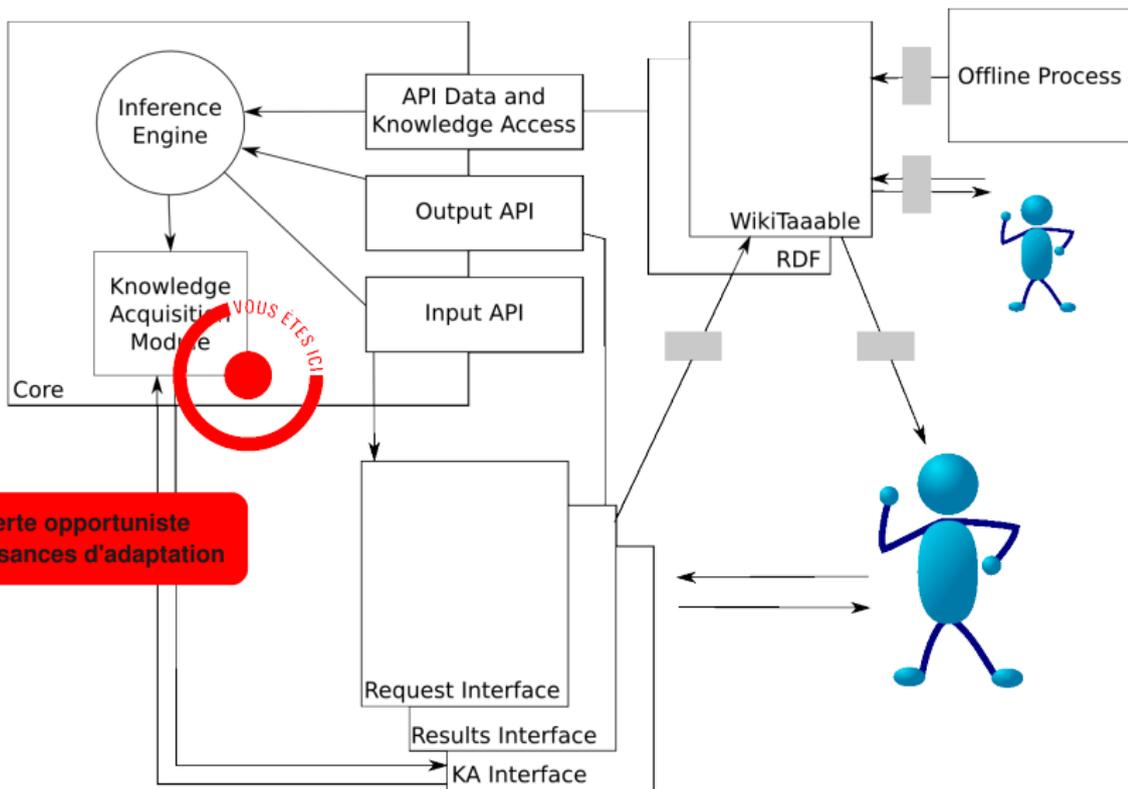


Découverte opportuniste de connaissances d'adaptation

Fadi Badra¹ Amélie Cordier² Jean Lieber¹

¹LORIA (CNRS, INRIA, Nancy université)
BP 239, 54506 Vandœuvre-lès-Nancy, France
{badra, lieber}@loria.fr

²LIRIS CNRS UMR 5202, Université Lyon 1, INSA Lyon, Université Lyon 2, ECL
43, bd du 11 novembre 1918, Villeurbanne, France
Amélie.Cordier@liris.cnrs.fr



Découverte opportuniste de connaissances d'adaptation



Contexte

Les connaissances d'adaptation

Un exemple culinaire.

« Réaliser une recette de pancake avec des piridions. »

« Réaliser une recette de pancake aux poires. »



« Suivre la recette *Apple pancakes from the townships.* »

?

Les connaissances d'adaptation

Un exemple culinaire.

