

Interopérabilité des systèmes d'information

5. Modéliser des connaissances

Mastère Management de la transition industrielle

Antoine Zimmermann

15 juin 2023

Modélisation formelle de connaissances

- Pour aller plus loin dans l'interopérabilité sémantique, on peut vouloir décrire des connaissances au sujet des entités d'intérêt, de façon à automatiser des raisonnements
- P.ex. :

Une émulsion est une substance qui contient un corps gras.

La mayonnaise est une émulsion.

Donc la mayonnaise contient un corps gras.

Introduction

- Pourquoi modéliser des connaissances ?
 - Rendre les connaissances explicites
 - De façon à ce qu'elle soit indépendante du système
 - réutilisable dans différents systèmes
 - Doit permettre de déduire des informations implicites à partir de connaissances et faits explicites

Introduction

Par exemple :

- Une application pour apprendre à cuisiner et un moteur de recherche culinaire requièrent tous deux des connaissances en matière de cuisine.
- Une application de gestion de stock et un site Web de commerce électronique requièrent des connaissances au sujet de produits et de services.

Introduction

Connaissance vs. données vs. information

Les **données** sont juste des valeurs (p.ex., 37.2, “il”, 2015-04-23), éventuellement des structures (p.ex., une table)

L'**information** est ce que les données signifient lorsqu'elles sont interprétées selon des connaissances (p.ex., 37,2 est la température en °C du corps d'une personne à un certain moment)

La **connaissance** est ce qui traduit les données en information (je sais que le capteur est un thermomètre qui vient de sortir de la bouche de Gustave)

Introduction

*Connaissances
générales*
(p.ex., un roman
est une
narration)

+

*Données
spécifiques*
(p.ex., *Harry
Potter and the
Philosophers's
Stone* est un
roman)

=

Conclusions
(p.ex., *Harry
Potter and the
Philosophers's
Stone* est une
narration)

Général vs. particulier (exemple)

Connaissances générales

Réponds à ces questions :

- Qu'est qu'un ingrédient ?
- Qu'est-ce qu'une recette ?
- Quelle est la relation entre une recette et un ingrédient ?

Connaissances particulières

Réponds à ces questions :

- Quelle quantité de beurre dans la purée de Robuchon ?
- Quels légumes utilise-t-on dans la ratatouille ?
- Quelle est la durée de cuisson du cake aux olives de Sophie ?

Systemes de classification

Supreme genus:

Differentiae:

Subordinate genera:

Differentiae:

Subordinate genera:

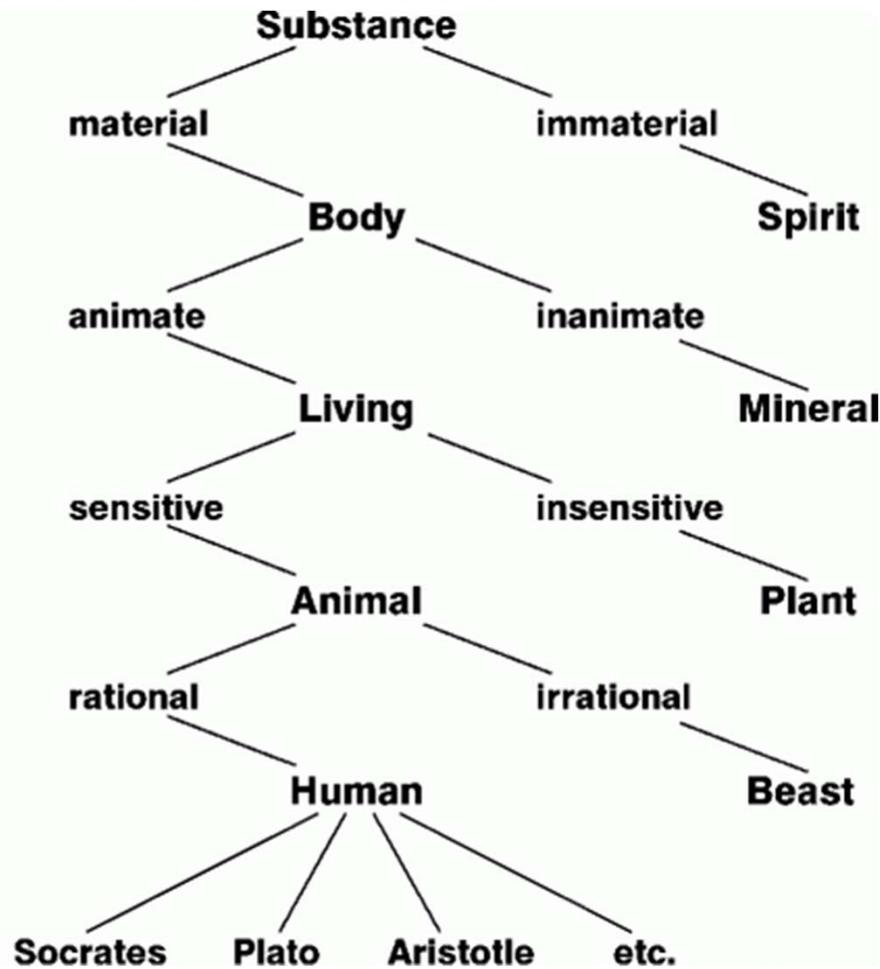
Differentiae:

