles de la MTL

## Logiques d'actions

- Event calculus
- Situation calculus

# LTL: Linear Temporal Logic

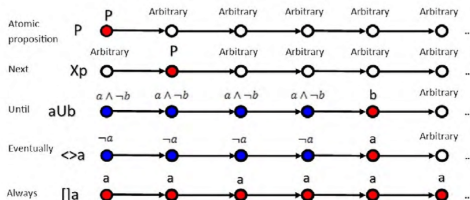
## Intuition de la LTL

- représentation discrète du temps : instants successifs
- chaque instant est décrit par des atomes d'un vocabulaire.

$\bigcirc \varphi$  La formule  $\varphi$  est vraie à l'instant suivant

$\Box \varphi$  La formule  $\varphi$  est vraie à tous les instants futurs (strict)

$\Diamond \varphi$  La formule  $\varphi$  sera vraie à au moins un instant futur (strict)



## Exemple de formule LTL

$$\Box((\neg \text{passport} \vee \neg \text{ticket}) \implies \bigcirc \neg \text{board\_flight})$$

# MTL: Metric Temporal Logic

## Intuition de la MTL

- représentation du temps continue : évènements datés
- chaque évènement est décrit par des atomes d'un vocabulaire.

$\Box_{[a,b]}\varphi$  La formule  $\varphi$  est vraie à tous les instants dans le délai  $[a, b]$   
 $\Diamond_{[a,b]}\varphi$  La formule  $\varphi$  sera vraie à au moins un instant dans le délai  $[a, b]$

## Exemple de formule LTL

$$\Diamond_{[0,10]}(p \wedge \Box_{[0,5]}\neg p)$$

→ il existe un moment dans les 10 UT où  $p$  est vraie, et qu'il ne sera pas vrai de nouveau pendant les 5 UT suivantes.

# Comparaison MTL/Chronicles [GM22]

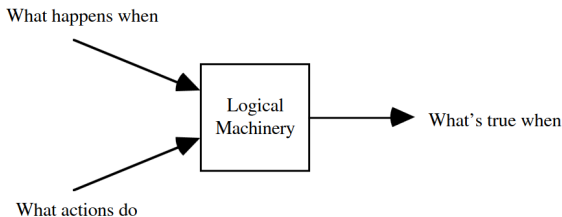
## Chroniques vs MTL

- Chronique  
 $(\{a, b, c\}, \{(a, 1)[1, 3](b, 2), (b, 2)[2, 4](c, 3), (a, 1)[5, 7](c, 3)\})$
- $\Box_{[0, \infty]}(a \wedge \Diamond_{[0, 4]}b \wedge \Diamond_{[2, 6]}b)$

- les chroniques sont théoriquement plus expressives que la MTL [GM22]
- en temps discret, il y a l'équivalence

# Event Calculus [KS89, Sha01]

- Introduit par Kowalski et Sergot [KS89], Mueller [Mue04] (DEC reasoner: EC+SAT)
- formalisme logique pour représenter les événements et leurs effets
- C'est une logique orientée vers des événements qui provoquent des *changements*



# Event Calculus [KS89, Sha01]

- Introduit par Kowalski et Sergot [KS89], Mueller [Mue04] (DEC reasoner: EC+SAT)
- formalisme logique pour représenter les événements et leurs effets
- C'est une logique orientée vers des événements qui provoquent des *changements*

- trois sortes d'objet:

- event  $e$
- fluent  $f$
- timepoint  $t$

- prédicats:

- $Happen(e, t)$
- $HoldsAt(f, t)$
- $Initiates(e, f, t)$
- $Releases(e, f, t)$
- $ReleasedAt(f, t)$
- $InitiatedAt(f, t)$

- des axiomes:

- $Happens(e, t) \wedge Initiates(e, f, t) \implies HoldsAt(f, t + 1)$
- $Happens(e, t) \wedge Terminates(e, f, t) \implies \neg HoldsAt(f, t + 1)$
- $Happens(e, t) \wedge Releases(e, f, t) \implies ReleasedAt(f, t + 1)$



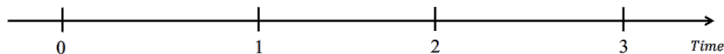
## Example: Wake up

- Events:  $WakeUp(a)$ ,  $FallSleep(a)$  où  $a$  est un agent
- Fluent:  $Awake(a)$
- Axiomes spécifiques:
  - $Initiates(WakeUp(a), Awake(a), t)$
  - $Terminates(FallAsleep(a), Awake(a), t)$
- Axiome générique :  
 $Happens(e, t) \wedge Initiates(e, f, t) \implies HoldsAt(f, t + 1)$

- Premises
  - $\neg HoldsAt(Awake(John), 0)$
  - $Happens(WakeUp(John), 1)$
- Question :  $HoldsAt(Awake(John), 3)$

$Happens(WakeUp(John), 1)$

$\neg HoldsAt(Awake(John), 0)$

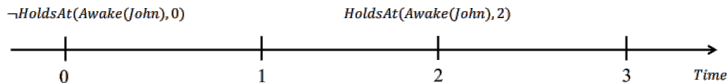


## Example: Wake up

- Events:  $WakeUp(a)$ ,  $FallSleep(a)$  où  $a$  est un agent
- Fluent:  $Awake(a)$
- Axiomes spécifiques:
  - $Initiates(WakeUp(a), Awake(a), t)$
  - $Terminates(FallAsleep(a), Awake(a), t)$
- Axiome générique :  
 $Happens(e, t) \wedge Initiates(e, f, t) \implies HoldsAt(f, t + 1)$

- Premises
  - $\neg HoldsAt(Awake(John), 0)$
  - $Happens(WakeUp(John), 1)$
- Question :  $HoldsAt(Awake(John), 3)$

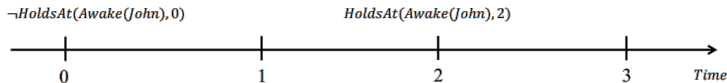
$Happens(WakeUp(John), 1)$



# Example: Wake up

- Events:  $WakeUp(a)$ ,  $FallSleep(a)$  où  $a$  est un agent
- Fluent:  $Awake(a)$
- Axiomes spécifiques:
  - $Initiates(WakeUp(a), Awake(a), t)$
  - $Terminates(FallAsleep(a), Awake(a), t)$
- Axiome générique :  
 $Happens(e, t) \wedge Initiates(e, f, t) \implies HoldsAt(f, t + 1)$

- Premises
  - $\neg HoldsAt(Awake(John), 0)$
  - $Happens(WakeUp(John), 1)$
- Question :  $HoldsAt(Awake(John), 3)$
- (Circumscription) hypothèse qu'aucun autre évènement n'existe  
 $Happens(WakeUp(John), 1)$

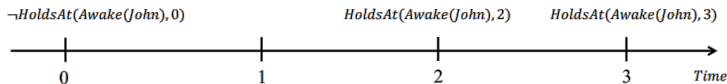


## Example: Wake up

- Events:  $WakeUp(a)$ ,  $FallSleep(a)$  où  $a$  est un agent
- Fluent:  $Awake(a)$
- Axiomes spécifiques:
  - $Initiates(WakeUp(a), Awake(a), t)$
  - $Terminates(FallAsleep(a), Awake(a), t)$
- Axiome générique :  
 $Happens(e, t) \wedge Initiates(e, f, t) \implies HoldsAt(f, t + 1)$

- Premises
  - $\neg HoldsAt(Awake(John), 0)$
  - $Happens(WakeUp(John), 1)$
- Question :  $HoldsAt(Awake(John), 3)$

$Happens(WakeUp(John), 1)$



# Résumé sur les représentations des informations temporelles

- Double dimension des événements du parcours : **nature** et **support temporel** des événements
- Nature des objets temporels : Ponctuels vs intervalles, relatifs vs absolus, qualitatifs vs quantitatifs
- Les formalismes associent **représentation** et **raisonnement** temporels
  - Approches algébrique
    - définition d'opérateurs algébriques entre les représentations temporelles (instants, intervalles)
    - représentation des connaissances sous la forme d'un graphe de contraintes
    - raisonnement 1: détection de la consistance du graphe
    - raisonnement 2: complémentation du graphe
    - raisonnement 3: vérification d'un scénario
  - Approches logiques
    - représentation des raisonnements sous forme logique (axiomes)
    - présentation de l'*event calculus*: événements et fluents

⇒ formalisation par des ontologies ?