« Réaliser une recette de pancake avec des piridions. »

« Réaliser une recette de pancake aux poires. »



« Suivre la recette *Apple pancakes from the townships*. »

« Suivre la recette *Apple pancakes from the townships*,
mais remplacer les pommes
par les poires. »

Les connaissances d'adaptation

Un exemple culinaire.

« Réaliser une recette de pancake avec des piridions. »

« Réaliser une recette de pancake aux poires. »



« Suivre la recette *Apple pancakes from the townships*. »

« Suivre la recette *Apple pancakes from the townships*,
mais remplacer les pommes
par les poires et remplacer la
cannelle par de la vanille. »

L'acquisition de connaissances d'adaptation

Principales questions posées

- De quelles connaissances le système a-t-il besoin pour réaliser l'adaptation ?
- À partir de quelles sources les acquérir (humain, connaissances du domaine, base de cas, le Web) ?
- Quand réaliser l'acquisition (hors ligne, en ligne) ?
- Quelle méthode utiliser pour les acquérir (manuelle, automatique, semi-automatique) ?

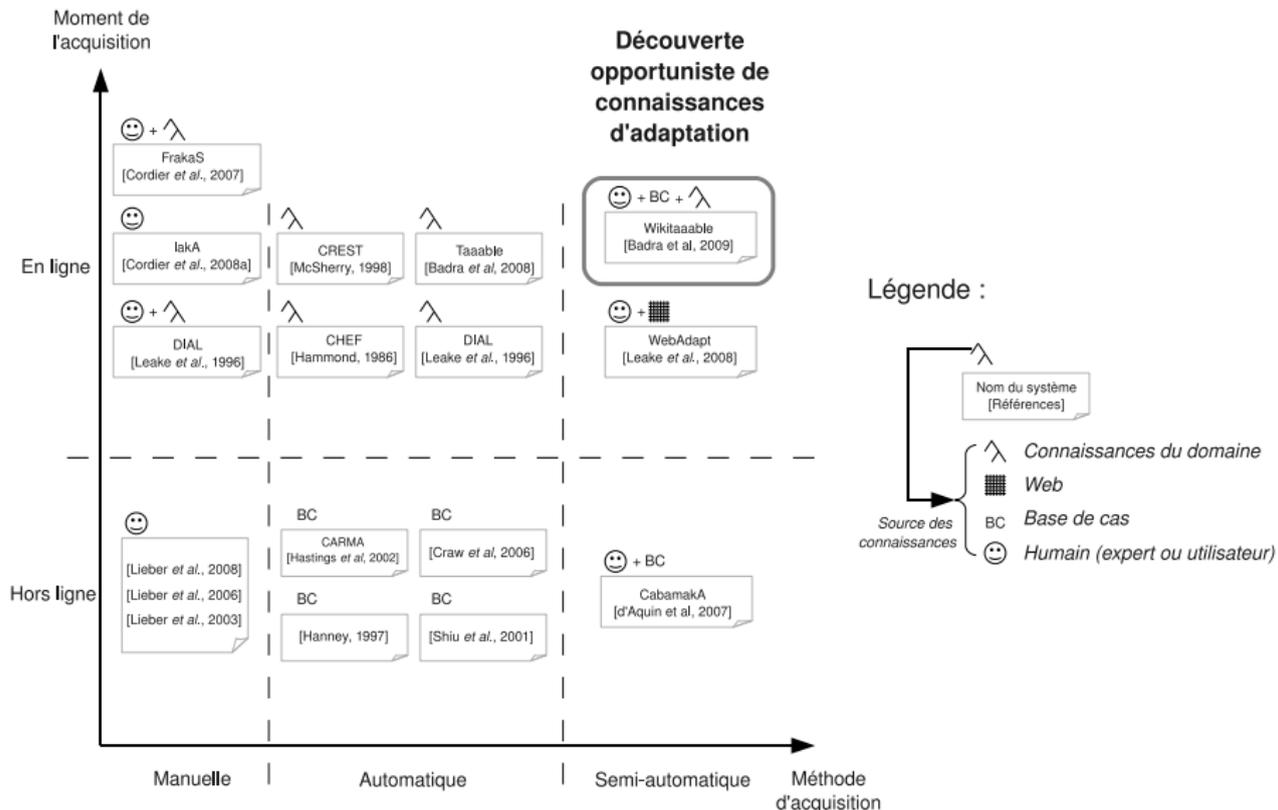
Contributions

- Une méthode d'**acquisition de connaissances d'adaptation**.

Celle-ci combine deux stratégies :

- 1 l'acquisition de connaissances est réalisée **à partir de la base de cas par des techniques d'extraction de connaissances**,
 - 2 l'acquisition de connaissances est déclenchée **en ligne et de manière opportuniste** au cours d'une session particulière de résolution de problème.
- Application au système de RÀPC culinaire TAAABLE.

Principales approches d'ACA

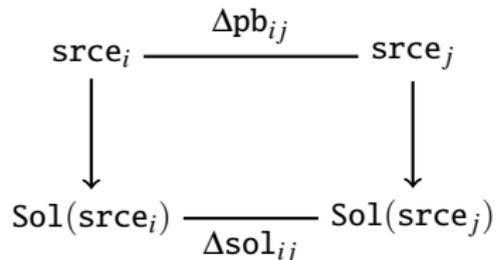


CabamakA

(CAsE-BAsE Mining for Adaptation Knowledge Acquisition)

Processus d'extraction de connaissances d'adaptation **à partir de la base de cas** (méthode **semi-automatique, hors-ligne**).

- On apprend des connaissances d'adaptation à partir des variations entre cas.



- On utilise des couples $(\Delta pb_{ij}, \Delta sol_{ij})$ comme ensemble d'apprentissage :

