

# Compatibilité de la sémantique Binding Trail avec GQL

Steven Sailly

Encadrante pédagogique : Cristina Sirangelo  
Encadrants de recherche : Nadime Francis & Victor Marsault

24/09/2024

# Introduction

# Base de données



# Base de données



Requête ?



# Base de données



Requête ?



Réponse !



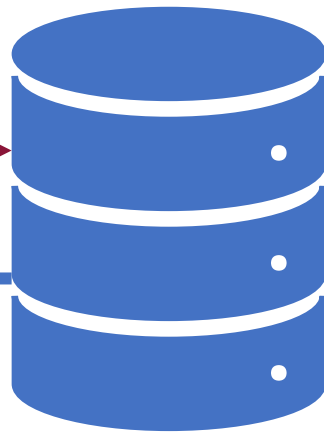
# Base de données



Requête ?



Réponse !



Relationnel  
Document

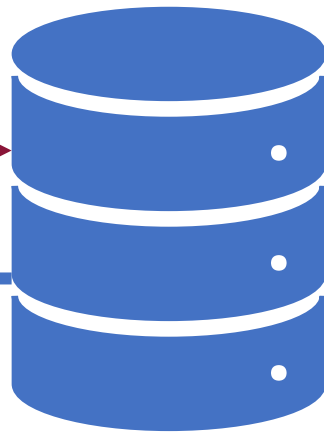
# Base de données



Requête ?



Réponse !



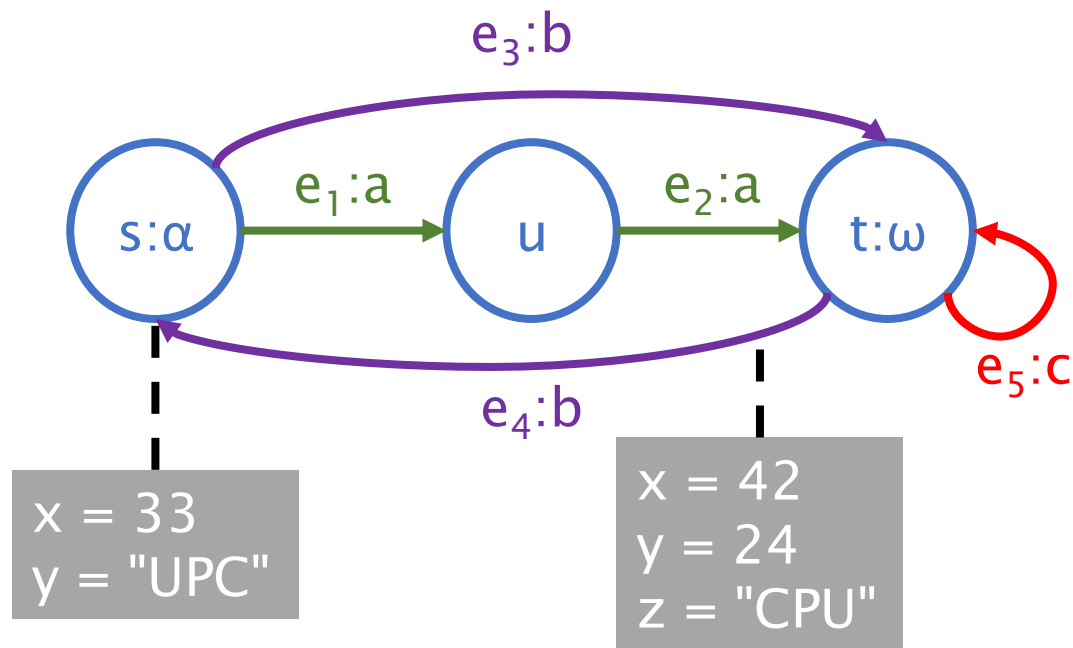
Relationnel  
Document  
Graphe

# Graphe à propriétés

Alphabet  $\Sigma = \{a, b, c, \alpha, \omega\}$

Nœuds et arêtes :

- Étiquettes dans  $\Sigma$
- Propriétés



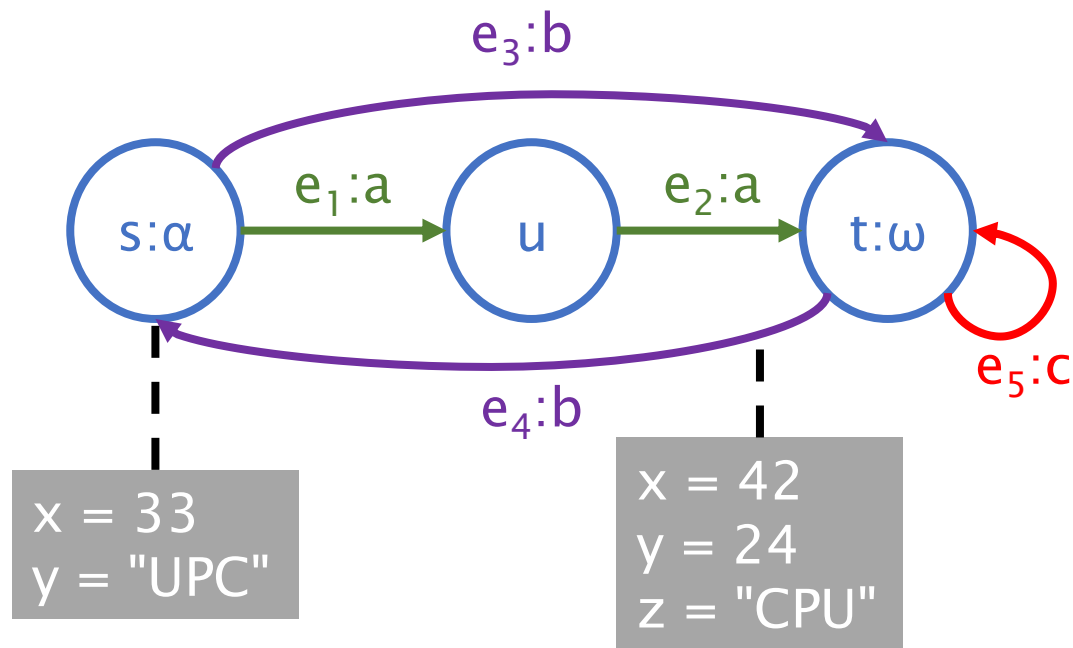
# Graphe à propriétés

Alphabet  $\Sigma = \{a, b, c, \alpha, \omega\}$

Nœuds et arêtes :

- Étiquettes dans  $\Sigma$
- Propriétés

Marche : suite de nœuds et d'arêtes



# Graphe à propriétés

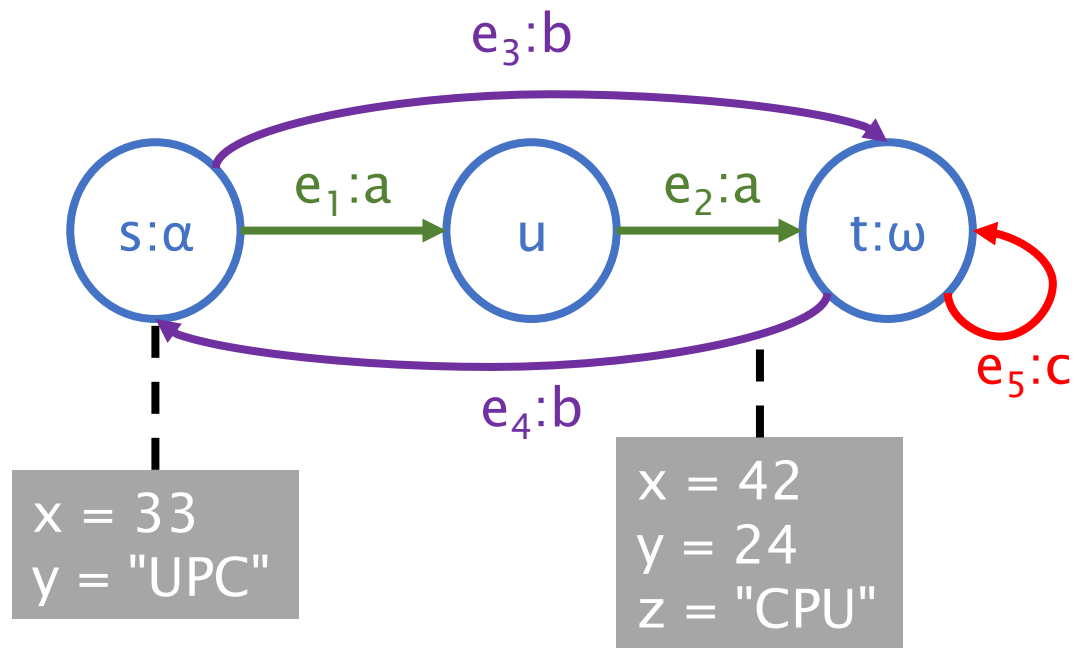
Alphabet  $\Sigma = \{a, b, c, \alpha, \omega\}$

Nœuds et arêtes :

- Étiquettes dans  $\Sigma$
- Propriétés

Marche : suite de nœuds et d'arêtes

Réponse à une requête : ensemble de marches



# Langages de requête

Standardisé  
par l'ISO

SQL  
[CB74]

# Langages de requête

Académique

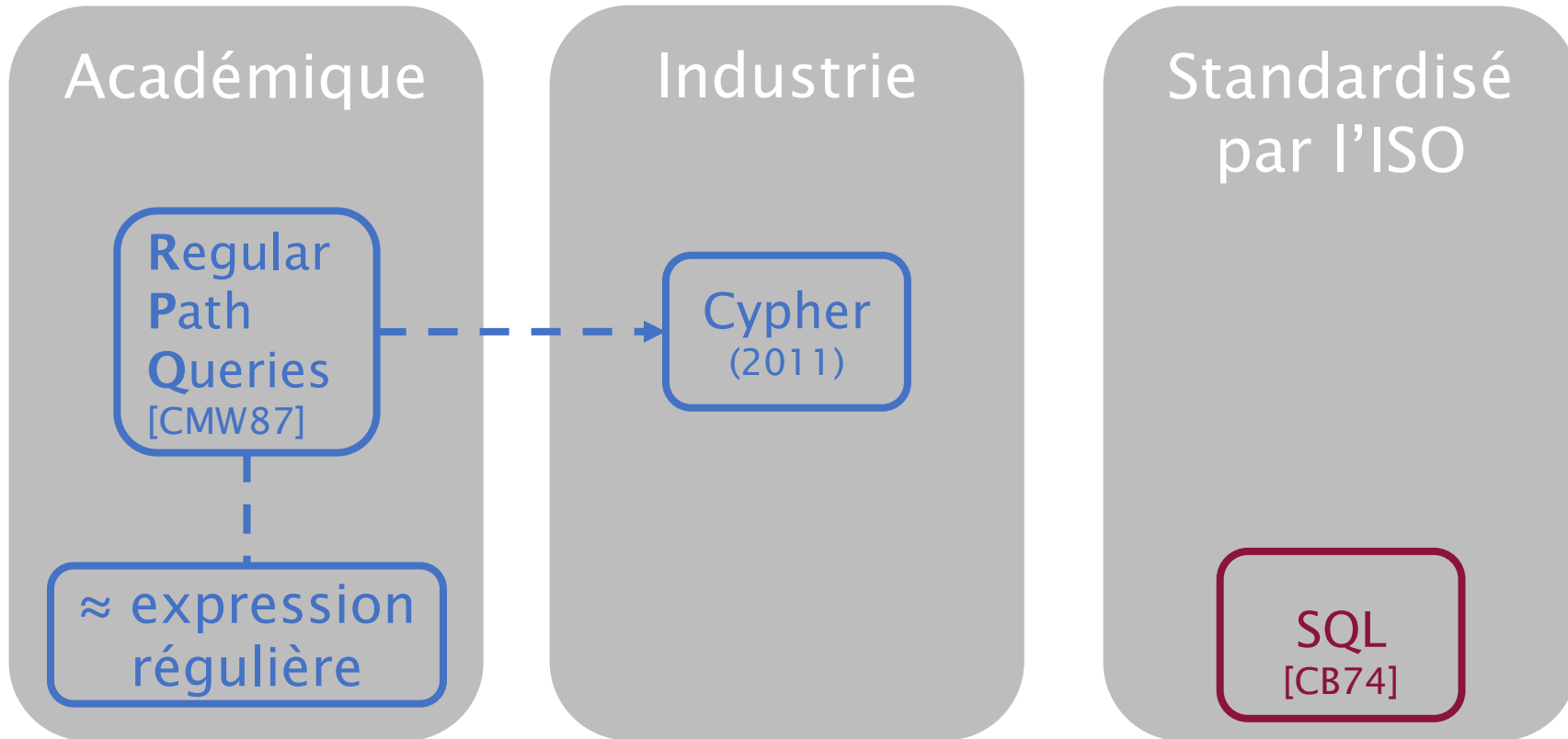
Regular  
Path  
Queries  
[CMW87]