# Ontologies du temps

- Time-OWL : standard W3C de la représentation du temps
- Timed-Event Ontology: approche événementielle
- autres ontologies (plus marginales) ...

# Time-OWL

## Objective of Time-OWL

- Represent knowledge of the temporal relationships between *events*
- Focused on the temporal ordering (qualitative relationship)

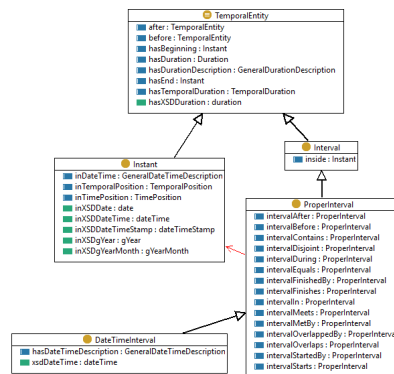
- Ontologie de référence de la W3C pour représenter l'information temporelle
- <https://www.w3.org/TR/owl-time>

## Overall organisation of the ontology

- Topological Temporal Relations
- Temporal reference systems, clocks, calendars
- Time position
- Duration

# Time-OWL: Topological Temporal Relations

- Formalized by Hobbs and Pan [HP04]
- Base class: **Temporal Entity**
  - Interval**
  - Instant**
- ProperInterval**: distinct beginning and end (disjoint from Instant)



$$\forall T, \text{ProperInterval}(T) \equiv \text{Interval}(T) \wedge (\exists(t_1, t_2), \text{begins}(t_1, T) \wedge \text{ends}(t_2, T) \implies t_1 \neq t_2)$$



# OWL-Time: Time position and duration

- Base class:

**TemporalPosition**:  
alternative to the  
`xsd:dateTime`

- Subclasses

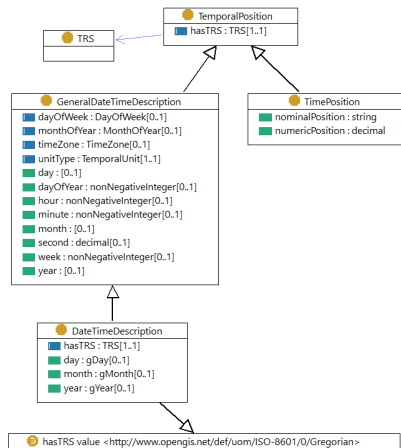
- TimePosition**

- 

- GeneralDateTimeDescription**

- DateTimeDescription**

- Can be used with alternative temporal reference systems



# OWL-Time: Time position example

`ex:meetingStart`

```
a                :Instant ;  
:inDateTime      ex:meetingStartDescription ;  
:inXSDDateTimeStamp 2017-04-12T10:30:00+10:00 .
```

`ex:meetingStartDescription`

```
a                :DateTimeDescription ;  
:unitType        :unitMinute ;  
:minute          30 ;  
:hour            10 ;  
:day             "---12"^^xsd:gDay ;  
:dayOfWeek       :Wednesday ;  
:dayOfYear       102 ;  
:week            15 ;  
:month           "--04"^^xsd:gMonth ;  
:monthOfYear     greg:April ;  
:timeZone        <https://www.timeanddate.com/time/zones/aest> ;  
:year            "2017"^^xsd:gYear .
```

Richer description of the time position

# OWL-Time: Time position and duration

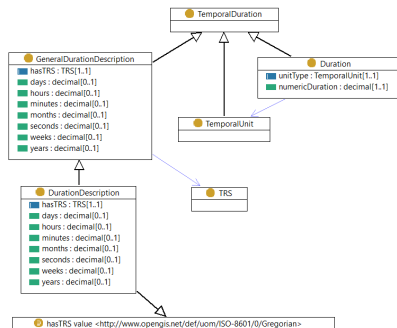
- An interval can be 1 day 2 hours, or 26 hours, or 1560 minutes, and so on.
- Base class: **TemporalDuration**
- Subclasses

- **Duration**

- 

- **GeneralDurationDescription**

- **DurationDescription**



# OWL-Time: à propos de son expressivité

## Peut-on “facilement” représenter la trajectoire de soin ?

- Hospitalisation: interval débutant le 24 mai, terminé le 2 juin
- Traitement: durée au moins 2 jours, pendant l'hospitalisation
- Réduction rougeur: Entre le début du traitement et le début des nausées
- Nausées: débute 2 jours après le début de traitement 1
- Augmentation: en même temps que nausées
- Diagnostic: après début des nausées et pendant l'hospitalisation

# À vous de jouer

## Peut-on “facilement” représenter la trajectoire de soins ?

- Hospitalisation: interval débutant le 24 mai, terminé le 2 juin
- Traitement: durée au moins 2 jours, pendant l'hospitalisation
- Réduction rougeur: Entre le début du traitement et le début des nausées
- Nausées: débute 2 jours après le début de traitement 1
- Diagnostic: après début des nausées et pendant l'hospitalisation

## Utilisation de Time-OWL

- Lancer Protégé
- Consulter la documentation <https://www.w3.org/TR/owl-time/>
- Représenter la situation de l'exemple en utilisant les ontologies de domaine
  - créer un nouveau projet
  - importer l'ontologie OWL-Time
  - utiliser des intervalles ou des instants
  - la nature des événements sera codée dans les labels

→ quelles informations sont difficiles à coder ?

→ le raisonneur enrichie-il les données de manière intéressante ?

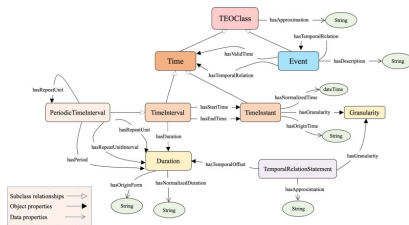
# OWL-Time: CCL

- Ontologie de référence de la W3C pour représenter l'information temporelle
- Ontologie très riche pour la description des positions temporelles et des durées
  - pas sûr que cela soit très utile pour représenter les trajectoires de soins
- Modélisation des entités temporelles intéressante mais limité
  - basé sur la représentation comme intervalle
  - représentation des relations qualitative entre intervalles
  - **pas de représentation des informations métriques** : “3 jours après” !
- Toutes les notions de l'ontologies ne sont pas vraiment implémentées dans l'ontologie (e.g. ProperInterval)

# TEO: Timed Event Ontology [LDH<sup>+</sup>20]

## TEO in a nutshell

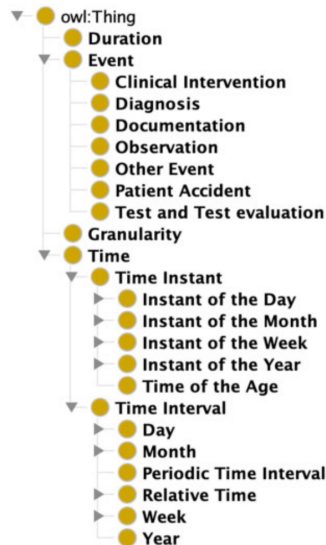
- Notion d'évènement temporisé au centre du modèle
- Redéfinition des notions temporelles de OWL-Time
- Ajoute la notion de granularité temporelle
- Réification des relations temporelles: une relation temporelle entre deux objets est représentée par une instance
  - permet un enrichissement sémantique
  - rend les raisonnements plus complexes



# TEO: Timed Event Ontology [LDH<sup>+</sup>20]

## TEO in a nutshell

- Notion d'évènement temporisé au centre du modèle
- Redéfinition des notions temporelles de OWL-Time
- Ajoute la notion de granularité temporelle
- Réification des relations temporelles: une relation temporelle entre deux objets est représentée par une instance
  - permet un enrichissement sémantique
  - rend les raisonnements plus complexes



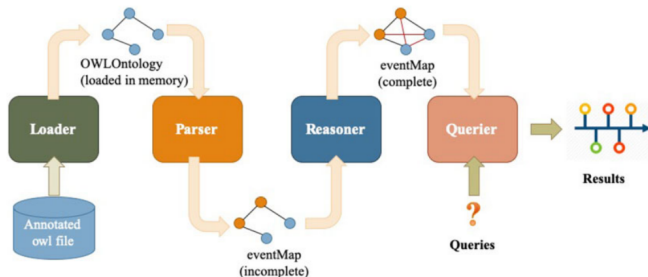


# TEO: Exemple de codage d'un évènement

- La Nausée début 2 jours après le début de traitement 1

```
evt1 rdf: type Event;  
      rdfs:label "Nausee";  
      after evt2.  
evt2 rdf: type Event;  
      rdfs:label "Traitement 1".  
duration rdf: type Duration;  
          rdfs:label "2 jours";  
          hasNormalizedDuration "2D".  
state1 rdf: type TemporalRelationStatement;  
        rdf:object evt1;  
        rdf: predicate "after";  
        rdf: subject evt2;  
        hasTimeOffset duration.
```

# TEO: raisonnement temporel



- Utilisation d'un raisonneur externe
  - Raisonnement temporel basé sur l'algèbre d'Allen
- ⇒ le raisonneur pourrait être enrichi!

