

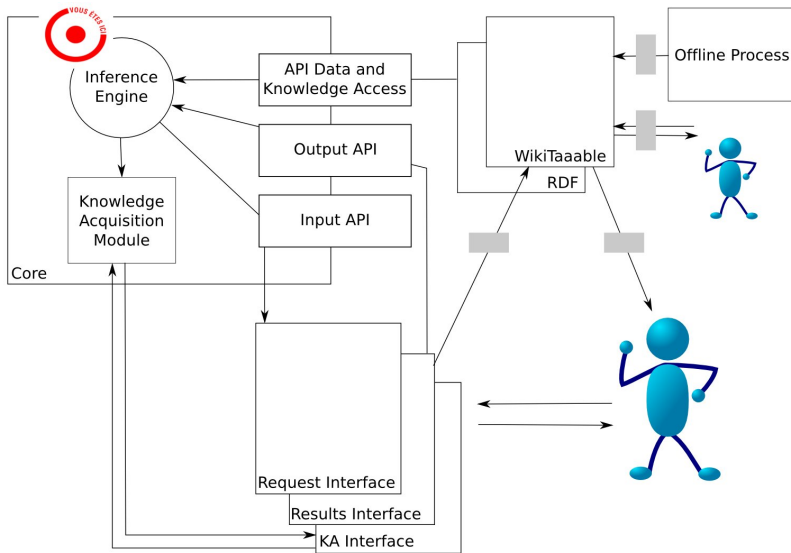
Le moteur de raisonnement à partir de cas de WikiTaaable

Comment l'approche de la remémoration-adaptation s'appuyant sur des reformulations peut s'appliquer sur une base de cas relativement volumineuse et peu organisée

Jean Lieber, équipe Orpailleur, LORIA

17^{ème} atelier RàPC, 29 et 30 juin 2009

Introduction



- ▶ Les reformulations, les chemins de similarité (en général et pour WikiTaaable)
- ▶ Les connaissances manipulées par WikiTaaable
- ▶ Choix des coûts
- ▶ Vers un nouveau moteur
- ▶ Conclusion et perspectives

Les reformulations, les chemins de similarité

Problème, solution, cas, remémoration et adaptation

- ▶ Problème : élément de Problèmes
- ▶ Solution : élément de Solutions
- ▶ Relation « a pour solution » sur Problèmes \times Solutions imparfaitement connue
- ▶ Mais connue pour chaque *cas source* :
(srce, Sol(srce)) \in Problèmes \times Solutions
tel que srce a pour solution Sol(srce)
- ▶ Remémoration et adaptation

Problème, solution, cas, remémoration et adaptation

- ▶ Problème : élément de Problèmes
- ▶ Solution : élément de Solutions
- ▶ Relation « a pour solution » sur Problèmes \times Solutions imparfaitement connue
- ▶ Mais connue pour chaque *cas source* :
(srce, Sol(srce)) \in Problèmes \times Solutions
tel que srce a pour solution Sol(srce)
- ▶ Remémoration et adaptation

cible

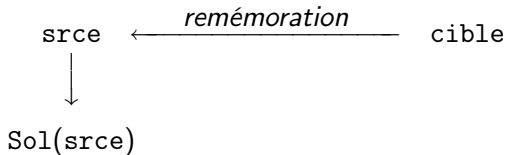
Problème, solution, cas, remémoration et adaptation

- ▶ Problème : élément de Problèmes
- ▶ Solution : élément de Solutions
- ▶ Relation « a pour solution » sur Problèmes \times Solutions imparfaitement connue
- ▶ Mais connue pour chaque *cas source* :
(srce, Sol(srce)) \in Problèmes \times Solutions
tel que srce a pour solution Sol(srce)
- ▶ Remémoration et adaptation

srce $\xleftarrow{\text{remémoration}}$ cible

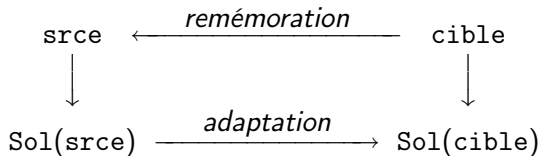
Problème, solution, cas, remémoration et adaptation

- ▶ Problème : élément de Problèmes
- ▶ Solution : élément de Solutions
- ▶ Relation « a pour solution » sur Problèmes \times Solutions imparfaitement connue
- ▶ Mais connue pour chaque *cas source* :
(srce, Sol(srce)) \in Problèmes \times Solutions
tel que srce a pour solution Sol(srce)
- ▶ Remémoration et adaptation