≈ expression  
régulière

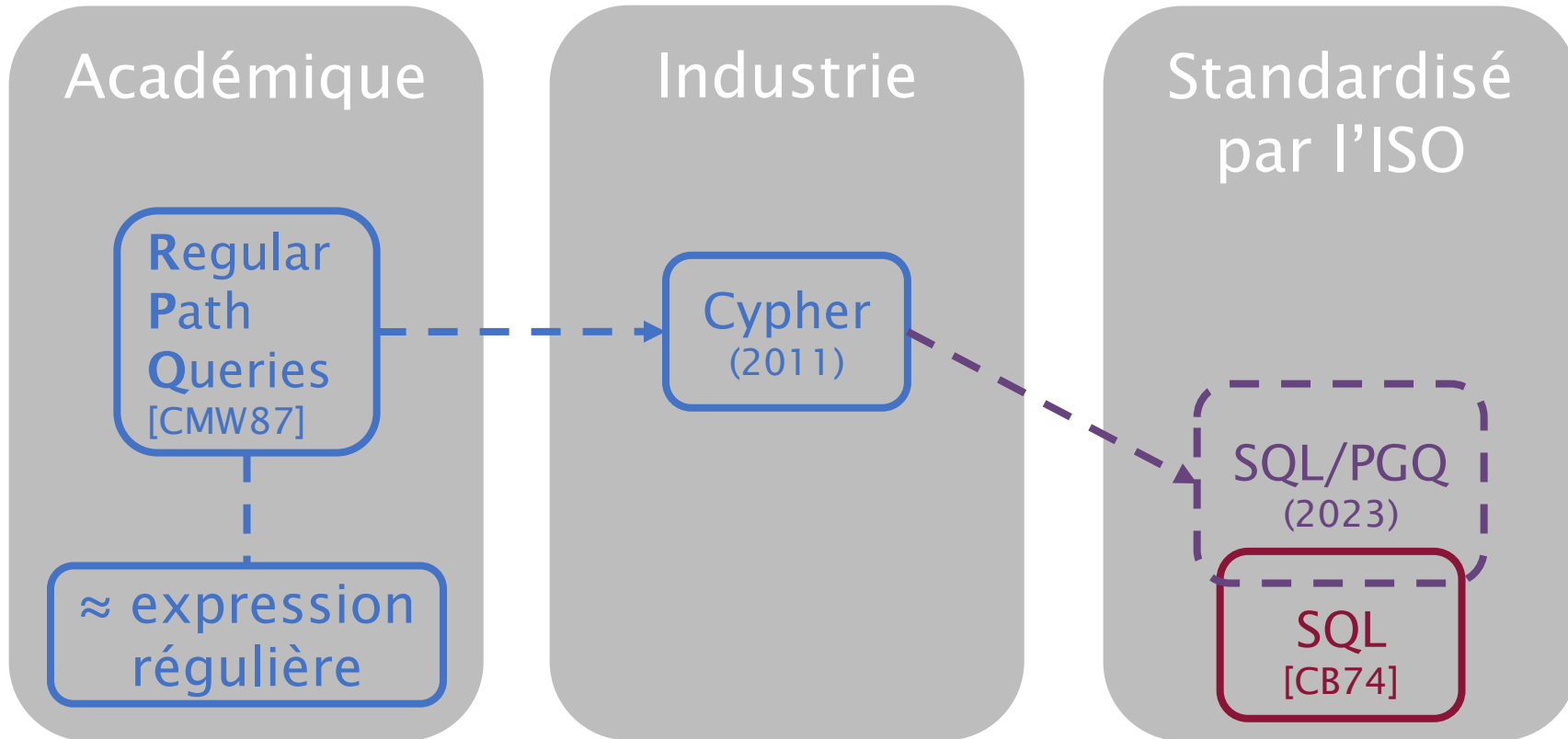
Standardisé  
par l'ISO

SQL  
[CB74]

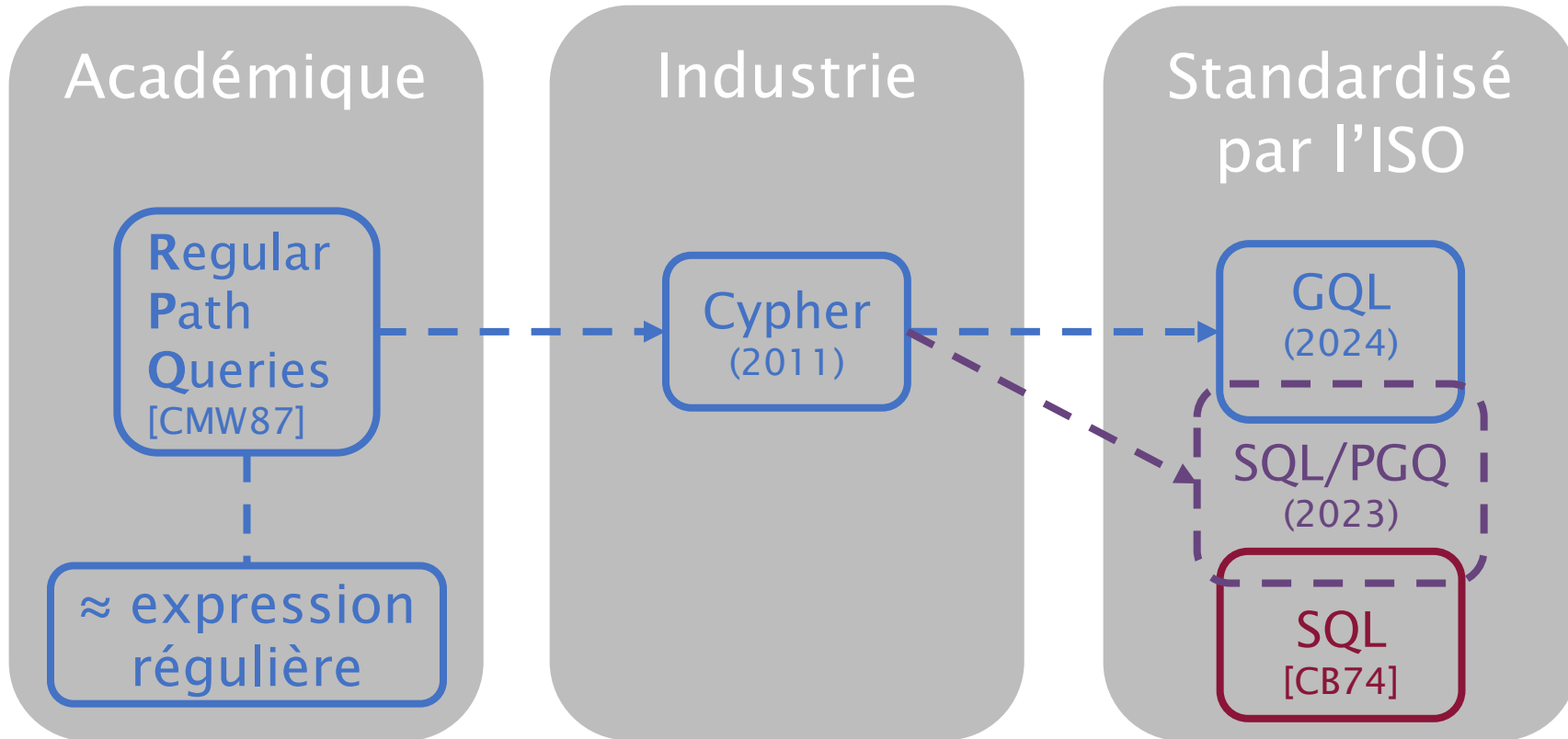
# Langages de requête



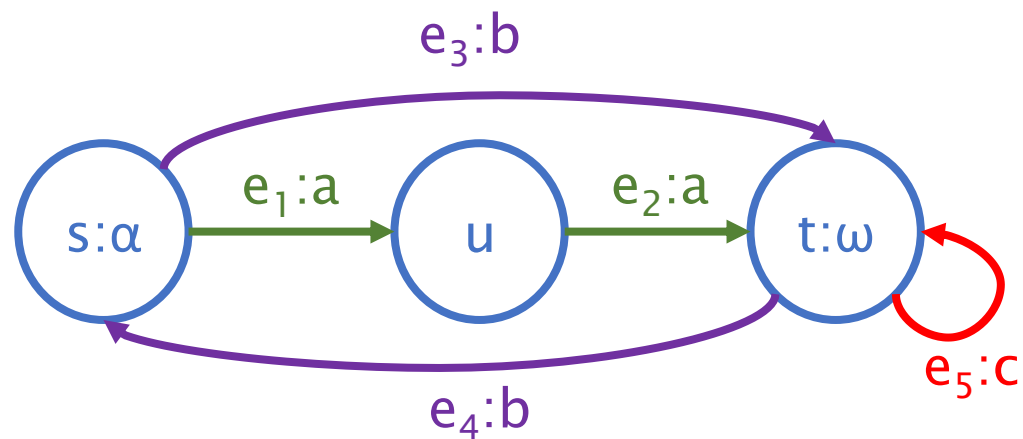
# Langages de requête



# Langages de requête

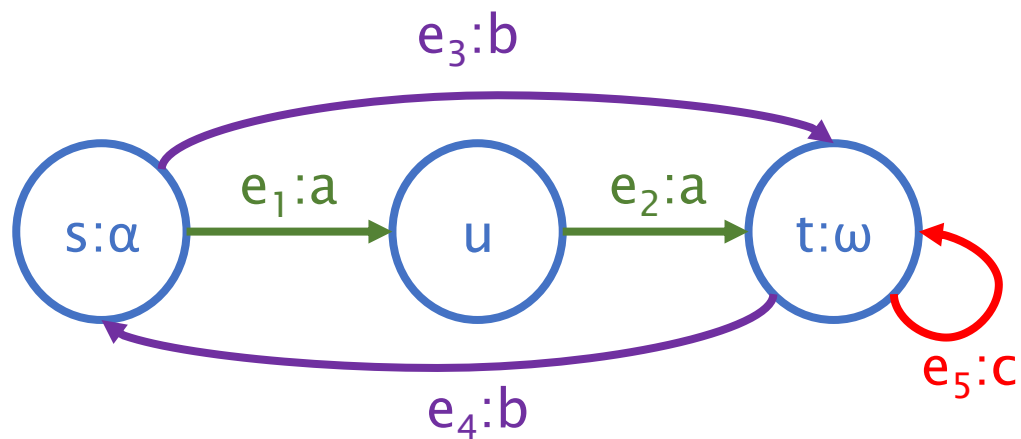


# GQL



# GQL

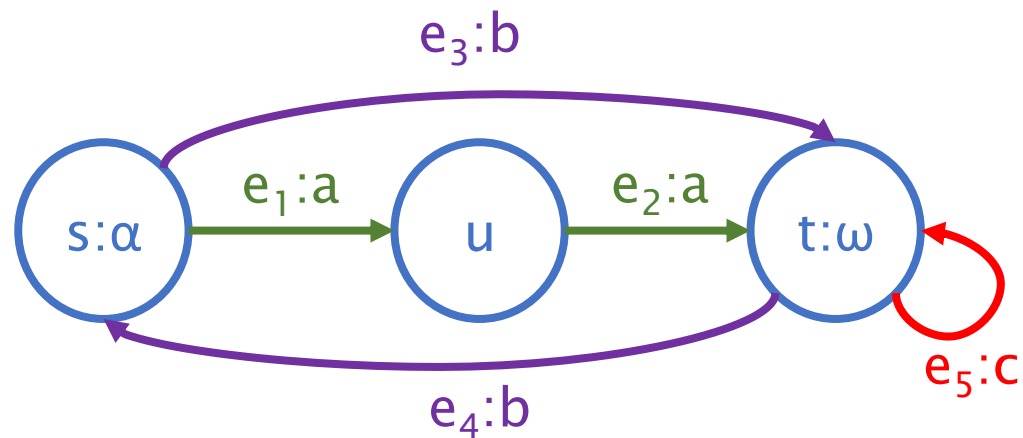
$$R = (: \alpha) \left[ \begin{array}{c} :a \\ \rightarrow \\ :b \\ \rightarrow \end{array} \right]^* \begin{array}{c} :c \\ \rightarrow \end{array} \left[ \begin{array}{c} :a \\ \rightarrow \\ :b \\ \rightarrow \end{array} \right]^* (: \omega)$$