# D'autres références d'ontologie du temps

## Autres ontologies

- CNTRO [TWS<sup>+</sup>10], très proche de TEO
- CHRONOS [PPB14] : outil de gestion de contraintes temporelles dans Protegee

## Temporal Knowledge Graphs

- Temporal knowledge graphs are graphs with a set of facts, information, or knowledge that have temporal features. These graphs can also be considered as dynamic, evolving, or time-varying graphs.
- Temporisation des instances et des relations
- e.g. EventKG [GD18]

- représentation riche => impossibilité d'utiliser les raisonnements ontologiques seul
- besoin de combinaison raisonneur ontologique / raisonnement temporel

# Conclusions

- Clarification: trajectoires vs parcours
  - Représentation et raisonnement temporels
    - notion multi-factorielle : instants/intervalles, granularité,
    - nombreux formalismes pour représenter l'information temporelle
    - expressivités variées
  - Intégration dans le cadre du web sémantique
    - représentation des données à l'aide d'ontologies déjà bien établies
    - les raisonnements temporels sont rarement compatibles avec la logique de description
      - besoin de combiner deux modes de raisonnements différents
- formalisation adaptée à représenter des parcours (plus que des trajectoires)

# Outline

- 1 Parcours de soins
  - Motivation
  - Représentation du parcours de soins
  - Les tâches
- 2 Représentation des parcours de soins
  - Représentation de l'information temporelle
  - Temps avec logiques modales
  - Ontologies about time
  - Conclusions
- 3 Requêtage de données temporelles
  - Contexte
  - Querying with temporal pattern matching
  - Querying with reasoning on care pathways: a web semantic approach
  - Conclusion
- 4 Conclusion

# Différents traitements informatiques des trajectoires

## Sur des trajectoires ou parcours individuels

- Représentation/visualisation
- Enrichir une trajectoire (inference): e.g. abstraction temporelle, reconstruction d'évènements
- Vérification de la consistance d'un parcours

## Confrontation de trajectoires avec un parcours

- Adhérence d'une trajectoire à un parcours (recherche)
- Conformance d'une trajectoire à un parcours (guidelines)

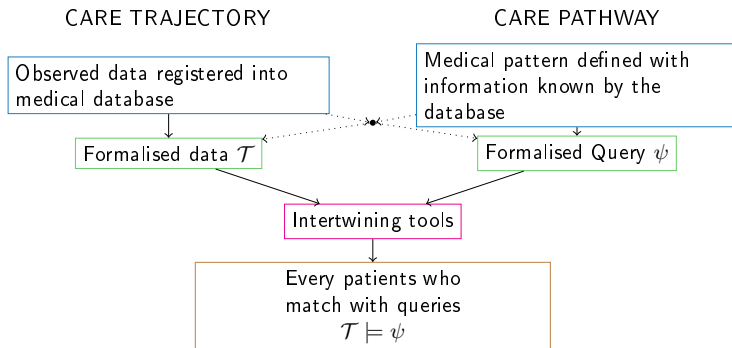
## Sur des ensembles de trajectoires

- Représentation/visualisation
- Clustering de trajectoires
- Généralisation d'un ensemble de trajectoires en parcours
- Préviation/classification de trajectoires

# Requêtage de données temporelles

## Confrontation de trajectoires avec un parcours

- Adhérence d'une trajectoire à un parcours (recherche)
- Conformance d'une trajectoire à un parcours (guidelines)



- ⇒ Motivation dans le contexte de l'épidémiologie à partir de grandes bases de données médico-administratives
- Temporal constraints are present in 38% of clinical research

# Données médico-administratives : qu'est ce que c'est ? I

## Données médico-administratives

### **Données produites dans le cadre de soins et collectées pour l'organisation et la prise en charge du patient**

- données collectées pour le suivi médical d'un patient/d'un service/d'un établissement
  - disponibilité des lits
  - suivi d'une pathologie
- données collectées pour la gestion administrative des patients
  - facturation des prestations de soins
  - gestion des remboursements

→ voir cours de Vendredi pour des aspects éthiques et légaux sur l'usage de ces données secondaires



# Données médico-administratives : qu'est ce que c'est ? II

## Deux sources typiques en France

- les données de l'assurance maladie (SNIIRAM)
  - très grande population
  - taux de couverture très élevé
- les données intra-hospitalières
  - très grands centres hospitaliers (APHP/APHM/HCL)
  - programme d'uniformisation assez précoce (PMSI)

## Leurs caractéristiques

- massives (++) , variées (?-)
- informatisée (++)
- disponibles facilement (+)
- information riche (+/-)
- "qualité" (+/-)

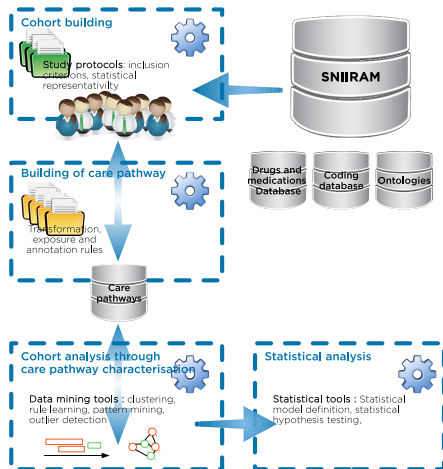
# Données médico-administratives : qu'est ce que c'est ? III

## Deux sources classiques en France

- les données de l'assurance maladie (SNDS/SNIIRAM)
  - très grande population
  - taux de couverture très élevé
  - vue longitudinale assez complète
  - complémenté avec CepiDC, données handicap
  - uniquement des données structurées
- les données intra-hospitalières
  - très grands centres hospitaliers (APHP/APHM/HCL)
  - programme d'uniformisation assez précoce (PMSI)
  - vue longitudinale aveugle hors de l'hôpital
  - données structurées et non-structurées

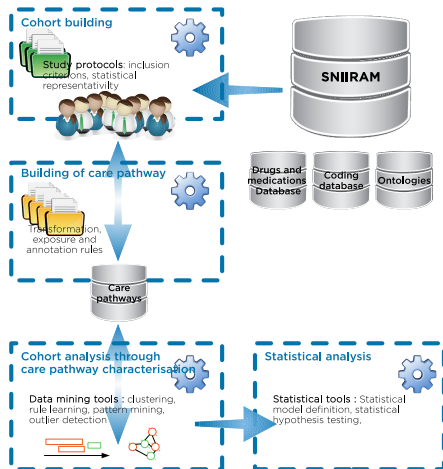
# Epidemiology with MADB

- 1 Sélection de patient : patients épileptique stable
- 2 Sélection des événements : délivrances de médicaments, crise
- 3 Mise en forme du jeu de données
- 4 Analyse de données
  - Algorithmes de machine learning
  - Méthode statistique
- 5 Retour à 1
- 6 Validation statistique



# Epidemiology with MADB

- 1 **Selection of patient : patients épileptique stable**
- 2 **Sélection des événements : délivrances de médicaments, crise**
- 3 Mise en forme du jeu de données
- 4 Analyse de données
  - Algorithmes de machine learning
  - Méthode statistique
- 5 Retour à 1
- 6 Validation statistique



# Patient selection [DBZ<sup>+</sup>13]

Patient  $p$  eligible iff:

- all inclusion criteria ( $I_i(p)$ ) are met
- none of the exclusion ( $E_j(p)$ ) are met

$$\text{patient eligible} \Leftrightarrow \left( \bigwedge I_i(p) \right) \wedge \left( \bigwedge \neg E_j(p) \right)$$

