Conclusion

Retours : PàP et classification évolutive

Faire évoluer l'implémentation pour rendre compte "linguistiquement" des évolutions de représentations mises en oeuvre

Le dispositif construit a permis la mise en place d'une série d'outils qui ont accompagné les évolutions de l'implémentation en cours de Self. La dernière version de Self permet ainsi un double regard sur les activités mises en place. La trace produite dans une session de travail fournit une présentation plus "linguistique" des savoirs associés aux résultats présentés. L'interface graphique de Self fournit quant à elle une présentation plus "technique" des objets construits. Le lien entre l'objet construit et le savoir qu'il représente n'est pas évident. Un objet de cette fenêtre graphique ne dit pas clairement ce qu'il représente (un mot ou un arbre par exemple). Il faut apprendre à maîtriser cet environnement de travail pour pouvoir identifier rapidement les objets. Cet apprentissage de Self peut d'ailleurs se révéler délicat. L'implémentation présentée ici porte d'ailleurs encore les traces de nos tâtonnements initiaux.

Nous souhaitons pouvoir réduire cette distance entre les savoirs représentés et les objets construits pour porter ces savoirs en utilisant le potentiel graphique que permet Self. Ce dernier est principalement utilisé dans le développement d'interfaces graphiques : la figure a-2-8.1 est une illustration de ce que peuvent produire les processus graphiques de Self. Ces processus n'ont pas été utilisés dans l'état actuel de `Gaspar`.

Prenons un exemple :

La multiplication des objets dans le monde Self (<- chapitre 8) est un problème crucial si l'on veut pouvoir évaluer l'état global du système représenté. En particulier dans notre cadre de classification évolutive où l'on peut produire un myriade de prototypes de mot (ou d'arbre) et de pôles de comportements partagés par des sous-ensembles de prototypes de mot (ou d'arbre). L'interface graphique permet ainsi d'atténuer cette impression de dispersion complète des savoirs représentés. L'objet zoom utilisé dans la figure a-2-8.1 permet de scruter une zone particulière du monde Self (comprenant un nombre limité de prototypes) tout en maintenant une présentation globale du monde Self. On peut ainsi avoir simultanément une vision locale de certains prototypes et la position de cette zone particulière dans l'état global du système.

Nous souhaitons aussi associer aux représentations d'unités de langue une représentation graphique qui rende plus transparente l'implémentation sous-jacente. La difficulté étant bien

évidemment de maintenir pour les objets construits une potentielle mise à jour dynamique des savoirs représentés. L'utilisateur doit pouvoir identifier plus facilement les savoirs représentés, il doit aussi pouvoir ajuster ces savoirs de manière "conviviale"; quitte à construire des outils qui guident la mise à jour éventuelle d'une représentation graphique des unités de langue ainsi représentées. La technicité graphique ne doit donc pas prendre le pas sur nos objectifs initiaux.

La PàP, un outil expérimental pour représenter la mouvance

Notre travail a consisté tout d'abord à une représentation progressive de catégorisation de formes complexes de type `N1 à N2` qui regroupent ces formes suivant la sémantique qu'elles développent. Cette phase de catégorisation s'appuie sur un processus de classification qui est suivie d'une phase de restructuration de la hiérarchie initiale adoptée. Cette évaluation de connaissances définies préalablement comme prototypiques a permis de produire une classification hiérarchisée, même si celle-ci reste évolutive. Pour déterminer les différents niveaux pour le classement de ces formes complexes, il est nécessaire de procéder à autant d'ajustements dans la hiérarchie à construire que les résultats fournis l'imposent. Il est donc nécessaire de pouvoir disposer de plusieurs niveaux de contrôle dans les processus d'analyse et d'évaluation des résultats produits qui permettent de résoudre les problèmes de classifications évolutives.