# GQL

$$R = (:\alpha) \left[ \begin{array}{c} :a \\ \rightarrow \\ :b \end{array} \right]^* \begin{array}{c} :c \\ \rightarrow \end{array} \left[ \begin{array}{c} :a \\ \rightarrow \\ :b \end{array} \right]^* (:\omega)$$

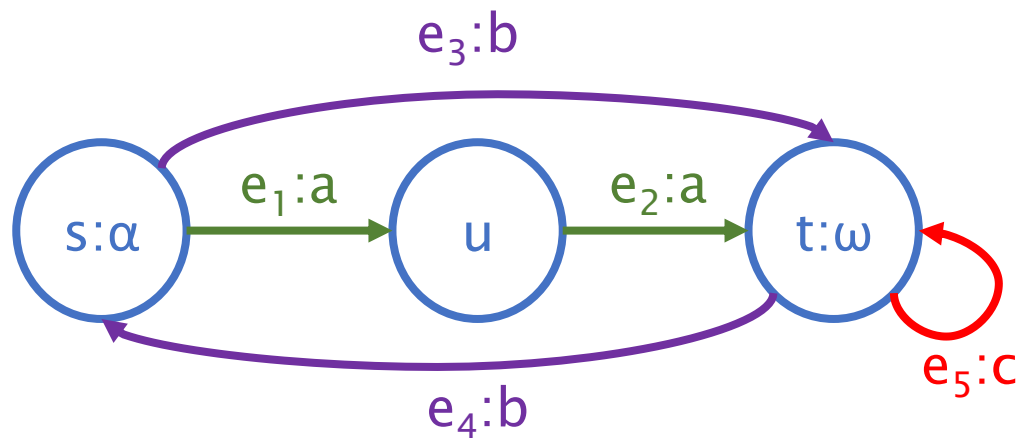
Motif atomique  
de nœud



# GQL

$$R = (: \alpha) \left[ \begin{array}{c} :a \quad :b \\ \rightarrow \quad | \rightarrow \end{array} \right]^* \begin{array}{c} :c \\ \rightarrow \end{array} \left[ \begin{array}{c} :a \quad :b \\ \rightarrow \quad | \rightarrow \end{array} \right]^* (: \omega)$$

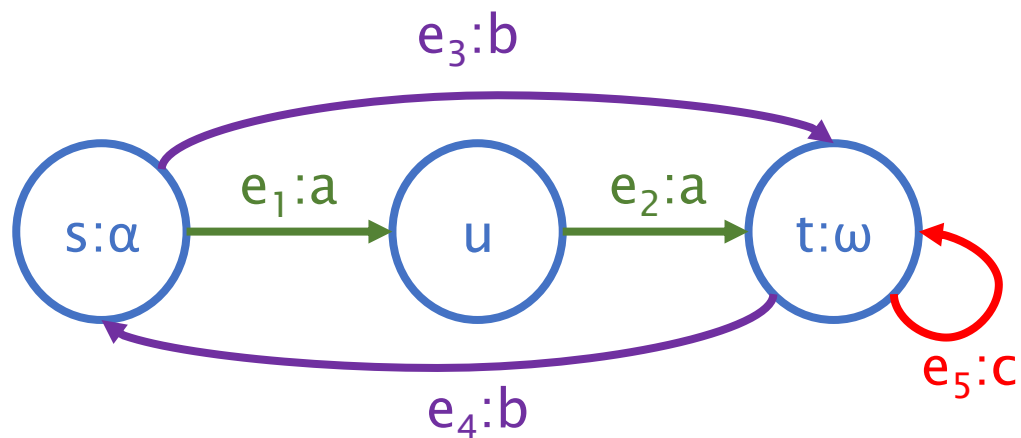
Motif atomique  
d'arête



# GQL

$$R = (: \alpha) \left[ \begin{array}{c|c} :a & :b \\ \hline \rightarrow & \rightarrow \end{array} \right]^* \begin{array}{c} :c \\ \rightarrow \end{array} \left[ \begin{array}{c|c} :a & :b \\ \hline \rightarrow & \rightarrow \end{array} \right]^* (: \omega)$$

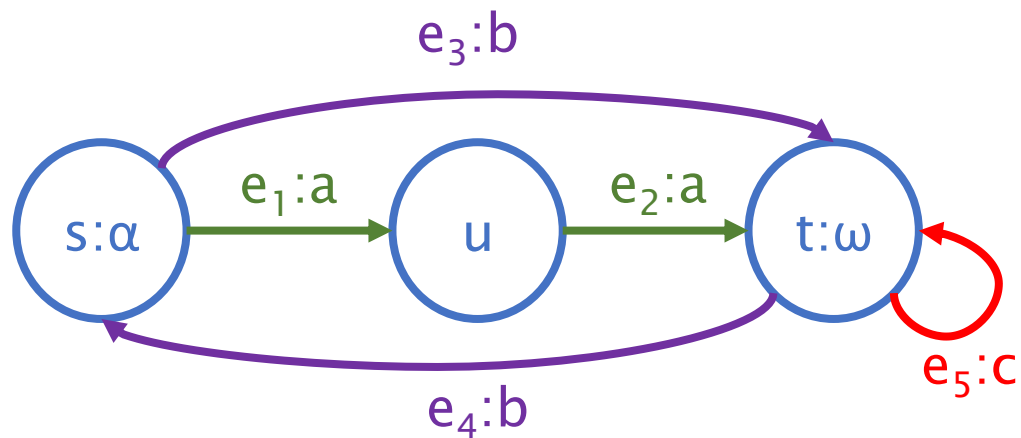
Répétition  
non bornée



# GQL

$$R = (: \alpha) \left[ \begin{array}{c} :a \quad :b \\ \rightarrow \quad \rightarrow \end{array} \right]^* \begin{array}{c} :c \\ \rightarrow \end{array} \left[ \begin{array}{c} :a \quad :b \\ \rightarrow \quad \rightarrow \end{array} \right]^* (: \omega)$$

Union



# GQL

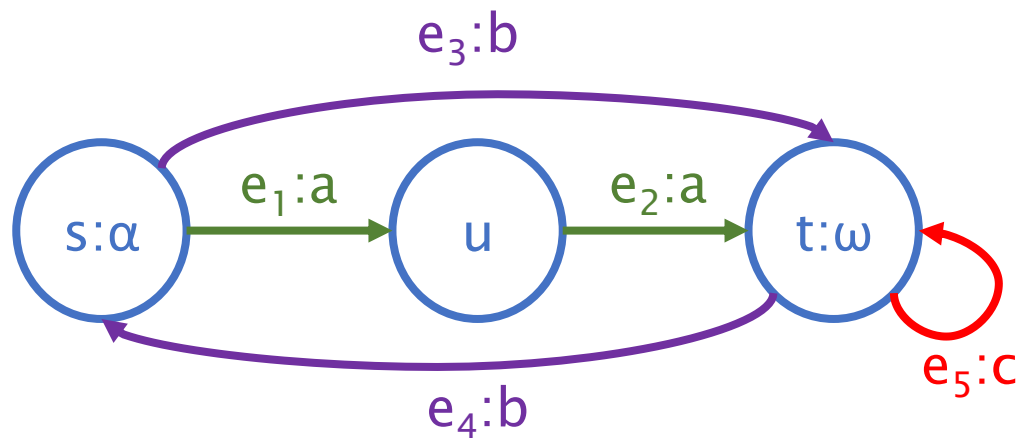
$$R = (: \alpha) \left[ \begin{array}{c} :a \\ \rightarrow \\ :b \\ | \\ \rightarrow \end{array} \right]^* \begin{array}{c} :c \\ \rightarrow \end{array} \left[ \begin{array}{c} :a \\ \rightarrow \\ :b \\ | \\ \rightarrow \end{array} \right]^* (: \omega)$$

$e_1$   $e_2$   $e_5$   $e_4$   $e_1$   $e_2$   
 $s \rightarrow u \rightarrow t \rightarrow t \rightarrow s \rightarrow u \rightarrow t$



$e_1$   $e_2$   $e_5$   $e_4$   $e_1$   $e_2$   
 $s \rightarrow u \rightarrow t \rightarrow t \rightarrow s \rightarrow u \rightarrow t$

$e_4$   $e_1$   $e_2$   $e_4$   $e_2$   
 $\rightarrow s \rightarrow u \rightarrow t \rightarrow \dots \rightarrow t$



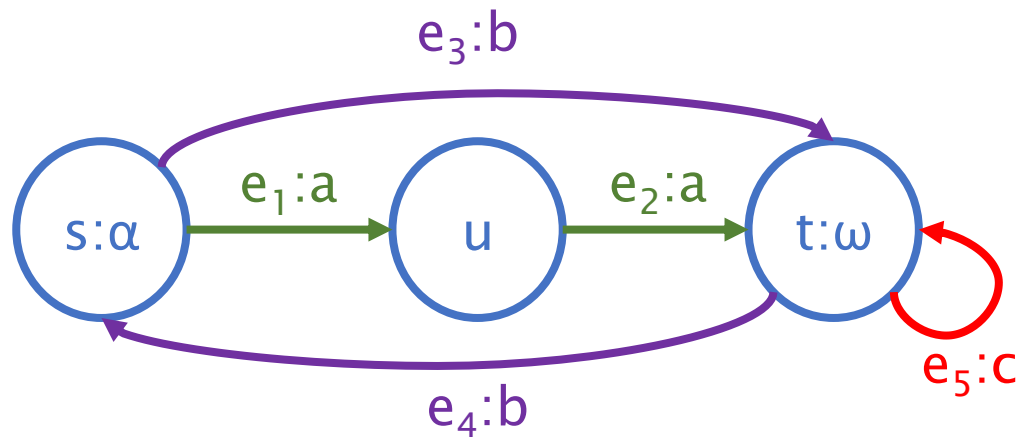
# GQL

$$R = (: \alpha) \left[ \begin{array}{c} :a \\ \rightarrow \\ :b \\ | \rightarrow \end{array} \right]^* \begin{array}{c} :c \\ \rightarrow \end{array} \left[ \begin{array}{c} :a \\ \rightarrow \\ :b \\ | \rightarrow \end{array} \right]^* (: \omega)$$

$s \xrightarrow{e_1} u \xrightarrow{e_2} t \xrightarrow{e_5} t \xrightarrow{e_4} s \xrightarrow{e_1} u \xrightarrow{e_2} t$  ✓

$s \xrightarrow{e_1} u \xrightarrow{e_2} t \xrightarrow{e_5} t \xrightarrow{e_4} s \xrightarrow{e_1} u \xrightarrow{e_2} t$

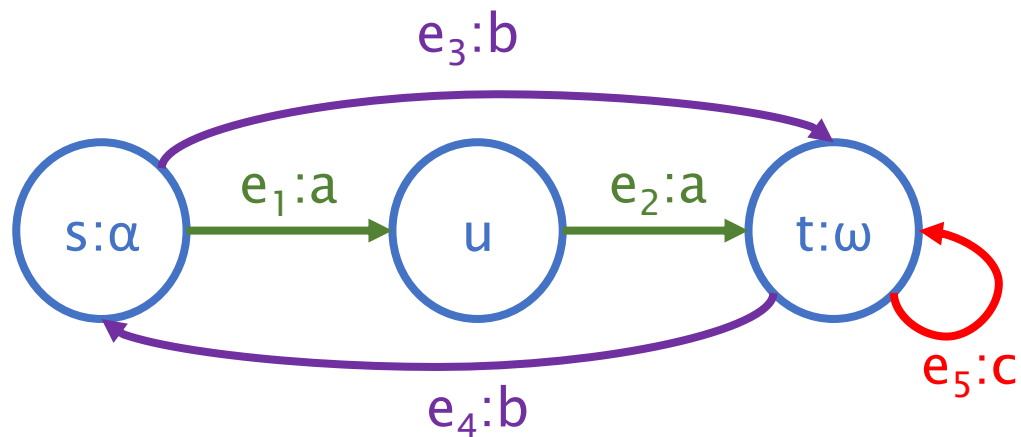
$\xrightarrow{e_4} s \xrightarrow{e_1} u \xrightarrow{e_2} t \xrightarrow{e_4} \dots \xrightarrow{e_2} t$  ✓