$$\{(\Delta pb_{ij}, \Delta sol_{ij})\}_{ij} \mapsto CA$$

- On apprend par généralisation (extraction de motifs fermés fréquents)

Application à Taaable

CABAMAKA permet d'apprendre des opérateurs de substitution d'ingrédients de la comparaison de deux ensembles de recettes.

$$R_k = \text{dessert} \wedge \text{pancake} \wedge \text{pomme} \wedge \text{piridion} \wedge \text{fruit} \wedge \text{cannelle} \wedge \dots$$
$$R_\ell = \text{dessert} \wedge \text{tarte} \wedge \text{poire} \wedge \text{piridion} \wedge \text{fruit} \wedge \text{vanille} \wedge \dots$$

↓ apprentissage par généralisation

$$\sigma = \text{dessert} \wedge \text{pomme} \wedge \text{cannelle} \rightsquigarrow \text{dessert} \wedge \text{poire} \wedge \text{vanille}$$

« Dans les recettes de dessert, on peut remplacer les pommes et la cannelle par des poires et de la vanille. »

Application à Taaable

Ensemble d'apprentissage \mathcal{E}

= ensemble de représentations Δ_{kl} de variations entre recettes R_k et R_ℓ

- $a^- \subseteq \Delta_{kl}$ si $R_k \models_{\mathcal{O}} a$ et $R_\ell \not\models_{\mathcal{O}} a$
- $a^+ \subseteq \Delta_{kl}$ si $R_k \not\models_{\mathcal{O}} a$ et $R_\ell \models_{\mathcal{O}} a$
- $a^\pm \subseteq \Delta_{kl}$ si $R_k \models_{\mathcal{O}} a$ et $R_\ell \models_{\mathcal{O}} a$

$R_k = \text{dessert} \wedge \text{pancake} \wedge \text{pomme} \wedge \text{piridion} \wedge \text{fruit} \wedge \text{cannelle} \wedge \dots$

$R_\ell = \text{dessert} \wedge \text{tarte} \wedge \text{poire} \wedge \text{piridion} \wedge \text{fruit} \wedge \text{vanille} \wedge \dots$

↓ représentation des variations

$\Delta_{kl} = \{\text{dessert}^\pm, \text{pancake}^-, \text{tarte}^+, \text{fruit}^\pm, \text{piridion}^\pm, \text{pomme}^-, \text{poire}^+, \text{cannelle}^-, \text{vanille}^+\}$

Application à Taaable

Apprentissage par généralisation

Extraction de motifs fréquents $\Delta \subseteq \Delta_{kl}$.