Proximate genera:

Differentiae:

Species:

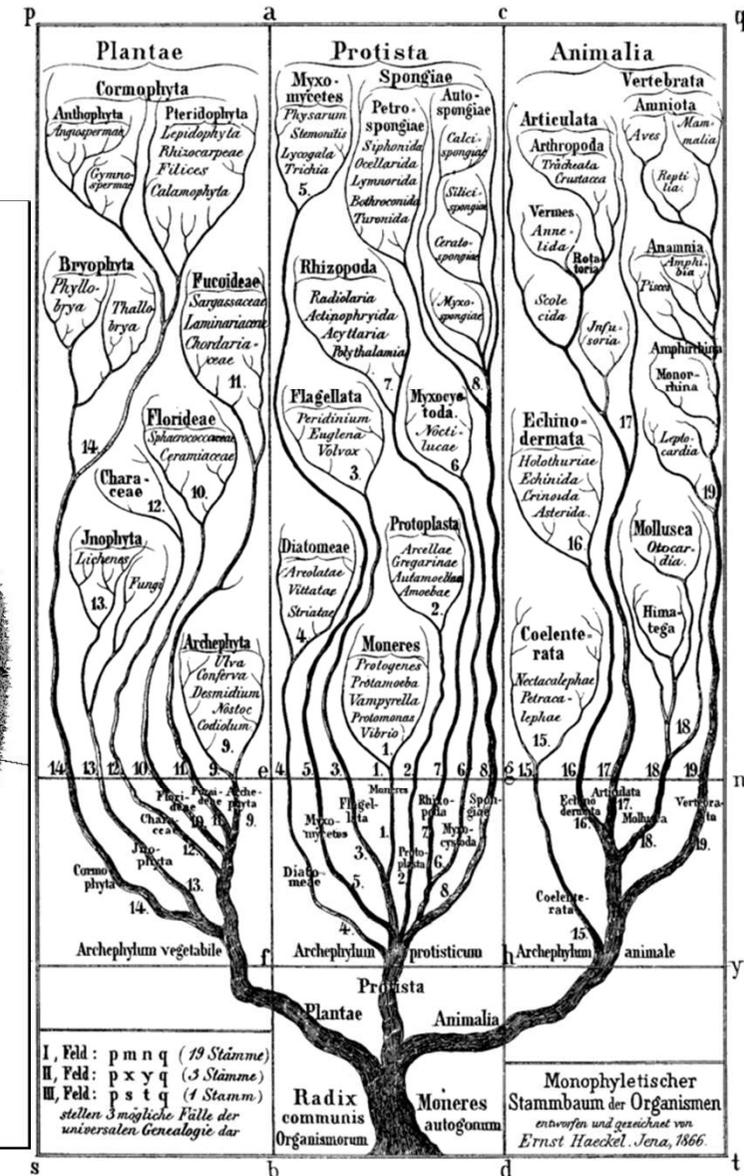
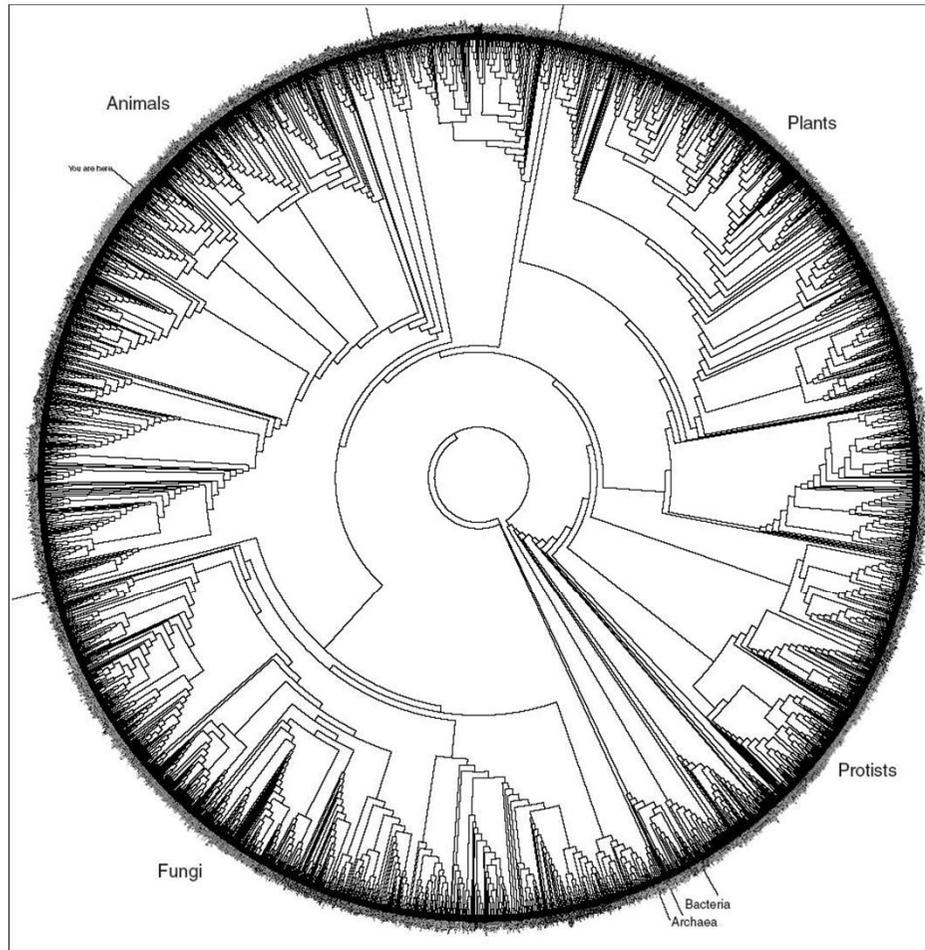
Individuals:



