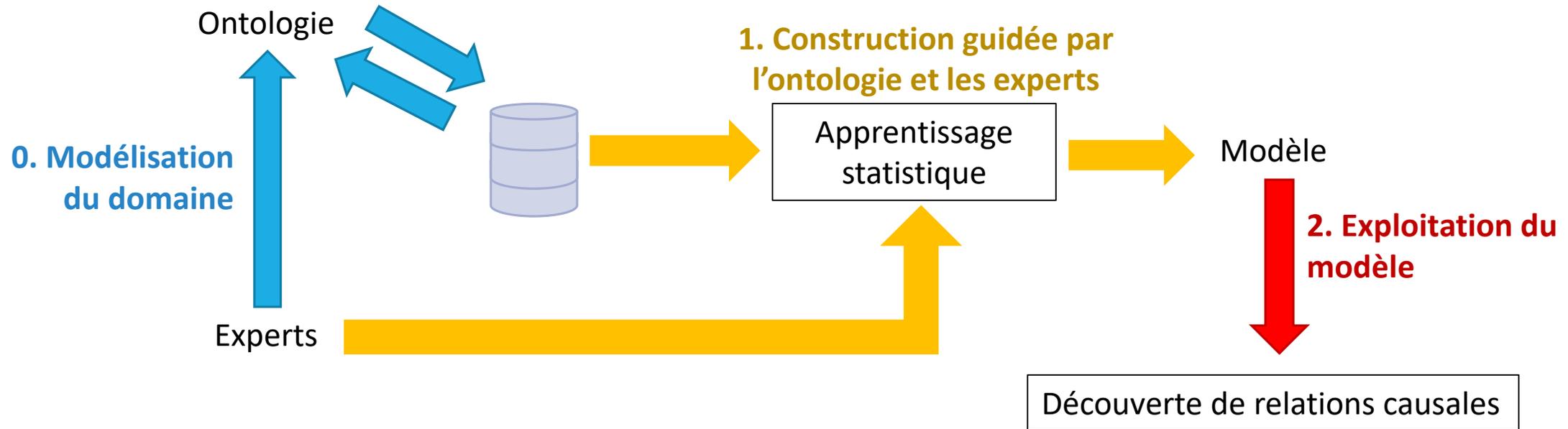


# Towards Interactive Causal Relation Discovery Driven by an Ontology

---

Melanie MUNCH, Juliette DIBIE, Pierre-Henri WUILLEMIN Cristina MANFREDOTTI

# Découverte causale dans des données hétérogènes



→ L'ontologie peut-elle aider à la découverte de relations causales?

# Plan

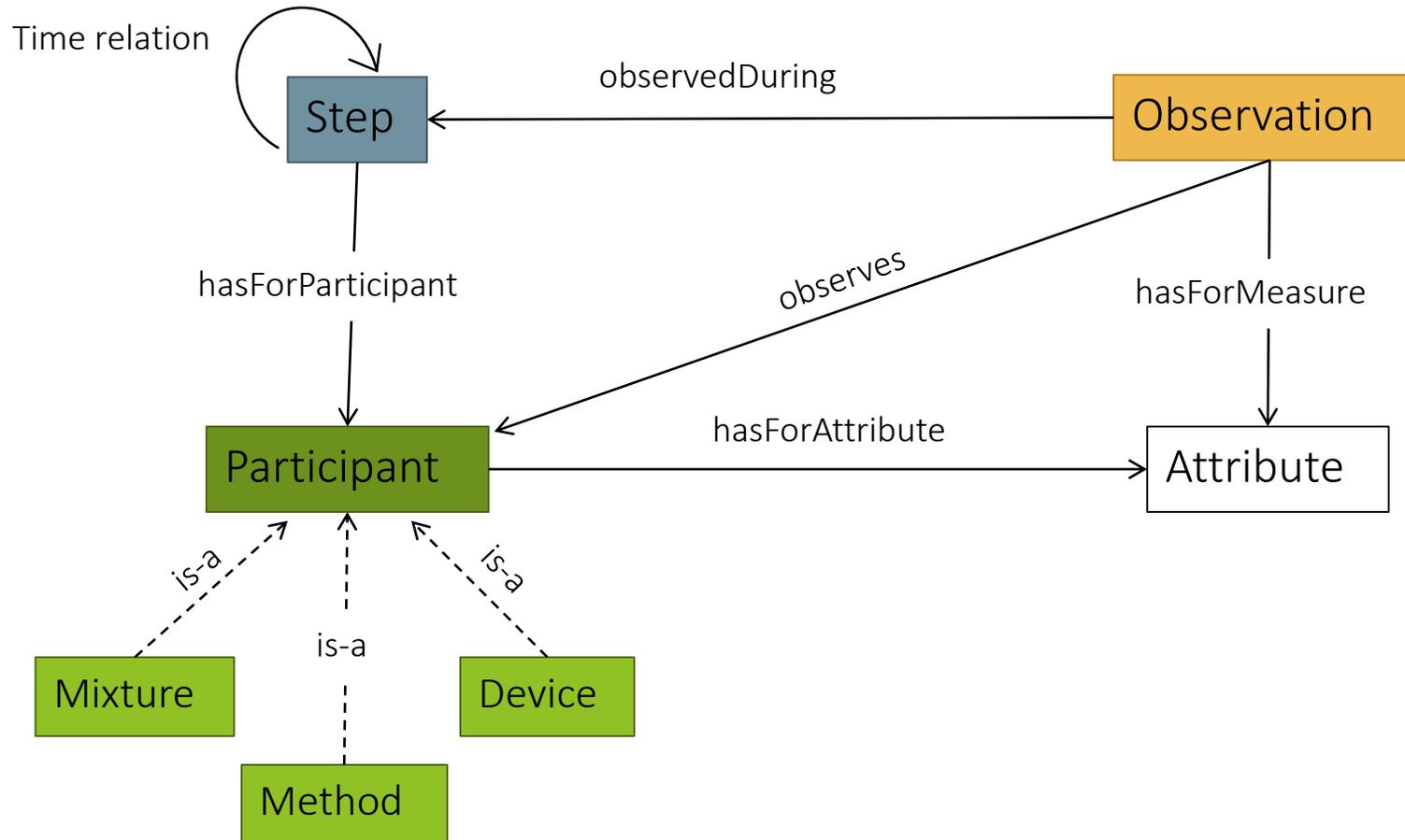
---

→ On souhaite **modéliser** les relations entre attributs dans un **domaine complexe** avec **des données hétérogènes** et de la **temporalité**

- 1.** Ontologies and PRMs
- 2.** Présentation de notre modèle
- 3.** Application et résultats

