

**Opérateurs naturels
en Programmation Logique Inductive**

Fabien Torre & Céline Rouveirol

Groupe Inférence et Apprentissage
Laboratoire de Recherche en Informatique
Université Paris-Sud, Orsay

Plan de la présentation

- Les opérateurs classiques ne permettent pas un élagage dynamique par rapport aux biais de langage.
- Définition de nouveaux pré-ordres, les *relations naturelles*, qui permettent cet élagage.
- Un opérateur idéal [van der Laag and Nienhuys-Cheng, 1994] peut-il exister pour les relations naturelles ?
- Conclusion, Perspectives.

Programmation Logique Inductive

Exemples et hypothèses sont des clauses définies.

[Muggleton and Raedt, 1994] : Étant donnés E^+ (des exemples positifs) et E^- (des exemples négatifs) pour le concept cible, et une théorie du domaine B , trouver une hypothèse H telle que

$$\forall e^+ \in E^+ : B \cup H \models e^+ \quad (H \text{ est complète}) ,$$

$$\forall e^- \in E^- : B \cup H \not\models e^- \quad (H \text{ est correcte}) .$$

Exemple : le concept grand-père

$pa(A,B) \leftarrow p(A,B)$ (*pa=parent, p=père, m=mère*)

$pa(A,B) \leftarrow m(A,B)$

$gp(abraham,bart) \leftarrow p(abraham,homer),p(homer,bart)$

$gp(grampa-bouvier,bart) \leftarrow p(grampa-bouvier,marge),m(marge,bart)$

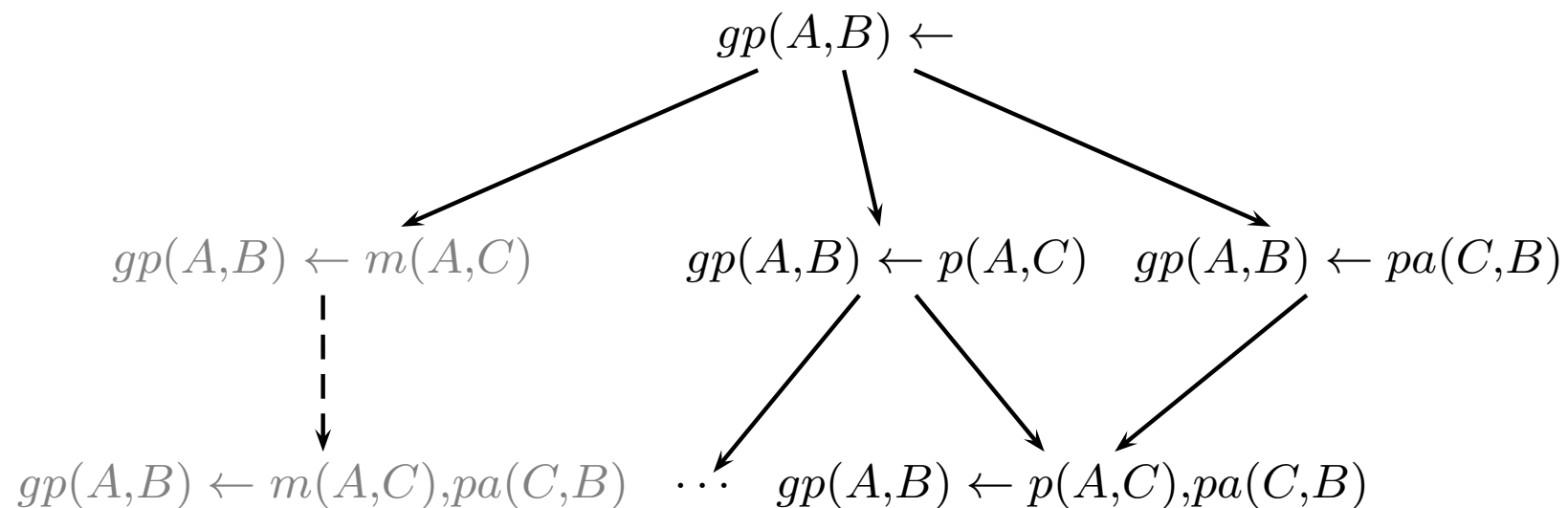
$\leftarrow gp(mona,bart),m(mona,homer),p(homer,bart)$

$\leftarrow gp(jackie,bart),m(jackie,marge),m(marge,bart)$

$gp(A,B) \leftarrow p(A,C),pa(C,B)$

Opérateurs de raffinement & Élagage

Exemple : ajout d'un littéral.