# Trail

Pas de répétition d'arête

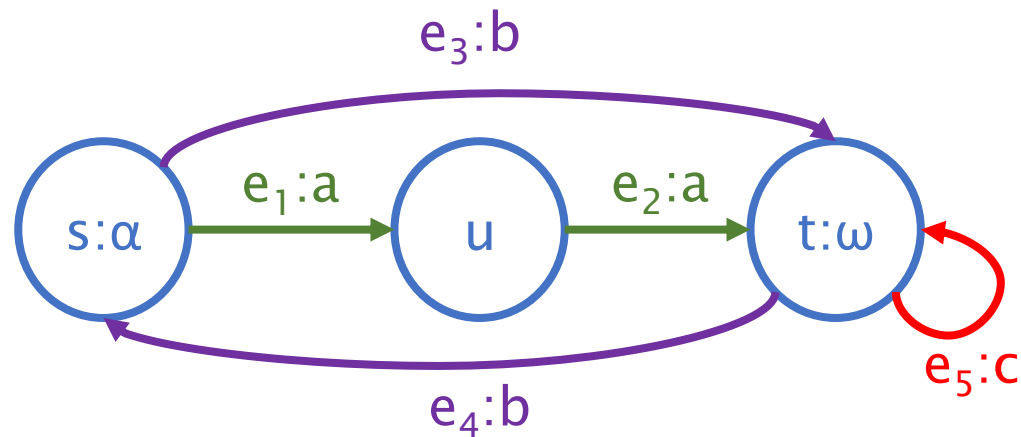
$$R = (: \alpha) \left[ \begin{array}{c|c} :a & :b \\ \hline \rightarrow & \rightarrow \end{array} \right]^* \begin{array}{c} :c \\ \rightarrow \end{array} \left[ \begin{array}{c|c} :a & :b \\ \hline \rightarrow & \rightarrow \end{array} \right]^* (: \omega)$$



# Trail

Pas de répétition d'arête

$$R = (: \alpha) \left[ \begin{array}{c} :a \\ \rightarrow \\ :b \end{array} \right]^* \begin{array}{c} :c \\ \rightarrow \end{array} \left[ \begin{array}{c} :a \\ \rightarrow \\ :b \end{array} \right]^* (: \omega)$$

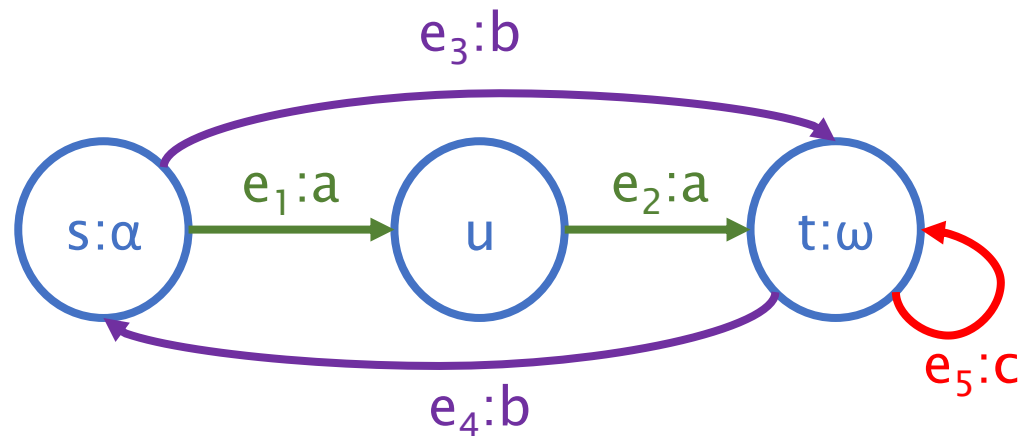


|                         |                         |                         |                         |                         |                         |   |
|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|---|
| $s \xrightarrow{e_1} u$ | $u \xrightarrow{e_2} t$ | $t \xrightarrow{e_5} t$ | $t \xrightarrow{e_4} s$ | $s \xrightarrow{e_3} t$ | ✓                       |   |
| $s \xrightarrow{e_1} u$ | $u \xrightarrow{e_2} t$ | $t \xrightarrow{e_5} t$ | $t \xrightarrow{e_4} s$ | $s \xrightarrow{e_1} u$ | $u \xrightarrow{e_2} t$ | ✗ |

# Trail

Pas de répétition d'arête

$$R = (: \alpha) \left[ \begin{array}{c} :a \\ \rightarrow \\ :b \\ \rightarrow \end{array} \right]^* \begin{array}{c} :c \\ \rightarrow \end{array} \left[ \begin{array}{c} :a \\ \rightarrow \\ :b \\ \rightarrow \end{array} \right]^* (: \omega)$$



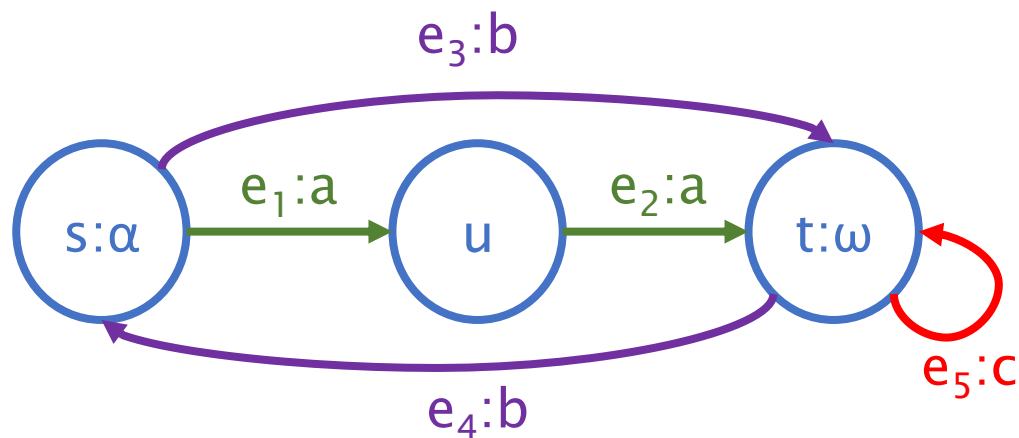
$s \xrightarrow{e_1} u \xrightarrow{e_2} t \xrightarrow{e_5} t \xrightarrow{e_4} s \xrightarrow{e_3} t$  ✓  
 $s \xrightarrow{e_1} u \xrightarrow{e_2} t \xrightarrow{e_5} t \xrightarrow{e_4} s \xrightarrow{e_1} u \xrightarrow{e_2} t$  ✗

Trouver un résultat  
est **NP-difficile** [MNP23]

# Shortest

La/les plus courte(s)

$$R = (: \alpha) \left[ \begin{array}{c} :a \\ \rightarrow \\ :b \\ \rightarrow \end{array} \right]^* \begin{array}{c} :c \\ \rightarrow \end{array} \left[ \begin{array}{c} :a \\ \rightarrow \\ :b \\ \rightarrow \end{array} \right]^* (: \omega)$$

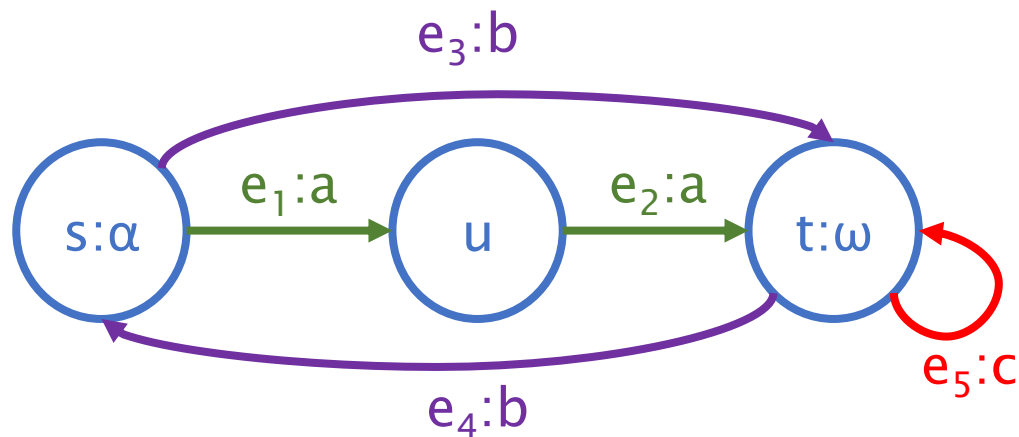


# Shortest

La/les plus courte(s)

$$R = (: \alpha) \left[ \begin{array}{c|c} :a & :b \\ \hline \rightarrow & \rightarrow \end{array} \right]^* \xrightarrow{:c} \left[ \begin{array}{c|c} :a & :b \\ \hline \rightarrow & \rightarrow \end{array} \right]^* (: \omega)$$

|   |   |
|---|---|
| $s \xrightarrow{e_3} t \xrightarrow{e_5} t$                     | ✓ |
| $s \xrightarrow{e_1} u \xrightarrow{e_2} t \xrightarrow{e_5} t$ | ✗ |

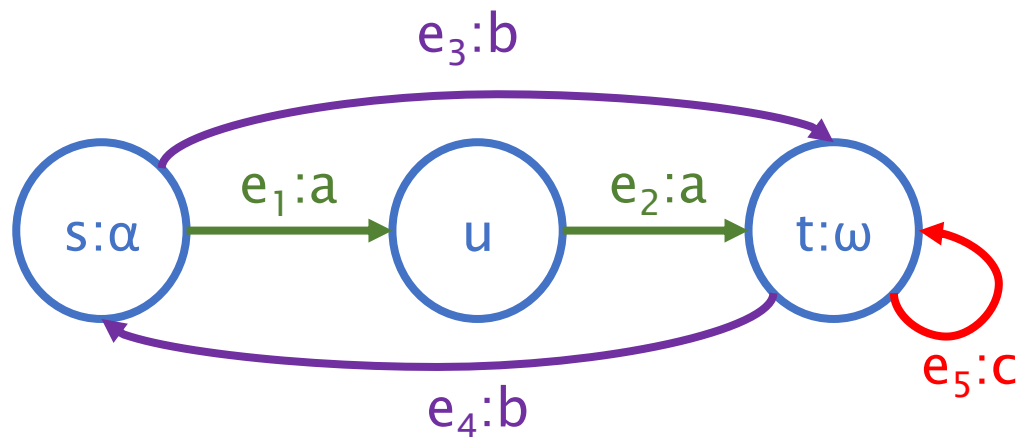


# Shortest

La/les plus courte(s)

$$R = (: \alpha) \left[ \begin{array}{c} :a \\ \rightarrow \\ :b \end{array} \middle| \begin{array}{c} :b \\ \rightarrow \\ :a \end{array} \right]^* \begin{array}{c} :c \\ \rightarrow \\ :c \end{array} \left[ \begin{array}{c} :a \\ \rightarrow \\ :b \end{array} \middle| \begin{array}{c} :b \\ \rightarrow \\ :a \end{array} \right]^* (: \omega)$$

$s \xrightarrow{e_3} t \xrightarrow{e_5} t$  ✓  
 $s \xrightarrow{e_1} u \xrightarrow{e_2} t \xrightarrow{e_5} t$  ✗