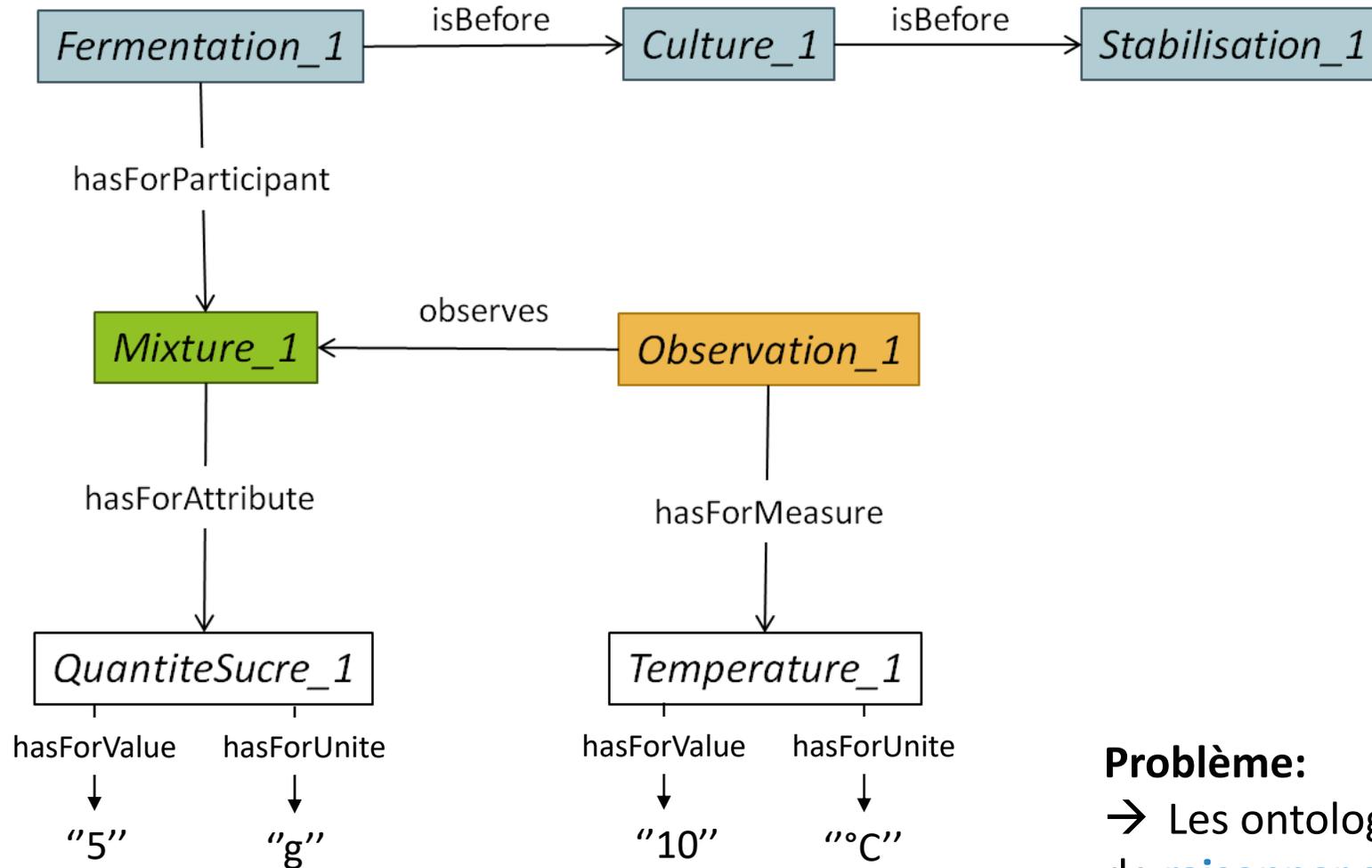
# Process and Observation Ontology PO<sup>2</sup>

**Ontologie:** Structure de représentation des connaissances dans un domaine donné



# Process and Observation Ontology PO<sup>2</sup>

Exemple d'un procédé particulier instancié

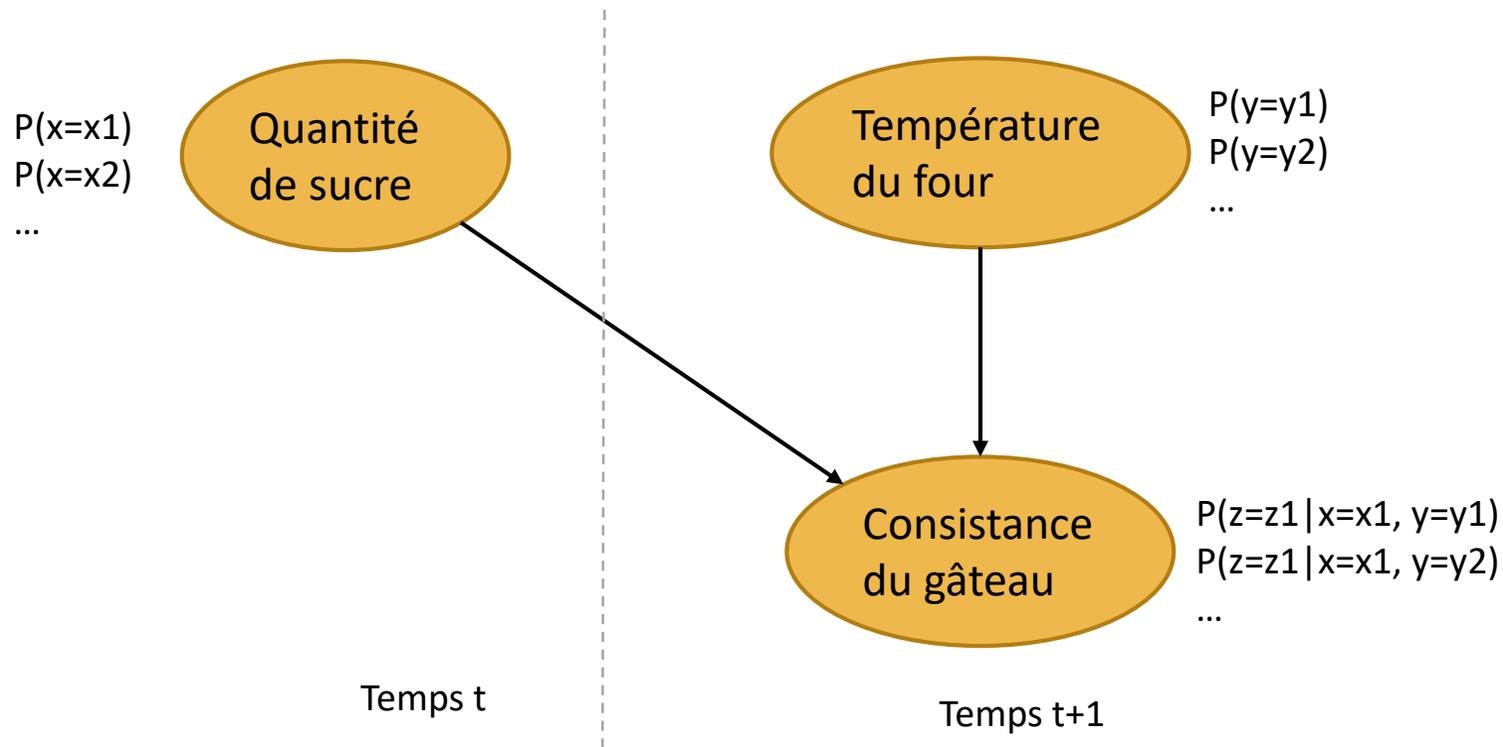


**Problème:**

→ Les ontologies ne permettent pas de raisonner sur l'incertain probabilisé.

# Raisonner dans l'incertain probabilisé (1)

**Réseau Bayésien (BN):** Graphe acyclique orienté représentant un espace probabilisé fini sur un ensemble de variables aléatoires