For example

- $I_0$ : evidence of a prostate adenocarcinoma
- $I_1$ : absence of metastasis
- $E_0$ : patient older than 70 years old
- $E_1$ : evidence of diabetes

$$\text{patient eligible} \Leftrightarrow I_0(p) \wedge I_1(p) \wedge \neg E_0(p) \wedge \neg E_1(p)$$

With medico-administrative databases

- patients are represented by care pathways  $\Rightarrow$  information richness
- high level medical events are not readily available

# Avantages et limites

## Avantages

- la donnée est déjà acquises et informatisée (données historisées)
- beaucoup de données structurées, normalisation des codages
- couverture large et population importante
- **données longitudinales** (données datées) : analyse des parcours de soins

## Difficultés

- difficultés techniques liées au sockage et à l'exploitation de données massives : systèmes informatiques robustes
- difficultés techniques liées à la conformité aux réglementations
- difficultés à fusionner des données : pour raison réglementaire, pour des raisons pratiques
- manque de standardisation
- **difficultés sémantiques : c'est une donnée administrative**

# Écart sémantique

## Changement des usages des données

- usage primaire : informations collectées pour des objectifs (administratifs)
  - usage secondaire : utilisation souhaitée pour un autre usage (médical)
- spécificité par rapport à la collecte de données : changement épistémologique important ... fortement lié à l'IA

## Écart sémantique

- on cherche à utiliser une information dans un champs sémantique différent de celui pour laquelle elle a été collectée
  - Exemple : identification d'une pathologie à partir des données du SNDS (délivrances de soins)
- ⇒ le raisonnement automatique permet de combler (partiellement) ces écarts

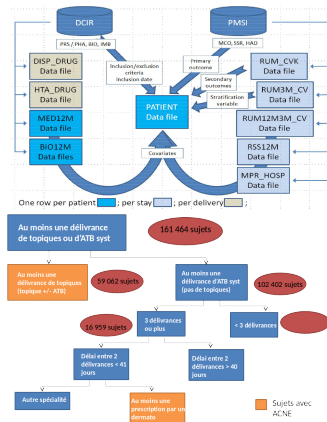
# From raw database to dataset to analyze

## Current approach: data management book

### Data Management Book (DMB)

- Write down the whole procedure of data manipulations
  - from raw data to statistical analysis
  - study validation purpose
- Two types of *queries*
  - *selection* of patients according to a criteria (disease, exposure)
  - *detection* of a medical procedure

⇒ requêtes sur des bases de données médico-administratives : classiquement faites en SQL (base de données relationnelle)





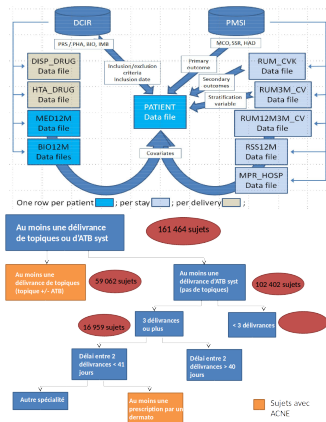
# From raw database to dataset to analyze

## Current approach: data management book

### Data Management Book (DMB)

- Write down the whole procedure of data manipulations
  - from raw data to statistical analysis
  - study validation purpose
- Two types of *queries*
  - *selection* of patients according to a criteria (**disease, exposure**)
  - *detection* of a **medical procedure**
- need to specify **medical events**

⇒ requêtes sur des bases de données médico-administratives : classiquement faites en SQL (base de données relationnelle)



# Use case: detecting patients having a Deep Venous Thrombosis (DVT) in the SNIIRAM

⇒ need for defining **high-level medical events** (*phenotypes*) from low level databases features (*semantic gaps*)

## Deep Venous Thrombosis (DVT)

- Deep vein thrombosis (DVT) is a medical condition that occurs when a blood clot forms in a deep vein
- ICD-10 Codes: I802
- SNOMED CT Concept Id: 128053003

## Detecting all patients with DVT in the SNIIRAM/DCIR

- ICD-10 codes are available only during hospitalisation and are not enough accurate (include suspicious DVT)
- Need to include additional features to better specify such a medical event (easily accessible features in SNIIRAM):
  - specific medical procedures: Doppler ultrasonography or CT-scan
  - antithrombotic deliveries / anticoagulant treatment

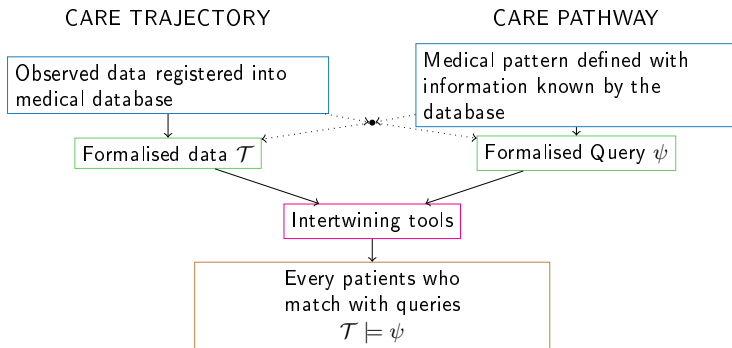
## Use case: detecting patients having a Deep Venous Thrombosis (DVT) in the SNIIRAM

*In clinical practice facing a suspicion of VTE physicians **first** prescribe antithrombotics and **then** confirm or not the diagnosis through **specific medical procedures**: e.g **Doppler ultrasonography** or **CT-scan**. Patients with suspected Pulmonary Embolism are often hospitalized whereas patients with suspected Deep Vein Thrombosis (DVT) are managed on an outpatient basis. On the one side, if the DVT suspicion is confirmed, **antithrombotic deliveries** continue **for 3 to 12 months (once per month)**. Hence, the diagnosis (through the same medical procedures as above) is preceded or followed by initiating an **anticoagulant treatment** within a time window of **at most 0 to 2 days**. On the other side, Pulmonary Embolism suspicion leads to hospitalization **during** which medical procedures are performed to confirm the diagnosis and then anticoagulant delivery is observed **only after** the patient comes back home.*

### Requirements for such complex queries

- **ontological reasoning** / ontology mediated query answering
- **temporal reasoning** / temporal query

# Requêtage de données temporelles



- ⇒ Représentation des trajectoires : données datées, évènements précis
- ⇒ Représentation des parcours (*computational phenotypes*) : généralisation de temps et d'évènements
- ⇒ Tâche de requêtage

# Our challenge

- Inputs:
  - A patient care pathway :  $\langle (\text{🩺}, 10), (\text{💊}, 20) \dots \rangle$
  - A computational phenotype (a *query*)
- Output: Yes or No ... the patient matches the phenotype

Challenge: Define a framework to specify and answer *computational phenotypes* (complex queries)

- expressive
  - use high-level medical concepts
  - handle temporal relations between events
- efficient to be applied on large numbers of care pathways

# Querying with temporal pattern matching

## Temporal sequence

- $E$ : a finite set of events,
- $T = \mathbb{Q}$  or  $\mathbb{R}$ : temporal domain
- $S = \langle (e_1, t_1), (e_2, t_2), \dots, (e_n, t_n) \rangle$  a finite temporal sequence where  $e_i \in E$  and  $t_i \in T$

## Temporal pattern

- A temporal pattern  $p$  represents a situation to recognize
  - $p$  is specified in a formal syntax
  - $p$  can *occur* in or *match* a sequence
- ⇒ a temporal pattern = a phenotype
- ⇒ recognize a temporal pattern = answer a query

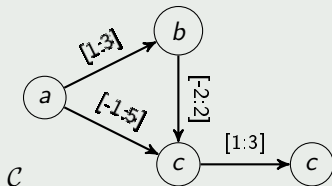
Many domains of temporal patterns: MTL, TPTL, Event Calculus, ...  
**chronicles**

# Chronicles

## Chronicle [DGG93]: Example and graphical representation

Chronicle  $(\mathcal{E}, \mathcal{T})$  where

- $\mathcal{E} = \{a, b, c, c\}$
- $\mathcal{T} = \{(a, 1)[1 : 3](b, 2),$   
 $(a, 1)[-1 : 5](c, 3),$   
 $(b, 2)[-2 : 2](c, 3),$   
 $(c, 3)[1 : 3](c, 4)\}$