Un motif fréquent Δ s'interprète comme une substitution σ .

$$\Delta_{kl} = \{\text{dessert}^-, \text{pancake}^-, \text{tarte}^+, \text{fruit}^-, \text{piridion}^-, \text{pomme}^-, \text{poire}^+, \text{cannelle}^-, \text{vanille}^+\}$$

↓ généralisation

$$\Delta = \{\text{dessert}^-, \text{pomme}^-, \text{cannelle}^-, \text{poire}^+, \text{vanille}^+\}$$

↓ interprétation

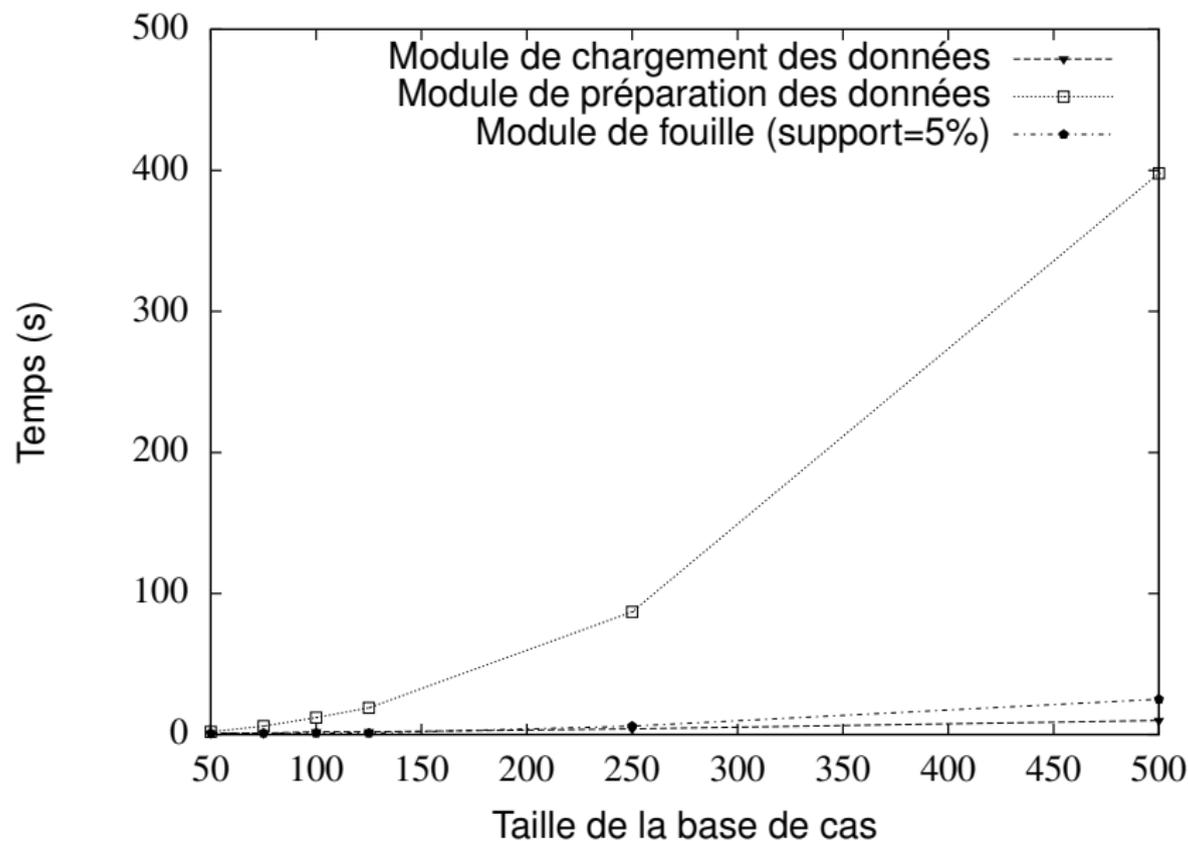
$$\sigma = \text{dessert} \wedge \text{pomme} \wedge \text{cannelle} \rightsquigarrow \text{dessert} \wedge \text{poire} \wedge \text{vanille}$$