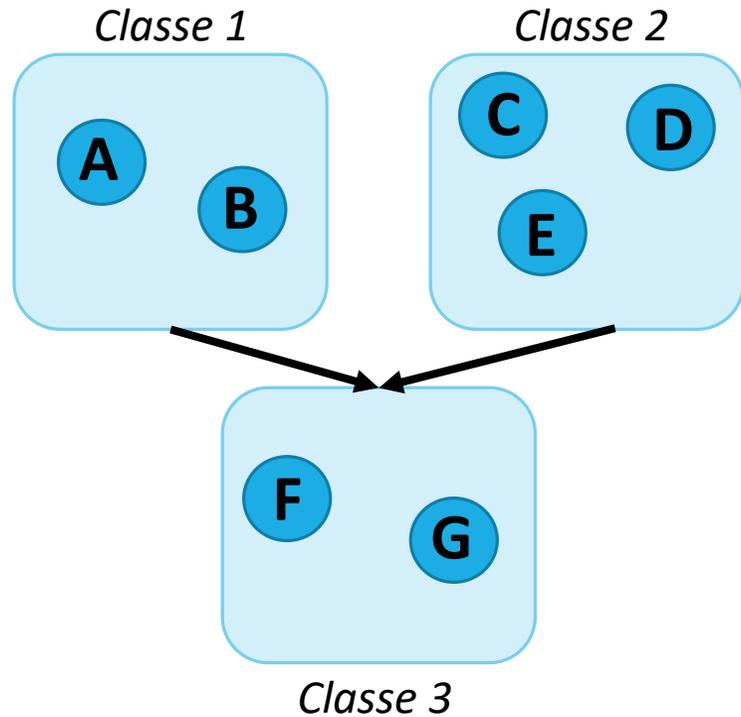


**Qualitatif:** « Est-ce que la quantité de sucre a un impact sur la température du four? »

**Quantitatif:** « Quel est l'impact de la la température du four sur la consistance du gâteau? »

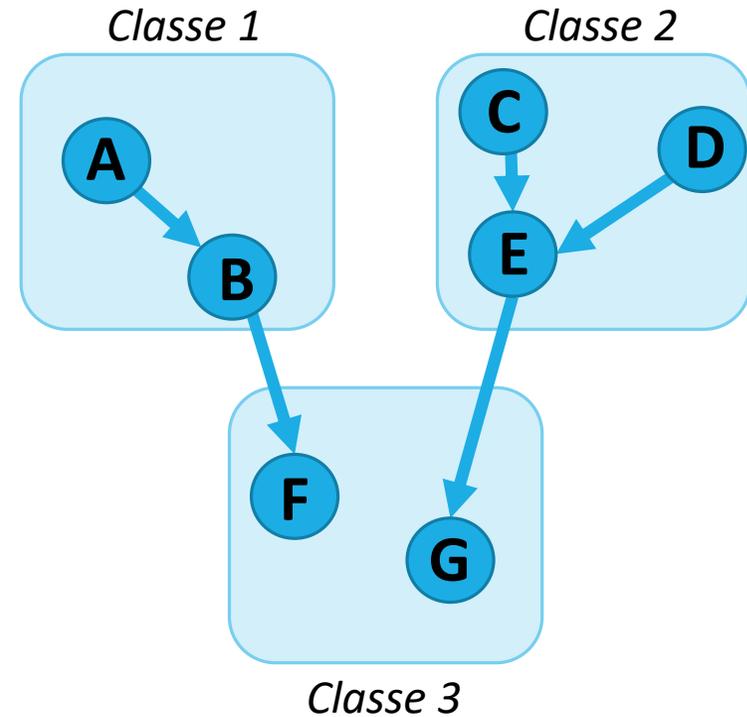
# Probabilistic Relational Models (PRMs)

**PRM:** Extension orientée objet des réseaux bayésiens (BN) définie en deux niveaux



**1. Schéma relationnel** → Construit

*Donne les relations entre classes*

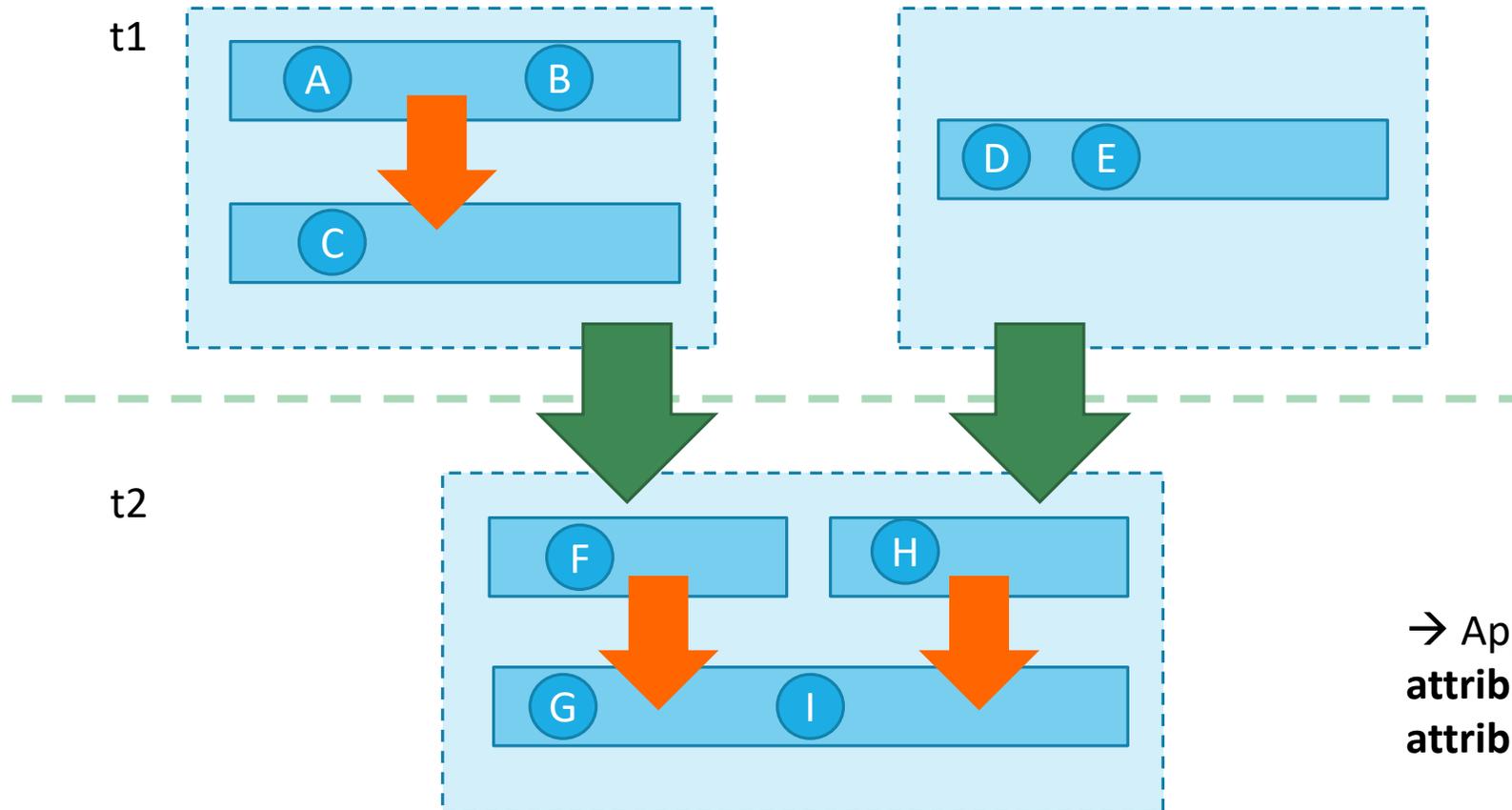


**2. Modèle relationnel** → Appris

*Donne les relations entre attributs*

# Définition du Stack Model: un schéma relationnel générique

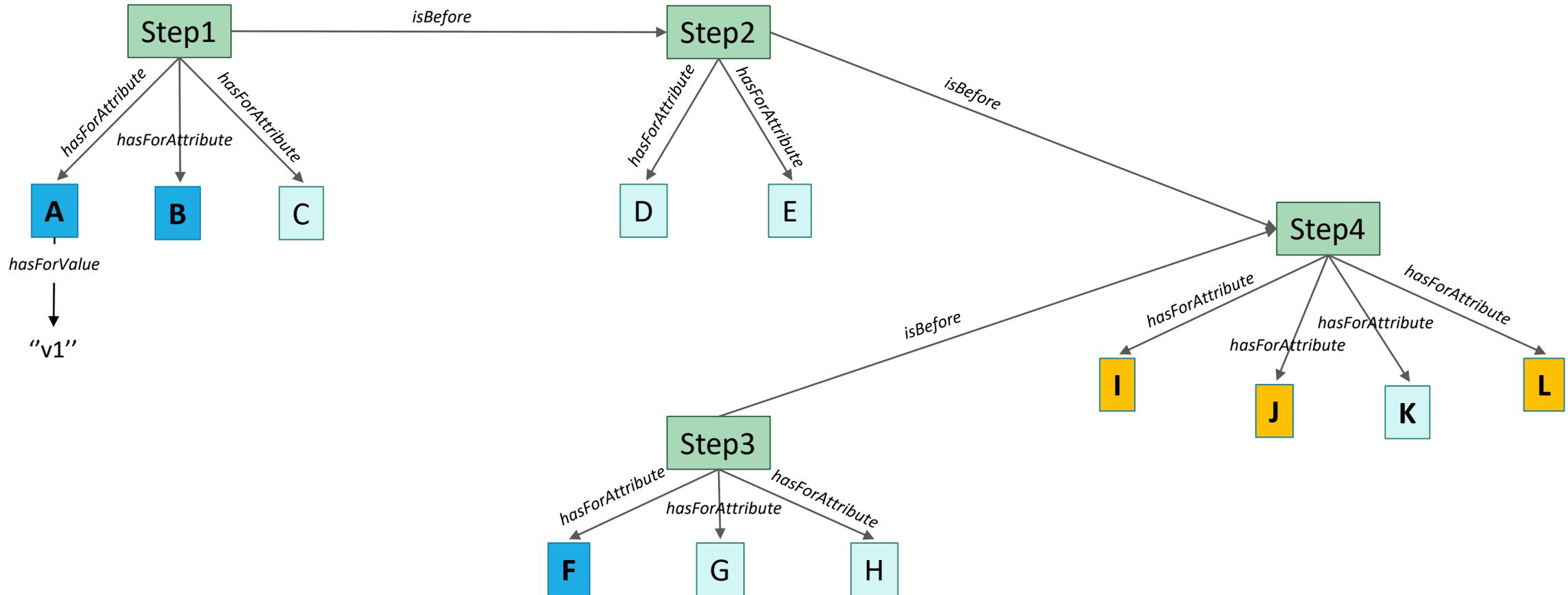
Défini par **deux contraintes**: **contraintes explicites**, données par l'ontologie  
**contraintes implicites**, données par l'expert



