Procédure d'optimisation pour Datalog disjonctif

Stage Master 2 IASD

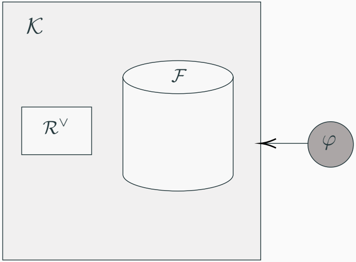
Pierre Bisquert, David Carral, Lucas Rouquette, Federico Ulliana

16 juin 2022

INRIA - Équipe GraphIK

Contexte et motivations

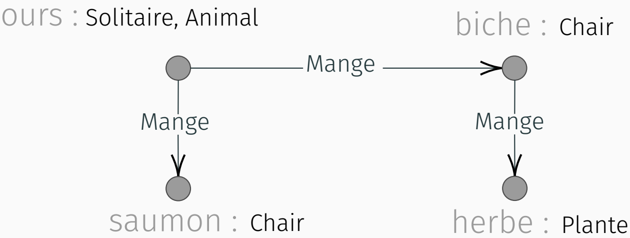
Contexte



- Règles disjonctives
- Inférence de fait
- Requêtes *atomiques instanciées*

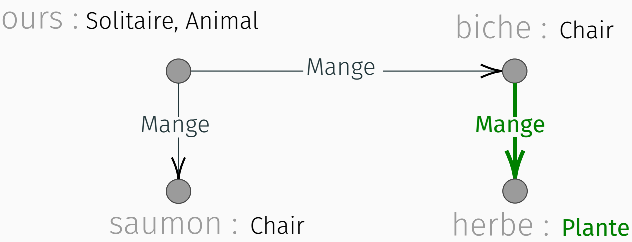
Chase disjonctif

$$\begin{aligned}
Animal(x) &\rightarrow Prédateur(x) \vee Proie(x)\\
Prédateur(x) &\rightarrow Carnivore(x)\\
Proie(x) &\rightarrow Herbivore(x) \vee Carnivore(x)\\
Mange(x,y) \wedge Chair(y) &\rightarrow Carnivore(x)\\
Mange(x,y) \wedge Plante(y) &\rightarrow Herbivore(x)\\
Solitaire(x) \wedge Mange(x,y) \wedge Chair(y) &\rightarrow Carnivore(x)
\end{aligned}$$



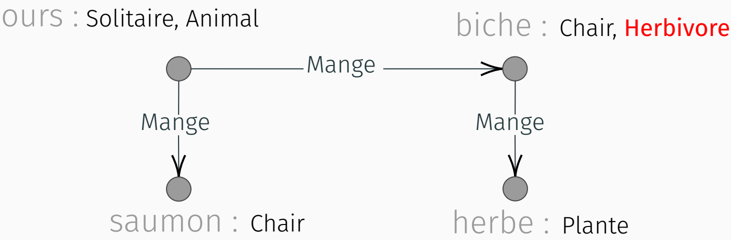
Chase disjonctif

$$\text{Mange(x,y)} \wedge \text{Plante(y)} \rightarrow \mathit{Herbivore}(x)$$



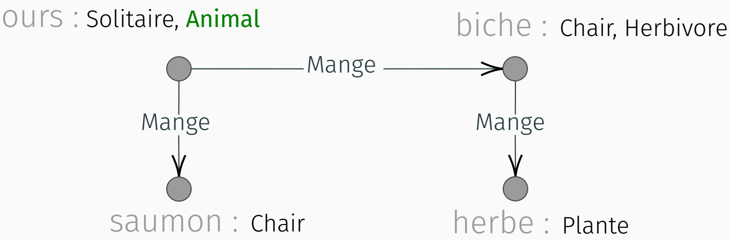
Chase disjonctif

$$Mange(x, y) \wedge Plante(y) \rightarrow \text{Herbivore(x)}$$



Chase disjonctif

$$\text{Animal}(x) \rightarrow \textit{Prédateur}(x) \vee \textit{Proie}(x)$$

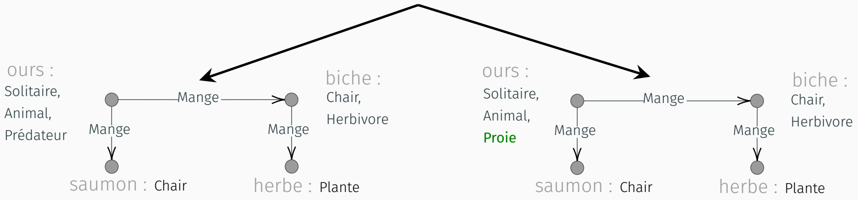


Chase disjonctif

$$Animal(x) \rightarrow \text{Prédateur(x)} \vee \text{Proie(x)}$$

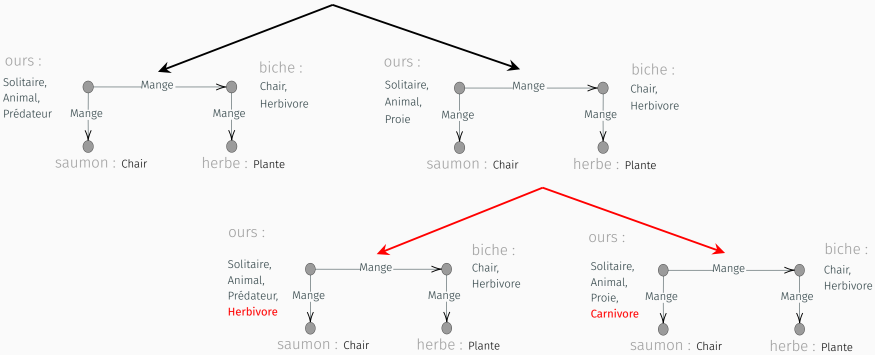
Chase disjonctif

$$\color{green}{\text{Proie(x)}} \rightarrow \mathit{Herbivore}(x) \vee \mathit{Carnivore}(x)$$



Chase disjonctif

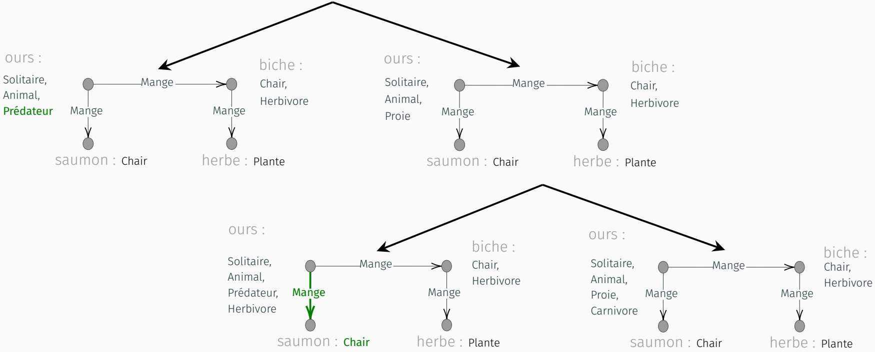
$Proie(x) \rightarrow \textbf{Herbivore(x)} \lor \textbf{Carnivore(x)}$



Chase disjonctif

$$\text{Prédateur(x)} \rightarrow Carnivore(x)$$

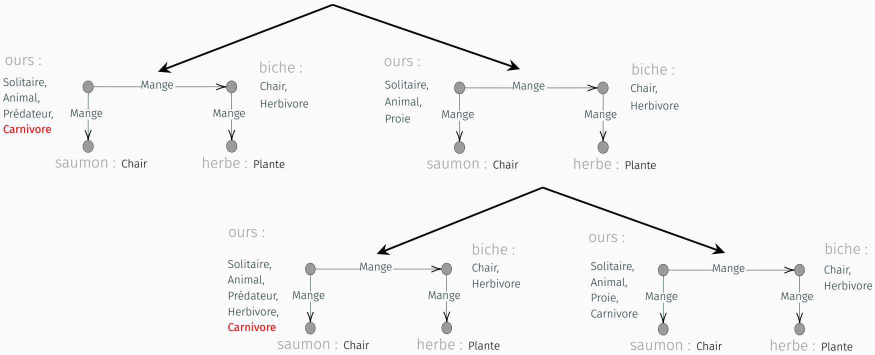
$$\text{Mange(x,y)} \wedge \text{Chair(y)} \rightarrow Carnivore(x)$$