Problème du choix de l'ensemble d'apprentissage

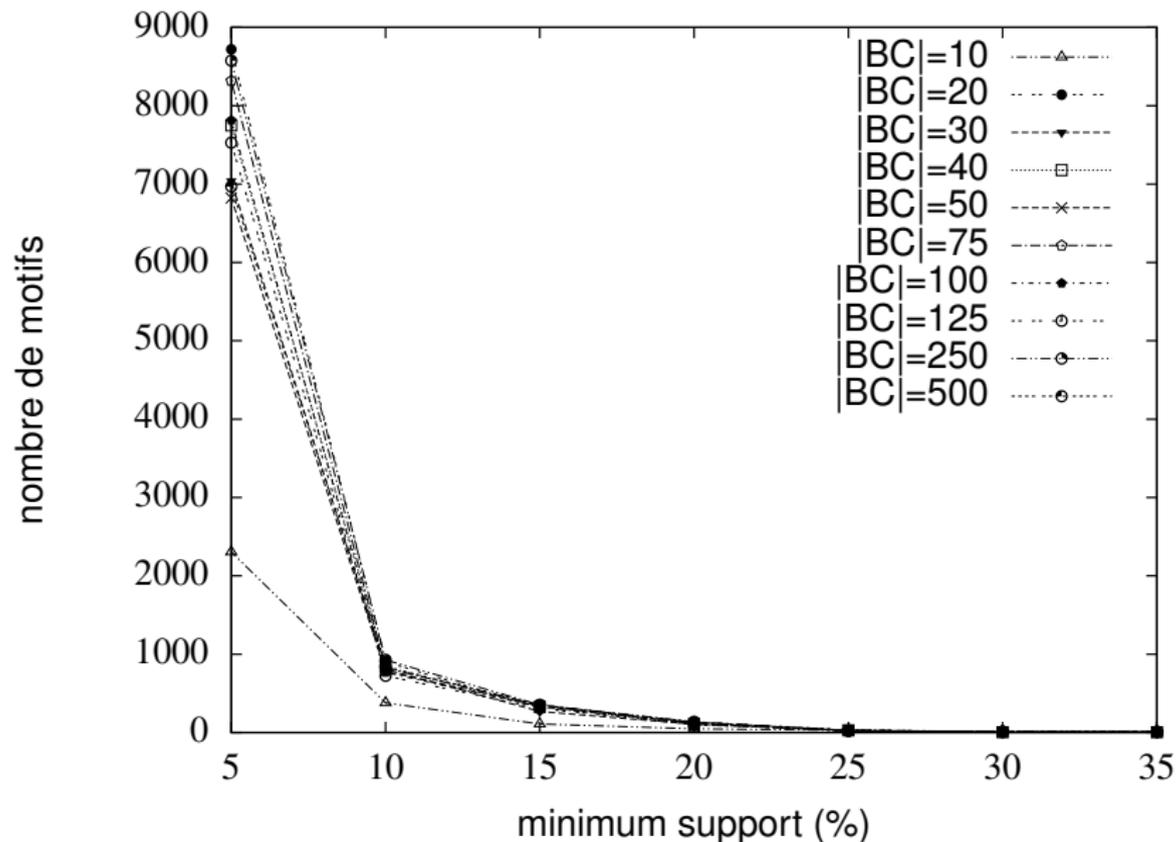
L'approche est impraticable si l'on ne restreint pas l'ensemble d'apprentissage !