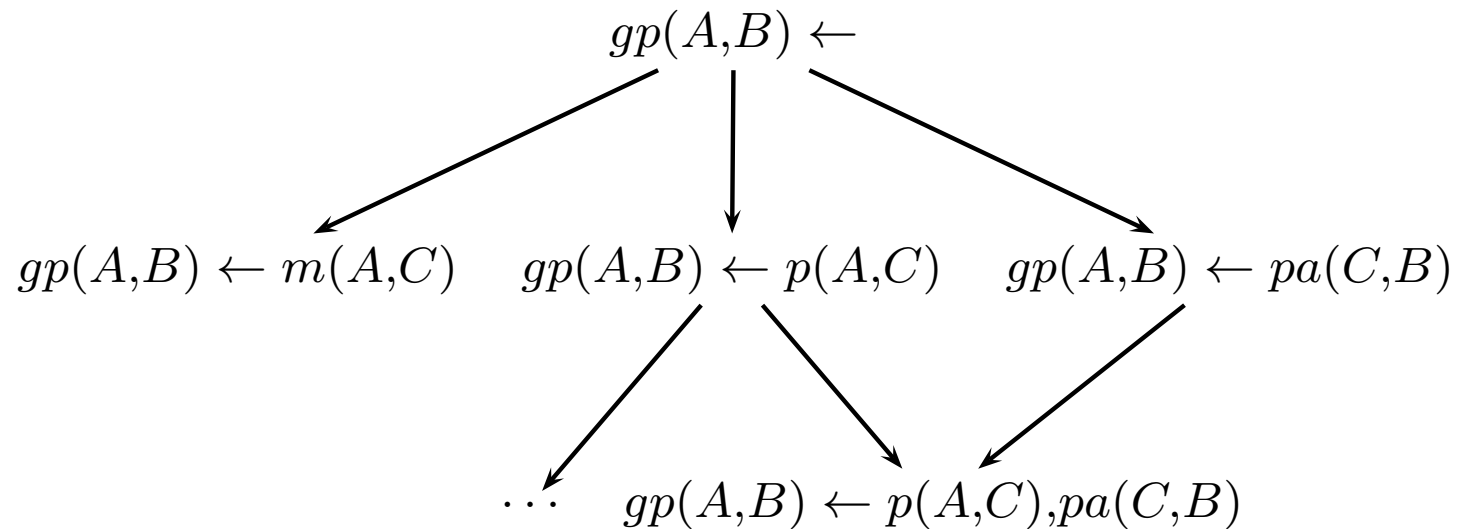


[Mitchell, 1982]: La recherche doit respecter un ordre de généralité pour permettre un élagage par rapport à la couverture des exemples.

Biais de langage

- [Mitchell, 1991] : Les biais sont *nécessaires* à l'apprentissage (qualité des résultats et efficacité).
- Biais de langage : contraintes sur la syntaxe des hypothèses.
 - champ réduit, connexion,
 - borne sur la taille, le nombre de variables, la profondeur, etc.
- Ces biais ne rendent pas la recherche plus efficace : l'élagage dynamique est généralement impossible par rapport à ces contraintes.