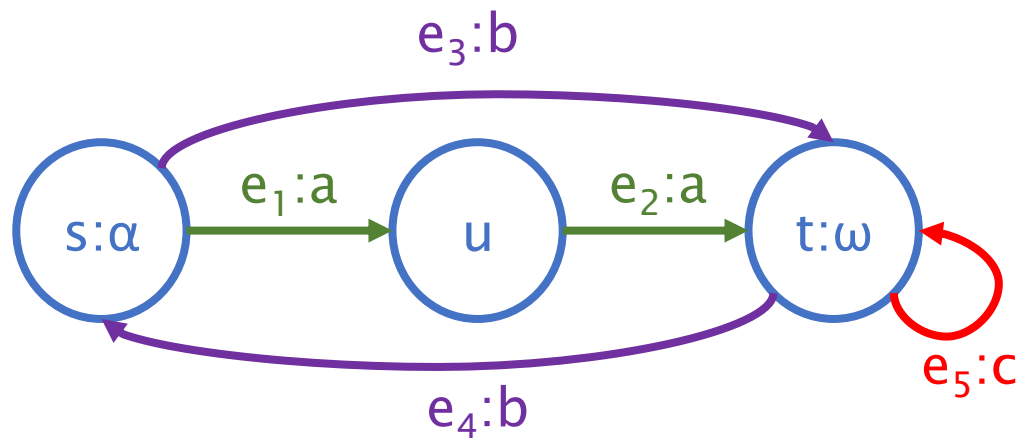


Ensemble très restreint

# Binding Trail ([DFM23] pour les RPQs)

Pas de répétition d'arête à la même position

$$R = (: \alpha) \left[ \begin{array}{c|c} \xrightarrow{:a} & \xrightarrow{:b} \\ \hline 1 & 2 \end{array} \right]^* \xrightarrow{:c} \left[ \begin{array}{c|c} \xrightarrow{:a} & \xrightarrow{:b} \\ \hline 4 & 5 \end{array} \right]^* (: \omega)$$



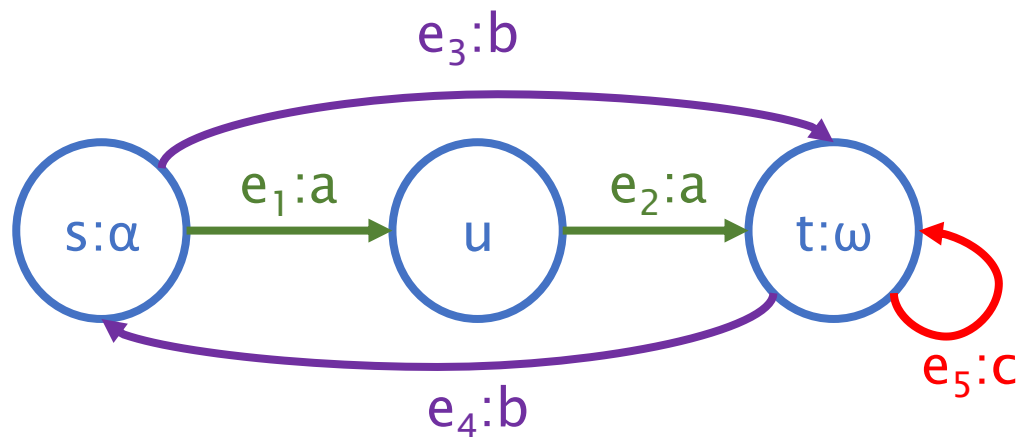
# Binding Trail ([DFM23] pour les RPQs)

Pas de répétition d'arête à la même position

$$R = (: \alpha) \left[ \begin{array}{c|c} \xrightarrow{:a} & \xrightarrow{:b} \\ \hline 1 & 2 \end{array} \right]^* \xrightarrow{:c} \left[ \begin{array}{c|c} \xrightarrow{:a} & \xrightarrow{:b} \\ \hline 4 & 5 \end{array} \right]^* (: \omega)$$

$s \xrightarrow[1]{e_1} u \xrightarrow{e_2} t \xrightarrow{e_5} t \xrightarrow{e_4} s \xrightarrow[4]{e_1} u \xrightarrow{e_2} t$ 
✓

$s \xrightarrow[1]{e_1} u \xrightarrow{e_2} t \xrightarrow{e_4} s \xrightarrow[1]{e_1} u \xrightarrow{e_2} t \xrightarrow{e_5} t$ 
✗



# Contributions

# Faire une preuve formelle :



dans le  
rapport



pendant la  
soutenance

# Tuple Membership sous la sémantique Binding Trail

- Entrées : une requête  $Q$ , un graphe à propriétés  $D$ , deux nœuds  $s, t \in D$
- Question : Existe-t-il une marche de  $s$  à  $t$  telle qu'elle satisfait  $Q$  sous la sémantique Binding Trail ?

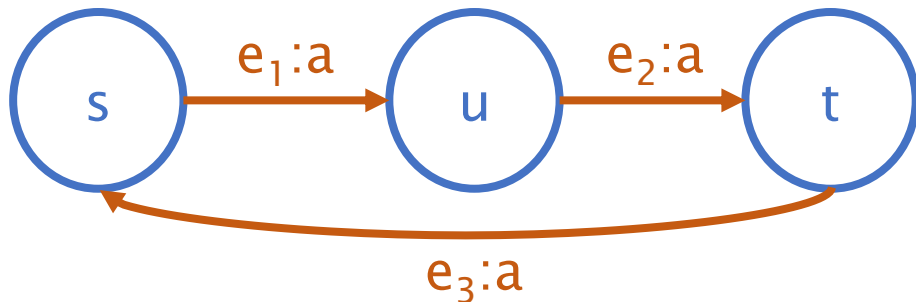
Ce qu'on sait : **PTIME** pour les RPQs [DFM23]

# Tuple Membership sous la sémantique Binding Trail

- Ce que je prouve :
  - **NP**-difficile avec une requête de la forme  $[\overset{:a}{\rightarrow} \overset{:b^2}{\rightarrow} \overset{:c}{\rightarrow}]^*$   
ou  $[\overset{:a}{\rightarrow} \overset{:b^*}{\rightarrow} \overset{:c}{\rightarrow}]^2$
  - **NP**-difficile avec une requête de la forme  $[(u) \overset{:a}{\rightarrow} \overset{:b}{\rightarrow} \overset{:c}{\rightarrow} (u) \overset{:d}{\rightarrow}]^*$  (si on a le temps)
  - **PTIME** dans la taille de la base de données sans variable sous étoile (pas aujourd'hui)

# Répétition fixe et Binding Trail

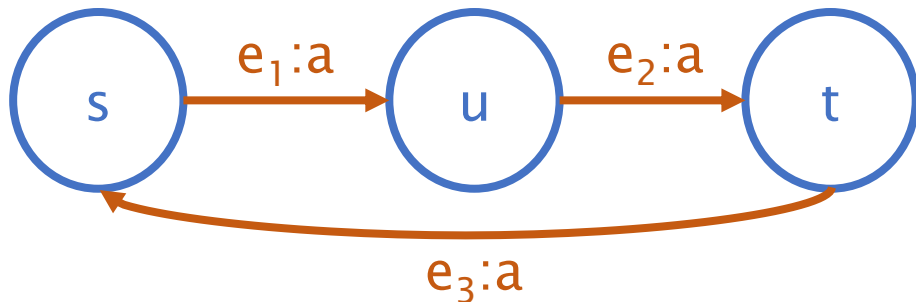
$$R = \left[ \begin{array}{c} :a^2 \\ \rightarrow \\ 1 \end{array} \right]^* \neq R' = \left[ \begin{array}{cc} :a & :a \\ \rightarrow & \rightarrow \\ 1 & 2 \end{array} \right]^*$$



# Répétition fixe et Binding Trail

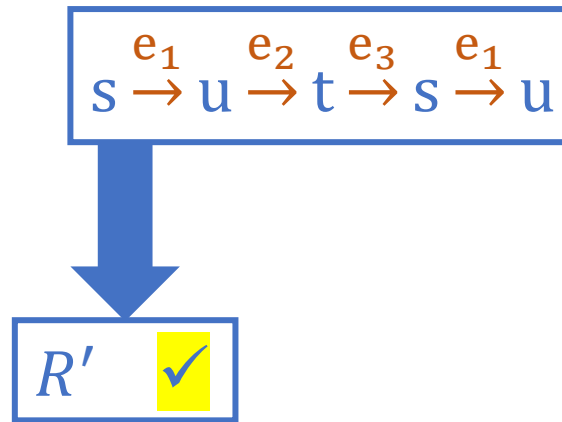
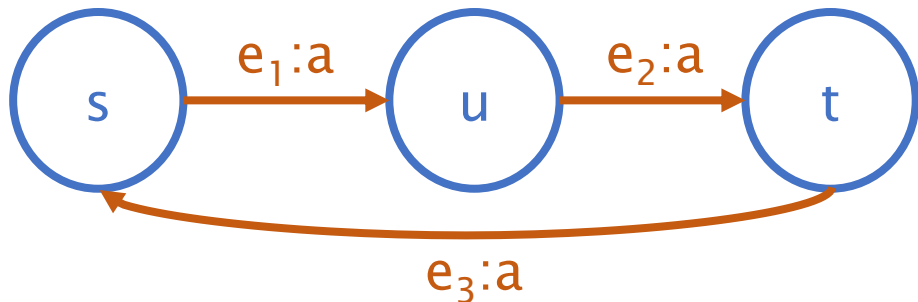
$$R = \left[ \begin{array}{c} :a^2 \\ \rightarrow \\ 1 \end{array} \right]^* \neq R' = \left[ \begin{array}{cc} :a & :a \\ \rightarrow & \rightarrow \\ 1 & 2 \end{array} \right]^*$$

$$s \xrightarrow{e_1} u \xrightarrow{e_2} t \xrightarrow{e_3} s \xrightarrow{e_1} u$$



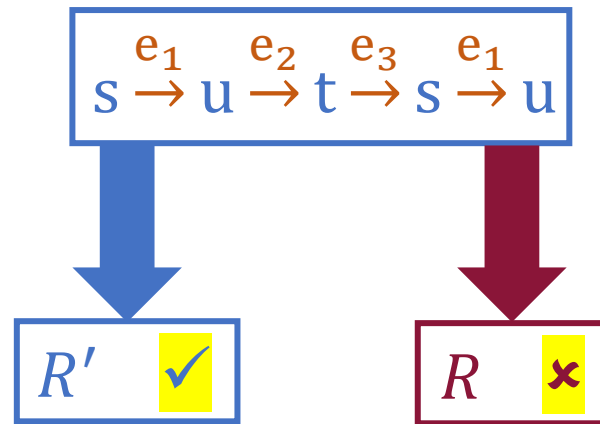
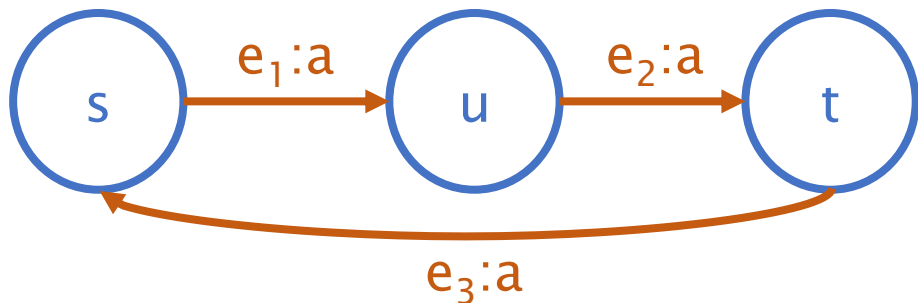
# Répétition fixe et Binding Trail

$$R = \left[ \begin{array}{c} :a^2 \\ \xrightarrow{1} \end{array} \right]^* \neq R' = \left[ \begin{array}{cc} :a & :a \\ \xrightarrow{1} & \xrightarrow{2} \end{array} \right]^*$$



# Répétition fixe et Binding Trail

$$R = \left[ \begin{array}{c} :a^2 \\ \xrightarrow{1} \end{array} \right]^* \neq R' = \left[ \begin{array}{cc} :a & :a \\ \xrightarrow{1} & \xrightarrow{2} \end{array} \right]^*$$



# Réduction : encoder 3-SAT

$I = c_1 \wedge \dots \wedge c_n$  une instance de 3-SAT  
Variables :  $x_1, \dots, x_m$