L'arbre de Porphyre, III^e siècles après J.-C., représentent les catégories d'Aristote

Systemes de classification

Taxonomie biologique



Systemes de classification

Classification documentaire dans les bibliothèques (p.ex., la classification Dewey)

500 *Sciences naturelles et mathématiques*

510 *Mathématiques*

516 *Géométrie*

516.3 *Géométrie analytique*

516.37 *Géométries différentielles métriques*

516.375 *Géométrie de Finsler*

Systemes de classification

Connaissances générales: *Les personnes sont des êtres vivants*

Données spécifiques: *Aristote est une personne*

Conclusion: *Aristote est un être vivant*

En logique du premier ordre :

$\forall x. \text{Personne}(x) \Rightarrow \hat{\text{Être-vivant}}(x)$

$\text{Personne}(\textit{Aristote})$

\models

$\hat{\text{Être-vivant}}(\textit{Aristote})$

Systemes de classification : instruction

- Un nom ne peut servir que pour **une** classe
- Un mot ne correspond pas forcément à une classe et vice-versa
- Utilisez des conventions de nommage cohérentes (p.ex., majucule, nom ou groupe nominal singulier)
- Définissez les classes par analogie : structure similaire, caractéristiques similaires

Options de classification

Classification binaire : est / n'est pas (a / n'a pas)

p.ex., Motorisé vs. NonMotorisé

Classification par ensemble fixe :

e.g., VéhiculeVolant, VéhiculeTerrestre, VéhiculeMarin

Des classes pour des valeurs discrètes

e.g., VéhiculeÀ1Roue, VéhiculeÀ2Roues, VéhiculeÀ3Roues, etc.

Des classes par plages de valeurs

e.g., 10à20ChevauxVapeurs, 20à100ChevauxVapeurs, etc.