→ Apprentissage des **attributs expliquant** vers les **attributs conséquences**

# Exemple

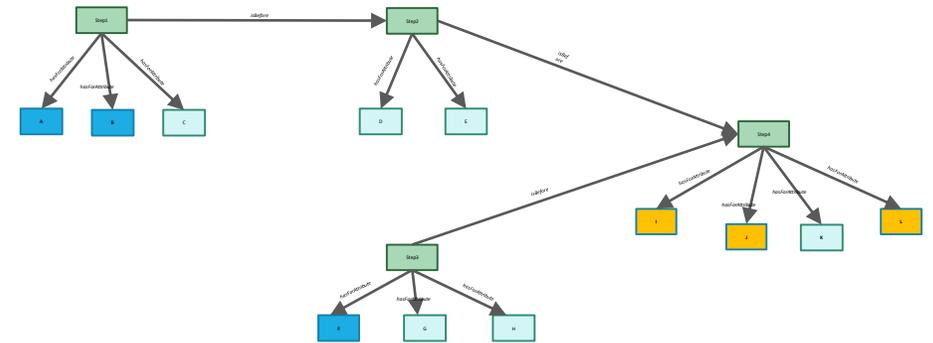
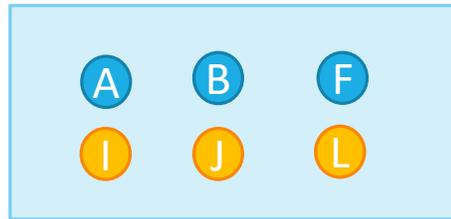
Question: Est-ce que **A**, **B** et **F** expliquent **I**, **J** et **L**?



# Exemple

**Question:** Est-ce que **A**, **B** et **F** sont des **paramètres de contrôle** pour **I**, **J** et **L**?

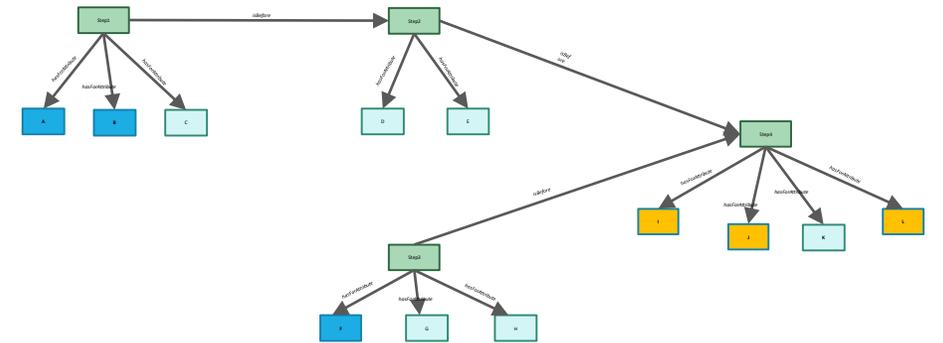
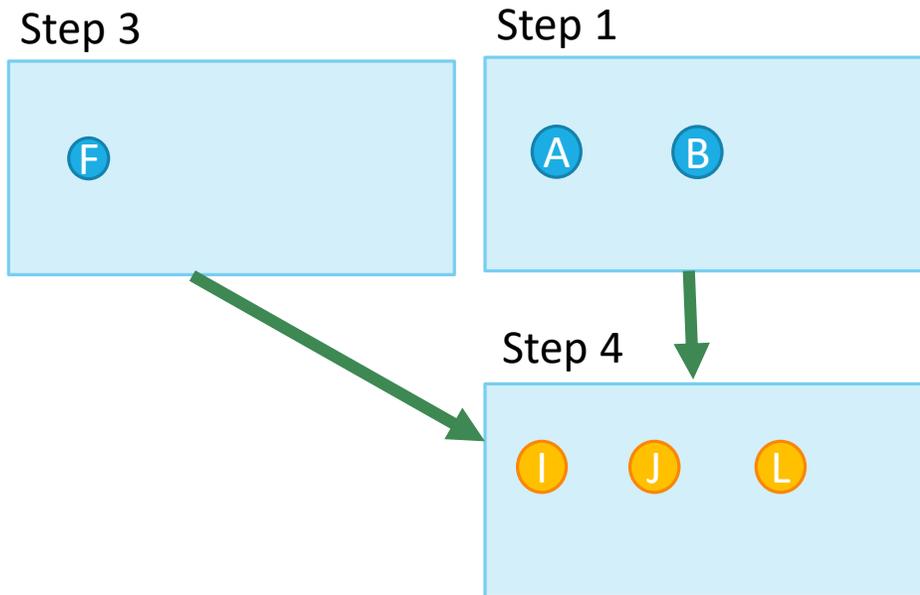
## 1. Sélection des DP intéressants d'après la question et l'ontologie



# Exemple

**Question:** Est-ce que **A**, **B** et **F** sont des **paramètres de contrôle** pour **I**, **J** et **L**?

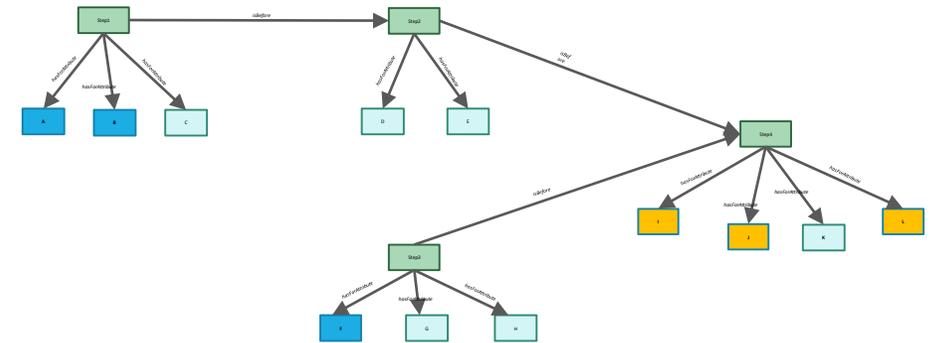
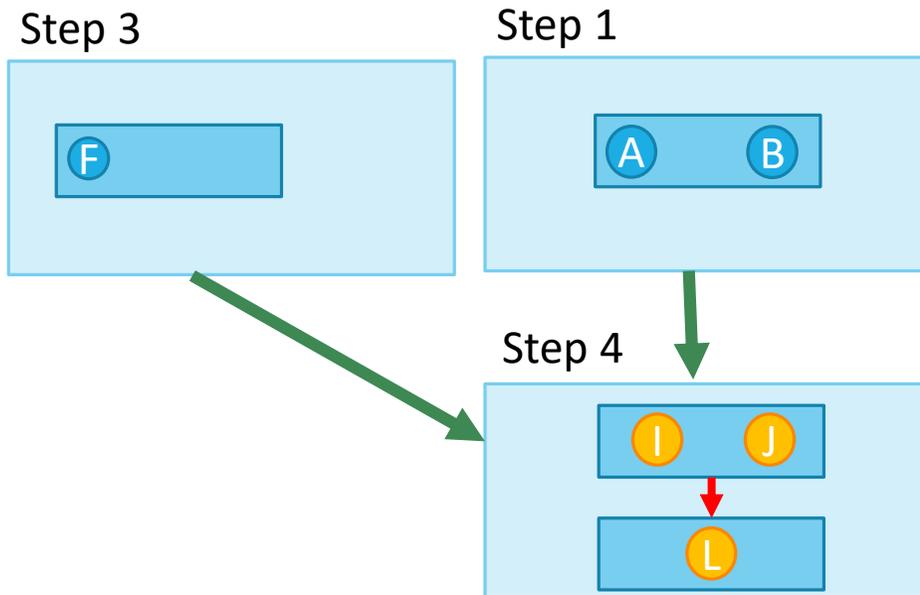
## 2. Séparation en utilisant les **contraintes explicites** de l'ontologie