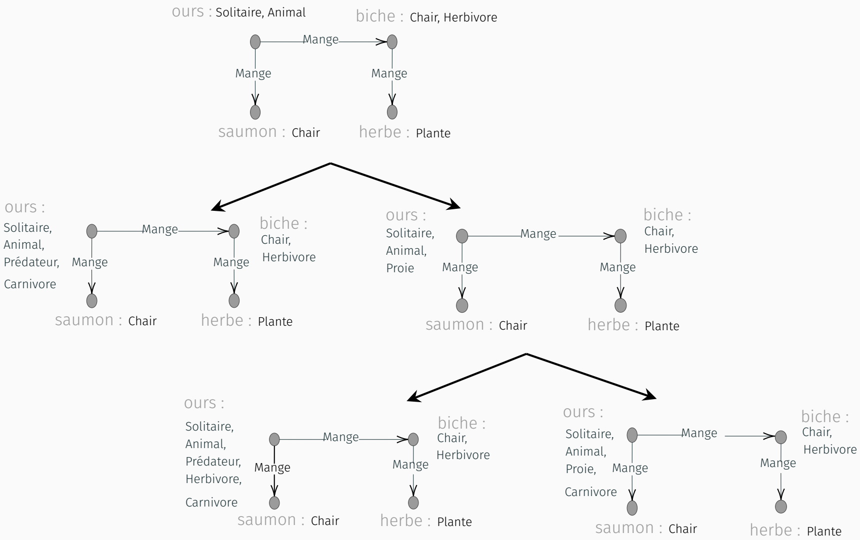
Chase disjonctif

$$Prédateur(x) \rightarrow \textcolor{red}{\text{Carnivore(x)}}$$

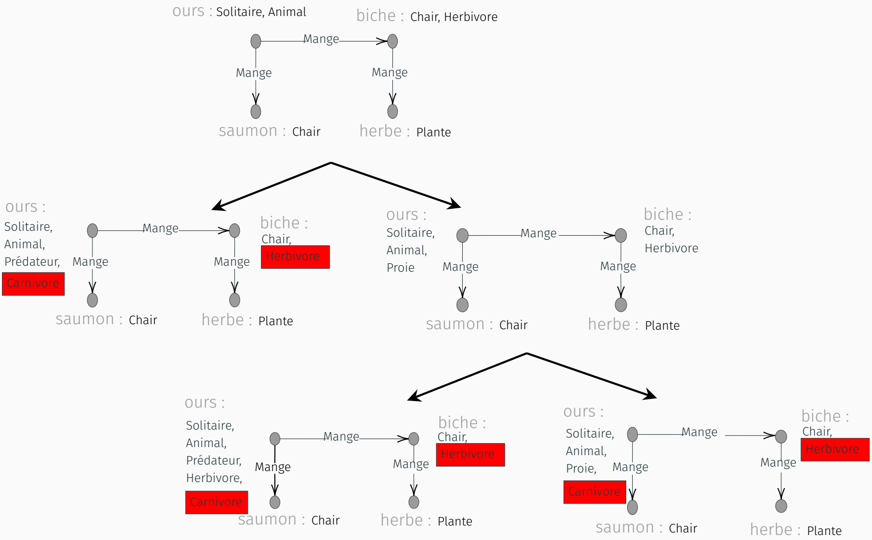
$$Mange(x, y) \wedge Chair(y) \rightarrow \textcolor{red}{\text{Carnivore(x)}}$$



Chase disjonctif



Inférence de fait

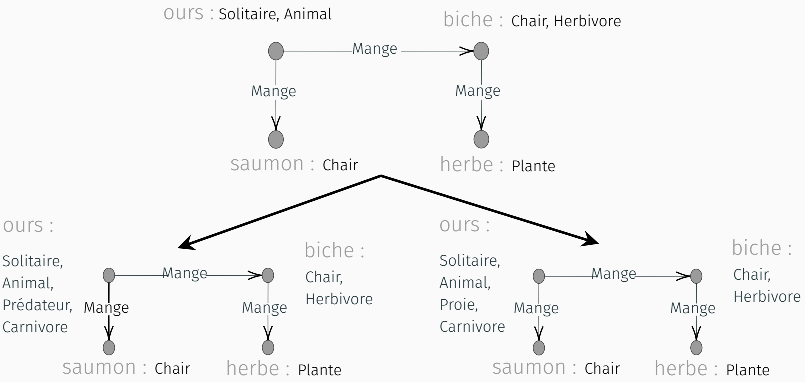


$$\mathcal{K} \models \mathcal{F} \cup \{Herbivore(biche), Carnivore(ours)\}$$

Différents arbres

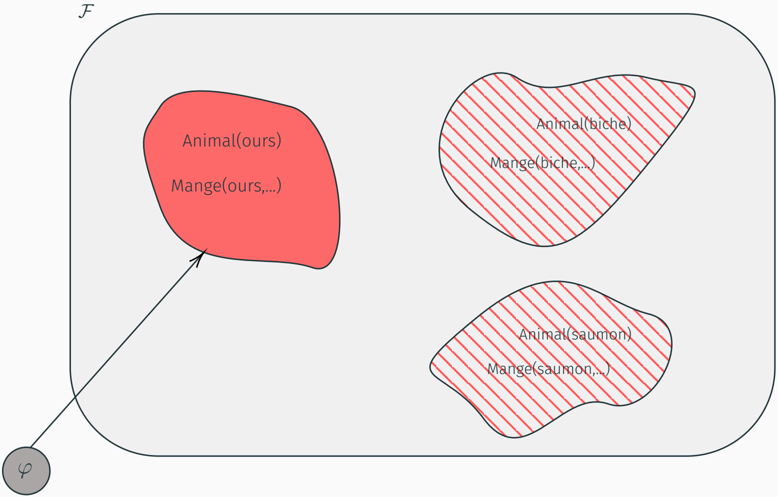
$\text{Mange(x,y)} \wedge \text{Chair(y)} \rightarrow \text{Carnivore(x)}$

$Proie(x) \rightarrow Herbivore(x) \vee Carnivore(x)$ *(non applicable)*



Faits utiles ?

Faits utiles ?



Pertinence des faits

Pertinence - définition sémantique

Définition

Étant donné une base de connaissances $\mathcal{K} = \langle \mathcal{R}, \mathcal{F} \rangle$ et un fait φ, on dit qu'un fait $\psi \in \mathcal{F}$ est *pertinent* pour l'inférence produisant φ s'il existe un ensemble $\mathcal{M} \subseteq \mathcal{F}$ tel que :

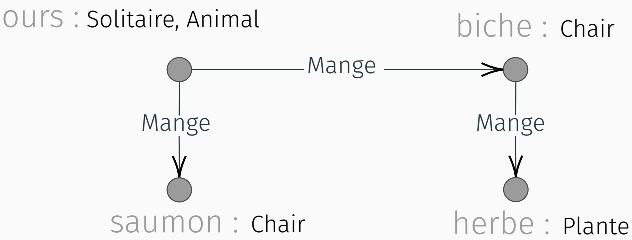
- $\psi \in \mathcal{M}$,
- $\langle \mathcal{R}, \mathcal{M} \rangle \models \varphi$, et
- $\langle \mathcal{R}, \mathcal{M}' \rangle \not\models \varphi$ pour tout $\mathcal{M}' \subset \mathcal{M}$.

Remarque

On dénote par Relevant$(\mathcal{K}, \varphi)$ l'ensemble des faits pertinents pour l'inférence de φ.

Pertinence - exemple

φ = Carnivore(ours) ?

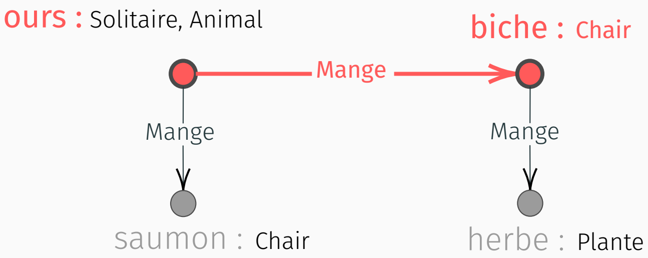


$\text{Relevant}(\mathcal{K}, \mathit{Carnivore}(\mathit{ours})) = \{\}$

Pertinence - exemple

φ = Carnivore(ours) ?

$$Mange(x, y) \wedge Chair(y) \rightarrow Carnivore(x)$$

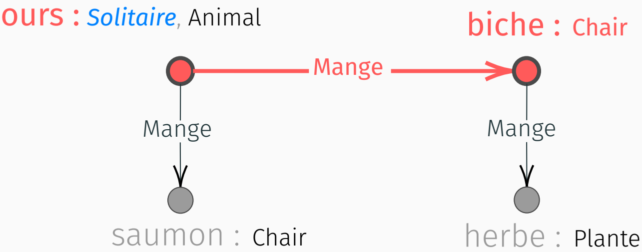


$$\text{Relevant}(\mathcal{K}, Carnivore(ours)) = \{\ Mange(ours, biche), Chair(biche)\ \}$$

Pertinence - exemple

φ = Carnivore(ours) ?

$$Mange(x,y) \wedge Chair(y) \rightarrow Carnivore(x)$$
$$\text{Solitaire(x)} \wedge Mange(x,y) \wedge Chair(y) \rightarrow Carnivore(x)$$

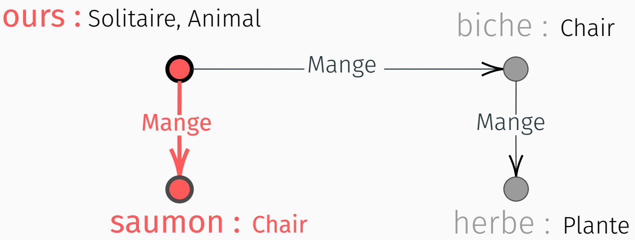


$$\text{Relevant}(\mathcal{K}, Carnivore(ours)) = \{\ Mange(ours, biche), Chair(biche)\ \}$$

Pertinence - exemple

φ = Carnivore(ours) ?

$$Mange(x, y) \wedge Chair(y) \rightarrow Carnivore(x)$$



$$\text{Relevant}(\mathcal{K}, \varphi) = \{\ Mange(ours, biche), Chair(biche), Mange(ours, saumon), Chair(saumon)\ \}$$

Fact Relevance($\mathcal{R}$)

Instance: Un ensemble de faits $\mathcal{F}$, un fait φ et un fait $\psi \in \mathcal{F}$.
Question: ψ est-il *pertinent* pour l'inférence de φ par rapport à $\langle \mathcal{R}, \mathcal{F} \rangle$?