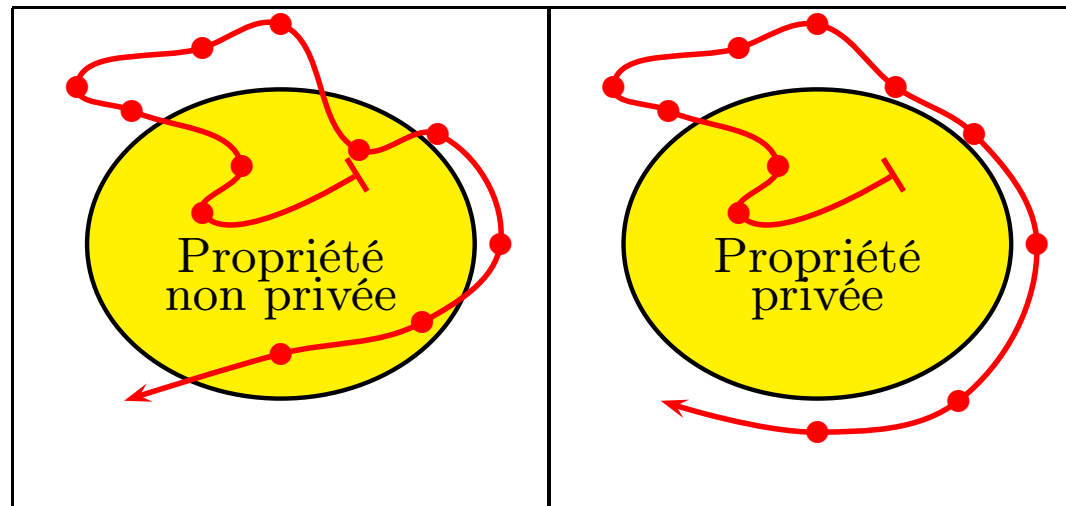
Exemple (suite)



Biais de langage : champ réduit.

P_1 : couverture de e^+ P_2 : rejet de e^- P_3 : champ réduit

Propriétés Privées



Définition 1 (propriété privée) Une propriété P est dite privée par rapport à une relation \mathcal{R} si et seulement si

$$H \mathcal{R} H' \wedge \overline{P(H)} \Rightarrow \overline{P(H')}$$

$\mathcal{R} = \text{ajout}$, $P = \text{couverture de } e^+$ ou $P = \text{taille bornée}$.

Relations naturelles

Relation naturelle d'une propriété P : une relation, la plus large possible, pour laquelle P est privée.

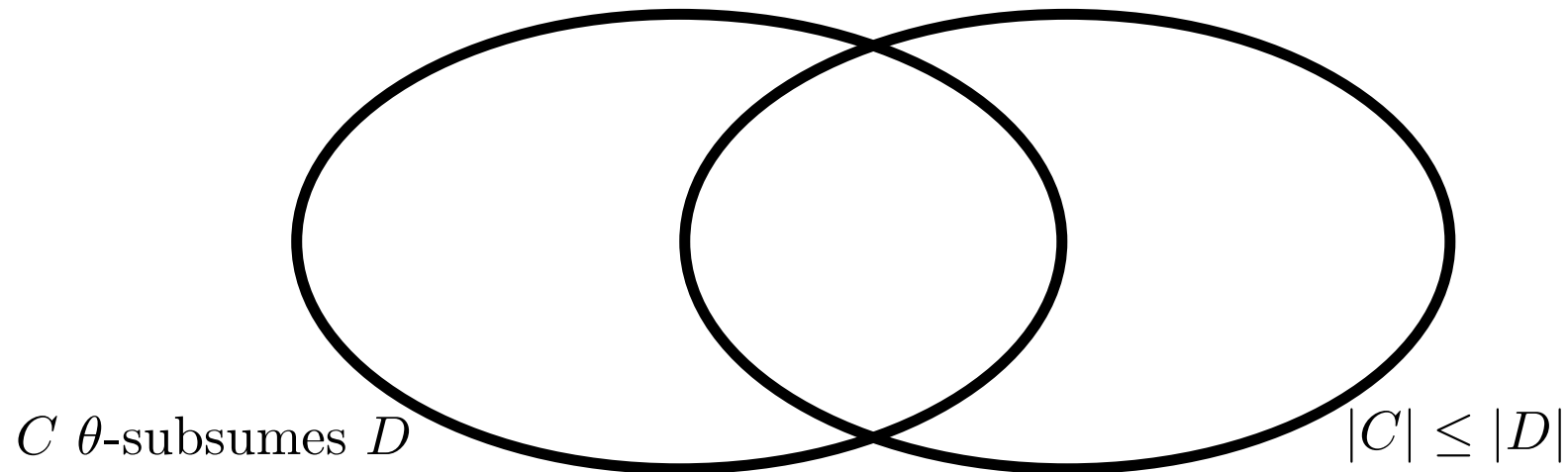
- La relation naturelle d'une propriété P est unique et toute relation qui rend P privée est incluse dans cette relation naturelle.
- Deux hypothèses C et D sont en relation naturelle pour une propriété $f(H) \mathcal{R} k$ si et seulement si $f(C) \mathcal{R} f(D)$.