Dans un deuxième temps nous avons présenté une approche qui vise à établir une représentation dynamique de savoirs lexicaux suivie d'une classification progressive des savoirs décrits. Là encore, la démarche proposée permet d'envisager une classification hiérarchisée, même si celle-ci reste évolutive. Il faut pour cela être capable de gérer la coexistence de connaissances parcellaires, morcellées et établir une classification progressive des acquis produits. La question du rapport entre différents types de savoir et le problème du contrôle ressurgit comme cruciale.

- L'évaluation de connaissances représentées préalablement comme prototypiques permet d'envisager une classification hiérarchisée des savoirs représentés, même si celle-ci reste évolutive. Les processus de classification proposés sont capables de gérer la coexistence de connaissances parcellaires et morcellées et d'établir une classification progressive des acquis produits (même si ceux-ci restent à enrichir). La PàP se révèle d'ailleurs être un outil cohérent par rapport à l'approche symbolique suivie.

- La PàP permet de construire la représentation du domaine visé puis de classer les savoirs représentés de manière progressive : les processus de représentation puis de classement permettent une approche incrémentale. La représentation du domaine étudié pouvant évoluer, sa représentation s'améliore en fonction des résultats obtenus. Il convient d'affiner cette représentation jusqu'à ce qu'une certaine forme de stabilité soit atteinte.

- Les limites de l'automatisation marquent le champ de travail qu'il reste à effectuer manuellement pour mener à bien le classement des unités lexicales. Il semble clair que le travail d'affinement des représentations obtenues va éclairer et guider la mise au point des éléments de savoirs pertinents que l'on souhaite mettre à jour pour aller vers une automatisation de plus en plus fine du classement escompté.

- La démarche suivie vise à définir des outils de classement et des processus de représentation qui permettent la prise en compte de l'évolution potentielle du domaine représenté. Elle s'accommode assez bien avec le modèle de représentation des connaissances retenu. La PàP encourage une approche de représentation faite de petits sauts successifs qui améliore la qualité de la représentation produite; d'autre part, cette démarche s'accorde aussi avec la nécessaire approche artisanale que constitue le travail du linguiste dans sa volonté de décrire les comportements des faits de langue (Milner 1989).

- Le système construit s'inscrit dans cette approche qui vise à établir un dialogue entre le dispositif informatique construit et les problèmes posés par les faits linguistiques étudiés. L'intervention manuelle du linguiste s'inscrit parfaitement dans une telle approche. La flexibilité de la PàP constitue d'ailleurs une solution satisfaisante pour développer des processus de représentation capables de rendre compte des aspects mouvants du domaine décrit.

Le travail présenté supra sur l'analyse automatique prototypique a mis en évidence la difficulté à cerner les modes de représentations à adopter pour modéliser un domaine de connaissances dans lequel le savoir n'est pas encore figé. Concrètement ce travail a permis de produire un premier dispositif capable de rendre compte des aspects évolutifs des savoirs représentés. Il a montré que le modèle de représentation choisi permet de traiter certains mécanismes évolutifs rencontrés dans les faits de langue étudiés. De nombreux prolongements restent à faire.

- Sur le plan technique, la couverture des gros corpus doit être réalisée; on pourrait ainsi, sur le plan linguistique, évaluer et analyser les résultats construits. Ces résultats doivent nous permettre d'améliorer la qualité du travail de représentation initié par les processus construits. La confrontation de ces résultats avec les connaissances du linguiste ou avec des bases de connaissances extérieures est prometteuse.

- Les aspects conceptuels n'ont pas été articulés aux phénomènes traités dans les processus construits; une étude théorique et technique de ces articulations avec les autres niveaux de savoirs représentés doit conduire à étendre et affiner le travail de représentation amorcé. Le traitement de connaissances extra-linguistiques peut aussi complèter et affiner les processus mis en place.

Ce travail a aussi permis de tracer des pistes pour une réflexion sur les problèmes complexes liés aux traitements automatiques des faits (évolutifs) de langue. Notre approche des problèmes de classification, d'apprentissage, de refléxivité doit être étendue. Ces notions dépassent très largement le cadre linguistique étudié ici et la confrontation de nos choix et de nos résultats avec les travaux établis par ailleurs doit constituer une voie de recherche fructueuse.