## Occurrences of a chronicle

SID	Sequence
s <sub>1</sub>	(a, 1), (b, 3), (a, 4), (b, 4.7), (c, 5), (c, 6), (d, 7)
s <sub>2</sub>	(b, 2), (d, 4), (a, 5), (c, 7)
s <sub>3</sub>	(a, 1), (b, 4), (c, 5), (b, 6), (c, 8), (d, 9)
s <sub>4</sub>	(b, 4), (a, 6), (e, 8), (c, 9)
s <sub>5</sub>	(b, 1), (a, 3), (c, 4)
s <sub>6</sub>	(c, 4), (a, 5), (c, 7), (b, 8), (c, 9)

# Why chronicles?

Sufficient for specifying most of the temporal relations needed in phenotype expressions

- “*for 3 to 12 months*”
- “*at most 0 to 2 days*”

## Good properties

- **expressiveness**
  - complex temporal arrangement can be specified
  - constraints on continuous temporal domain
  - does not enforce to have a strict order on the sequence of events
- **interpretability**: the graphical representation makes them easy to manipulate / interpret
- **efficiency**: There are efficient algorithms to recognize/enumerate occurrences of a chronicle in long sequences [GBS<sup>+</sup>20]



# Chronicle occurrence enumeration algorithm

## Main principles

- an *on-going occurrence* of a chronicle maps each event of  $\mathcal{E}$  with an interval of admissible positions

$$\forall e \in \mathcal{E}, \text{adm}(e) = [a^-, a^+]$$

- the algorithm transverses the set of events  $e \in \mathcal{E}$ 
  - progressively narrows intervals to single position (within admissible ones) corresponding to occurrence of  $e$  in  $S$ 
    - if does not exist, then this occurrence is discarded
    - if multiple admissible, duplicate on-going occurrences
- for each (not yet used) constraint  $(e, o_e)[t^-, t^+](e', o_{e'})$ , the admissible positions are updated:

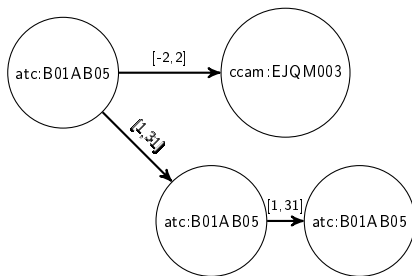
$$\text{adm}(e') = [a'^-, a'^+] := [a'^-, a'^+] \cap [t_e + t^-, t_e + t^+],$$

as soon as  $\text{adm}(e') = \emptyset$  discard the on-going occurrence!

# Use case: Deep Venous Thrombosis

Traitement par *enoxaparin* pour au moins 3 mois, *demarré* 2 jours avant ou après un echo-Doppler pour thrombose.

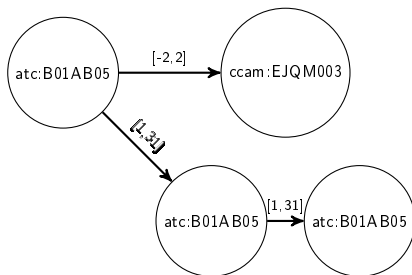
- *atc* : B01AB05: enoxaparin (anticoagulant injectable)
- *ccam* : EJQM003: echo Doppler



# Use case: Deep Venous Thrombosis

Traitement par *enoxaparin* pour au moins 3 mois, *demarré* 2 jours avant ou après un echo-Doppler pour thrombose.

- *atc* : B01AB05: enoxaparin (anticoagulant injectable)
- *ccam* : EJQM003: echo Doppler



## Limites

- *enoxaparin* is a type of anticoagulant ... *miss the others*
  - *echo Doppler* is not the only procedure to diagnose DVT ...
- ⇒ need for more expressiveness about the event types

# Alternative du web sémantique

## Semantic web

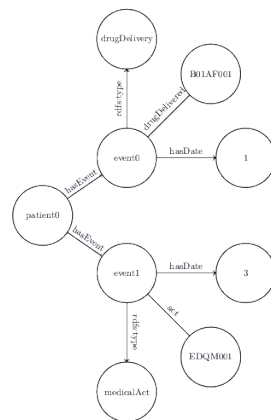
- Adapté pour représenter des données et raisonner sur des connaissances formalisées
- langage de requête expressif (SPARQL)
- Peu adapté pour manipuler des données et des requêtes temporelles (*nativement*) [ZWLC19]

## Proposition (J. Bakalara [BGD<sup>+</sup>19])

- Représentation des parcours de soins du SNDS en RDF
- “chronique” comme requête SPARQL
- Mécanisme de requêtage efficace pour l’identification des patients adhérents au phénotype

# Représentation RDF des parcours de soins du SNDS

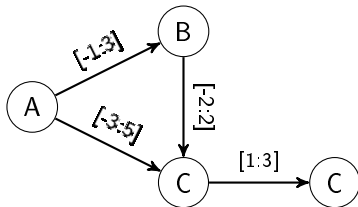
- Modèle de représentation des parcours de soins en RDF
  - Adaptation du travail de Y. Rivault [RDLM19]
  - Dérive de la notion de séquence d'événements
    - un type
    - une date (nombre entier)
  - Modélisation plus complète des notions du SNDS
- Outil de transformation des données SNDS en format RDF
  - Expérimentées sur données synthétiques
  - En cours sur données réelles





# Reconnaissance de chronique par une requête SPARQL

- Chronique : exprimable en SPARQL



```
SELECT DISTINCT * WHERE{
```

```
GRAPH ?pat {
```

```
  ?evt1 snds:has_time ?date1 .
```

```
  ?evt1 snds:deliver A .
```

```
  ?evt2 snds:has_time ?date2 .
```

```
  ?evt2 snds:deliver B .
```

```
  ?evt3 snds:has_time ?date3 .
```

```
  ?evt3 snds:deliver C .
```

```
  ?evt4 snds:has_time ?date4 .
```

```
  ?evt4 snds:deliver C .
```

```
}
```

```
FILTER ( ofn:asDays(?date2 - ?date1) >= -1)
```

```
FILTER ( ofn:asDays(?date2 - ?date1) <= 3)
```

```
FILTER ( ofn:asDays(?date3 - ?date1) >= -3)
```

```
FILTER ( ofn:asDays(?date3 - ?date1) <= 5)
```

```
FILTER ( ofn:asDays(?date3 - ?date2) >= -2)
```

```
FILTER ( ofn:asDays(?date3 - ?date2) <= 2)
```

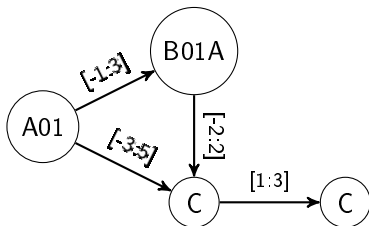
```
FILTER ( ofn:asDays(?date4 - ?date3) >= 1)
```

```
FILTER ( ofn:asDays(?date4 - ?date3) <= 3)
```

```
}
```

# Reconnaissance de chronique par une requête SPARQL

- extensible avec du raisonnement ontologique
- illustration avec relation `subClassOf*`
- plus d'expressivité possible pour définir des événements en exploitant le schéma des données
  - "acte d'un angiologue"
  - "délivrance d'un moins 3 boîtes d'un antithrombotique"



```
SELECT DISTINCT * WHERE{
```

```
GRAPH ?pat {
  ?evt1 snds:has_time ?date1 .
  ?evt1 snds:deliver ?atc1 .
  ?evt2 snds:has_time ?date2 .
  ?evt2 snds:deliver ?atc2 .
  ?evt3 snds:has_time ?date3 .
  ?evt3 snds:deliver ?atc3 .
  ?evt4 snds:has_time ?date4 .
  ?evt4 snds:deliver ?atc4 .
}
?atc1 rdfs:subClassOf* atc:A01 .
?atc2 rdfs:subClassOf* atc:B01A .
?atc3 rdfs:subClassOf* atc:C .
?atc4 rdfs:subClassOf* atc:C .
```

```

FILTER ( ofn:asDays(?date2 - ?date1) > 0 )
FILTER ( ofn:asDays(?date2 - ?date1) < 3 )
FILTER ( ofn:asDays(?date3 - ?date1) > 0 )
FILTER ( ofn:asDays(?date3 - ?date1) < 3 )
FILTER ( ofn:asDays(?date3 - ?date2) > 0 )
FILTER ( ofn:asDays(?date3 - ?date2) < 3 )
FILTER ( ofn:asDays(?date4 - ?date3) > 0 )
FILTER ( ofn:asDays(?date4 - ?date3) < 3 )
}