$$H \models e^+ \dots\dots\dots C \models D$$

$$|H| \leq_{\mathbb{N}} 5 \dots\dots\dots |C| \leq_{\mathbb{N}} |D|$$

Conjonctions de propriétés

H θ -subsume e^+ et $|H| \leq k$.



Définition 2 C θ -subsume D si et seulement si il existe une substitution θ telle que $C\theta \subseteq D$.

Opérateurs idéaux

[van der Laag and Nienhuys-Cheng, 1994]

Définition 3 (idéauté) *Un opérateur \mathcal{O} est idéal s'il est à la fois localement fini, strict et complet.*

Localement fini : $\mathcal{O}(H)$ est calculable ;

Strict : $\mathcal{O}(H)$ ne contient aucune clause équivalente à H ;

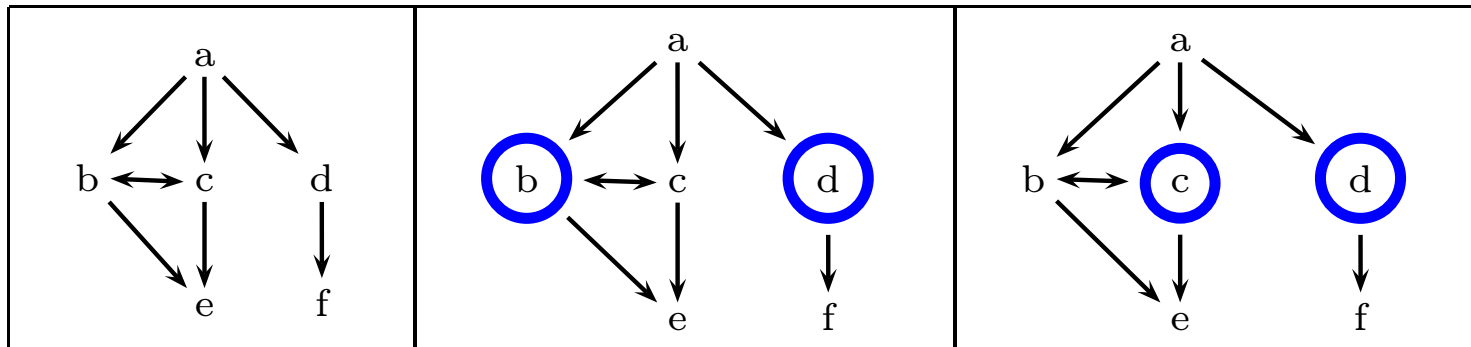
Complet : $\mathcal{O}^*(H)$ contient toutes les clauses comparables à H .

Les opérateurs idéaux n'existent pas sous θ -subsumption ou implication logique.

Couvertures

Définition 4 (couverture) C couvre D si $C > D$ et il n'existe pas de E telle que $C > E > D$. C est une couverture supérieure de D , D un couverture inférieure de C .

Définition 5 (ensemble couvrant) Un ensemble couvrant inférieur (resp. supérieur) d'une clause C est un ensemble maximal de couvertures inférieures (resp. supérieures) de C , incomparables entre elles.