FACT RELEVANCE($\mathcal{R}$)

INSTANCE: Un ensemble de faits $\mathcal{F}$, un fait φ et un fait $\psi \in \mathcal{F}$.
QUESTION: ψ est-il *pertinent* pour l'inférence de φ par rapport à $\langle\mathcal{R}, \mathcal{F}\rangle$?

Théorème

FACT RELEVANCE($\mathcal{R}$) est $\mathrm{NP}^{\mathrm{NP}}$-complet.

(réduction TRUE QBF)

Théorème

FACT RELEVANCE($\mathcal{R}^\wedge$) est NP-complet.

(réduction SAT)

Approximation

Graphe des déclencheurs

Graphe des déclencheurs

$\mathcal{F}$

$Animal(ours)$

$Solitaire(ours)$

$Mange(ours, saumon)$

$Mange(ours, biche)$

$Chair(biche)$

$Chair(saumon)$

$Mange(biche, herbe)$

$Plante(herbe)$

Graphe des déclencheurs

$$\text{Animal(x)} \rightarrow \textit{Prédateur}(x) \vee \textit{Proie}(x)$$

$\mathcal{F}$

Animal(ours)

$\textit{Solitaire}(\textit{ours})$

$\textit{Mange}(\textit{ours}, \textit{saumon})$

$\textit{Mange}(\textit{ours}, \textit{biche})$

$\textit{Chair}(\textit{biche})$

$\textit{Chair}(\textit{saumon})$

$\textit{Mange}(\textit{biche}, \textit{herbe})$

$\textit{Plante}(\textit{herbe})$

Graphe des déclencheurs

$$Animal(x) \rightarrow \text{Prédateur(x)} \vee \text{Proie(x)}$$

$\mathcal{F}$

$Animal(ours)$ → Proie(ours)

$Animal(ours)$ → Prédateur(ours)

$Solitaire(ours)$

$Mange(ours, saumon)$

$Mange(ours, biche)$

$Chair(biche)$

$Chair(saumon)$

$Mange(biche, herbe)$

$Plante(herbe)$

Graphe des déclencheurs

$$\text{Mange(x,y)} \wedge \text{Chair(y)} \rightarrow \mathit{Carnivore}(x)$$

$\mathcal{F}$

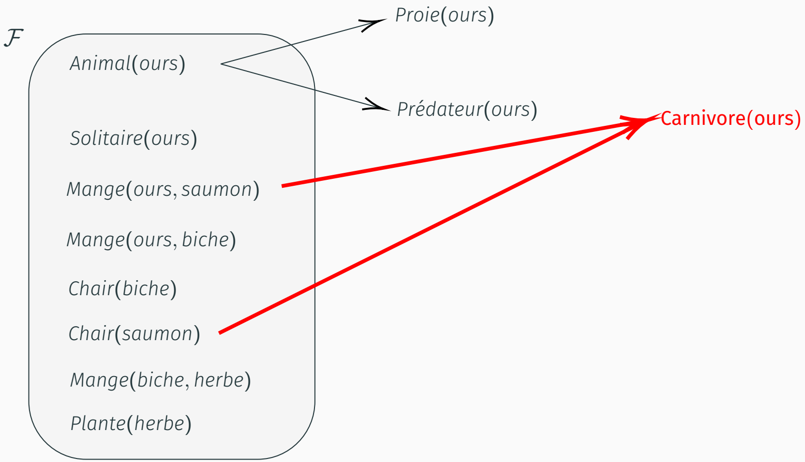
- $\mathit{Animal}(\mathit{ours})$
- $\mathit{Solitaire}(\mathit{ours})$
- $\text{Mange}(\text{ours}, \text{saumon})$
- $\mathit{Mange}(\mathit{ours}, \mathit{biche})$
- $\mathit{Chair}(\mathit{biche})$
- Chair(saumon)
- $\mathit{Mange}(\mathit{biche}, \mathit{herbe})$
- $\mathit{Plante}(\mathit{herbe})$

$\mathit{Proie}(\mathit{ours})$

$\mathit{Prédateur}(\mathit{ours})$

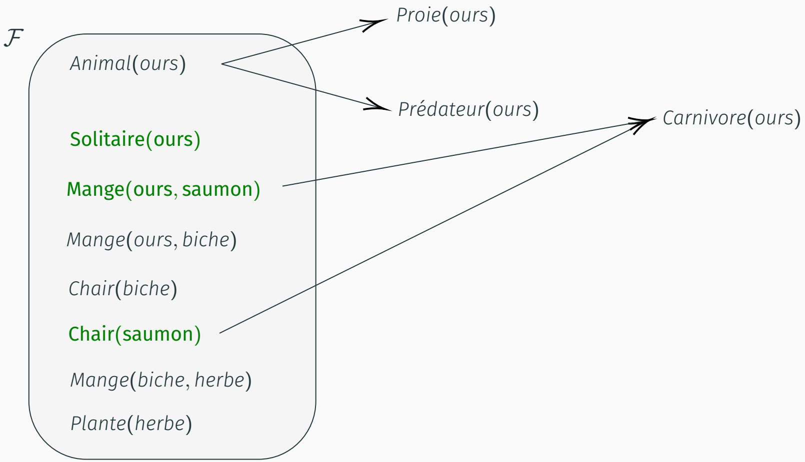
Graphe des déclencheurs

$$Mange(x, y) \wedge Chair(y) \rightarrow \text{Carnivore(x)}$$



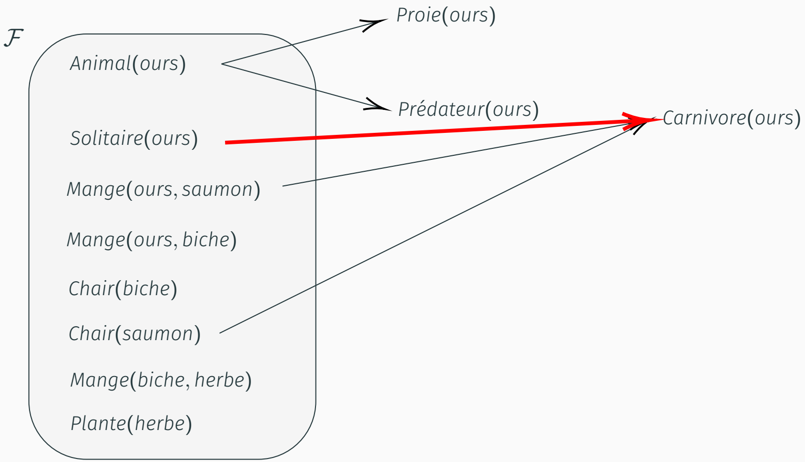
Graphe des déclencheurs

$$\text{Solitaire(x)} \wedge \text{Mange(x,y)} \wedge \text{Chair(y)} \rightarrow \mathit{Carnivore}(x)$$



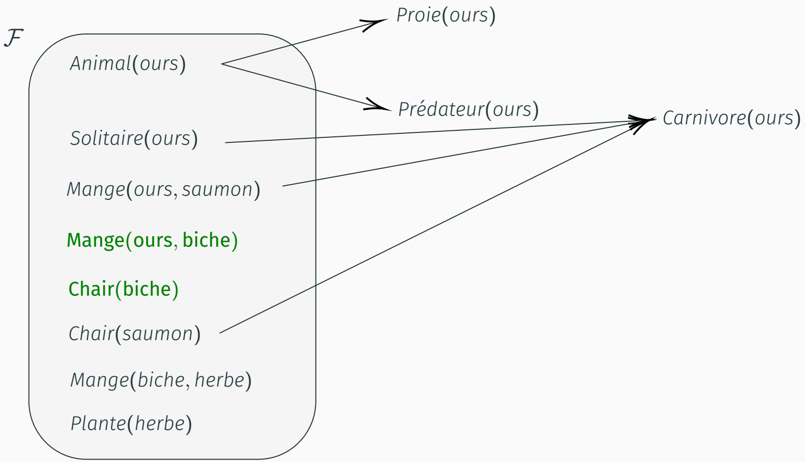
Graphe des déclencheurs

$$Solitaire(x) \wedge Mange(x, y) \wedge Chair(y) \rightarrow \text{Carnivore(x)}$$