# Exemple

**Question:** Est-ce que **A**, **B** et **F** sont des **paramètres de contrôle** pour **I**, **J** et **L**?

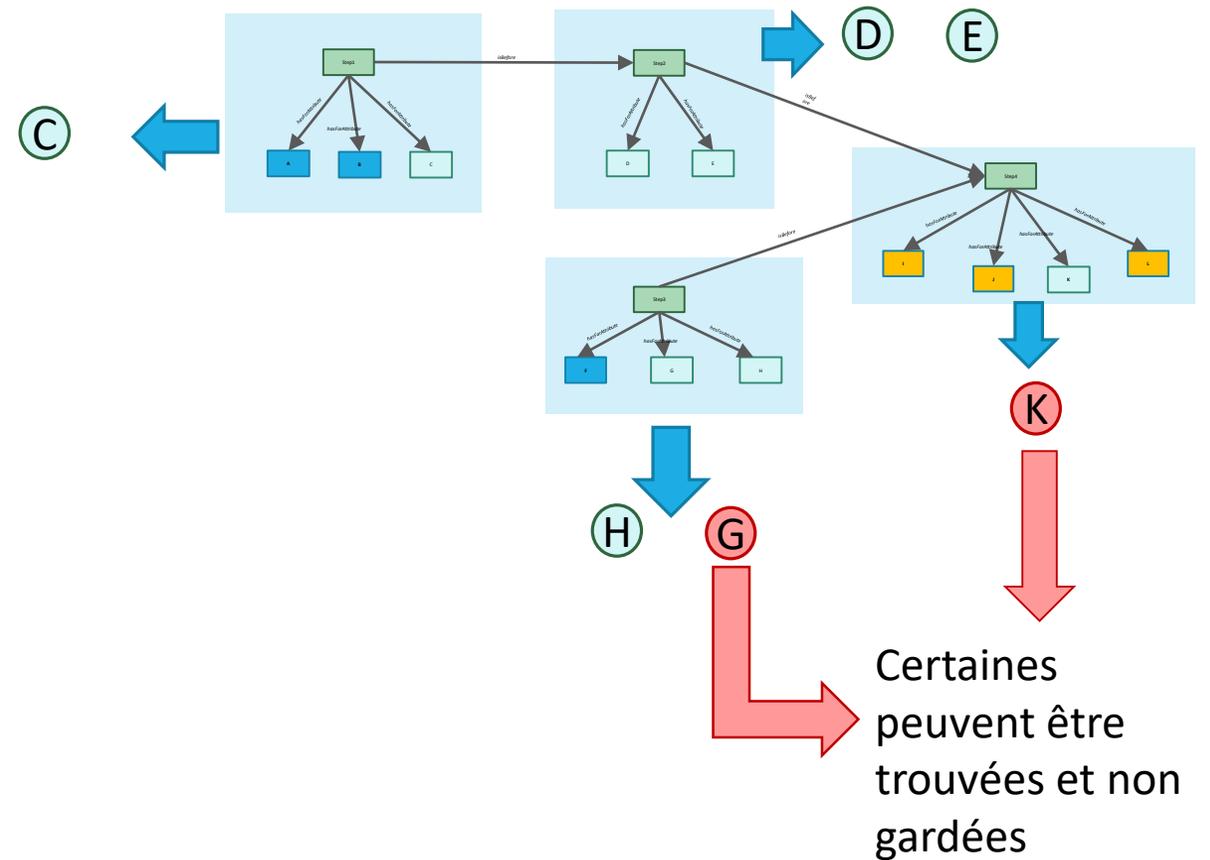
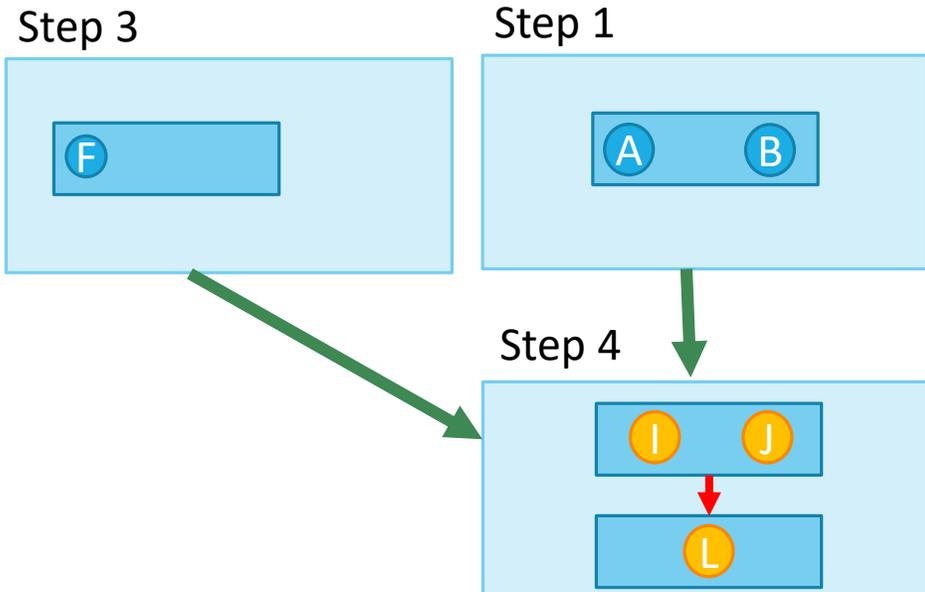
## 3. Séparation en utilisant les **contraintes implicites** de l'expert



# Exemple

**Question:** Est-ce que **A**, **B** et **F** sont des **paramètres de contrôle** pour **I**, **J** et **L**?

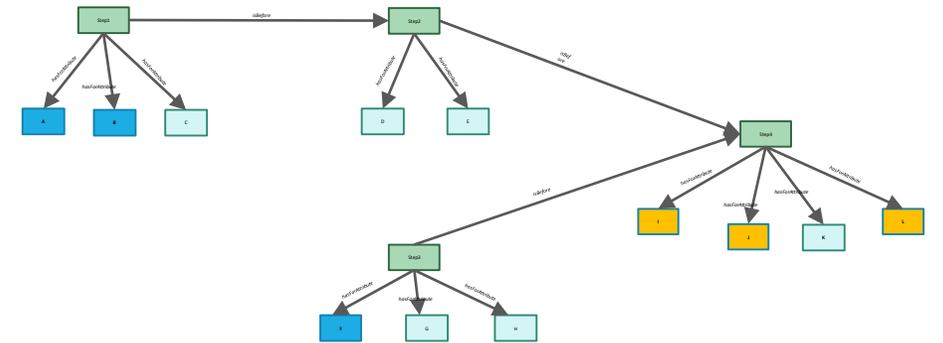
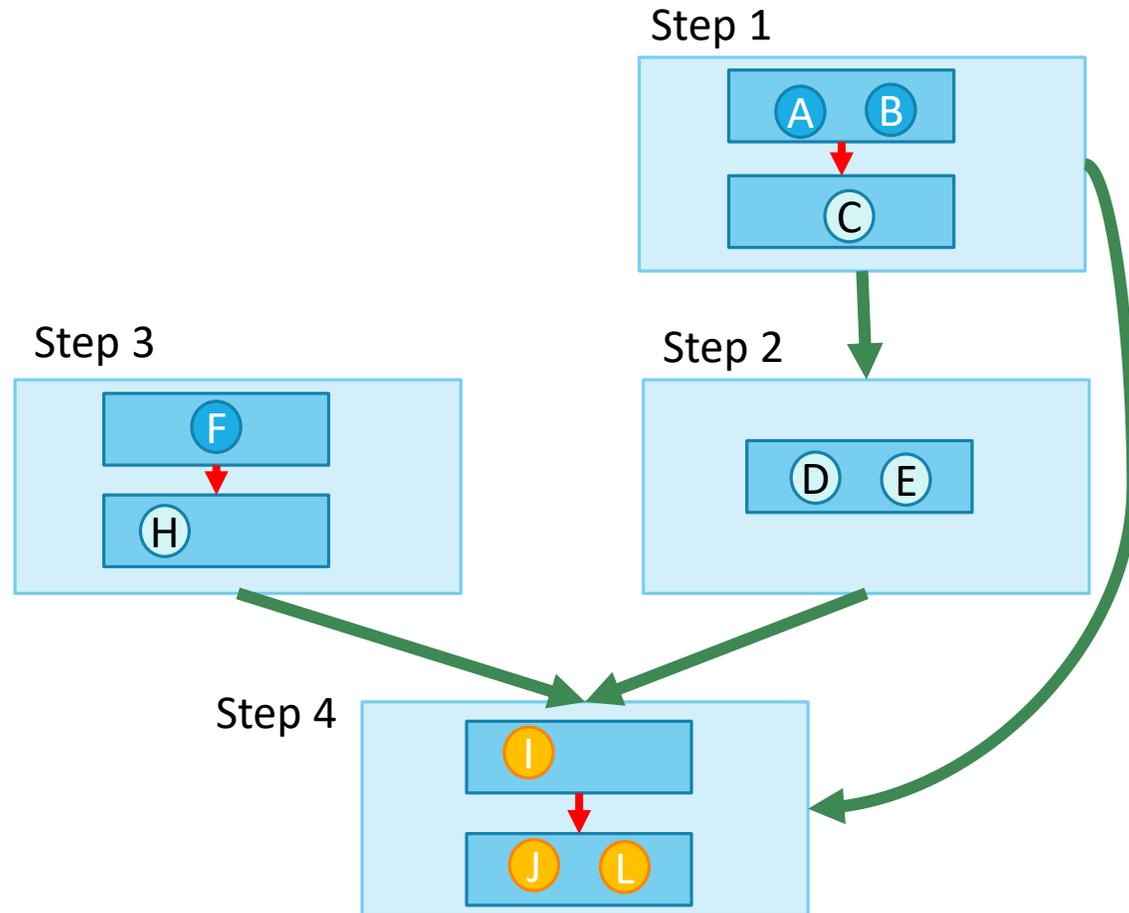
## 4. Enrichissement avec les autres DPs possibles