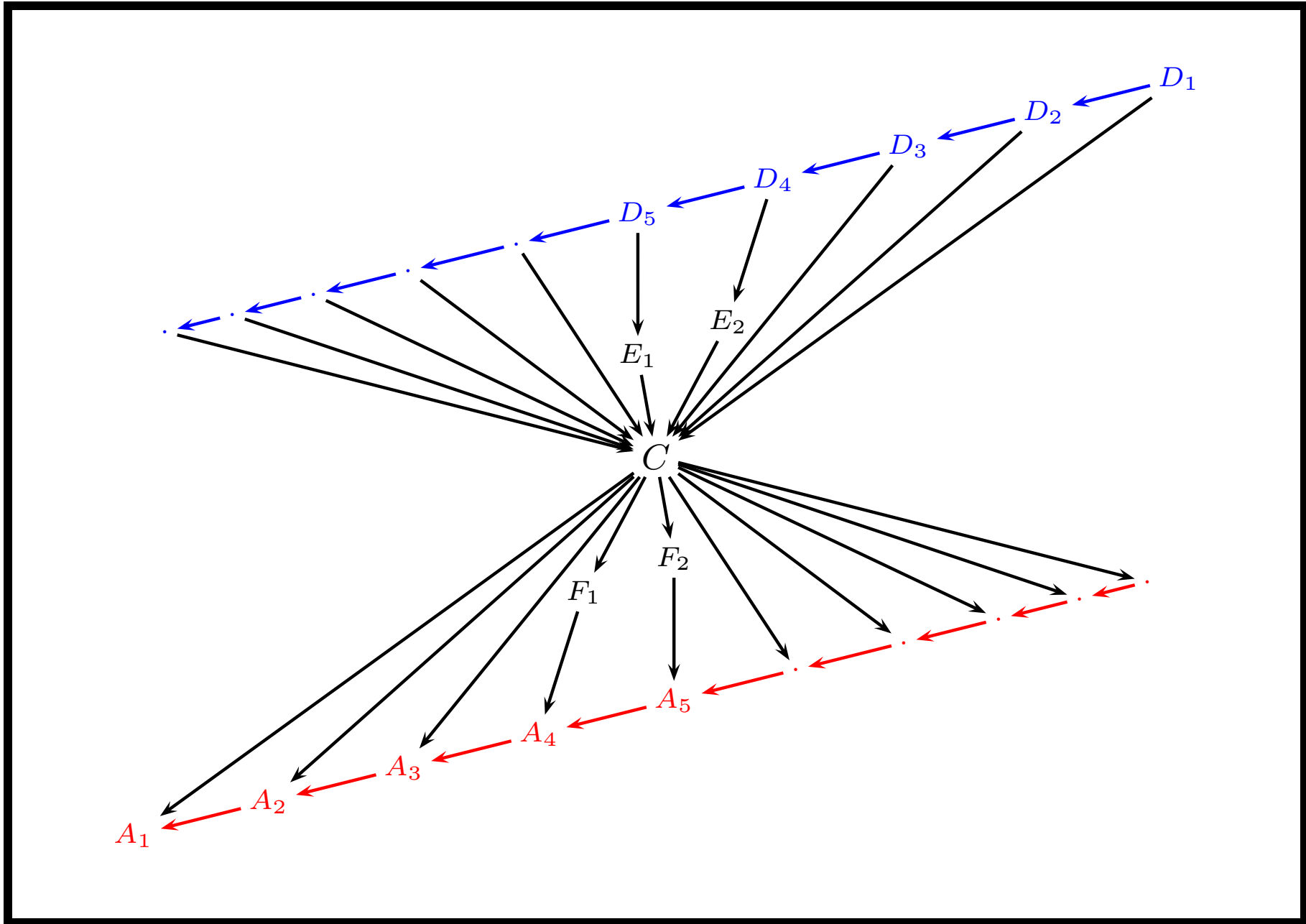


Non-existence

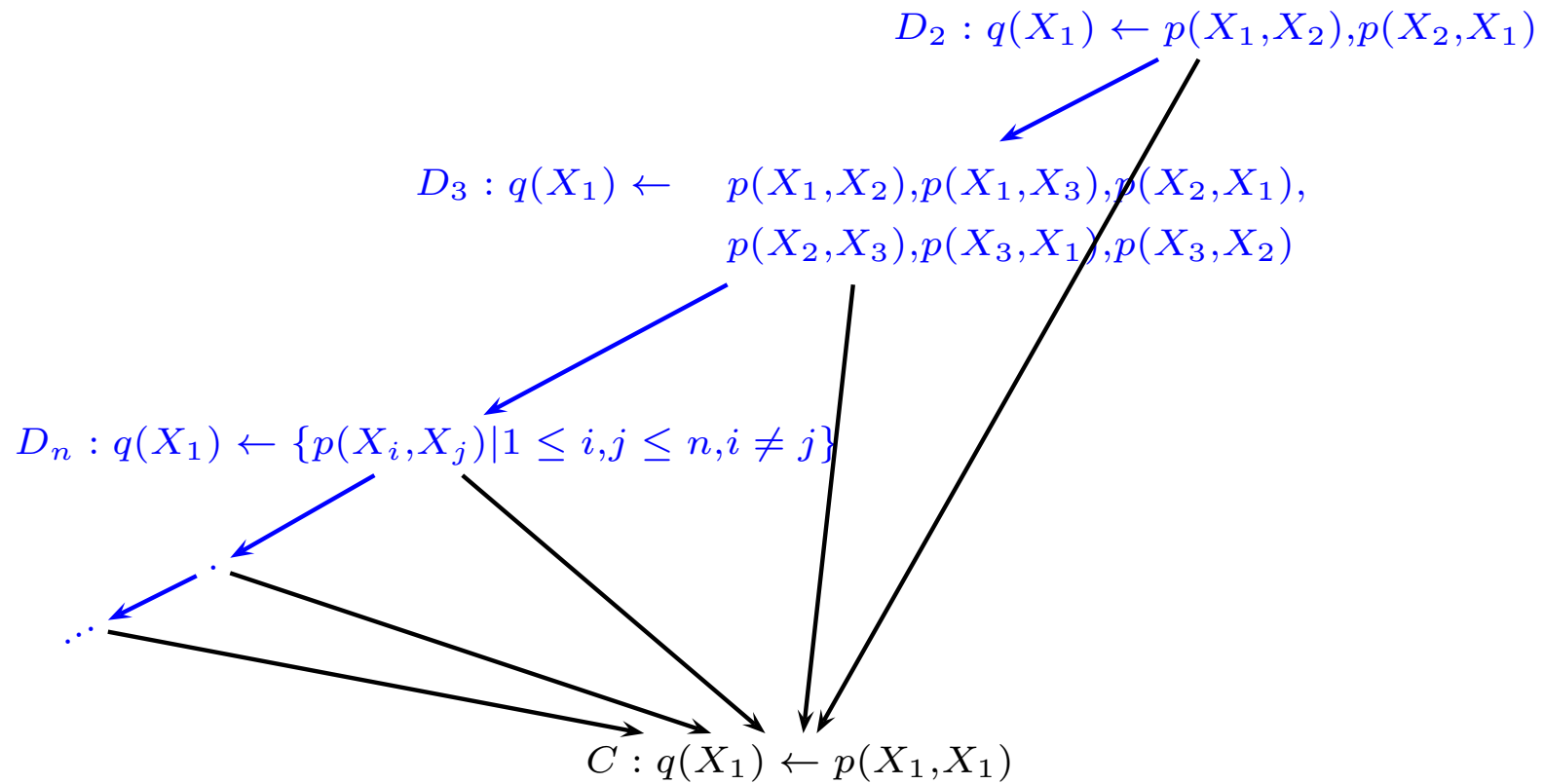
Un opérateur idéal calcule au moins un ensemble couvrant.