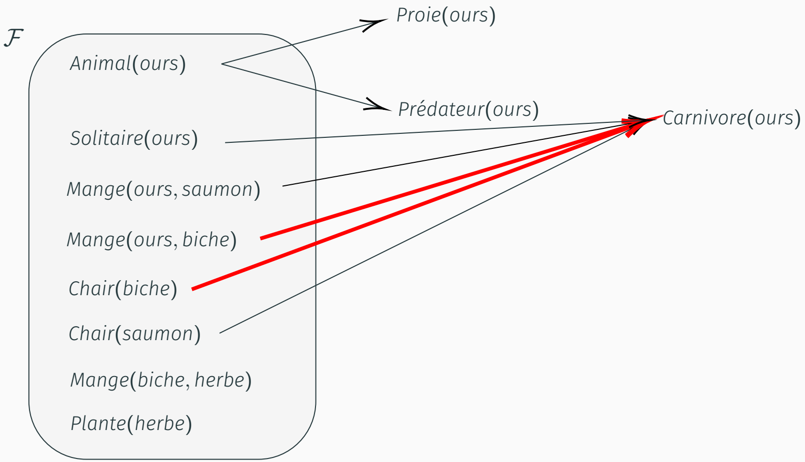
Graphe des déclencheurs

$$\textcolor{green}{\text{Mange(x,y)}} \wedge \textcolor{green}{\text{Chair(y)}} \rightarrow Carnivore(x)$$



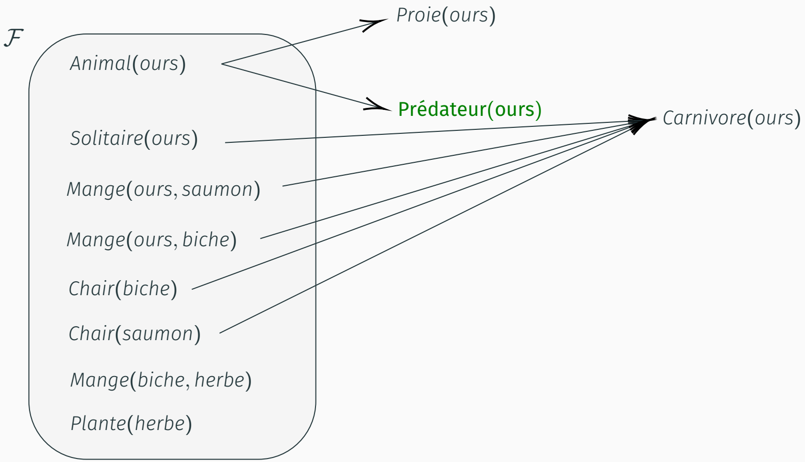
Graphe des déclencheurs

$$Mange(x, y) \wedge Chair(y) \rightarrow \color{red}{\text{Carnivore(x)}}$$



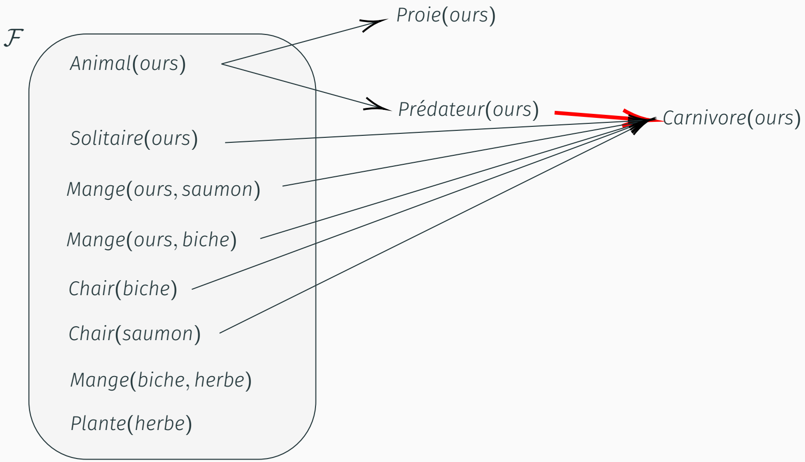
Graphe des déclencheurs

$$\text{Prédateur(x)} \rightarrow Carnivore(x)$$



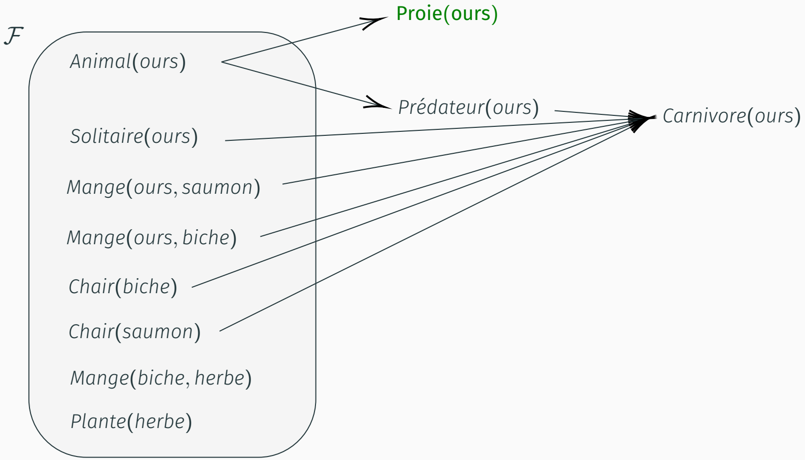
Graphe des déclencheurs

$$\mathit{Prédateur}(x) \rightarrow \color{red}{\text{Carnivore(x)}}$$



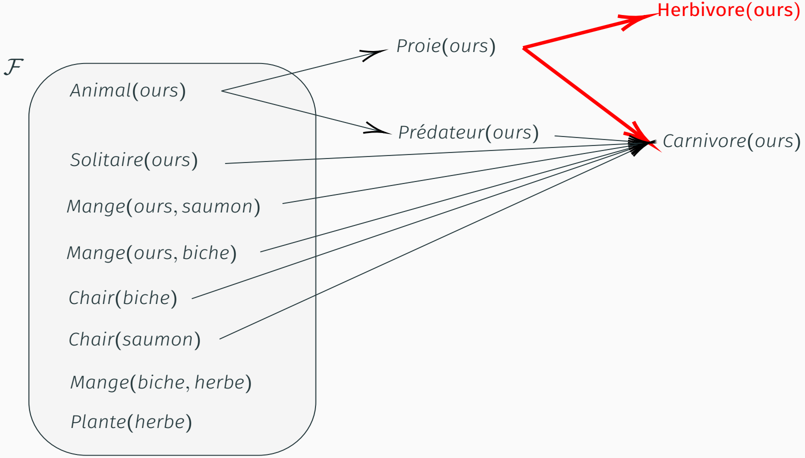
Graphe des déclencheurs

$$\text{Proie(x)} \rightarrow Herbivore(x) \vee Carnivore(x)$$



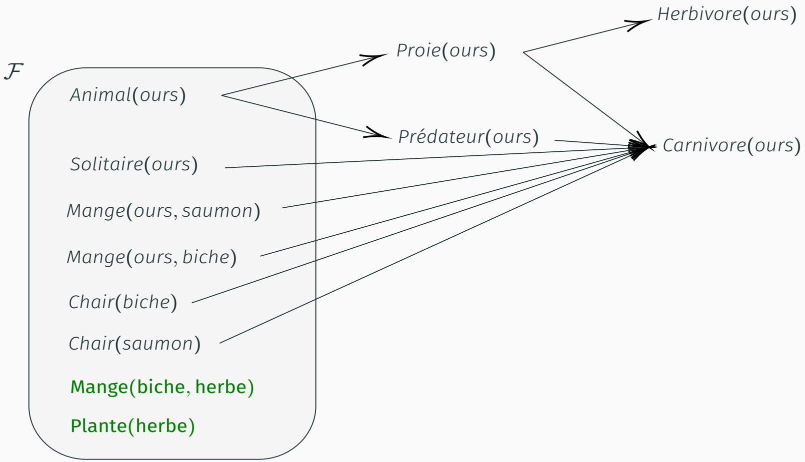
Graphe des déclencheurs

$$\mathit{Proie}(x) \rightarrow \textbf{Herbivore(x)} \vee \textbf{Carnivore(x)}$$