```



# Reconnaissance de chronique par une requête SPARQL

## Chronicle occurrence query

- Let  $\mathcal{C} = (\mathcal{E}, \mathcal{T})$  be a chronicle, the chronicle query is defined as follows:

$$q_{\mathcal{C}}(p) = \exists(t_1, \dots, t_n), \bigwedge_{e_i \in \mathcal{E}} q_{e_i}(t_i, p) \wedge \bigwedge_{(e_i, i)[l:u](e_j, j) \in \mathcal{T}} (t_j - t_i \leq u \wedge t_j - t_i \geq l)$$

- $\forall e \in \mathcal{E}$ ,  $q_e(t, p)$  est une requête où  $t$  est un instant et  $p$  une séquence (patient)

## Problème majeur

- Ces requêtes prennent beaucoup de temps
- Les contraintes temporelles sont peu efficacement traitées

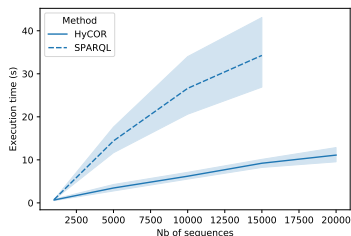
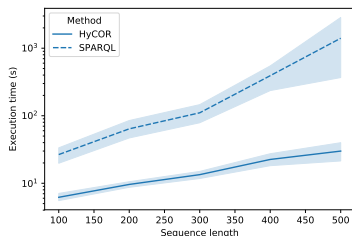
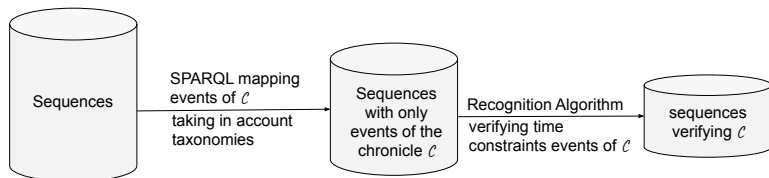
# Recherche efficace à base de chroniques sémantiques

## Rewriting of a chronicle query

$$\begin{aligned}
 q_C(p) = & \bigwedge_{e_i \in \mathcal{E}, t} S(t, p, i) \leftarrow q_{e_i}(t, p) \wedge \\
 & \exists (t_1, \dots, t_n), \bigwedge_i S(t_i, p, i) \wedge \\
 & \bigwedge_{(e_i, i)[l:u](e_j, j) \in \mathcal{T}} (t_j - t_i \leq u \wedge t_j - t_i \geq l)
 \end{aligned}$$

- $S$ : représentation intermédiaire : structure du séquence temporelle
- ligne 2 et 3: il s'agit d'une reconnaissance de chronique "classique"  
→ peut être traité par l'algorithme vu précédemment
- Technique inspirée des approches OMQA ("*Ontology Mediated Query Answering*" [Bie16, AKK<sup>+</sup>17])

# HyCor: reconnaissance efficace de chroniques “sémantiques”



**Figure:** Execution times (in seconds) of SPARQL and HyCOR wrt sequences length on 10 000 sequences (on the left) and wrt number of sequences (with length 100).

# HyCor: intérêt et limites

## Intérêt de HyCor

- Outil permettant d'exécuter efficacement des requêtes temporelles complexes
- Répond largement aux besoins rencontrés dans nos cas applicatifs
- Utilisation des modèles de chroniques très expressif

## Limites de HyCor

- étendre sémantiquement la nature des contraintes temporelles (opérateurs d'Allen)
- prise en compte de la négation

## Perspectives

- Comparaison de différents raisonneurs (incluant Neo4j)
- Construction d'un éditeur graphique de chroniques sémantiques

# À vous de jouer

## Exploration de données médico-administratives

- Consulter le notebook `requetes_graphdb.ipynb`
- Objectifs
  - mettre en place une base de données avec graphdb
  - comprendre un modèle de trajectoire de soins en web sémantique
  - tester des requêtes sur les données (parcours de soins)
  - tester l'écriture de chroniques

# Conclusion : requêtes temporelles sémantiques pour l'épidémiologie

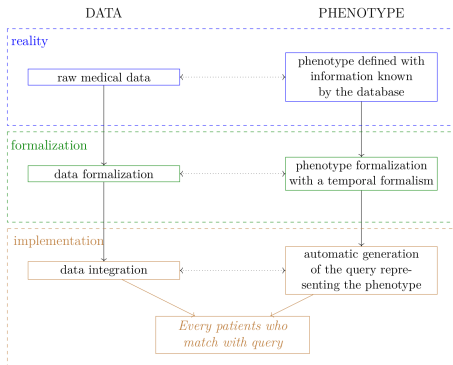
## SNDS/données hospitalières pour l'épidémiologie

- source de données intéressante pour certaines études
- données complexes et usage secondaire => fossé sémantique
- préparation des données déterminantes pour les résultats d'apprentissage automatique

## Modéliser et raisonner sur les parcours de soins pour combler (en partie) le fossé sémantique

- Parcours de soin : la dimension temporelle pour améliorer la précision des requêtes
- Besoin d'outils pour faciliter la création de requêtes
  - Flexibles et expressives : pour exploiter au mieux les données
  - Réutilisables et automatisables : pour focaliser les épidémiologistes sur les questions difficiles

# Cadre général de l'approche



## Notre question : choix de formalisation

- Chronique: efficace, pas de raisonnement ontologique
- WS: raisonnement sémantique, contraintes temporelles peu efficaces
- HyCor: meilleur des deux mondes

⇒ D'autres pistes potentielles

# Outline

- 1 Parcours de soins
  - Motivation
  - Représentation du parcours de soins
  - Les tâches
- 2 Représentation des parcours de soins
  - Représentation de l'information temporelle
  - Temps avec logiques modales
  - Ontologies about time
  - Conclusions
- 3 Requêtage de données temporelles
  - Contexte
  - Querying with temporal pattern matching
  - Querying with reasoning on care pathways: a web semantic approach
  - Conclusion
- 4 Conclusion



# Conclusion générale: THM 1

## Sur des trajectoires ou parcours individuels

- Représentation/visualisation
- Enrichir une trajectoire (inference): e.g. abstraction temporelle, reconstruction d'évènements
- Vérification de la consistance d'un parcours

## Confrontation de trajectoires avec un parcours

- Adhérence d'une trajectoire à un parcours (recherche)
- Conformance d'une trajectoire à un parcours (guidelines)

## Sur des ensembles de trajectoires

- Représentation/visualisation
- Clustering de trajectoires
- Généralisation d'un ensemble de trajectoires en parcours
- Préviation/classification de trajectoires

# Conclusion générale: THM 2

## Intérêt des ontologies: aider à combler les écarts sémantiques

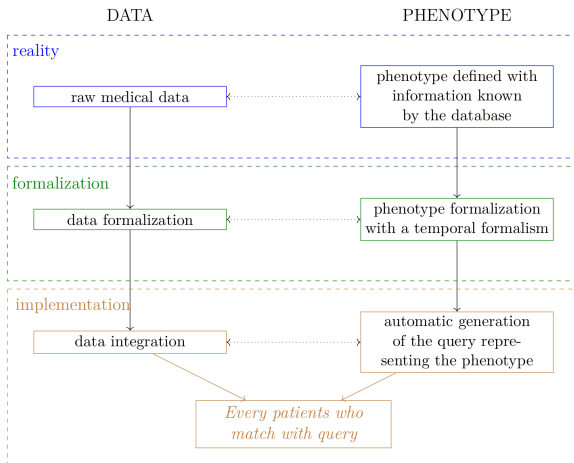
- proposer une représentation des données plus intelligible
- structurer une information
- permettre de représenter des parcours

## Représentation et raisonnement temporel

- Idéalement expressif et permettant des raisonnements riches
- Représentation de l'information temporelle :
  - OWL-Time: beaucoup de limites, focalisation Allen un peu secondaire
  - alternatives ?
- Raisonnement temporel
  - raisonneurs DL souvent limités (e.g. OWL-Time)
  - solution : combiner raisonneurs temporelles et raisonneurs ontologiques !