Représentation des connaissances en graphes

- Les classes ne sont pas suffisantes pour représenter la plupart des connaissances
- Des relations entre entités sont nécessaires
 - *Quelles parties a un véhicule ? Une voiture a un moteur*
 - *Quelle valeur a un attribut ? Une voiture fait 5,2 m de longueur*

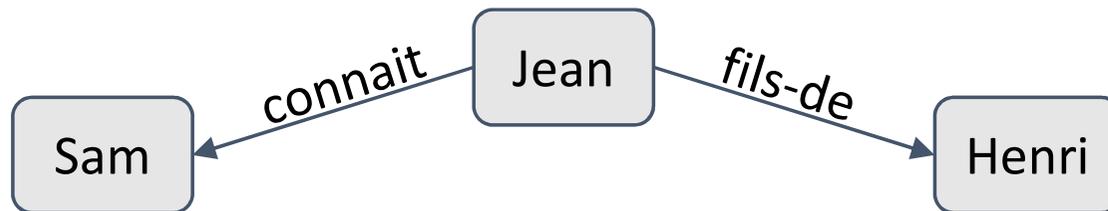
Représentation des connaissances en graphes

Représenter des entités et leurs relations :

- Cartes mentales
- Cartes topiques (ou carte de thèmes)
- Réseaux sémantiques
- Graphes conceptuels
- Le cadre de description de ressources (*Resource Description Framework* ou *RDF*)

Représentation des connaissances en graphes

Réseaux sémantiques : décrivent les entités particulières
p.ex., Jean connaît Sam et est le fils de Henri



En logique du premier ordre :

$\text{connaît}(\textit{Jean}, \textit{Sam}) \wedge \text{fils-de}(\textit{Jean}, \textit{Henri})$

Représentation des connaissances en graphes

Exercices :

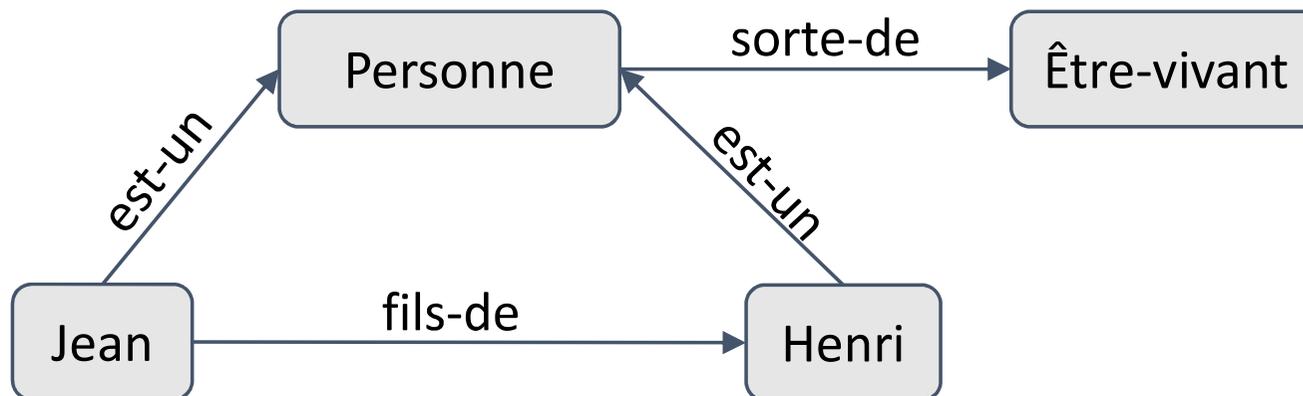
- *Décrire l'omelette nature que j'ai eu à midi avec un graphe conceptuel*
- *Décrire une recette d'omelette nature ?*

Représentation des connaissances en graphes

Ajoutons 2 relations spéciales

est-un (entre 1 entité et 1 classe à laquelle elle appartient)