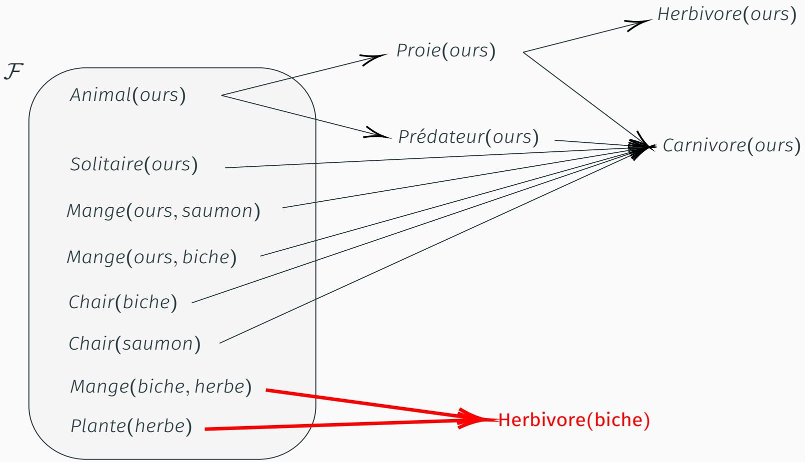
Graphe des déclencheurs

$$\text{Mange(x,y)} \wedge \text{Plante(y)} \rightarrow \mathit{Herbivore}(x)$$



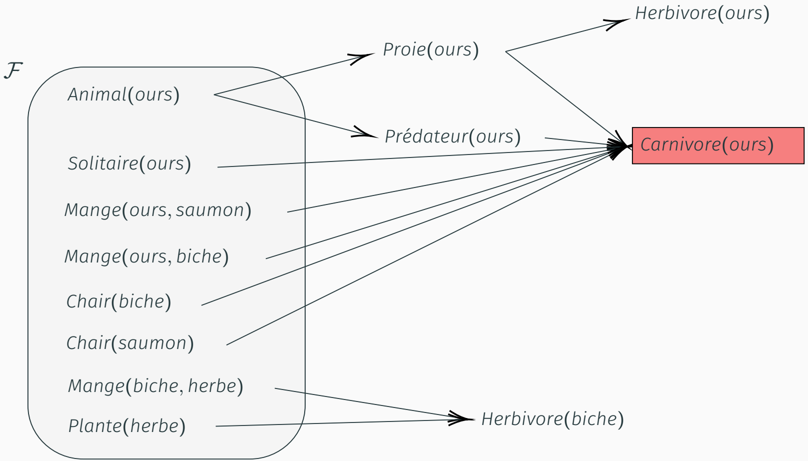
Graphe des déclencheurs

$$Mange(x, y) \wedge Plante(y) \rightarrow \textcolor{red}{\text{Herbivore(x)}}$$

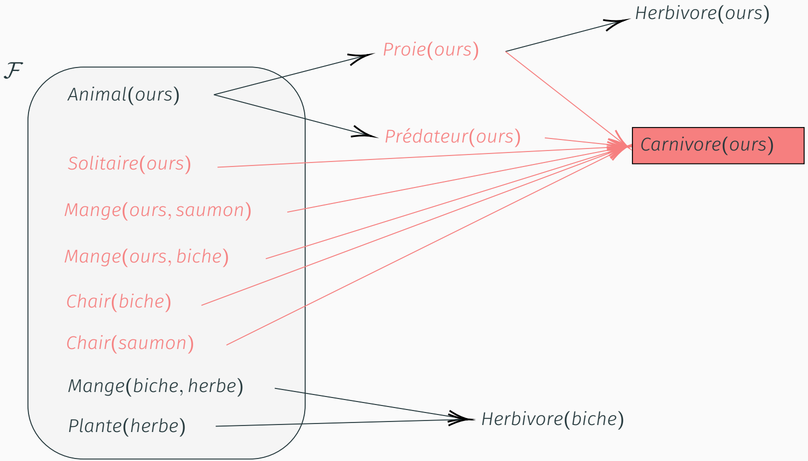


Filtrage et approximation

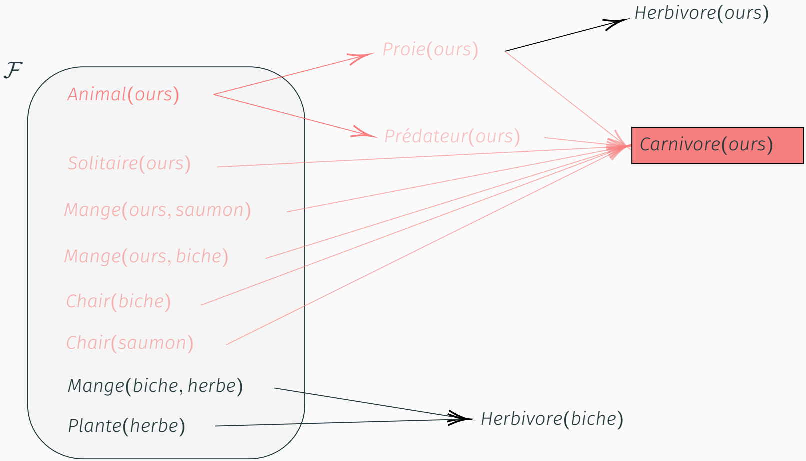
Requête φ = Carnivore(ours) ?



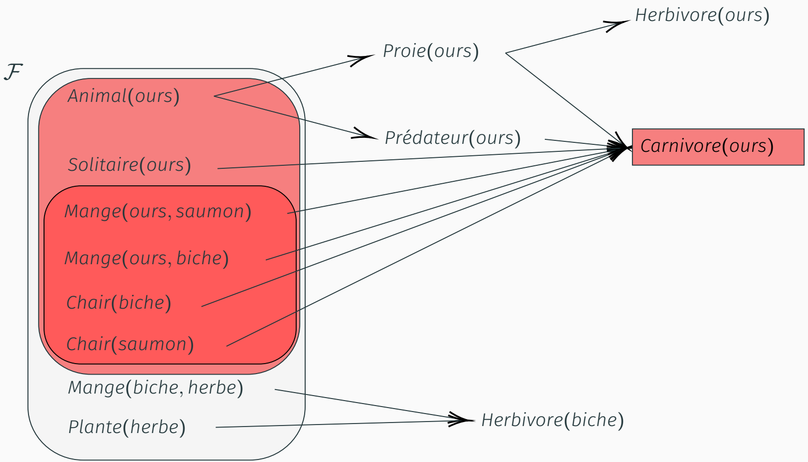
Requête φ = Carnivore(ours) ?



Requête φ = Carnivore(ours) ?



Filtrage et approximation



Remarque

L'approximation SupRelevant($\mathcal{K}, \varphi$) est calculable à l'aide d'une traduction vers un ensemble de règles déterministes et un chase $\rightarrow$ *temps polynomial.*

Évaluation

Benchmark et algorithmes

- Ontologie *BioPaX* (242 règles, 17 disjonctives).
- 10 jeux de données de taille croissante (100K à 1M).
- 10 jeux de 40 requêtes équilibrés.

Base de connaissances $\mathcal{K} = \langle \mathcal{R}, \mathcal{F} \rangle$, jeu de requêtes $\mathcal{Q}$:

AllModels :

1. Calcule tous les modèles de $\mathcal{K}$.
2. Pour tout $\varphi \in \mathcal{Q}$, vérifie si φ appartient à tous les modèles de $\mathcal{K}$.

Unsat :

1. Pour tout $\varphi \in \mathcal{Q}$, crée la KB $\mathcal{K}_\varphi^\perp = \langle \mathcal{R} \cup \{\varphi \rightarrow \perp\}, \mathcal{F} \rangle$.
2. Vérifie l'existence d'un modèle de $\mathcal{K}_\varphi^\perp$.

Benchmark et algorithmes

Base de connaissances $\mathcal{K} = \langle \mathcal{R}, \mathcal{F} \rangle$, jeu de requêtes $\mathcal{Q}$:

AllModels(f) :

1. Calcule le graphe des déclencheurs.
2. Pour tout φ, calcule tous les modèles de $\langle \mathcal{R}, \mathrm{SupRelevant}(\mathcal{K}, \varphi) \rangle$.
3. Vérifie si φ appartient à tous les modèles.

Unsat(f) :

1. Calcule le graphe des déclencheurs.
2. Pour tout $\varphi \in \mathcal{Q}$, crée la KB $\mathcal{K}_\varphi^\perp = \langle \mathcal{R} \cup \{\varphi \to \perp\}, \mathrm{SupRelevant}(\mathcal{K}, \varphi) \rangle$.
3. Vérifie l'existence d'un modèle de $\mathcal{K}_\varphi^\perp$.