- Si l'on compare les $n = 1482$ recettes de la base de cas entre elles, la taille de l'ensemble d'apprentissage est $n(n - 1) = 2\,194\,842$!
- On comparerait des recettes de pancake et des recettes de cocktail !
- On ne pourrait apprendre que des connaissances très générales, les connaissances les plus spécifiques seraient noyées dans l'ensemble des résultats.

Problème du choix de l'ensemble d'apprentissage



Nécessité de restreindre l'espace des hypothèses



Limitations de l'apprentissage hors ligne

- apprentissage très coûteux
 - ▶ besoin d'un expert
 - ▶ besoin d'un ingénieur de la connaissance pour sélectionner dans les résultats des règles à présenter à l'expert et les interpréter,
- Il faudrait anticiper toutes les connaissances nécessaires au système : le système n'apprend pas à adapter avec l'expérience. . .

On se tourne vers les approches interactives : le processus d'extraction de connaissances doit être déclenché en ligne !

IACA [Cordier, 2008]

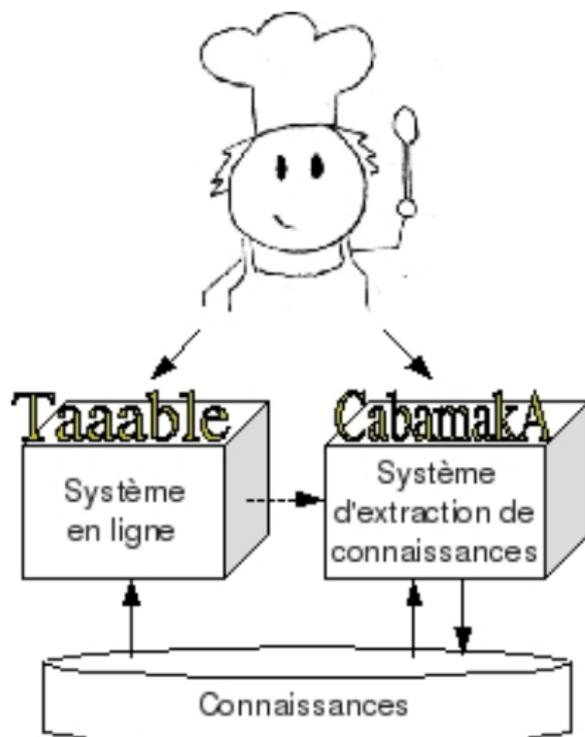
Interactive Adaptation Knowledge Acquisition

Acquisition de connaissances d'adaptation :

- **en ligne**
- **interactive** : les connaissances sont apprises à travers l'interaction avec l'utilisateur
- **opportuniste** : les connaissances sont apprises suite à un échec d'adaptation, et dans le but de le réparer