Problème, solution, cas, remémoration et adaptation

- ▶ Problème : élément de Problèmes
- ▶ Solution : élément de Solutions
- ▶ Relation « a pour solution » sur Problèmes \times Solutions imparfaitement connue
- ▶ Mais connue pour chaque *cas source* :
(srce, Sol(srce)) \in Problèmes \times Solutions
tel que srce a pour solution Sol(srce)
- ▶ Remémoration et adaptation



Problème, solution, cas, remémoration et adaptation pour WikiTaaable

- Problème : requête Q

$$Q = \text{scarole} \wedge \text{jus_citron} \wedge \neg \text{oignon}$$

- Solution : une recette R

$$R = \text{scarole} \wedge \text{jus_citron} \wedge \text{huile_olive} \wedge \text{sel} \wedge \text{Rien d'autre}$$

- $\left\{ \begin{array}{l} R \text{ répond à } Q \\ Q \text{ a pour solution } R \end{array} \right\}$ noté $R \models_{\mathcal{O}} Q$

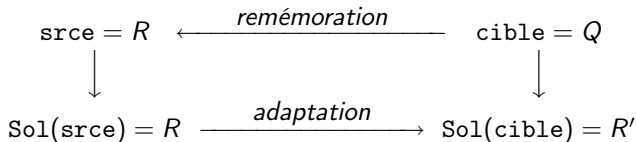
- Cas source

- $\text{Sol}(\text{srce}) = R$

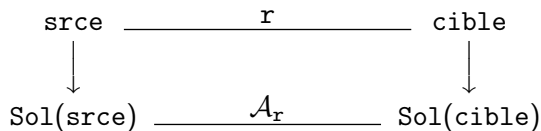
- srce : une des requêtes auquel répond $\text{Sol}(\text{srce})$

On choisit $\text{srce} = R$

- $(\text{srce}, \text{Sol}(\text{srce})) = (R, R)$



► Reformulation (r, \mathcal{A}_r)

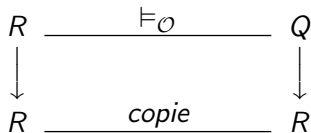


► Interprétation

Si $\text{srce } r \text{ cible}$
alors $\mathcal{A}_r : (\text{srce}, \text{Sol}(\text{srce}), \text{cible}) \mapsto \text{Sol}(\text{cible})$

Les reformulations pour WikiTaaable

► Reformulation ($\models_{\mathcal{O}}$, *copie*)



Les reformulations pour WikiTaaable

- ▶ Reformulation $(\models_{\mathcal{O}}, copie)$
- ▶ Reformulations $(r, \mathcal{A}_r) = (s, s)$ (s : une substitution)

$$\begin{array}{ccc} R & \xrightarrow{s} & Q \\ \downarrow & & \downarrow \\ R & \xrightarrow{s} & s(R) \end{array}$$

Les reformulations pour WikiTaaable

- ▶ Reformulation ($\models_{\mathcal{O}}, copie$)
- ▶ Reformulations $(r, \mathcal{A}_r) = (s, s)$ (s : une substitution)
- ▶ Deux types : σ et γ^{-1}
- ▶ σ : connaissances d'adaptation en cuisine

$$\sigma = \left| \begin{array}{l} \text{salade} \wedge \neg p_de_terre \\ \wedge \text{vinaigre} \end{array} \right| \rightsquigarrow \left| \begin{array}{l} \text{salade} \wedge \neg p_de_terre \\ \wedge \text{jus_de_citron} \wedge \text{sel} \end{array} \right|$$

$$\begin{array}{ccc} R & \xrightarrow{\sigma} & Q \\ \downarrow & & \downarrow \\ R & \xrightarrow{\sigma} & \sigma(R) \end{array}$$

Les reformulations pour WikiTaaable

- ▶ Reformulation ($\models_{\mathcal{O}}, copie$)
- ▶ Reformulations $(r, \mathcal{A}_r) = (s, s)$ (s : une substitution)
- ▶ Deux types : σ et γ^{-1}