Résultats

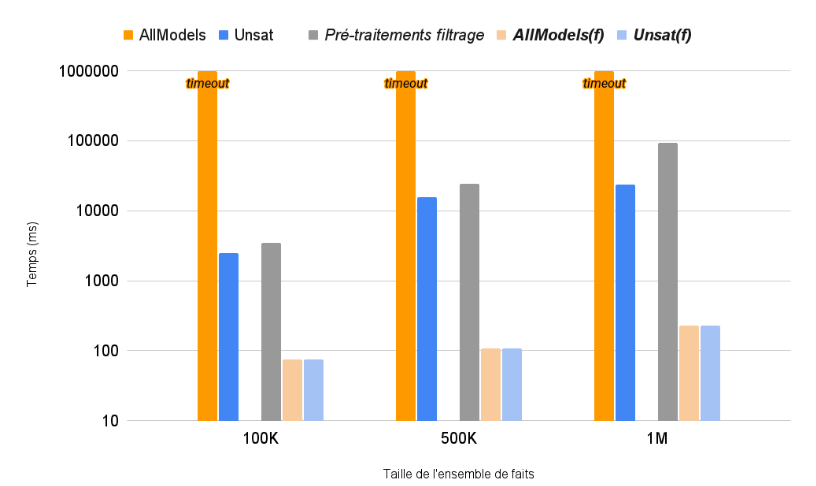
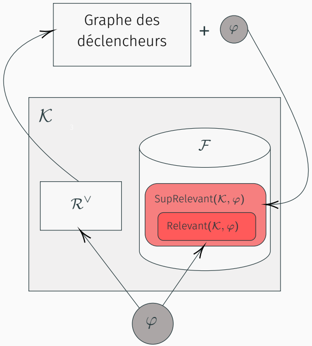


Figure 1: Coûts d'évaluation du jeu de requêtes avec filtrage (et pré-traitements) et sans en fonction de la taille de l'ensemble de faits

Conclusion et perspectives

Conclusion



1. Chase disjonctif
   - coûteux
   - filtrage
2. *Pertinence* des faits
   - problème difficile
3. Sur-approximation
   - graphe des déclencheurs

$\rightarrow$ impact sur le calcul des *ensembles minimaux* de faits et règles pour une inférence (basé sur SAT)

Merci de votre attention !

Graphe des déclencheurs

Définition

Soit une base de connaissances $\mathcal{K} = \langle \mathcal{R}, \mathcal{F} \rangle$, alors TgGraph($\mathcal{K}$) = $\langle V, E \rangle$ est le graphe orienté tel que:

- V est l'union des ensembles dans le résultat du chase sur $\mathcal{K}$.
- $E \subseteq V \times V$, et $(\psi, \phi) \in E$ s'il y a un déclencheur $\rho = (r, h)$ sur V tel que $\psi \in h(body(r))$ et il existe une conjonction α dans $head(r)$ telle que $\phi \in h(\alpha)$.

Un fait ψ est un *ancêtre* de ϕ s'il y a un chemin (possiblement vide) de ψ à ϕ dans le graphe des déclencheurs.

Remarque

On dit que ψ est un *i-ancêtre* de ϕ si le chemin est de longueur i, en particulier on utilise le terme prédécesseurs pour dénoter les 1-ancêtres.

Calcul du GDD et de l'approximation

Définition

Soit un ensemble de règles $\mathcal{R}$, $TransFunc(\mathcal{R})$ est l'ensemble minimal de règles contenant :

- $P(\bar{x}) \rightarrow P'(\bar{x}, f_P(\bar{x}))$, pour tout prédicat P figurant dans $\mathcal{R}^\wedge$
- $\bigwedge_{i=1}^{k} B'_i(\bar{x}_i, y_i) \ \rightarrow \ H'(\bar{x}, f_H(\bar{x})) \wedge \left(\bigwedge_{i=1}^{k} \boldsymbol{E}(y_i, f_H(\bar{x})) \right)$, pour toute règle déterministe $r = \bigwedge_{i=1}^{k} B_i(\bar{x}_i) \rightarrow H(\bar{x}) \in \mathcal{R}^\wedge$
- $\boldsymbol{\mathtt{Rel}}(x) \wedge \boldsymbol{E}(y, x) \rightarrow \boldsymbol{\mathtt{Rel}}(y)$

Remarque

Pour tout prédicat P dans $\mathcal{R}$, P' est un nouveau prédicat dans $TransFunc(\mathcal{R})$ d'arité $arit\acute{e}(P) + 1$.

Résultats d'évaluation - *Expérience 1 (coûts statiques)*

	Traduction (ms)	GDD (ms)	*Total (ms)*
100K	1	3 453	*3 454*
200K	1	5 906	*5 907*
300K	1	12 697	*12 698*
400K	3	17 577	*17 580*
500K	1	24 348	*24 349*
600K	3	28 632	*28 635*
700K	1	40 422	*40 423*
800K	2	58 972	*58 974*
900K	1	63 657	*63 658*
1M	2	94 662	*94 664*

Table 1: Coût des pré-traitements du filtrage en fonction de la taille de l'ensemble de faits

Résultats d'évaluation - *Expérience 1 (coûts dynamiques)*

	AllModels	AllModels^f				Unsat			Unsat^f			
	Exec. clingo (ms)	Traçage (ms)		Exec. clingo (ms)	*Total (ms)*	Exec. clingo (ms)		*Total (ms)*	Traçage (ms)		Exec. clingo (ms)	*Total (ms)*
		$\mathcal{Q}^{\models}$	$\mathcal{Q}^{\not\models}$			$\mathcal{Q}^{\models}$	$\mathcal{Q}^{\not\models}$		$\mathcal{Q}^{\models}$	$\mathcal{Q}^{\not\models}$		
100K	*timeout 20min*	19	19	38	***76***	620	1 883	***2 503***	19	19	38	***76***
200K	*timeout 20min*	20	20	40	***80***	1 604	4 668	***6 272***	20	20	40	***80***
300K	*timeout 20min*	20	20	40	***80***	2 184	6 255	***8 439***	20	20	40	***80***
400K	*timeout 20min*	20	20	40	***80***	3 088	8 649	***11 737***	20	20	40	***80***
500K	*timeout 20min*	49	20	40	***109***	4 801	10 693	***15 494***	49	20	40	***109***
600K	*timeout 20min*	91	20	40	***151***	4 303	11 841	***16 144***	91	20	40	***151***
700K	*timeout 20min*	52	29	40	***121***	5 013	13 781	***18 794***	52	29	40	***121***
800K	*timeout 20min*	109	43	40	***192***	6 306	16 786	***23 092***	109	43	40	***192***
900K	*timeout 20min*	120	24	40	***184***	5 972	16 470	***22 442***	120	24	40	***184***
1M	*timeout 20min*	139	53	40	***232***	6 893	16 813	***23 706***	139	53	40	***232***

Table 2: Expérience 1 : coûts d'évaluation du jeu de requêtes des algorithmes avec et sans filtrage en fonction de la taille de l'ensemble de faits

Résultats d'évaluation - *Expérience 2*

	Explications			Explications[f]					
	Pré-traitement	Eval. requêtes	*Total (ms)*	Pré-traitement		Eval. requêtes			*Total (ms)*
	Grounding (ms)	Exec. MARCO (ms)		Trad. (ms)	GDD (ms)	Grounding (ms)	Traçage (ms)	Exec. MARCO (ms)	
100K	2 335	613 324	***615 659***	1	3 453	41	38	2 776	***6 309***
200K	6 091	2 695 205	***2 701 296***	1	5 906	51	40	3 717	***9 715***
300K	10 592	4 250 504	***4 261 096***	1	12 697	66	40	3 456	***16 260***

Table 3: Expérience 2 : coût de pré-traitements et de calcul des MUSes avec et sans filtrage en fonction de la taille de l'ensemble de faits

SAT $\leq_p$ Fact Relevance($\mathcal{R}^\wedge$)

Formule CNF $\Phi = C_1 \wedge \ldots \wedge C_m$ définie en utilisant les variables $\{V_1, \ldots, V_n\}$.

Base de connaissances déterministe $\mathcal{K} = \langle \mathcal{R}, \mathcal{F} \rangle$ avec les prédicats d'arité 0 Source et Target.

Φ est satisfiable si et seulement si Source est pertinent pour l'inférence de Target.

$$\mathcal{F} = \{\mathrm{T}(v_i), \mathrm{F}(v_i) \mid 1 \leq i \leq n\} \cup \{\mathrm{Source}\}$$

$$\mathcal{R} = \{\mathrm{T}(v_k) \rightarrow \mathrm{T}(c_i) \mid 1 \leq i \leq m \text{ et pour tout littéral positif } V_k \text{ dans } C_i\} \cup \tag{1}$$

$$\{\mathrm{F}(v_k) \rightarrow \mathrm{T}(c_i) \mid 1 \leq i \leq m \text{ et pour tout littéral négatif } \neg V_k \text{ dans } C_i\} \cup \tag{2}$$

$$\{\mathrm{Source} \wedge \bigwedge_{i=1}^{m} \mathrm{T}(c_i) \rightarrow \mathrm{Target}\} \cup \tag{3}$$

$$\{\mathrm{T}(x) \wedge \mathrm{F}(x) \rightarrow \mathrm{Target}\} \tag{4}$$

Preuves *difficulté* - Déterministe (2)

(Correction) On montre que si Source est pertinent pour l'inférence de Target avec $\mathcal{K}$ alors la formule CNF Φ est satisfiable.

ⓐ Par hypothèse il y a un ensemble minimal $\mathcal{A} \subseteq \mathcal{F}$ tel que Source $\in \mathcal{A}$ et $\langle \mathcal{R}, \mathcal{A} \rangle \models$ Target.

ⓑ On montre d'abord que pour toute constante v_i figurant dans $\mathcal{A}$, on a $\mathrm{T}(v_i) \in \mathcal{A}$ si et seulement si $\mathrm{F}(v_i) \notin \mathcal{A}$. En effet, supposons que $\mathrm{T}(v_i), \mathrm{F}(v_i) \in \mathcal{A}$ pour une constante v_i. Alors, avec la règle (4) on a $\langle \mathcal{R}, \{\mathrm{T}(v_i), \mathrm{F}(v_i)\} \rangle \models$ Target. Ce qui contredit le fait que $\mathcal{A}$ est un ensemble minimal qui, avec $\mathcal{R}$, infère Target.