Découverte opportuniste de connaissances d'adaptation

Principe



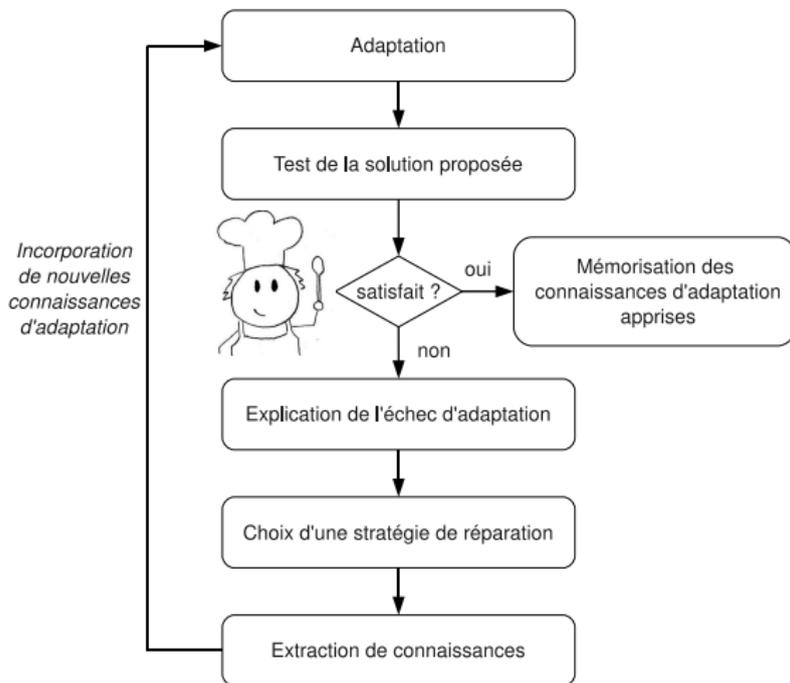
Rôle du processus d'extraction de connaissances

- CABAMAKA **fournit une assistance** à l'utilisateur dans la formulation de nouvelles connaissances,
- CABAMAKA **fait de la base de cas une source additionnelle** de connaissances d'adaptation pour le système.

Questions clés

- Quand déclencher le processus d'extraction de connaissances ?
- Comment paramétrer le processus d'extraction de connaissances par un contexte donné de résolution de problèmes ?

Quand déclencher l'extraction de connaissances ?



CABAMAKA recherche dans la base de cas les informations nécessaires à l'instanciation d'une stratégie de réparation.

Comment le paramétrer ?

stratégie de réparation + utilisateur \longrightarrow variation $\Delta_{\mathcal{E}}$

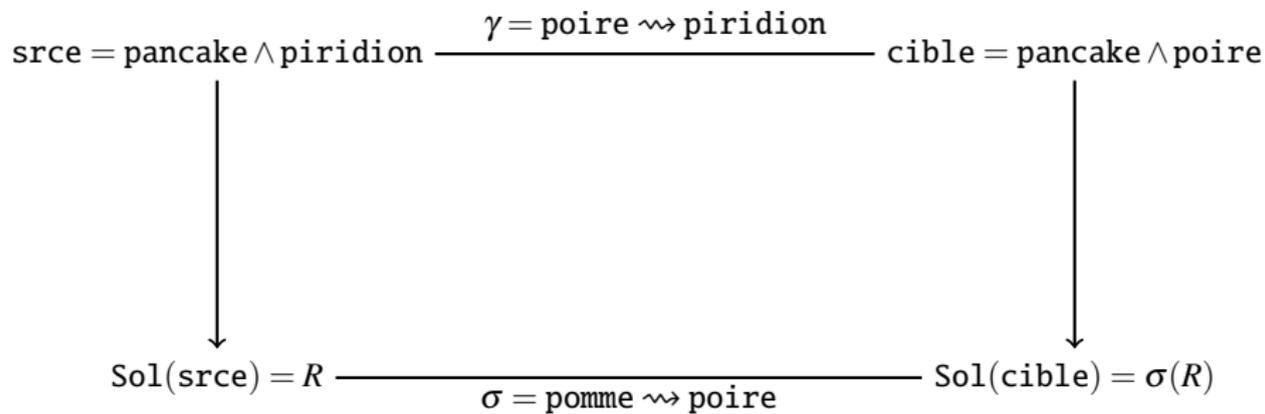
Le paramétrage se fait :

- par un filtre sur l'ensemble d'apprentissage : $\Delta_{\mathcal{E}} \subseteq \Delta_{kl}$
- en réglant le support minimum,
- par deux filtres sur les résultats de la fouille :
 - ▶ Δ satisfait les contraintes données par $\Delta_{\mathcal{E}}$: $\Delta_{\mathcal{E}} \subseteq \Delta$
 - ▶ Δ correspond à une substitution σ' applicable pour modifier la recette remémorée $\text{Sol}(\text{srce})$

Application à Taaable

Application à Taaable

Adaptation proposée



avec $R = \text{pancake} \wedge \text{pomme} \wedge \text{piridion} \wedge \text{fruit} \wedge \text{cannelle} \wedge \dots$
représente la recette *Apple pancakes from the townships*.