- ▶ σ : connaissances d'adaptation en cuisine

$$\sigma = \left| \begin{array}{l} \text{salade} \wedge \neg \text{p_de_terre} \\ \wedge \text{vinaigre} \end{array} \right| \rightsquigarrow \left| \begin{array}{l} \text{salade} \wedge \neg \text{p_de_terre} \\ \wedge \text{jus_de_citron} \wedge \text{sel} \end{array} \right|$$

- ▶ γ^{-1} : spécialisation issue de l'ontologie

$$\gamma^{-1} = \text{agrumes} \rightsquigarrow \text{citron} \quad \text{i.e., } \gamma = \text{citron} \rightsquigarrow \text{agrumes}$$

$$\begin{array}{ccc} R = \gamma(Q) & \xrightarrow{\text{agrumes} \leftarrow \text{citron}} & Q \\ \downarrow & & \downarrow \\ R & \xrightarrow{\text{agrumes} \rightsquigarrow \text{citron}} & \gamma^{-1}(R) \end{array}$$

Les chemins de similarité

- ▶ On peut composer les reformulations :

$$\begin{array}{ccccccc} \text{srce} & \xrightarrow{r_1} & \text{pb}_1 & \xrightarrow{r_2} & \text{pb}_2 & \xrightarrow{r_3} & \text{cible} \\ \downarrow & & \downarrow & & \downarrow & & \downarrow \\ \text{Sol}(\text{srce}) & \xrightarrow{\mathcal{A}_{r_1}} & \text{Sol}(\text{pb}_1) & \xrightarrow{\mathcal{A}_{r_2}} & \text{Sol}(\text{pb}_2) & \xrightarrow{\mathcal{A}_{r_3}} & \text{Sol}(\text{cible}) \end{array}$$

- ▶ Chemin de similarité de srce à cible :

$$\text{srce} = \text{pb}_0 \ r_1 \ \text{pb}_1 \ \dots \ \text{pb}_{n-1} \ r_n \ \text{pb}_n = \text{cible}$$

- ▶ Chemin d'adaptation :

$$\text{Sol}(\text{srce}) \xrightarrow{\mathcal{A}_{r_1}} \text{Sol}(\text{pb}_1) \dots \text{Sol}(\text{pb}_{n-1}) \xrightarrow{\mathcal{A}_{r_n}} \text{Sol}(\text{cible})$$

- ▶ Coût d'un chemin de similarité

- ▶ Mesure de l'« effort d'adaptation »
- ▶ Supposé additif et ne dépendant que des r_i : $\sum_i \text{coût}(r_i)$
- ▶ Utilisé pour effectuer une préférence entre chemins

Les chemins de similarité pour WikiTaaable

- ▶ S'appuient sur les σ , les γ et $\models_{\mathcal{O}}$:

$$R \xrightarrow{\sigma_1} \dots \xrightarrow{\sigma_p} \Sigma(R) \models_{\mathcal{O}} \Gamma(Q) \xleftarrow{\gamma_q} \dots \xleftarrow{\gamma_1} Q$$

- ▶ Chemin d'adaptation correspondant :

$$R \xrightarrow{\sigma_1} \dots \xrightarrow{\sigma_p} \Sigma(R) \xrightarrow{\text{copie}} \Sigma(R) \xrightarrow{\gamma_q^{-1}} \dots \xrightarrow{\gamma_1^{-1}} \Gamma^{-1}(\Sigma(R))$$

- ▶ Algorithme d'appariement

- ▶ A* paramétré par la fonction coût
- ▶ État initial $[R ? Q]$
- ▶ État final $\left[R \xrightarrow{\sigma_1} \dots \xrightarrow{\sigma_p} \Sigma(R) \models_{\mathcal{O}} \Gamma(Q) \xleftarrow{\gamma_q} \dots \xleftarrow{\gamma_1} Q \right]$

Remémoration guidée par l'adaptation (RGA)

- ▶ RGA [Smyth et Keane] :
prise en compte de l'adaptation et de l'adaptabilité
lors de la remémoration
- ▶ Une remémoration qui retourne
 - ▶ Un cas source ($srce, Sol(srce)$)
 - ▶ Un chemin de similarité de $srce$ à $cible$est une RGA.
- ▶ L'idéal : trouver le chemin de similarité optimal
- ▶ Mais chercher à la fois à modifier tous les cas sources et le problème $cible$ est coûteux en temps.
- ▶ Solution 1 : travailler sur une organisation de la base de cas et faire les modifications sur cette organisation
- ▶ Solution 2 : ne faire que les modifications sur $cible$
(solution pour [WikiTaaable](#))