ⓒ Soit $\sigma_{\mathcal{A}}$ l'assignation telle que $\sigma_{\mathcal{A}}(V_i) = \mathit{Vrai}$ si et seulement si $\mathrm{T}(v_i) \in \mathcal{A}$ pour toute variable V_i dans Φ et sa constante correspondante v_i dans $\mathcal{K}$.

ⓓ Par ⓑ, on sait que (4) ne peut pas être appliquée sur $\mathcal{A}$. Donc, la règle (3) a été appliquée pour inférer Target.
Donc, $\langle \mathcal{R}, \mathcal{A} \rangle \models \mathrm{T}(c_i)$ pour tout $1 \leq i \leq m$.

ⓔ Pour $1 \leq i \leq m$ on sait que $\mathrm{T}(c_i)$ a été inféré soit par une règle dans (1) ou dans (2).
  - Si une règle de la forme (1) a été appliquée alors il y a une constante v_k tel que $\mathrm{T}(v_k) \in \mathcal{A}$ et $\mathrm{T}(v_k) \rightarrow \mathrm{T}(c_i) \in \mathcal{R}$. Soit V_k la variable correspondant à v_k et C_i la clause correspondant à c_i dans Φ, alors par la construction des règles dans (1) on a $V_k \in C_i$ et par ⓒ on sait que $\sigma_{\mathcal{A}}(V_k) = \mathit{Vrai}$, donc $\sigma_{\mathcal{A}}(C_i) = \mathit{Vrai}$.
  - Si une règle de la forme (2) a été appliquée alors il y a une constante v_k tel que $\mathrm{F}(v_k) \in \mathcal{A}$ et $\mathrm{F}(v_k) \rightarrow \mathrm{T}(c_i) \in \mathcal{R}$. Soit V_k la variable correspondant à v_k et C_i la clause correspondant à c_i dans Φ, alors par la construction des règles dans (2) on a $\neg V_k \in C_i$ et par ⓒ on sait que $\sigma_{\mathcal{A}}(V_k) = \mathit{Faux}$ donc $\sigma_{\mathcal{A}}(C_i) = \mathit{Vrai}$.

Donc, $\sigma_{\mathcal{A}}(C_i) = \mathit{Vrai}$ pour toute clause $C_i \in \Phi$, ainsi $\sigma(\Phi) = \mathit{Vrai}$ et Φ est satisfiable.

Preuves *difficulté* - Déterministe (3)

(Complétude) On montre que si Φ est satisfiable alors Source est pertinent pour Target avec $\mathcal{K}$.

ⓐ Par hypothèse il existe une assignation σ telle que σ est un modèle de Φ. Soit l'ensemble de faits $\mathcal{A} = \{\text{Source}\} \cup \{\text{T}(v_i) \mid \sigma(V_i) = \textit{Vrai}\} \cup \{\text{F}(v_i) \mid \sigma(V_i) = \textit{Faux}\} \subseteq \mathcal{F}$.

ⓑ Par hypothèse, on a $\sigma(C_i) = \textit{Vrai}$ pour toute clause C_i dans Φ et on distingue deux cas.

- Il y a un littéral positif $V_k \in C_i$ tel que $\sigma(V_k) = \textit{Vrai}$. Dans ce cas, soient v_k et c_i les constantes correspondant à V_k et C_i respectivement. Par ⓐ on a $\text{T}(v_k) \in \mathcal{A}$, et $\text{T}(v_k) \rightarrow \text{T}(c_i) \in \mathcal{R}$ par la construction des règles dans (1).
- Il y a un littéral négatif $\neg V_k \in C_i$ tel que $\sigma(V_k) = \textit{Faux}$. Soient v_k et c_i les constantes correspondant à V_k et C_i respectivement. Par ⓐ on a $\text{F}(v_k) \in \mathcal{A}$, et $\text{F}(v_k) \rightarrow \text{T}(c_i) \in \mathcal{R}$ par la construction des règles dans (2).

Dans les deux cas, $\langle \mathcal{R}, \mathcal{A} \rangle \models \text{T}(c_i)$ pour tout $1 \leq i \leq m$.

ⓒ Soit $\mathcal{A}'$ un ensemble minimal de $\mathcal{A}$ tel que $\langle \mathcal{R}, \mathcal{A}' \rangle \models$ Target. On sait que (3) et (4) sont les seules règles pouvant inférer Target. En revanche, (4) ne peut pas être appliquée sur $\mathcal{A}'$. En effet, $\mathcal{A}$ est construit à partir d'une assignation σ et ainsi soit $\text{T}(v_i)$ ou $\text{F}(v_i)$ appartient à $\mathcal{A}'$ pour toute constante v_i correspondant à une variable V_i dans Φ. Donc la règle (3) a été appliquée pour dériver Target.

ⓓ Pour conclure, notons que pour appliquer (3) l'atome Source doit appartenir à $\mathcal{A}'$. Donc, Source est pertinent pour Target.

(Polynomialité) On montre que la taille de $\mathcal{K}$ est polynomiale en fonction de la taille de Φ. Dans l'ensemble de faits $\mathcal{F}$ on a $2n + 1$ atomes avec n étant le nombre de variables propositionnelles dans Φ. Ensuite l'ensemble de règles contient au plus $k \times m + 2$ règles où m est le nombre de clauses dans Φ et k le plus grand nombre de littéraux dans une clause. Donc l'instance $\mathcal{K}$ est polynomiale par rapport à Φ et la traduction peut être calculée en temps polynomial.

Preuves *difficulté* - Non-déterministe (1)

$\exists\forall$-TQBF

INSTANCE: Une formule $\exists\forall$-TQBF Φ de la forme $\exists\bar{x_1}\forall\bar{x_2}.\Phi[\bar{x_1}, \bar{x_2}]$.
QUESTION: Est-ce que Φ est *satisfiable*, i.e., existe-t-il une assignation $\sigma_\exists$ des variables de $\bar{x_1}$ telle que pour toute assignation $\sigma_\forall$ des variables de $\bar{x_2}$, on a $\sigma_\exists \cup \sigma_\forall(\Phi) = \mathit{Vrai}$?

Preuves *difficulté* - Non-déterministe (2)

$\exists\forall$-TQBF $\leq_p$ Fact Relevance($\mathcal{R}$)

Formule $\exists\forall$-TQBF Φ sur les variables $\{V_1, \ldots, V_n\}$ et clauses $\{C_1, \ldots, C_m\}$.

Base de connaissances non-déterministe $\mathcal{K} = \langle \mathcal{R}, \mathcal{F} \rangle$ avec les prédicats d'arité 0 Source et Target.

Φ est satisfiable si et seulement si Source est pertinent pour l'inférence de Target.

$$\mathcal{F} = \{\mathrm{T}(v_i), \mathrm{F}(v_i) \mid \text{ pour toute variable } V_e \text{ quantifiée existentiellement dans } \boldsymbol{\Phi}\} \cup \{\mathrm{Source}\}$$

$$\mathcal{R} = \{\mathrm{T}(v_k) \rightarrow \mathrm{T}(c_i) \mid 1 \leq i \leq m \text{ et pour tout littéral positif } V_k \text{ dans } C_i\} \ \cup \tag{5}$$

$$\{\mathrm{F}(v_k) \rightarrow \mathrm{T}(c_i) \mid 1 \leq i \leq m \text{ et pour tout littéral négatif } \neg V_k \text{ in } C_i\} \ \cup \tag{6}$$

$$\{\mathrm{Source} \wedge \bigwedge_{i=1}^{m} \mathrm{T}(c_i) \rightarrow \mathrm{Target}\} \ \cup \tag{7}$$

$$\{\mathrm{T}(x) \wedge \mathrm{F}(x) \rightarrow \mathrm{Target}\} \ \cup \tag{8}$$

$$\{ \rightarrow \mathrm{T}(v_u) \vee \mathrm{F}(v_u) \mid \text{ pour toute variable universelle } V_u \text{ dans } \boldsymbol{\Phi} \ \} \tag{9}$$

Preuves *difficulté* - Non-déterministe (3)

(Correction) On montre que si Source est pertinent pour l'inférence de Target avec $\mathcal{K}$ alors la formule $\exists\forall$-TQBF Φ est satisfiable :

ⓐ Par la définition de pertinence il y a un ensemble minimal $\mathcal{A} \subseteq \mathcal{F}$ tel que Source $\in \mathcal{A}$ et $\langle \mathcal{R}, \mathcal{A} \rangle \models$ Target.