# Conclusion générale: THM 3

→ Interrogation de données par des "Parcours d'intérêt": utile pour l'exploration des données temporelles



# Questions ?

Contributeurs: J. Bakalara, N. P. Sawadogo, O. Dameron, A. Happe. E. Oger, E. Audureau, N. Markey.

Intéressé.e.s par cette présentation ?

⇒ Poste ingénieur junior sur  
modélisation ontologique des parcours  
de soins ouvert en Novembre (Lyon,  
Fr)!

# References |



Rajeev Alur and Thomas A Henzinger.

Real-time logics: complexity and expressiveness.

[Information and Computation](#), 104(1):35–77, 1993.



Alessandro Artale, Roman Kontchakov, Alisa Kovtunova, Vladislav Ryzhikov, Frank Wolter, and Michael Zakharyashev.

Ontology-mediated query answering over temporal data: A survey.  
2017.



James F Allen.

Maintaining knowledge about temporal intervals.

[Communications of the ACM](#), 26(11):832–843, 1983.



Patricia Bouyer, Fabrice Chevalier, and Nicolas Markey.

On the expressiveness of tptl and mtl.

[Information and Computation](#), 208(2):97–116, 2010.



Johanne Bakalars, Thomas Guyet, Olivier Dameron, Emmanuel Oger, and André Happe.

Temporal models of care sequences for the exploration of medico-administrative data.

In [Workshop IA&Santé, PFIA](#), 2019.



Meghyn Bienvenu.

Ontology-mediated query answering: Harnessing knowledge to get more from data.

In [IJCAI: International Joint Conference on Artificial Intelligence](#), 2016.



Marcello M Bersani, Matteo Rossi, and Pierluigi San Pietro.

A tool for deciding the satisfiability of continuous-time metric temporal logic.

[Acta Informatica](#), 53(2):171–206, 2016.

# References II



Olivier Dameron, Paolo Besana, Oussama Zekri, Annabel Bourdé, Anita Burgun, and Marc Cuggia.

**Owl model of clinical trial eligibility criteria compatible with partially-known information.**

[Journal of Biomedical Semantics](#), 4:1–10, 2013.



Christophe Dousson, Paul Gaborit, and Malik Ghallab.

**Situation recognition: Representation and algorithms.**

In Ruzena Bajcsy, editor, [Proc. of the 13th Int. Joint Conf. on Artificial Intelligence \(IJCAI 1993\)](#), pages 166–172. Morgan Kaufmann, August 28 - September 3, 1993.



Rina Dechter, Itay Meiri, and Judea Pearl.

**Temporal constraint networks.**

[Artificial intelligence](#), 49(1-3):61–95, 1991.



Thomas Guyet, Philippe Besnard, Ahmed Samet, Nasreddine Ben Salha, and Nicolas Lachiche.

**Énumération des occurrences d'une chronique.**

In [Actes de la conférence Extraction et Gestion des Connaissances \(EGC\)](#), pages 253–260, 2020.



Simon Gottschalk and Elena Demidova.

**Eventkg: A multilingual event-centric temporal knowledge graph.**

In [The Semantic Web: 15th International Conference, ESWC 2018, Heraklion, Crete, Greece, June 3–7, 2018, Proceedings 15](#), pages 272–287. Springer, 2018.



Thomas Guyet and Nicolas Markey.

**Logical forms of chronicles.**

In [TIME 2022-29th International Symposium on Temporal Representation and Reasoning](#), pages 1–15, 2022.

## References III



Jerry R Hobbs and Feng Pan.

An ontology of time for the semantic web.

[ACM Transactions on Asian Language Information Processing \(TALIP\)](#), 3(1):66–85, 2004.



George Hripcsak, Li Zhou, Simon Parsons, Amar K. Das, and Stephen B. Johnson.

Modeling Electronic Discharge Summaries as a Simple Temporal Constraint Satisfaction Problem.

[Journal of the American Medical Informatics Association](#), 12(1):55–63, 2005.



R Kowalski and M Sergot.

A logic-based calculus of events. foundations of knowledge base management, 1989.



Fang Li, Jingcheng Du, Yongqun He, Hsing-Yi Song, Mohdine Madkour, Guozheng Rao, Yang Xiang,

Yi Luo, Henry W Chen, Sijia Liu, Liwei Wang, Hongfang Liu, Hua Xu, and Cui Tao.

Time event ontology (TEO): to support semantic representation and reasoning of complex temporal relations of clinical events.

[Journal of the American Medical Informatics Association](#), 27(7):1046–1056, 2020.



Zhihui Luo, Stephen B Johnson, Albert M Lai, and Chunhua Weng.

Extracting temporal constraints from clinical research eligibility criteria using conditional random fields.

In [AMIA annual symposium proceedings](#), volume 2011, page 843. American Medical Informatics Association, 2011.



Drew McDermott.

A temporal logic for reasoning about processes and plans.

[Cognitive science](#), 6(2):101–155, 1982.



Itay Meiri.

Combining qualitative and quantitative constraints in temporal reasoning.

[Artificial Intelligence](#), 87(1-2):343–385, 1996.



# References IV



**Erik T Mueller.**

**Event calculus reasoning through satisfiability.**  
[\*Journal of Logic and Computation\*, 14\(5\):703–730, 2004.](#)



**Amy L. Olex and Bridget T. McInnes.**

**Review of temporal reasoning in the clinical domain for timeline extraction: Where we are and where we need to be.**  
[\*Journal of Biomedical Informatics\*, 118:103784, 2021.](#)



**Alexandros Preventis, Euripides GM Petrakis, and Sotirios Batsakis.**

**Chronos ed: A tool for handling temporal ontologies in protege.**  
[\*International Journal on Artificial Intelligence Tools\*, 23\(04\):1460018, 2014.](#)



**Yann Rivault, Olivier Dameron, and Nolwenn Le Meur.**

**queryMed: Semantic web functions for linking pharmacological and medical knowledge to data.**  
[\*Bioinformatics\*, 2019.](#)



**Murray Shanahan.**

**The event calculus explained.**  
 In [\*Artificial intelligence today: Recent trends and developments\*](#), pages 409–430. Springer, 2001.



**Cui Tao, Wei-Qi Wei, Harold R Solbrig, Guergana Savova, and Christopher G Chute.**

**Cntro: a semantic web ontology for temporal relation inferencing in clinical narratives.**  
 In [\*AMIA annual symposium proceedings\*](#), volume 2010, page 787. American Medical Informatics Association, 2010.



**Marc B Vilain and Henry A Kautz.**

**Constraint propagation algorithms for temporal reasoning.**  
 In [\*Aai\*](#), volume 86, pages 377–382, 1986.

# References V



Fu Zhang, Ke Wang, Zhiyin Li, and Jingwei Cheng.

**Temporal data representation and querying based on RDF.**

[IEEE Access](#), 7:85000–85023, 2019.

# Chronicle occurrence enumeration algorithm

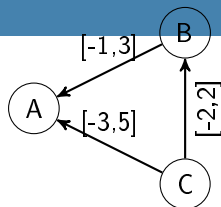
## Algorithm “idea”

order event types by increasing frequency in sequence  $S$   
 for all occurrence of  $e_0 \in \mathcal{E}$

- ➊ create a on-going occurrence  $o = ([t_0, t_0], [-\infty, \infty], [-\infty, \infty], \dots)$
- ➋ for events  $e_i \in [e_1, \dots, e_n]$ 
  - ➊ for each on-going occurrence  $o$ 
    - ➊ look for event  $e_i$  within interval  $o_i = [t_i^-, t_i^+]$
    - ➋ duplicate  $o_i$  for each of the admissible occurrences of  $e_i$
    - ➌ propagate constraints

# Algorithm example

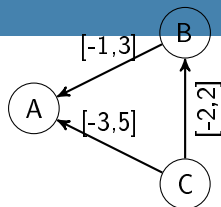
$\langle (E, 1) (B, 2) (A, 3) (E, 4) (C, 5)$   
 $(B, 6) (A, 7) (E, 8) (A, 9) (A, 10) \rangle$



- Event processing order :  $C$ ,  $B$  and  $A$
- Processing of  $C$ 
  - generate a single occurrence  $o = ([5, 5], [-\infty, \infty], [-\infty, \infty])$