Remémoration guidée par l'adaptation (RGA) pour WikiTaaable

- ▶ Algorithme :
 1. $\Gamma := \text{identité}$
 2. **tant que** $\neg \exists R, R \models_{\mathcal{O}} \Gamma(Q)$ **faire**
 3. $\Gamma := \text{prochaineModification}()$
 4. **fin tant que**
- ▶ Le test $\exists R, R \models_{\mathcal{O}} \Gamma(Q)$
est effectué par une classification hiérarchique de $\Gamma(Q)$
- ▶ Γ parcourt l'espace des modifications de Q
par $\text{coût}(\Gamma)$ croissant
- ▶ Résultat :
des recettes R telles que $R \models_{\mathcal{O}} \Gamma(Q)$ à $\text{coût}(\Gamma)$ minimal

Les connaissances manipulées par WikiTaaable

- ▶ \mathcal{O} : ensemble d'axiomes de la forme $a \Rightarrow b$
(citron \Rightarrow agrume) $\in \mathcal{O}$
 - ▶ D'où la relation $\models_{\mathcal{O}}$ entre conjonctions de littéraux
 - ▶ D'où les substitutions par généralisation $\gamma = a \rightsquigarrow b$
 $\gamma = \text{citron} \rightsquigarrow \text{agrume}$
- ▶ Représentée sous la forme d'une hiérarchie de racine \top
- ▶ $\mathcal{E}(\top)$ = l'univers des recettes
- ▶ $\mathcal{E}(\text{citron})$ = les recettes avec du citron
- ▶ $\mathcal{E}(\text{Italie})$ = les recettes d'origine italienne
- ▶ $\mathcal{E}(\text{dîner})$ = les recettes de plats qu'on peut manger au dîner

- ▶ Dans cet exposé : R représentation formelle d'une recette
- ▶ (Dans le papier : $idx(R)$)

Les connaissances d'adaptation

Elles sont données par :

- ▶ Les substitutions σ
- ▶ Les substitutions γ issues de l'ontologie
- ▶ La fonction de coût

D'où viennent ces connaissances ?

- ▶ Acquisition et extraction de connaissances (initiales et opportunistes)
- ▶ Annotations des recettes
- ▶ Stockées sur le Wiki sémantique

Choix des coûts

Additivité du coût

- Coût d'un chemin de similarité :

$$\text{coût}(\text{pb}_0 \text{ r}_1 \text{ pb}_1 \dots \text{pb}_{n-1} \text{ r}_n \text{ pb}_n) = \sum_i \text{coût}(\text{r}_i)$$

- Donc

$$\text{coût}(s_2 \circ s_1) = \text{coût}(s_1) + \text{coût}(s_2)$$

- D'où

$$\text{coût}(a \rightsquigarrow c) = \text{coût}(a \rightsquigarrow b) + \text{coût}(b \rightsquigarrow c)$$

- S'applique à a, b, c tels que $c \models_{\mathcal{O}} b \models_{\mathcal{O}} a$

$$\begin{aligned} \text{coût}(\text{fruit} \rightsquigarrow \text{citron}) &= \text{coût}(\text{fruit} \rightsquigarrow \text{agrumes}) \\ &\quad + \text{coût}(\text{agrumes} \rightsquigarrow \text{citron}) \end{aligned}$$

- À l'encontre du choix $\text{coût}(b \rightsquigarrow a) = 1$ pour $(a \Rightarrow b) \in \mathcal{O}$
 - Rendrait le coût dépendant de l'ajout d'un concept dans \mathcal{O} !
L'ajout de *agrumes* dans \mathcal{O} ne devrait pas augmenter l'effort d'adaptation *fruit* \rightsquigarrow *citron* !

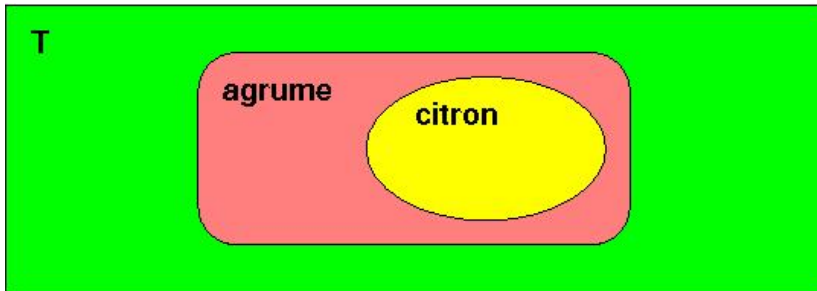
Première idée pour coût(γ^{-1})