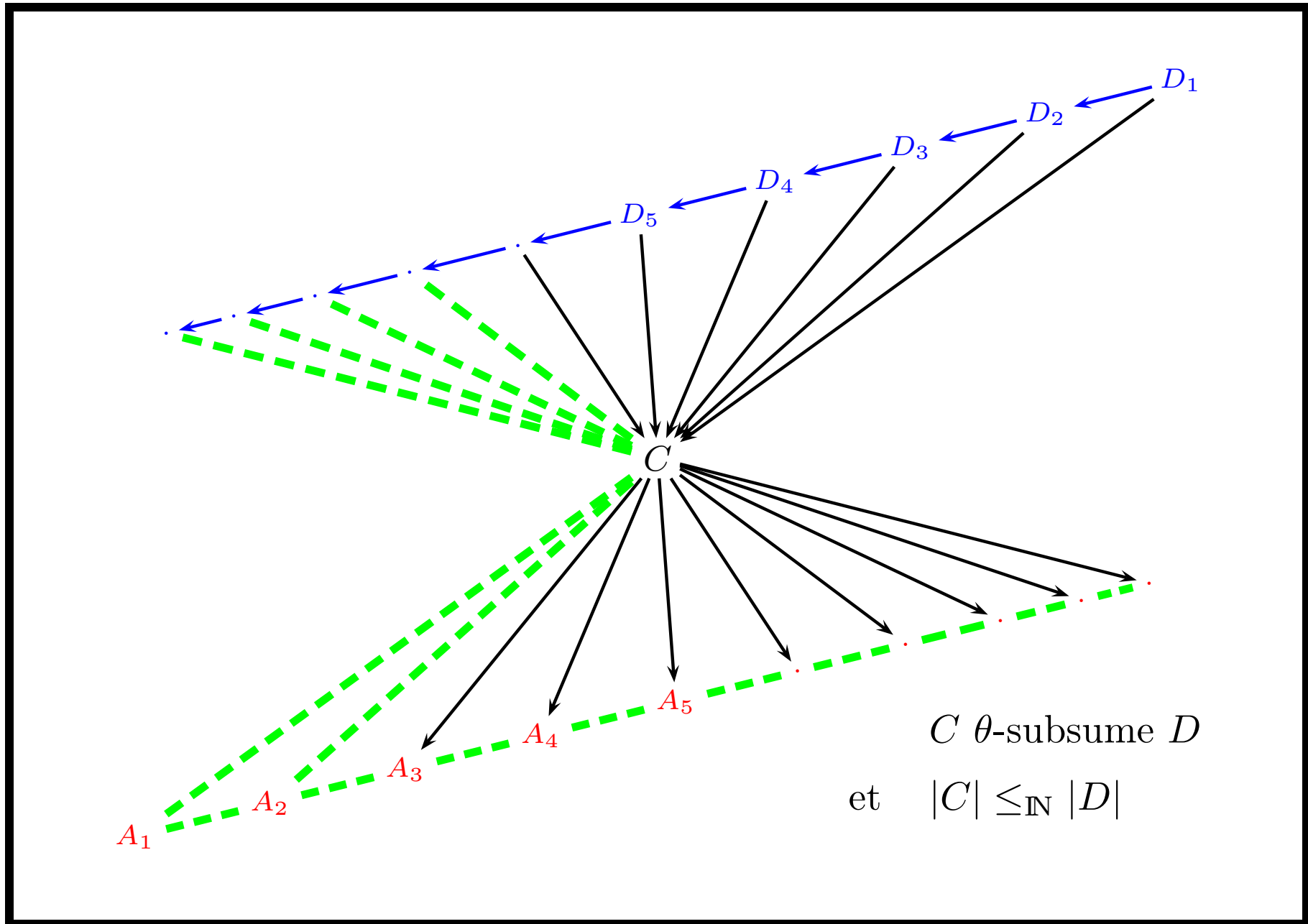
Par suite, il y a deux problèmes possibles.

1. L'ensemble couvrant n'est pas défini, il y a une chaîne infinie non couverte.
2. L'ensemble couvrant est infini, un opérateur complet n'est pas calculable.



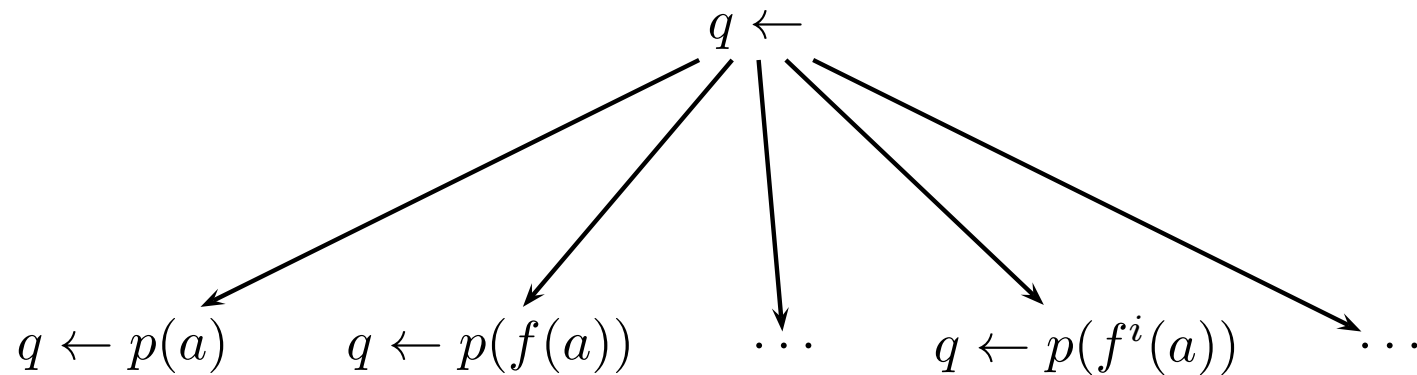
$$\begin{cases} C & : q(X_1) \leftarrow p(X_1, X_1) \text{ ,} \\ D_n & : q(X_1) \leftarrow \{p(X_i, X_j) \mid 1 \leq i, j \leq n, i \neq j\} \text{ .} \end{cases}$$