NP-  
difficile

# Réduction : encoder 3-SAT

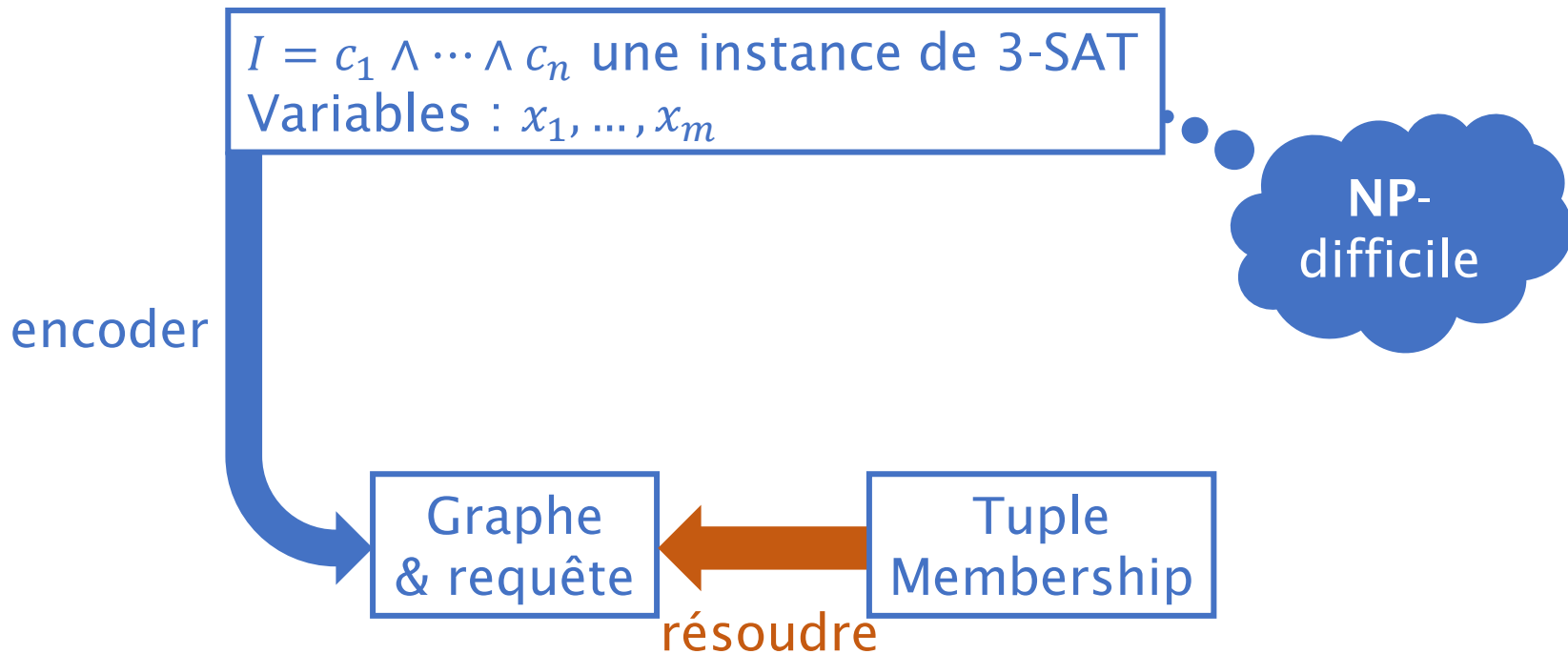
$I = c_1 \wedge \dots \wedge c_n$  une instance de 3-SAT  
Variables :  $x_1, \dots, x_m$

NP-  
difficile

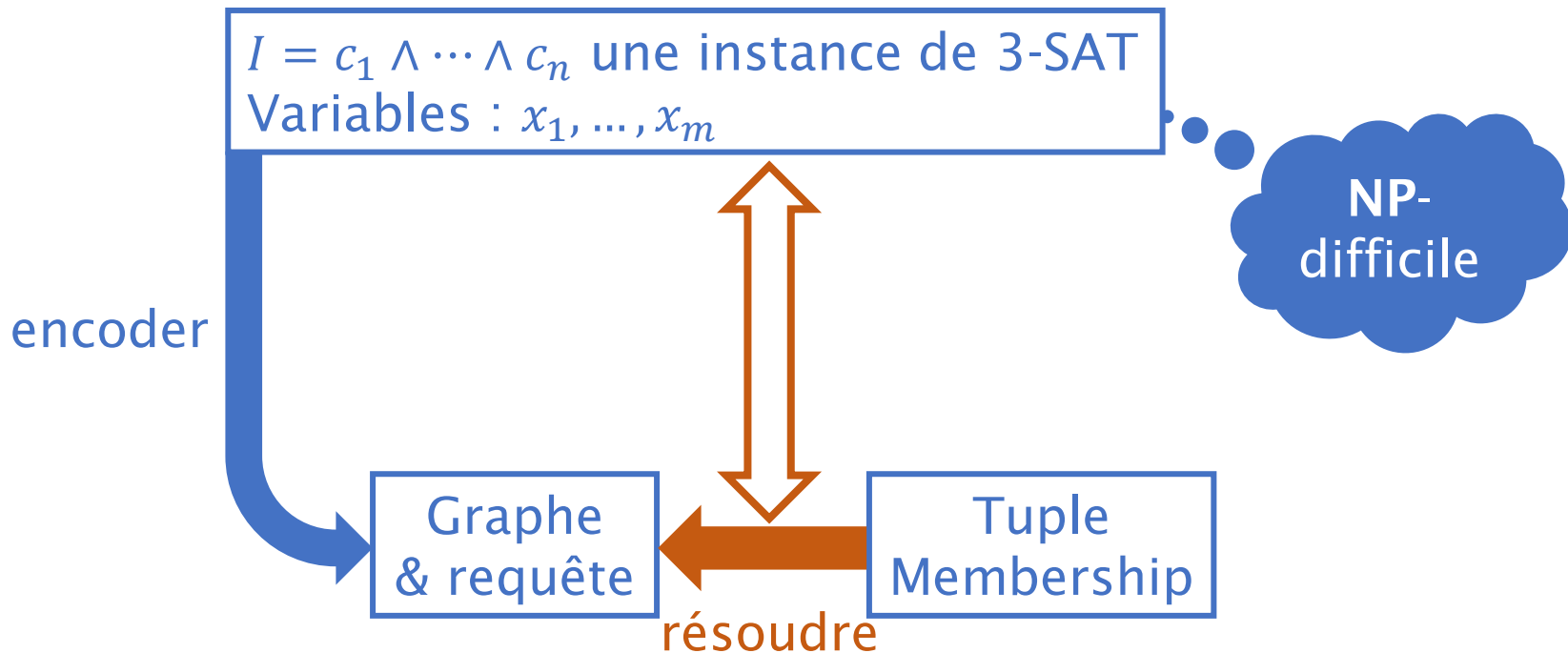
encoder

Graphe  
& requête

# Réduction : encoder 3-SAT



# Réduction : encoder 3-SAT



# Répétition fixe sous étoile

$I = c_1 \wedge \dots \wedge c_n$  une instance de 3-SAT  
Variables :  $x_1, \dots, x_m$

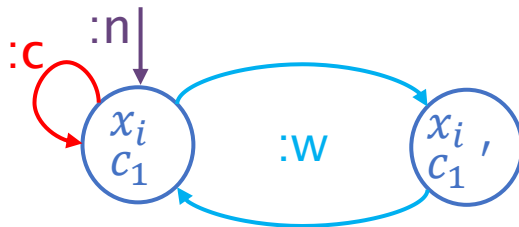
$R = [ \xrightarrow{\text{:next}} \xrightarrow{\text{:witness}^2} \xrightarrow{\text{:check}} ]^*$

# Répétition fixe sous étoile

$I = c_1 \wedge \dots \wedge c_n$  une instance de 3-SAT

Variables :  $x_1, \dots, x_m$

$R = [ \xrightarrow{\text{:next}} \xrightarrow{\text{:witness}^2} \xrightarrow{\text{:check}} ]^*$

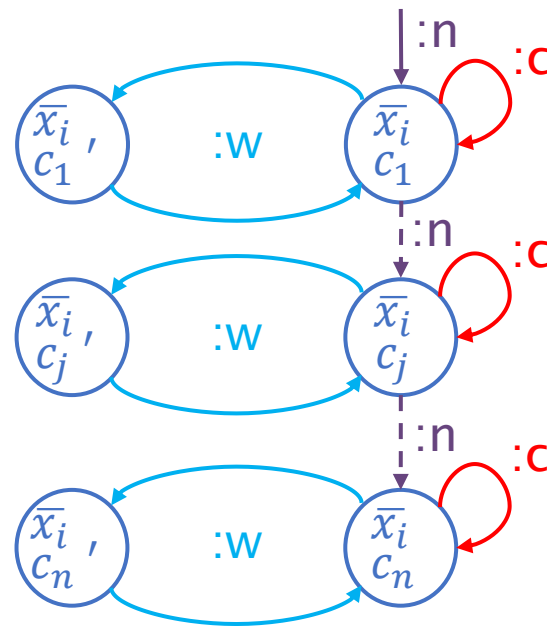
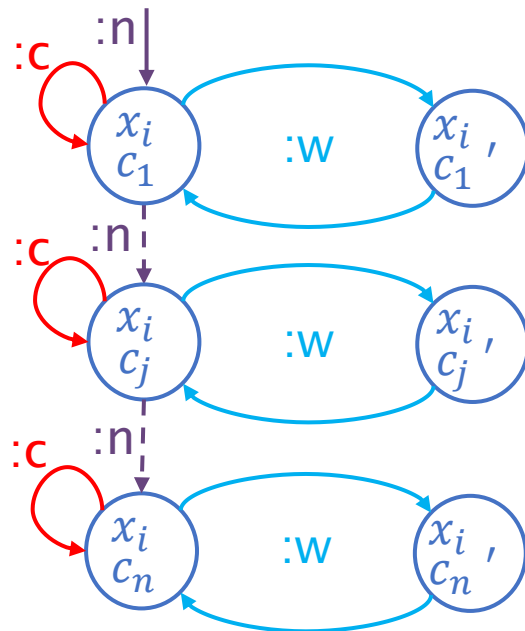


# Répétition fixe sous étoile

$I = c_1 \wedge \dots \wedge c_n$  une instance de 3-SAT

Variables :  $x_1, \dots, x_m$

$R = [ \xrightarrow{\text{:next}} \xrightarrow{\text{:witness}^2} \xrightarrow{\text{:check}} ]^*$



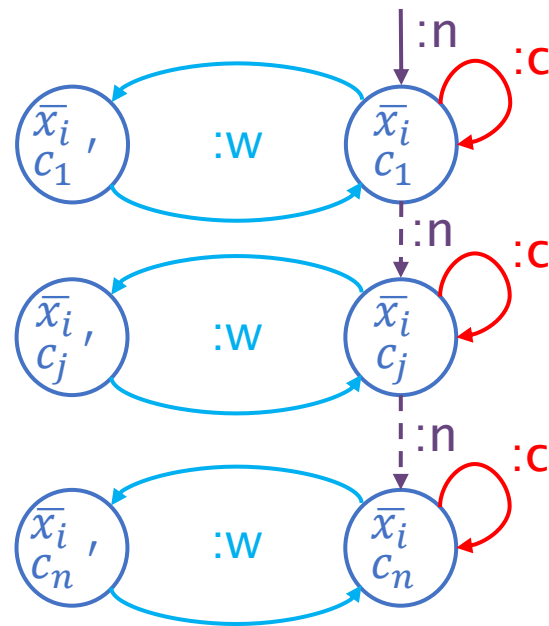
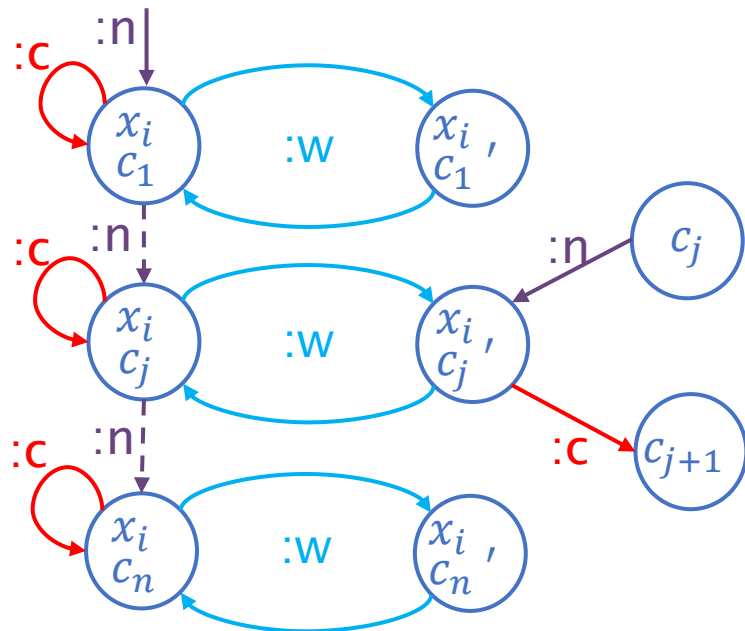
# Répétition fixe sous étoile