# Algorithm example

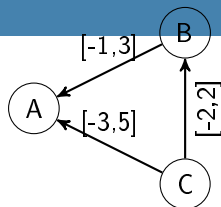
$\langle (E, 1) (B, 2) (A, 3) (E, 4) (C, 5)$   
 $(B, 6) (A, 7) (E, 8) (A, 9) (A, 10) \rangle$



- Event processing order :  $C$ ,  $B$  and  $A$
- Processing of  $C$ 
  - generate a single occurrence  $o = ([5, 5], [-\infty, \infty], [-\infty, \infty])$
  - propagate constraints:
    - $C \rightarrow B$ :  $o = ([5, 5], [3, 7], [-\infty, \infty])$
    - $C \rightarrow A$ :  $o = ([5, 5], [3, 7], [2, 10])$

# Algorithm example

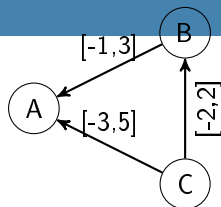
$\langle (E, 1) (B, 2) (A, 3) (E, 4) (C, 5)$   
 $(B, 6) (A, 7) (E, 8) (A, 9) (A, 10) \rangle$



- Event processing order :  $C$ ,  $B$  and  $A$
- Processing of  $C$ 
  - generate a single occurrence  $o = ([5, 5], [-\infty, \infty], [-\infty, \infty])$
  - propagate constraints:
    - $C \rightarrow B$ :  $o = ([5, 5], [3, 7], [-\infty, \infty])$
    - $C \rightarrow A$ :  $o = ([5, 5], [3, 7], [2, 10])$
- Processing of  $B$ 
  - narrow intervals with occurrences:  $(B, 2)$  is invalid ( $2 \notin [3, 7]$ ), so  $o = ([5, 5], [6, 6], [2, 10])$
  - propagate constraints:
    - $B \rightarrow A$ :  $o = ([5, 5], [6, 6], [2, 10] \cap [5, 9]) = ([5, 5], [6, 6], [5, 9])$

# Algorithm example

$\langle (E, 1) (B, 2) (A, 3) (E, 4) (C, 5)$   
 $(B, 6) (A, 7) (E, 8) (A, 9) (A, 10) \rangle$



- Event processing order :  $C$ ,  $B$  and  $A$
- Processing of  $C$ 
  - generate a single occurrence  $o = ([5, 5], [-\infty, \infty], [-\infty, \infty])$
  - propagate constraints:
    - $C \rightarrow B$ :  $o = ([5, 5], [3, 7], [-\infty, \infty])$
    - $C \rightarrow A$ :  $o = ([5, 5], [3, 7], [2, 10])$
- Processing of  $B$ 
  - narrow intervals with occurrences:  $(B, 2)$  is invalid ( $2 \notin [3, 7]$ ), so  $o = ([5, 5], [6, 6], [2, 10])$
  - propagate constraints:
    - $B \rightarrow A$ :  $o = ([5, 5], [6, 6], [2, 10] \cap [5, 9]) = ([5, 5], [6, 6], [5, 9])$
- Processing of  $A$ 
  - narrow intervals with occurrences:  $(A, 3)$  and  $(A, 10)$  are invalid, but  $(A, 7)$  and  $(A, 9)$  are valid, two occurrences
    - $([5, 5], [6, 6], [7, 7])$  and  $([5, 5], [6, 6], [9, 9])$

# (Rappel de) notations de logique du premier order

## Exemple de formule de premier ordre

$$\varphi : \forall x \exists y, \text{amis}(x, y) \wedge \text{amis}(y, \text{Harry})$$

- $x, y$  et  $z$  sont des **variables**
- *amis* est un **prédicat** d'arité 2 (une relation)
- Harry est une **constante**
- $\forall$  et  $\exists$  sont des **quantificateurs**
- $\wedge, \vee, \implies, \neg$  sont des opérateurs logiques

## Interprétation

- Modèle: ensemble de domaines pour les variables d'une formule
- Une formule est **satisfaisable** s'il existe un modèle  $\mathcal{M}$  qui la rend vraie
- Notation usuelle

$$\mathcal{M} \models \varphi$$



# Logique de description: une logique du premier ordre spécifique

## Logique de description

- logique pour représenter et raisonner sur des connaissances
- fondement pour les outils du web sémantique tels que protégé, SPARQL, etc.
- utilisation de la syntaxe de la logique du premier ordre

## Structuration en deux ensembles de formules

- ABox (*assertion box*): description de l'ensemble des faits connus du *monde*
- TBox (*terminological box*): description des règles qui régissent la description du *monde*

# Logique de description: ABox

La ABox contient des faits

Un **fait** est une formule sans variables (uniquement des constantes) : ce sont des connaissances **vraies**.

- Prédicats unaires: propriétés d'individus

*Personne*(Thomas)

*Personne*(Harry)

*Chat*(Minou)

- Thomas est un individu qui a la propriété d'être une personne

- Prédicats binaires : relation entre individus

*amis*(Thomas, Harry)

*eduque*(Harry, Minou)

- Harry eduque Minou

Représentation  
en graphe

→ On se limite à des prédicats binaires

# Logique de description: ABox

La ABox contient des faits

Un **fait** est une formule sans variables (uniquement des constantes) : ce sont des connaissances **vraies**.

- Prédicats unaires: propriétés d'individus  
*Personne*(Thomas)  
*Personne*(Harry)  
*Chat*(Minou)
  - Thomas est un individu qui a la propriété d'être une personne
- Prédicats binaires : relation entre individus  
*amis*(Thomas, Harry)  
*eduque*(Harry, Minou)
  - Harry eduque Minou

Représentation  
en graphe

→ La ABox utilise les concepts/rôles définis dans la TBox

# Logique de description: TBox

La TBox contient des *règles* de description du monde

- Notion de *concept*: à propos des prédicats unaires
- Notion de *rôle*: à propos des prédicats binaires

Concept : dénote un ensemble d'individus

- Concepts atomiques: *Masculin*, *Feminine*, *Parent*, *Personne*, *Chat*
- Concepts complexes : définis en terme d'autres concepts

$Femme \equiv Personne \sqcap Feminine$  :  $\forall x Femme(x) \leftrightarrow Personne(x) \wedge Feminine(x)$

$Personne \subset EtreVivant$  :  $\forall x, Personne(x) \rightarrow EtreVivant(x)$

- operateurs boolean :  $\neg R, R \sqcap R', R \sqcup R'$
- autres operateurs : composition, inverse, etc.

Rôle : dénote un ensemble de couples d'individus

# Logique de description: TBox

La TBox contient des *règles* de description du monde

- Notion de *concept*: à propos des prédicats unaires
- Notion de *rôle*: à propos des prédicats binaires

Concept : dénote un ensemble d'individus

Rôle : dénote un ensemble de couples d'individus

- Rôles atomiques: *mereDe*
- Rôles complexes:  $grandMere \equiv mereDe \circ mereDe$ 
  - opérateurs boolean :  $\neg R, R \sqcap R', R \sqcup R'$
  - autres opérateurs de l'algèbre des relations : composition, itération (Kleene) inverse, etc.

# Logique de description: TBox

La TBox contient des *règles* de description du monde

- Notion de *concept*: à propos des prédicats unaires
- Notion de *rôle*: à propos des prédicats binaires

Concept : dénote un ensemble d'individus

Rôle : dénote un ensemble de couples d'individus

La logique de description ne parle que de concepts et de rôles

- seule des raisonnements sur les concepts et les rôles peuvent être faits
- les outils du web sémantique sont faits pour manipuler (efficacement) ces notions de concepts et rôles
- que se passe-t-il lorsqu'on veut ajouter de nouvelles "règles" de fonctionnement du monde dans le web sémantique

# SWRL: Semantic Web Rule Language

- Le SWRL est un langage qui permet d'exprimer des règles "personnalisées"
  - Une règle est une formule de la forme :

$$A_1(x) \wedge A_2(y) \wedge \dots \wedge R(x, y) \wedge \dots \implies C(x) \wedge \dots R'(x, y)$$

- Des moteurs d'inférence de règle permettent d'exploiter ces règles pour inférer de nouveaux faits
  - Pellet
  - Hermitt
  - Fact++
  - ...

- L'utilisation de SWRL permet d'enrichir (dans une certaine mesure) les capacités de raisonnement de la logique de description en ajouter de nouvelles *théories*.