# Exemple

**Question:** Est-ce que **A**, **B** et **F** sont des paramètres de contrôle pour **I**, **J** et **L**?

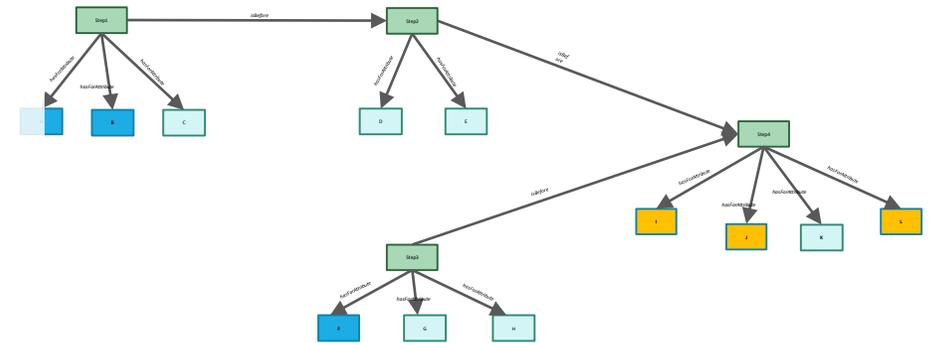
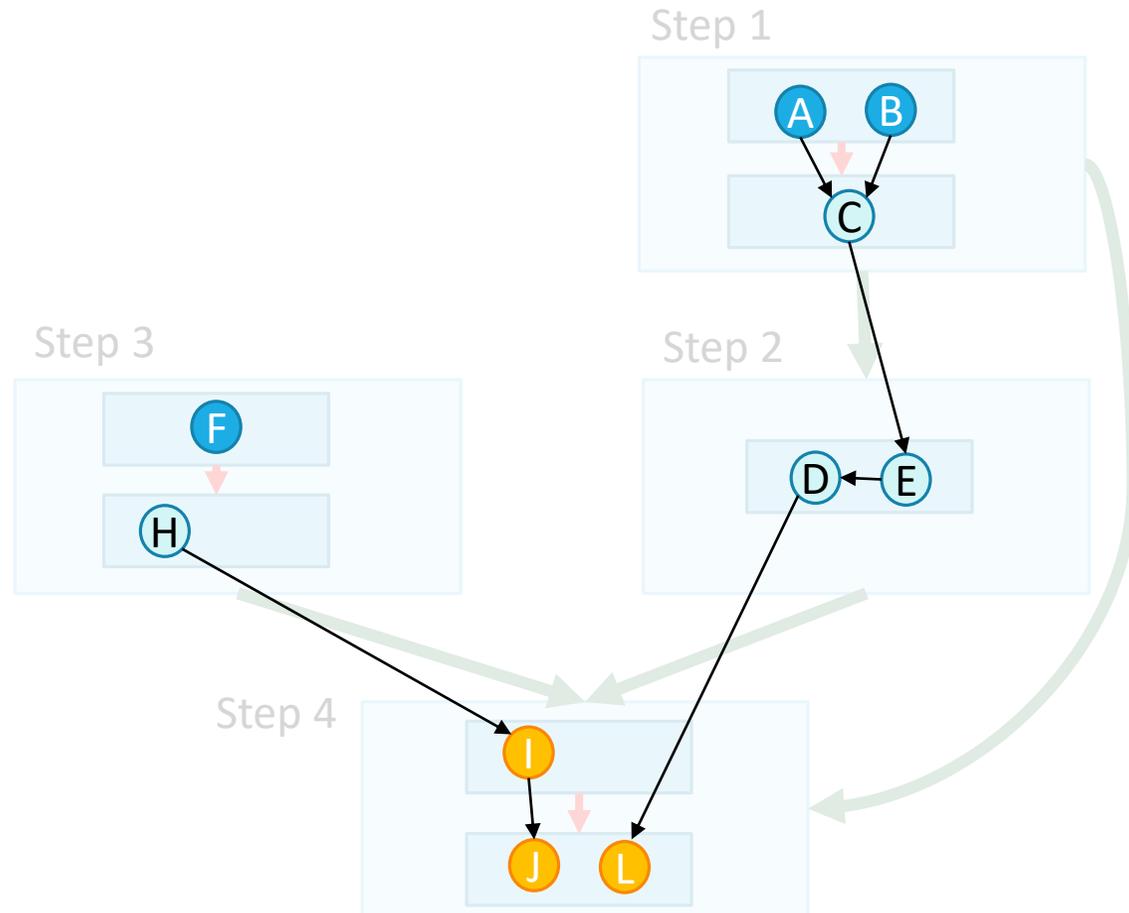
5. On intègre les DP's à l'aide des contraintes **implicites** et **explicites** jusqu'à satisfaction



# Exemple

**Question:** Est-ce que **A**, **B** et **F** sont des **paramètres de contrôle** pour **I**, **J** et **L**?

## 6. Apprentissage à l'aide d'un algorithmes classiques d'apprentissage des réseaux bayésiens



# Question de la causalité

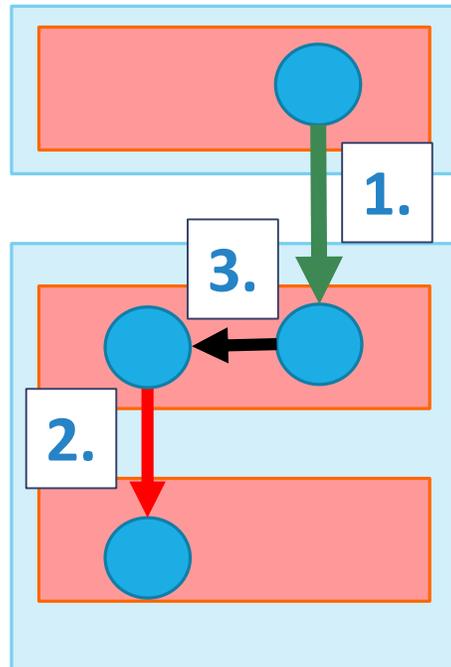
---

Apprendre un modèle probabiliste



Apprendre un modèle causal

→ Comment peut-on déduire la causalité pour répondre à la question?