$$\text{coût}(a \rightsquigarrow b) = K (\mu(a) - \mu(b))$$

où $K > 0$ est une constante,

$$\text{où } \mu(x) = \frac{\mathcal{N}(x)}{\mathcal{N}(\mathbb{T})} \text{ pour toute classe } x \text{ de } \mathcal{O}$$

et où $\mathcal{N}(x)$ = le nombre de recettes indexées par x
= $\text{card} \{R \text{ du livre de recettes} \mid R \models_{\mathcal{O}} x\}$



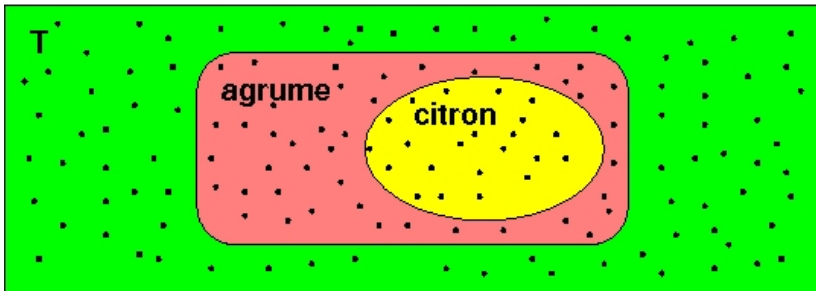
Première idée pour coût(γ^{-1})

$$\text{coût}(a \rightsquigarrow b) = K (\mu(a) - \mu(b))$$

où $K > 0$ est une constante,

$$\text{où } \mu(x) = \frac{\mathcal{N}(x)}{\mathcal{N}(\top)} \text{ pour toute classe } x \text{ de } \mathcal{O}$$

et où $\mathcal{N}(x)$ = le nombre de recettes indexées par x
= $\text{card} \{R \text{ du livre de recettes} \mid R \models_{\mathcal{O}} x\}$



Deuxième idée pour coût(γ^{-1})

Cf. le papier

Vers un nouveau moteur

Des cas-recettes spécifiques aux cas-recettes généraux

Cf. le papier

Extension du langage à l'aide d'attributs

- Prise en compte de relations entre classes et des quantités

$C(R_{\text{tap}})$ avec $C \equiv \text{Tarte} \sqcap \text{Dessert} \sqcap \exists.\text{prNbPersonnes}.=4$
 $\sqcap \exists\text{ingrédient}.(\text{Pomme} \sqcap \exists\text{quantité}.=3)$
 $\sqcap \exists\text{ingrédient}.(\text{Sucre} \sqcap \exists\text{masseG}.=40)$
 $\sqcap \exists\text{ingrédient}.\text{PâteSablée}$

$\text{Pomme} \sqsubseteq \text{Fruit}$

$Q \equiv \text{Tarte} \sqcap \exists.\text{prNbPersonnes}.=8$
 $\sqcap \exists\text{ingrédient}.\text{Poire}$
 $\sqcap \exists\text{ingrédient}.(\text{Sucre} \sqcap \exists\text{masseG}.\leq 60)$

- Sorte de $\mathcal{EL}(\mathcal{D})$
 - Mais intégrant l'hypothèse du monde clos
 $\models_{\mathcal{O}} (\neg \exists\text{ingrédient}.\text{Poisson})(R_{\text{tap}})$
 - Et la non-monotonie...

Gérer la non-monotonie et la typicalité

- ▶ Non monotonie, typicalité, valeurs par défaut.
 - ▶ Le lait en poudre est du lait.
 - ▶ Le lait est liquide.
 - ▶ Le lait en poudre n'est pas liquide.

Gérer la non-monotonie et la typicalité

- ▶ Non monotonie, typicalité, valeurs par défaut.
 - ▶ Le lait en poudre est du lait.
 - ▶ Le lait est *typiquement* liquide.
 - ▶ Le lait en poudre n'est pas liquide.