$I = c_1 \wedge \dots \wedge c_n$  une instance de 3-SAT

Variables :  $x_1, \dots, x_m$

$R = [ \xrightarrow{\text{:next}} \xrightarrow{\text{:witness}^2} \xrightarrow{\text{:check}} ]^*$

Si  $x_i$  apparaît positivement dans  $c_j$

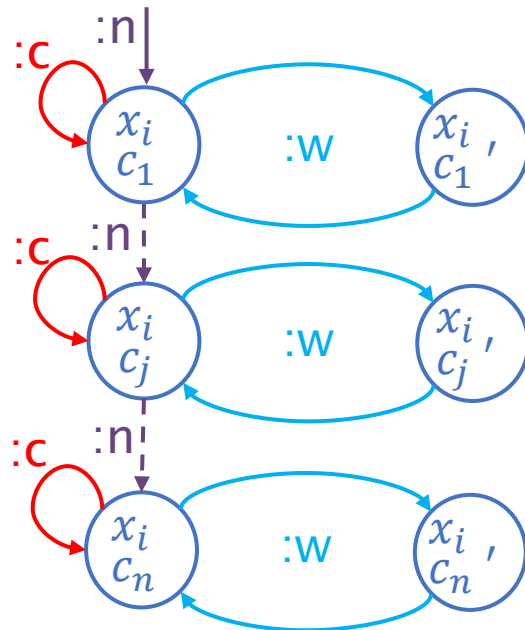


# Répétition fixe sous étoile

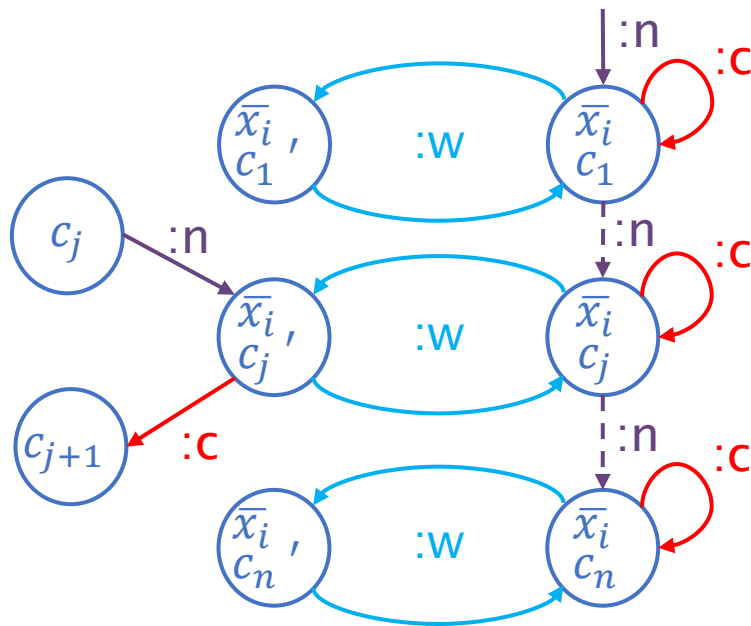
$I = c_1 \wedge \dots \wedge c_n$  une instance de 3-SAT

Variables :  $x_1, \dots, x_m$

Si  $x_i$  apparaît  
négativement  
dans  $c_j$

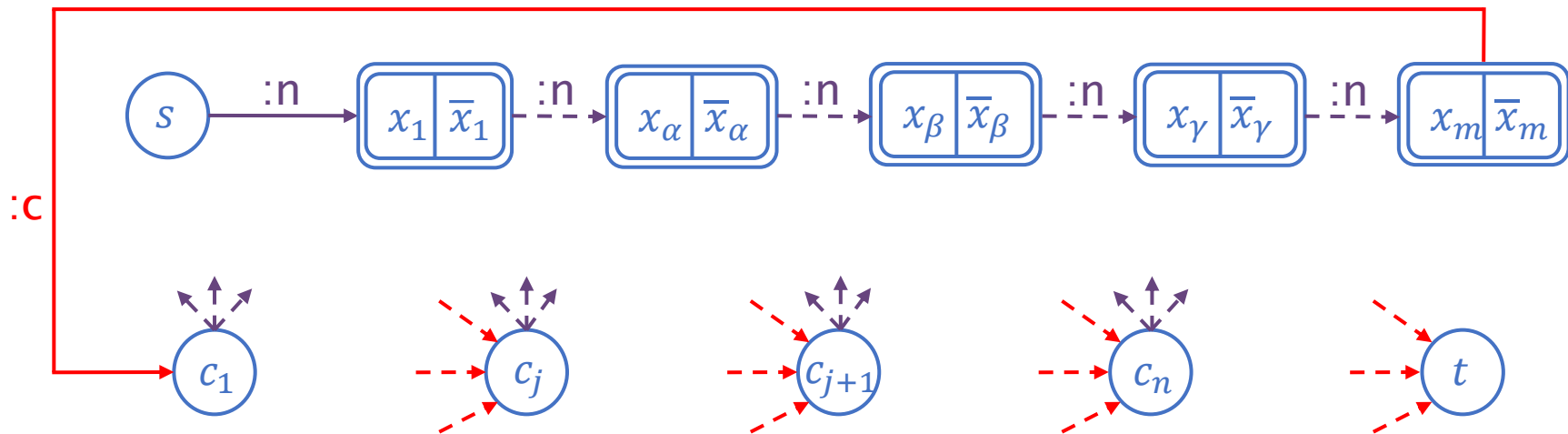


$R = [ \xrightarrow{\text{:next}} \xrightarrow{\text{:witness}^2} \xrightarrow{\text{:check}} ]^*$



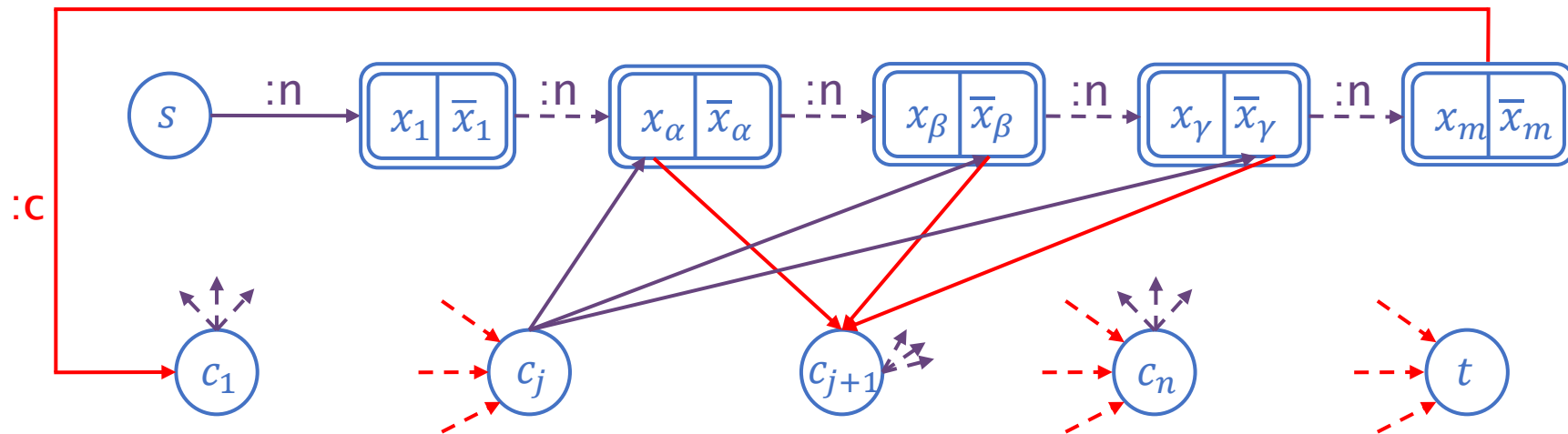
# Répétition fixe sous étoile

$$R = [ \xrightarrow{\text{:next}} \xrightarrow{\text{:witness}^2} \xrightarrow{\text{:check}} ]^*$$



# Répétition fixe sous étoile

$$R = \left[ \xrightarrow{\text{:next}} \xrightarrow{\text{:witness}^2} \xrightarrow{\text{:check}} \right]^*$$



$$\text{Avec } c_j = (x_\alpha \vee \neg x_\beta \vee \neg x_\gamma)$$

# Bilan

## Bilan de l'adaptation de Binding Trail à GQL :

- **NP**-difficile pour GQL :
  - avec répétition fixe
  - avec variable sous étoile
- **PTIME** en données pour GQL sans variable sous étoile  
⇒ Énumération des résultats en **DelayP** en données

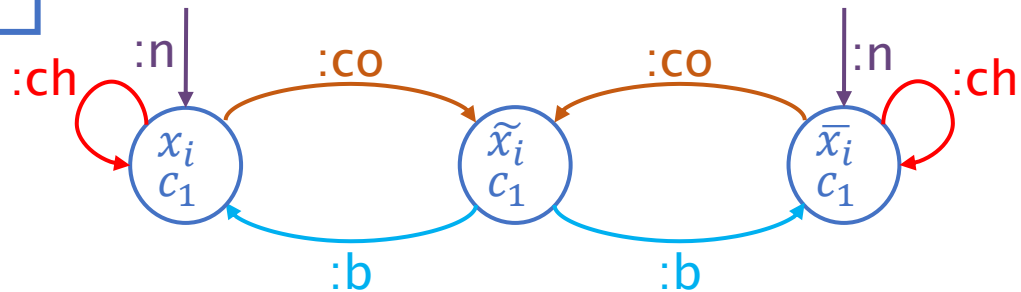
## Bibliographie

- [CB74] Chamberlin, D. D., & Boyce, R. F. (1974, May). “SEQUEL: A structured English query language”. In *Proceedings of the 1974 ACM SIGFIDET (now SIGMOD) workshop on Data description, access and control* (pp. 249-264).
- [CMW87] Cruz, I. F., Mendelzon, A. O., & Wood, P. T. (1987). “A graphical query language supporting recursion”. *ACM SIGMOD Record*, 16(3), 323-330.
- [PGQ] International Organization for Standardization. (2023). *SQL — Part 16: SQL Property Graph Queries (SQL/PGQ)*. <https://www.iso.org/standard/79473.html>. Standard ISO/IEC 9075-16:2023.
- [GQL] International Organization for Standardization. (2024). *GQL*. <https://www.iso.org/standard/76120.html>. Standard ISO/IEC 39075:2024
- [MNP23] Martens, W., Niewerth, M., & Popp, T. (2023). “A trichotomy for regular trail queries”. *Logical Methods in Computer Science*, 19.
- [DFM23] David, C., Francis, N., & Marsault, V. (2023). “Run-based semantics for RPQs”. *Proceedings of the 20th International Conference on Principles of Knowledge Representation and Reasoning*.