Application à Taaable

Explication de l'échec

La **recherche d'une explication** s'effectue en trois étapes :

- ① sélection d'une substitution $A \rightsquigarrow B$ qui pose problème

$\sigma = \text{pomme} \rightsquigarrow \text{poire}, A = \text{pomme}, B = \text{poire}$

- ② choix d'un "patron d'explication" (abstrait) :

"Un ingrédient x de B est incompatible avec un ingrédient y de la recette obtenue."

- ③ instanciation du patron (identification des ingrédients)

$x = \text{poire}, y = \text{cannelle}$

Application à Taaable

Choix d'une stratégie de réparation

Pour une substitution $A \rightsquigarrow B$

Explications possibles	Stratégies de réparation
Un ingrédient x de B nécessite un autre ingrédient y qui est absent de la recette obtenue.	<ul style="list-style-type: none">● si cible $\neq_{\emptyset} x$, supprimer x● si cible $\neq_{\emptyset} x$, trouver un substitut à x● si cible $\neq_{\emptyset} \neg y$, ajouter y● si cible $\neq_{\emptyset} \neg y$, ajouter un substitut à y
Un ingrédient x de la recette obtenue nécessite un ingrédient y de A qui vient d'être supprimé.	
Un ingrédient x de B est incompatible avec un ingrédient y de la recette obtenue.	<ul style="list-style-type: none">● si cible $\neq_{\emptyset} x$, supprimer x● si cible $\neq_{\emptyset} x$, trouver un substitut à x● si cible $\neq_{\emptyset} y$, supprimer y● si cible $\neq_{\emptyset} y$, trouver un substitut à y

Application à Taaable

Choix d'une stratégie de réparation

- On cherche une substitution de la forme :

$$\sigma' = \text{pomme} \wedge \text{cannelle} \rightsquigarrow \text{poire} \wedge \textit{quelque chose}$$

- donc on forme la variation :

$$\Delta_{\mathcal{E}} = \{\text{pomme}^-, \text{cannelle}^-, \text{poire}^+\}$$

Application à Taaable

Paramétrage de CABAMAKA

- On compare les recettes contenant des pommes et de la cannelle avec les recettes contenant des poires :

$$\mathcal{E} \longleftarrow \Delta_{kl} \text{ tels que } \Delta_{\mathcal{E}} \subseteq \Delta_{kl}$$

- On ne garde que les variations extraites Δ qui satisfont $\Delta_{\mathcal{E}}$:

$$\Delta_{\mathcal{E}} \subseteq \Delta$$

et qui correspondent à des substitutions σ' applicables pour modifier la recette remémorée $\text{Sol}(\text{srce})$:

$$\sigma' = A' \rightsquigarrow B' \text{ telle que } \begin{cases} \text{Sol}(\text{srce}) \vDash_{\theta} A' \\ \text{Sol}(\text{srce}) \not\vDash_{\theta} a \text{ si } A' \not\vDash_{\theta} a \text{ et } B' \vDash_{\theta} a \end{cases}$$

Conclusion et perspectives

Résumé de l'approche :

- méthode **semi-automatique** mais déclenchée **en ligne**, de manière **interactive** et **opportuniste**
- CABAMAKA génère « à la volée » des connaissances à partir de la base de cas
- la méthode a été appliquée à Taaable et la faisabilité de l'approche a été démontrée sur des exemples

Perspectives :

- validation quantitative !
- tests avec une base de recettes plus volumineuse
- gestion des connaissances apprises (coûts, stockage, conflits)
- interface graphique / ergonomie