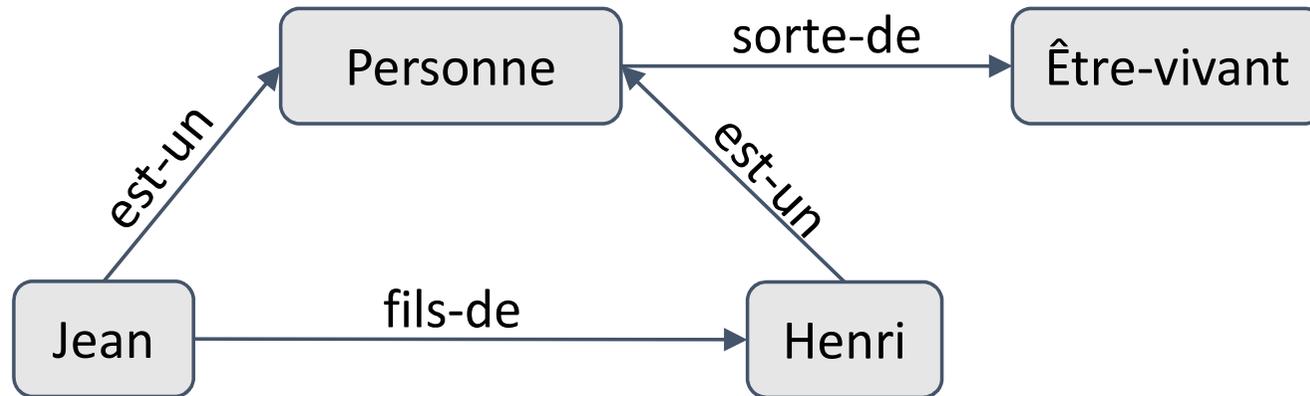
sorte-de (entre 1 classe et une super-classe)



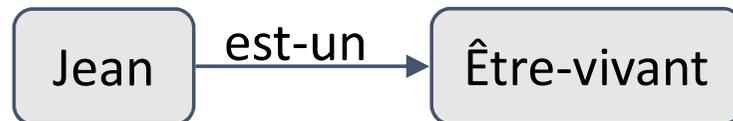
Représentation des connaissances en graphes

Signification de **est-un** et **sorte-de**

Nous voudrions que ce graphe :



implique logiquement :



Représentation des connaissances en graphes

Exercice :

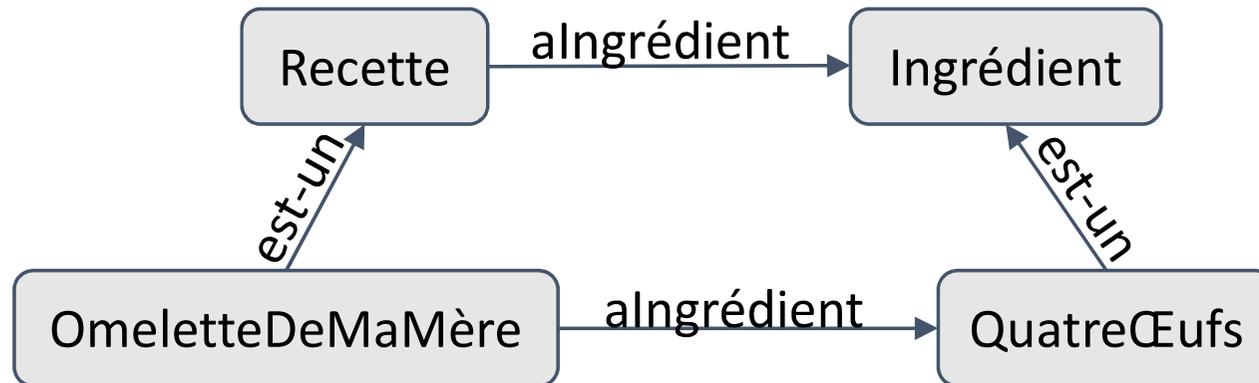
*Complétez votre description du modèle de connaissances des recettes avec **est-un** et **sorte-de***

Utilisation visible des graphes de connaissances

- Le *Knowledge Graph* de Google
- Versions similaires chez Bing, Yahoo, Yandex, etc.
- See <https://bit.ly/2HbppFz>, <https://binged.it/2vqsriF>, <https://bit.ly/4dwazET>, etc.

Représentation en graphe : la généralisation

- Nous avons vu un formalisme qui peut exprimer qu'une recette particulière (p.ex., l'omelette de ma mère) a des ingrédients particuliers (p.ex., 4 œufs). Comment représenter le fait que les recettes, en général, ont des ingrédients ?



?