Trois façons

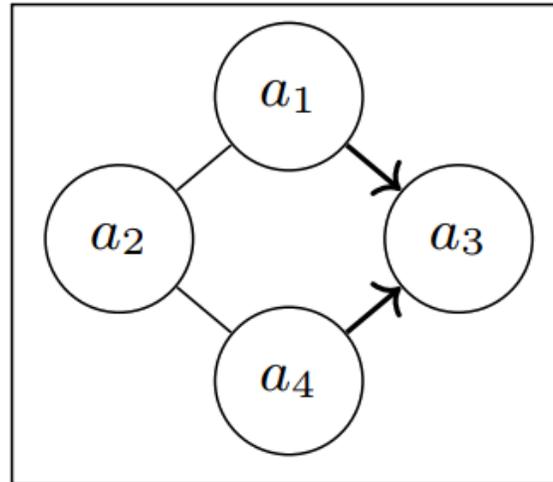
1. Contraintes explicites
2. Contraintes implicites
3. On récupère les arcs essentiellement causaux

# Graphe essentiel

---

**Graphe essentiel:** graphe semi-dirigé qui montre les arcs orientés obligatoires

*Graphe essentiel*

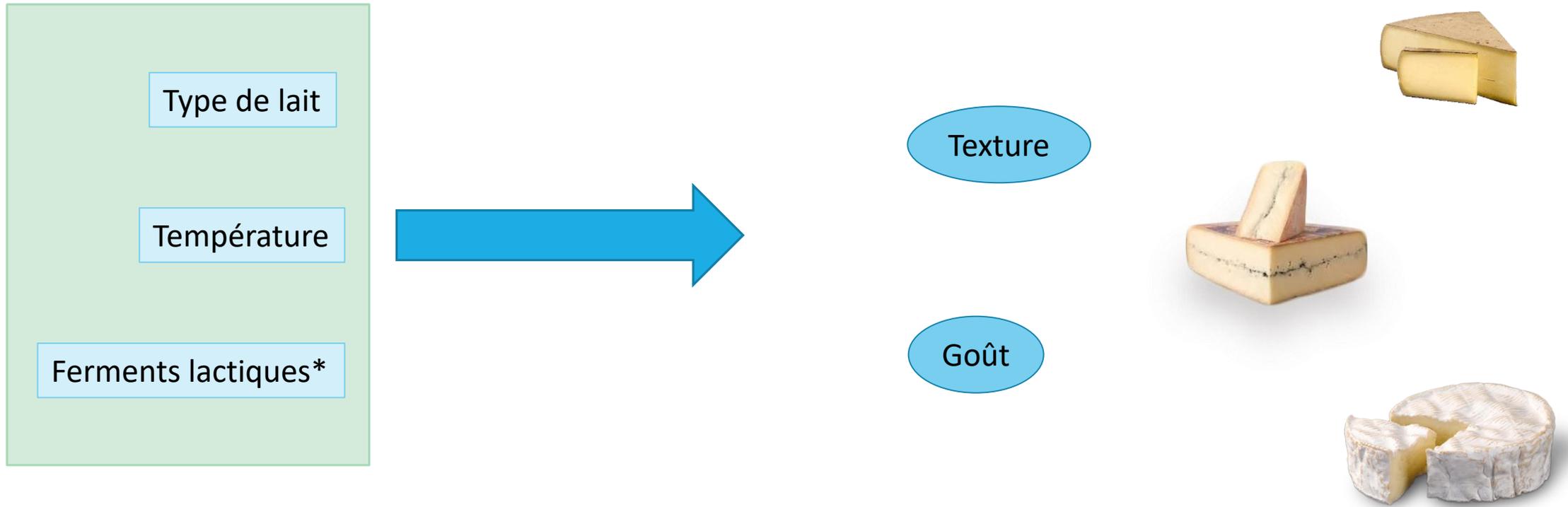


Puisque l'on apprend sous contraintes causales, on suppose que les arcs orientés sont probablement causaux

**→ Très probable, mais non garanti**

# Application à la modélisation de fabrication de fromages

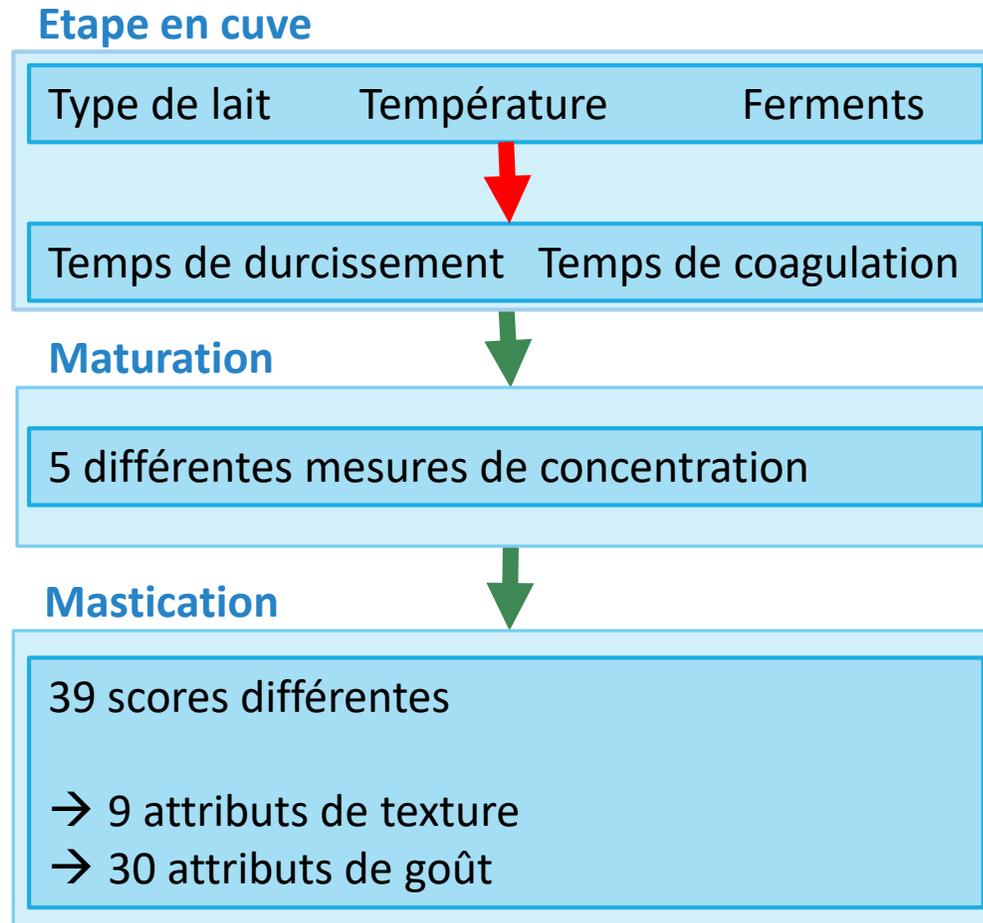
Ces paramètres sont-ils des **paramètres de contrôle**?