Ensembles couvrants infinis

C θ -subsume D et $N\text{Vars}(C) \leq_{\mathbb{N}} N\text{Vars}(D)$



Un opérateur idéal : $\rho_{||}^{\theta}(C)$

1. Ajouter à C un littéral dont les arguments sont de nouvelles variables distinctes.
2. Unifier deux variables X_1 et X_2 de C , telles que $C >^{\theta} C\{X_1/X_2\}$. Si la taille de la clause-résultat a diminué, ajouter un littéral ne contenant que des nouvelles variables.
3. Appliquer l'une des précédentes transformations (1, 2) sur une clause équivalente à C (θ -équivalente et de même taille).
4. Appliquer la transformation 1 sur un sous-ensemble de C θ -équivalent à C , l'un des nouveaux littéraux utilisant un symbole de prédicat qui n'apparaissait pas dans C .

Conclusion

- Élagage dynamique par rapport aux biais de langage basé sur l'utilisation de nouveaux pré-ordres : les *relations naturelles*.
- Existence d'opérateurs idéaux pour des espaces non réduits ordonnés par les relations naturelles.

Autres travaux : • [Shapiro, 1981]
• [Champesme et al., 1995, Esposito et al., 1996]

Perspectives : considérer d'autres familles d'opérateurs (opérateurs *optimaux* de [De Raedt and Bruynooghe, 1993]).

Références

- [Champesme et al., 1995] Champesme, M., Brézellec, P., and Soldano, H. (1995). Empirically conservative search space reductions. In Raedt, L. D., editor, *Proceedings of the 5th International Workshop on Inductive Logic Programming*, pages 387–402. Department of Computer Science, Katholieke Universiteit Leuven.
- [De Raedt and Bruynooghe, 1993] De Raedt, L. and Bruynooghe, M. (1993). A theory of clausal discovery. pages 1058–1063. Morgan Kaufmann.
- [Esposito et al., 1996] Esposito, F., Laterza, A., Malerba, D., and Semeraro, G. (1996). Refinement of datalog programs. In *Proceedings of the MLnet Familiarization Workshop on Data Mining with Inductive Logic Programming (ILP for KDD)*, pages 73–94.
- [Mitchell, 1982] Mitchell, T. M. (1982). Generalization as search. *Artificial Intelligence*, 18:203–226.
- [Mitchell, 1991] Mitchell, T. M. (1991). The need for biases in learning generalizations. In *Readings in Machine Learning*. Morgan Kaufmann.
- [Muggleton and Raedt, 1994] Muggleton, S. and Raedt, L. D. (1994). Inductive logic programming: Theory and methods. *Journal of Logic Programming*, 19:629–679.
- [Shapiro, 1981] Shapiro, E. Y. (1981). Inductive inference of theories from facts. Technical Report 192, Yale University Department of Computer Science.
- [van der Laag and Nienhuys-Cheng, 1994] van der Laag, P. and Nienhuys-Cheng, S. (1994). Existence and nonexistence of complete refinement operators. In Bergadano, F. and de Raedt, L., editors, *Proceedings of the 7th European Conference on Machine Learning*, volume 784 of *Lecture Notes in Artificial Intelligence*, pages 307–322. Springer-Verlag.