Représentation en graphe : la généralisation

- Il y a plusieurs interprétations de cette relation



- 1) La classe *Recette* elle-même est lié à la classe *Ingrédient* par la relation *aIngrédient* $aIngrédient(Recette, Ingrédient)$
- 2) Toutes les recettes ont des ingrédients
 $\forall x. \exists y. Recette(x) \Rightarrow Ingrédient(y) \wedge aIngrédient(x, y)$ ou
 $\forall x. \exists y. est-un(x, Recette) \Rightarrow est-un(y, Ingrédient) \wedge aIngrédient(x, y)$
- 3) Les recettes peuvent avoir des ingrédients

Représentation en graphe : la généralisation

- Il y a plusieurs interprétations de cette relation



- 4) La relation *aIngrédient* ne relie que des recettes à des ingrédients
 $\forall x. \forall y. aIngrédient(x, y) \Rightarrow Recette(x) \wedge Ingrédient(y)$ ou
 $\forall x. \forall y. aIngrédient(x, y)$
 $\Rightarrow est-un(x, Recette) \wedge est-un(y, Ingrédient)$

Les graphes conceptuels (John F. Sowa, 1979)

- Une représentation sous forme de graphe très expressive



- L'omelette de ma mère est une recette, quatre œufs sont un ingrédient et L'omette de ma mère a quatre œufs.

$aIngrédient(OmeletteDeMaman, quatreŒufs) \wedge$
 $Recette(OmeletteDeMaman) \wedge Ingrédient(quatreŒufs)$

Les graphes conceptuels (John F. Sowa, 1979)



- Il y a une recette qui a quelque ingrédient.

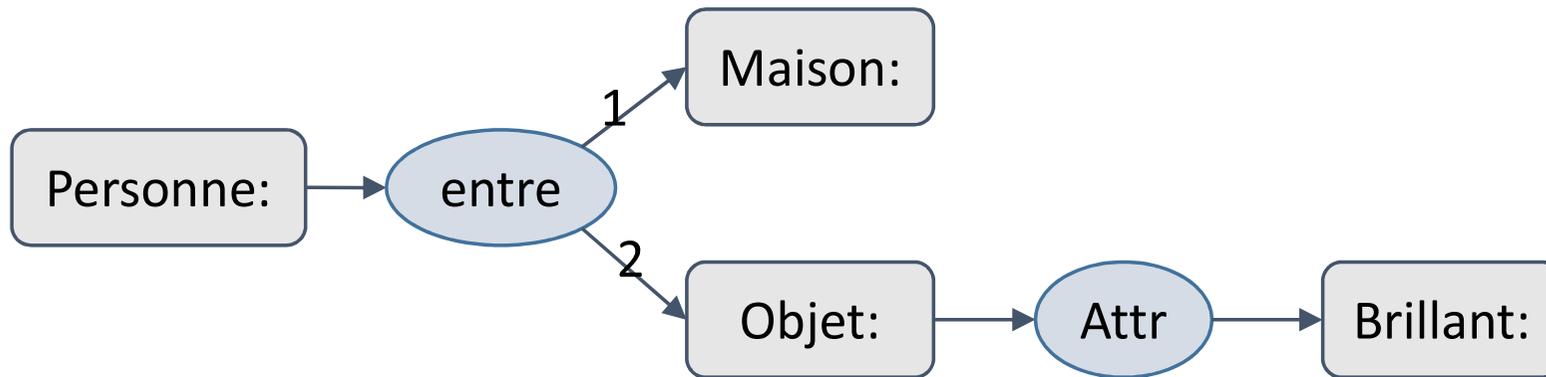
$\exists x \exists y. a\text{Ingrédient}(x, y) \wedge \text{Recette}(x) \wedge \text{Ingrédient}(y)$



- Toutes les recettes ont au moins un ingrédient.

$\forall x \exists y. \text{Recette}(x) \Rightarrow a\text{Ingrédient}(x, y) \wedge \text{Ingrédient}(y)$

Les graphes conceptuels (John F. Sowa, 1979)



- Quelqu'un entre une maison et un objet brillant.

$$\exists w \exists x \exists y \exists z. \text{Personne}(w) \wedge \text{Maison}(x) \wedge \text{Objet}(y) \wedge \text{Brillant}(z) \wedge \text{Entre}(w, x, y) \wedge \text{Attr}(y, z)$$