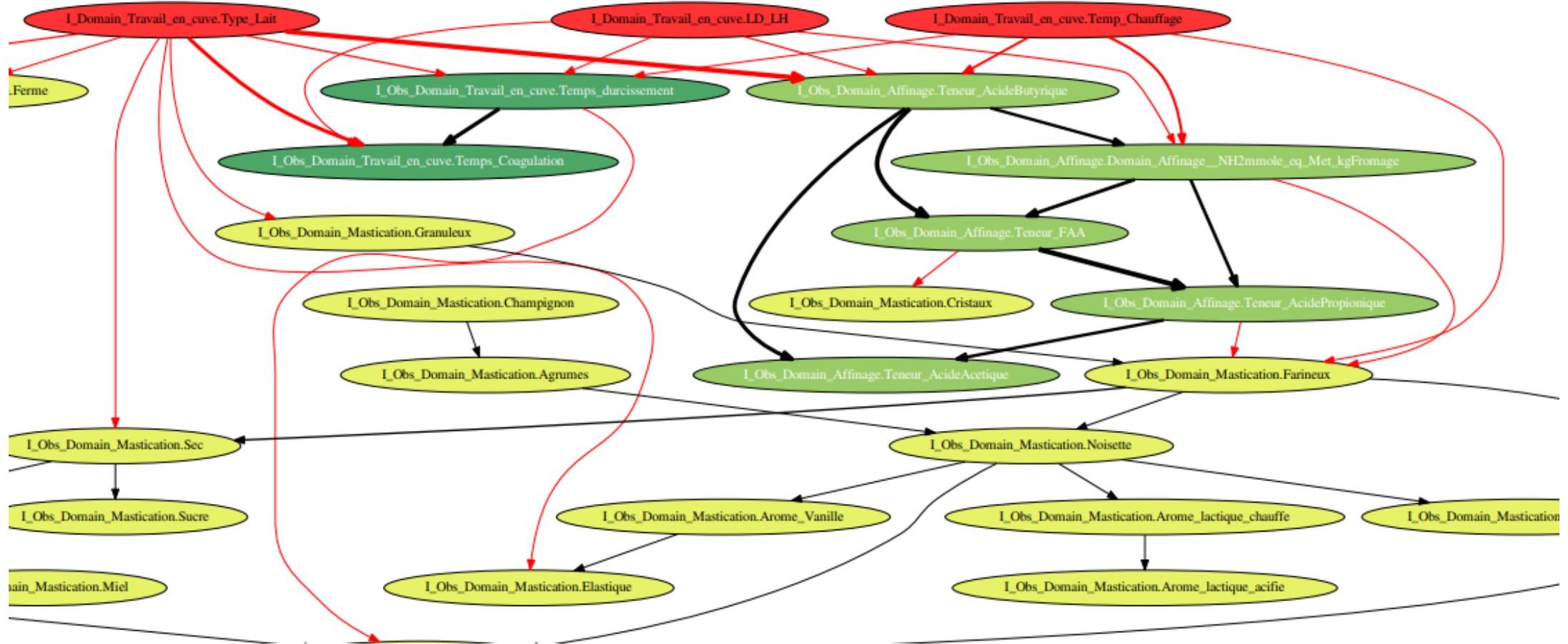
\*Thermophile lactic bacteria

# Application à la modélisation de fabrication de fromages

## Schématisation du Stack Model

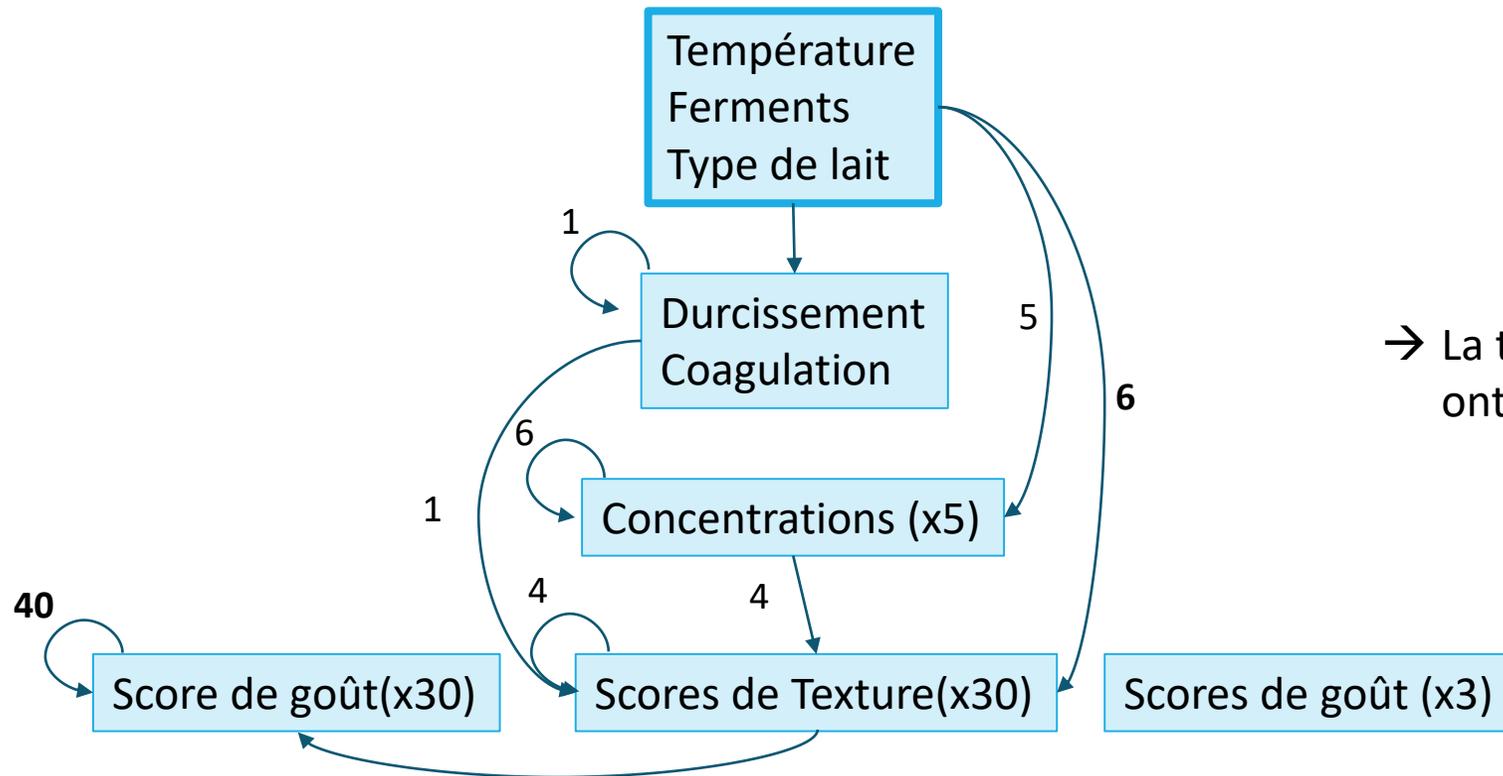


# Résultats



# Interprétation (1)

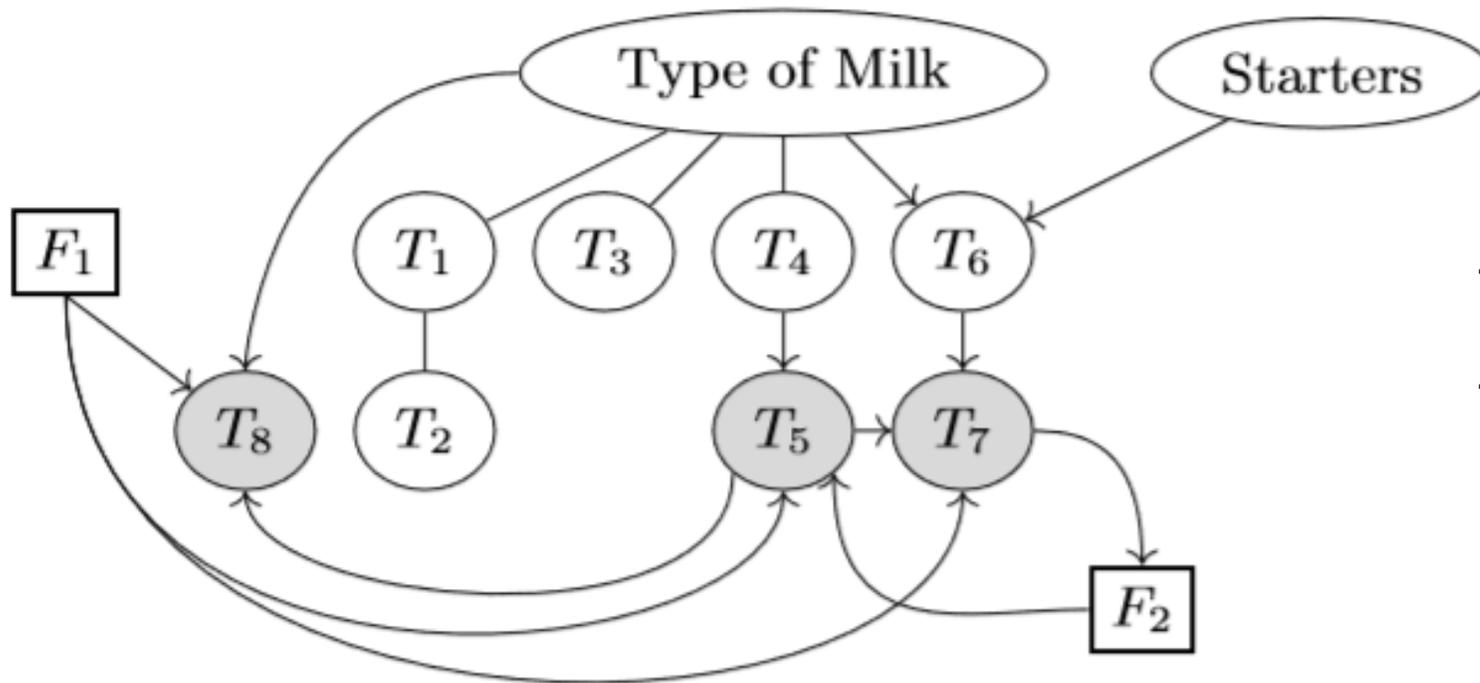
Résumé des relations inter et intra classes observées



→ La température, les ferments et le type de lait ont un lien direct sur 6 scores de texture

# Interprétation (2)

**Extrait** de l'essential graphe avec T un attribut de texture et F un groupe d'attributs de goût



- Le type de lait explique tous les **attributs de texture**
- Les **attributs de goût** sont difficilement explicables par les trois paramètres testés

# Conclusion

---

Méthode basée sur **l'intégration** du savoir expert

Nécessite: → Une question

→ Des connaissances causales

Permet une construction très proche de ce que l'expert veut exprimer

S'adapte à n'importe quelle ontologie

**Futurs travaux:** Exploiter les informations causales découvertes pour enrichir l'ontologie