Gérer la non-monotonie et la typicalité

- ▶ Non monotonie, typicalité, valeurs par défaut.
 - ▶ Le lait en poudre est du lait.
 - ▶ Le lait est *typiquement* liquide.
 - ▶ Le lait en poudre n'est pas liquide.
- ▶ Comment représenter/encoder cela ?
 - ▶ LaitEnPoudre sous-classe de Lait
 - ▶ laitPrototype instance de Lait et de Liquide
 - ▶ laitEnPoudrePrototype instance de \neg Liquide et de LaitEnPoudre

Gérer la non-monotonie et la typicalité

- ▶ Non monotonie, typicalité, valeurs par défaut.
 - ▶ Le lait en poudre est du lait.
 - ▶ Le lait est *typiquement* liquide.
 - ▶ Le lait en poudre n'est pas liquide.
- ▶ Comment représenter/encoder cela ?
 - ▶ LaitEnPoudre sous-classe de Lait
 - ▶ laitPrototype instance de Lait et de Liquide
 - ▶ laitEnPoudrePrototype instance de \neg Liquide et de LaitEnPoudre
- ▶ Oui, mais les inférences ?

Conclusion et perspectives

Conclusion

Y a eu des trucs de faits.

Y a encore du boulot.

Cimetière de transparents

Critique de cette première idée pour coût(γ^{-1})

- ▶ Coût de agrume \rightsquigarrow citron :
mesure les (mal)chances que la substitution de
agrume $\wedge \neg$ citron par citron donne une bonne recette.
 - ▶ On fait une substitution et on espère que le résultat sera acceptable.
- ▶ Coût de Asie \rightsquigarrow Chine :
mesure les (mal)chances qu'une recette indexée par Asie mais
pas par Chine passe pour une recette chinoise.
 - ▶ On ne fait pas de substitution et on espère que la copie de la
recette (vietnamienne ? indienne ?) donnera satisfaction à
l'utilisateur...
- ▶ Deux processus de natures différentes :
les coûts devraient être très différents.

Deuxième idée pour coût(γ^{-1})

- On associe à chaque classe de l'ontologie des points de vue.

$\text{pdv}(\text{citron}) = \text{pdv}(\text{agrume}) = \text{pdv_ingrédient}$

$\text{pdv}(\text{Chine}) = \text{pdv}(\text{Asie}) = \text{pdv_géographie}$

- On s'en sert pour définir la fonction de coût :

$$\text{coût}(a \rightsquigarrow b) = \begin{cases} K_{\text{pdv}(a)} (\mu(a) - \mu(b)) & \text{si } \text{pdv}(a) = \text{pdv}(b) \\ +\infty & \text{sinon} \end{cases}$$

où $K_{\text{pdv}} > 0$ ne dépend que du point de vue pdv

- C'est cette idée qui est actuellement implantée.

Des cas-recettes spécifiques aux cas-recettes généraux

- $R_{\text{tap}} = \text{tarte} \wedge \text{pomme} \wedge \text{p\^ate} \wedge \text{sucres} \wedge \text{Rien d'autre}$

Actuellement

- Une recette R est interprétée comme une instance.
- Hypothèse du monde clos : si $R \not\models_{\mathcal{O}} a$ alors $R \models_{\mathcal{O}} \neg a$
 $R_{\text{tap}} \models_{\mathcal{O}} \text{Rien d'autre} \equiv \neg \text{poisson} \wedge \neg \text{rainette} \wedge \dots$
- Donc, R_{tap} n'est pas une solution de $Q = \text{rainette}$!
- Il faut adapter R_{tap} pour résoudre $Q = \text{rainette}$!

$$R_{\text{tap}} \models_{\mathcal{O}} \Gamma(Q) \xrightarrow{\text{pomme} \rightsquigarrow \text{rainette}} Q$$

Bientôt

- Une recette R sera interprétée comme une classe.
- Recherche des R telles que $R \wedge Q \wedge \mathcal{O}$ est satisfiable
- Pour $Q = \text{rainette}$, $R_{\text{tap}} \wedge Q \wedge \mathcal{O}$ est satisfiable
- Application directe de R_{tap} pour une recette avec des rainettes
- Plus généralement, chemin d'adaptation :

$$R \mapsto \Sigma(R) \mapsto \boxed{\Sigma(R) \wedge \Gamma(Q)} \mapsto \Gamma^{-1}(\Sigma(R) \wedge \Gamma(Q))$$

- Problème ouvert : nouvelle hypothèse du monde clos ?

$$R_{\text{tap}} \models_{\mathcal{O}} \neg \text{poisson} \quad \text{mais} \quad R_{\text{tap}} \not\models_{\mathcal{O}} \neg \text{rainette}$$