ⓑ De plus, pour toute constante v_i dans $\mathcal{A}$, $\mathsf{T}(v_i) \in \mathcal{A}$ si et seulement si $\mathsf{F}(v_i) \notin \mathcal{A}$. En effet, par définition de $\mathcal{F}$ si v_i apparaît dans $\mathcal{A}$ alors seulement un parmi $\mathsf{T}(v_i)$ et $\mathsf{F}(v_i)$ appartient à $\mathcal{A}$. Maintenant supposons que $\mathsf{T}(v_i)$ et $\mathsf{F}(v_i)$ appartiennent à $\mathcal{A}$. Dans ce cas, par la règle (8), on a $\langle \mathcal{R}, \{\mathsf{T}(v_i), \mathsf{F}(v_i)\} \rangle \models$ Target. Ceci contredit le fait que $\mathcal{A}$ est minimal et que Source est pertinent pour Target.

ⓒ Soit $\sigma^{\exists}_{\mathcal{A}}$ l'assignation associée à $\mathcal{A}$. Elle est définie comme suit. Soit v_e une constante dans $\mathcal{F}$ et V_e sa variable existentiellement quantifiée correspondante dans Φ. Alors, $\sigma^{\exists}_{\mathcal{A}}(V_e) = \mathit{Vrai}$ si et seulement si $\mathsf{T}(v_e) \in \mathcal{A}$. Il s'ensuit que $\sigma^{\exists}_{\mathcal{A}}(V_e) = \mathit{Faux}$ si $\mathsf{F}(v_e) \in \mathcal{A}$ ou v_e n'apparaît pas dans $\mathcal{A}$.

ⓓ Soit R le résultat du chase sur $\langle \mathcal{R}, \mathcal{A} \rangle$.
On montre que pour toute assignation totale $\sigma \supseteq \sigma^{\exists}_{\mathcal{A}}$ des variables dans Φ il y a un $\mathcal{S}_\sigma \in R$ tel que, pour toute variable universelle V_u dans Φ, soit v_u sa constante correspondante dans $\mathcal{R}$ et il est vrai que $\sigma(V_u) = \mathit{Vrai}$ si et seulement si $\mathsf{T}(v_u) \in \mathcal{S}_\sigma$ et $\sigma(V_u) = \mathit{Faux}$ si et seulement si $\mathsf{F}(v_u) \in \mathcal{S}_\sigma$.
Par les règles dans (9), pour toute variable universelle V_u et tout noeud n dans l'arbre du chase, soit n contient déjà $\mathsf{T}(v_u)$ ou $\mathsf{F}(v_u)$, ou il y a une application de règle créant deux noeuds enfants avec l'un contenant $\mathsf{T}(v_u)$ et l'autre $\mathsf{F}(v_u)$. De plus, pour tout $\mathcal{S} \in R$, aucune application de règle n'est possible. Donc pour tout assignation σ il existe un $\mathcal{S}_\sigma \in R$ correspondant.

Preuves *difficulté* - Non-déterministe (4)

ⓔ Par hypothèse, Target $\in \mathcal{S}$ pour tout $\mathcal{S} \in R$ et donc Target $\in \mathcal{S}_\sigma$ pour toute assignation totale σ. Par ⓑ on sait que (8) n'a pas été appliquée. Donc Target a été dérivé avec la règle (7). Ainsi pour tout $\mathcal{S}_\sigma$ on a $\mathsf{T}(c_i) \in \mathcal{S}_\sigma$ pour toute c_i correspondant à une clause C_i dans Φ.

ⓕ Si $\mathsf{T}(c_i) \in \mathcal{S}_\sigma$ alors soit une règle du type (5) ou (6) a été appliquée.

- Si une règle dans (5) a été appliquée alors il existe une constante v_k tel que $\mathsf{T}(v_k)$ appartient à $\mathcal{S}_\sigma$ et la règle $\mathsf{T}(v_k) \to \mathsf{T}(c_i)$ appartient à $\mathcal{R}$. Par définition de (5) on sait que v_k correspond à une variable V_k étant un littéral positif dans la clause C_i correspondant à c_i. Donc $\sigma(V_k) = \mathit{Vrai} = \sigma(C_i)$.
- Si une règle dans (6) a été appliquée alors il existe une constante v_k tel que $\mathsf{F}(v_k)$ appartient à $\mathcal{S}_\sigma$ et la règle $\mathsf{F}(v_k) \to \mathsf{T}(c_i)$ appartient à $\mathcal{R}$. Par définition de (6) on sait que v_k correspond à une variable V_k étant un littéral négatif dans la clause C_i correspondant à c_i. Donc $\sigma(V_k) = \mathit{Faux}$ et $\sigma(C_i) = \mathit{Vrai}$.

Donc, $\sigma(\Phi) = \mathit{Vrai}$ pour tout assignation totale $\sigma \supseteq \sigma^{\exists}_{\mathcal{A}}$ des variables dans Φ.

Preuves *difficulté* - Non-déterministe (5)

(Complétude) On montre que si Φ est satisfiable alors Source est pertinent pour Target avec $\mathcal{K}$:

ⓐ Par hypothèse il y a une assignation $\sigma_\exists$ telle que pour toute assignation totale $\sigma \supseteq \sigma_\exists$ des variables dans Φ on a $\sigma(\Phi) = \mathit{Vrai}$.

ⓑ Soit l'ensemble de faits
$\mathcal{A} = \{\mathrm{Source}\} \cup \{\mathrm{T}(v) \mid \sigma_\exists(V) = \mathit{Vrai}\} \cup \{\mathrm{F}(v) \mid \sigma_\exists(V) = \mathit{Faux}\} \subseteq \mathcal{F}$.

ⓒ Soit R le résultat du chase sur $\langle \mathcal{R}, \mathcal{A} \rangle$. Puisque le chase est complet, pour tout $\mathcal{S} \in R$ les règles dans (9) sont appliquées et soit $\mathrm{T}(v_u) \in \mathcal{S}$ ou $\mathrm{F}(v_u) \in \mathcal{S}$ pour toute constante v_u correspondant à une variable universelle V_u dans Φ.

ⓓ Soit $\sigma \supseteq \sigma_\exists$ une assignation totale des variables dans Φ et $\mathcal{S}_\sigma \in R$ telle que $\mathrm{T}(v) \in \mathcal{S}_\sigma$ si et seulement si $\sigma(V) = \mathit{Vrai}$ pour toute variable V dans Φ et sa constante correspondante v dans $\mathcal{K}$.

ⓔ Par hypothèse $\sigma(C_i) = \mathit{Vrai}$ pour tout $1 \leq i \leq m$ et on distingue deux cas.

- Il y a un littéral positif $V \in C_i$ et $\sigma(V) = \mathit{Vrai}$. Soit v et c_i les constantes correspondantes de V et C_i respectivement, par ⓓ on a $\mathrm{T}(v) \in \mathcal{S}_\sigma$ et par la construction des règles dans (5) on a $\mathrm{T}(v) \to \mathrm{T}(c_i) \in \mathcal{R}$. Donc, $\mathrm{T}(c_i) \in \mathcal{S}_\sigma$.
- Il y a un littéral négatif $\neg V \in C_i$ et $\sigma(V) = \mathit{Faux}$. Soit v et c_i les constantes correspondantes de V et C_i respectivement, par ⓓ on a $\mathrm{F}(v) \in \mathcal{S}_\sigma$ et par la construction des règles dans (6) on a $\mathrm{F}(v) \to \mathrm{T}(c_i) \in \mathcal{R}$. Donc, $\mathrm{T}(c_i) \in \mathcal{S}_\sigma$.

Donc, $\mathrm{T}(c_i) \in \mathcal{S}_\sigma$ pour tout $1 \leq i \leq m$. Ainsi, par la règle (7) Target $\in \mathcal{S}_\sigma$.

Preuves *difficulté* - Non-déterministe (6)

ⓕ Par définition de $\mathcal{R}$ les seules règles pouvant inférer Target sont (7) et (8). Mais (8) ne peut pas être appliquée sur $\mathcal{A}$ car tout $\mathcal{S}_\sigma \in R$ est associé à une assignation totale correcte, c'est-à-dire, pour toute constante v_i représentant une variable V_i de Φ alors soit $\mathrm{T}(v_i) \in \mathcal{S}_\sigma$ ou $\mathrm{F}(v_i) \in \mathcal{S}_\sigma$. Donc l'unique règle applicable est (7).

ⓖ On conclut en montrant que Source est pertinent pour Target. Par contradiction, supposons qu'il existe $\mathcal{A}' \subseteq \mathcal{A}$ tel que Source $\notin \mathcal{A}'$ et $\langle \mathcal{R}, \mathcal{A}' \rangle \models$ Target. Alors la règle (7) n'est pas applicable car Source $\notin \mathcal{A}'$. Donc $\langle \mathcal{R}, \mathcal{A}' \rangle \models$ Target n'est pas vrai. Finalement tout ensemble minimal de $\mathcal{A}' \subseteq \mathcal{A}$ tel que $\langle \mathcal{R}, \mathcal{A}' \rangle \models$ Target contient Source et ainsi Source est pertinent pour Target.

(Polynomialité) Montrer que la traduction est polynomiale peut être fait facilement de façon similaire à la preuve dans le cadre déterministe.