# Annexe

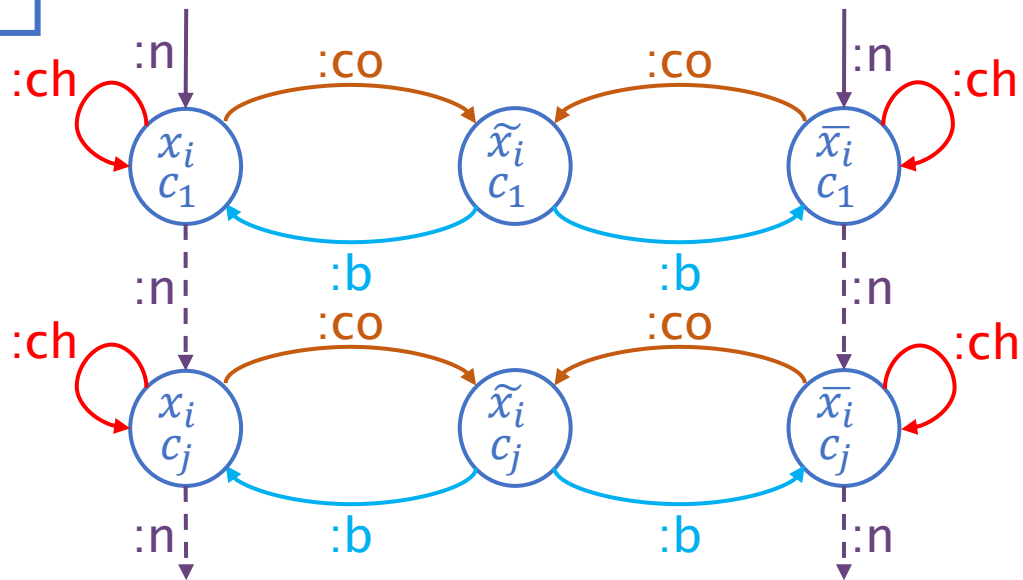
# Variable sous étoile

$$R = [(u) \xrightarrow{\text{:choose}} \xrightarrow{\text{:commit}} \xrightarrow{\text{:back}} (u) \xrightarrow{\text{:next}} ]^*$$



# Variable sous étoile

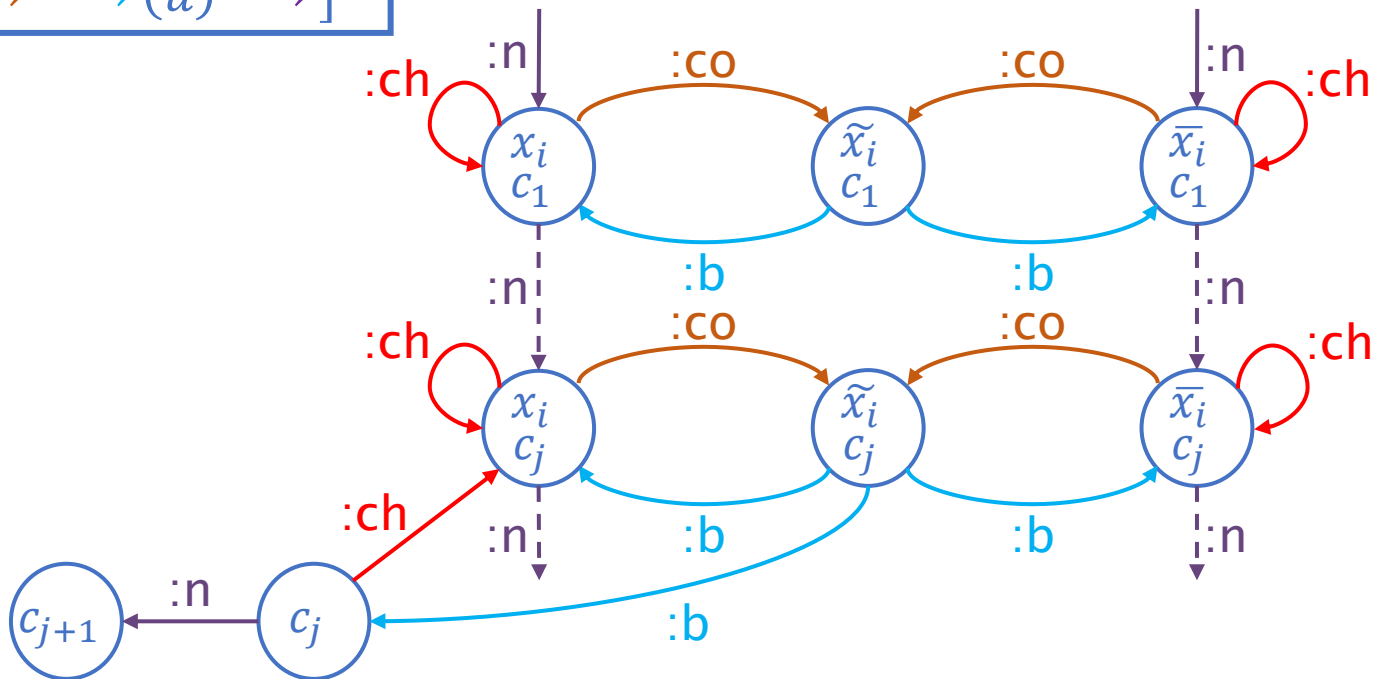
$$R = [(u) \xrightarrow{\text{:choose}} \xrightarrow{\text{:commit}} \xrightarrow{\text{:back}} (u) \xrightarrow{\text{:next}} ]^*$$



# Variable sous étoile

$$R = [(u) \xrightarrow{\text{:choose}} \xrightarrow{\text{:commit}} \xrightarrow{\text{:back}} (u) \xrightarrow{\text{:next}} ]^*$$

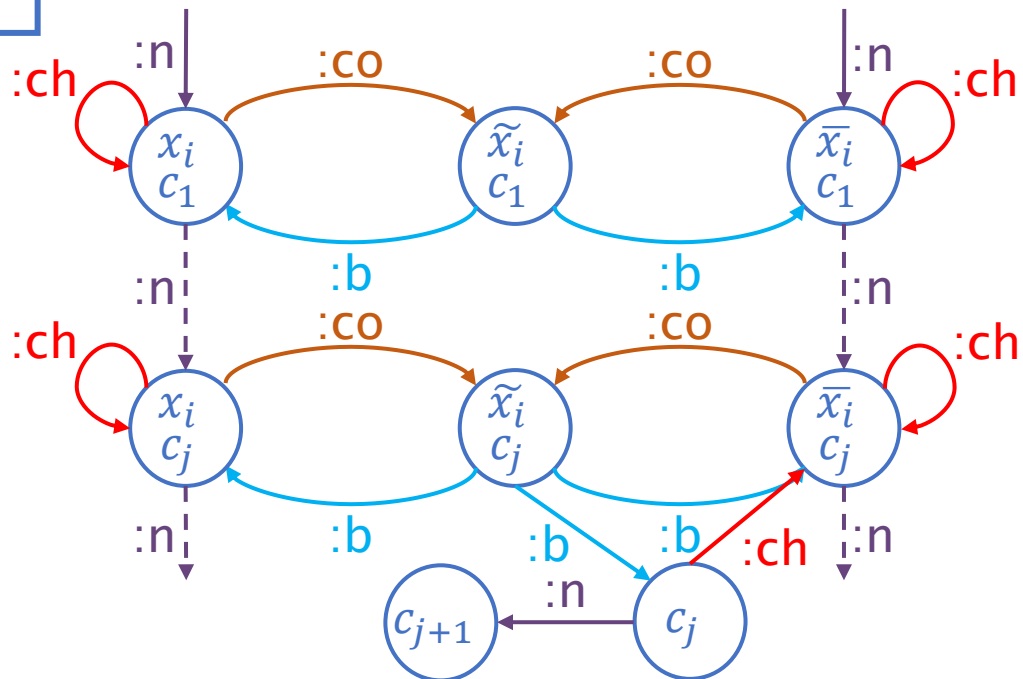
Si  $x_i$  apparaît positivement dans  $c_j$



# Variable sous étoile

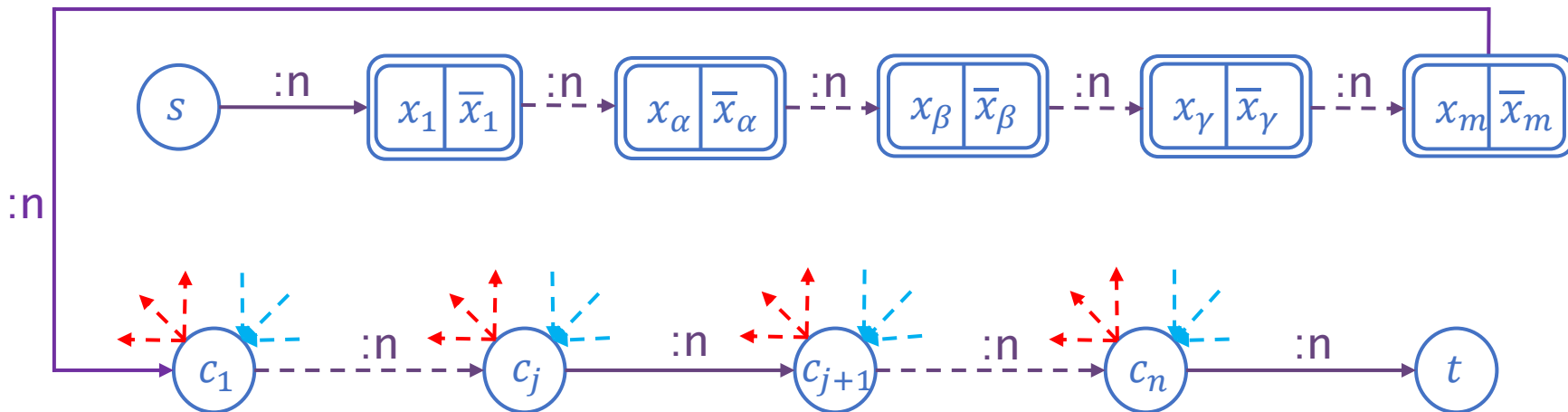
$$R = [(u) \xrightarrow{\text{:choose}} \xrightarrow{\text{:commit}} \xrightarrow{\text{:back}} (u) \xrightarrow{\text{:next}} ]^*$$

Si  $x_i$  apparaît négativement dans  $c_j$



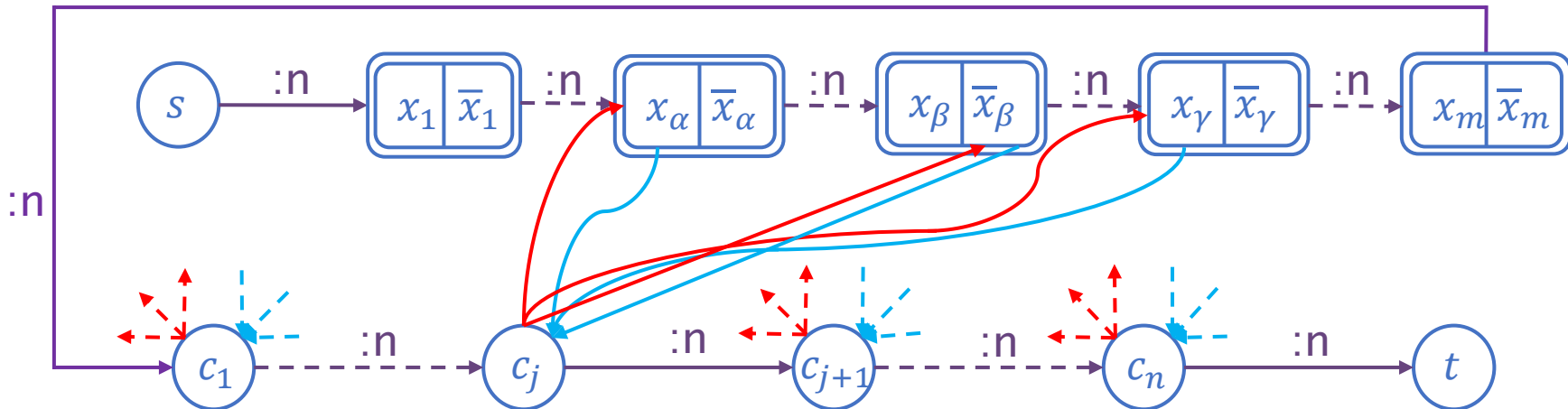
# Variable sous étoile

$$R = [(u) \xrightarrow{\text{:choose}} \xrightarrow{\text{:commit}} \xrightarrow{\text{:back}} (u) \xrightarrow{\text{:next}} ]^*$$



# Variable sous étoile

$$R = [(u) \xrightarrow{\text{choose}} \xrightarrow{\text{commit}} \xrightarrow{\text{back}} (u) \xrightarrow{\text{next}} ]^*$$



$$\text{Avec } c_j = (x_\alpha \vee \neg x_\beta \vee x_\gamma)$$