Raisonnement sur les normes

Farid Nouioua
LIPN UMR 7030 du C.N.R.S.
Institut Galilée – Univ. Paris-Nord
farid.nouioua@lipn.univ-paris13.fr

Dans de nombreuses situations, l'être humain peut imaginer ce qui se passera normalement à partir de l'état actuel du monde. Cette capacité inférentielle remarquable fondée sur la notion de norme permet d'aller beaucoup plus loin que les théories classiques basées exclusivement sur la notion de vérité, pour résoudre le problème d'une compréhension automatique de la langue.

Dans cette optique, l'extraction des normes d'un domaine donné est une étape primordiale pour le développement d'outils logiciels capables de simuler la compréhension de textes écrits en langage naturel. La question qui se pose alors est d'effectuer cette extraction à partir d'un corpus textuel. Un texte ne s'intéresse pas souvent à décrire le déroulement habituel et normal des événements mais se focalise plutôt sur la description des déviations par rapport aux normes. A partir de là, nous découvrons les normes en cherchant à détecter dans chacun des textes de notre corpus (constats d'accidents) les violations de normes auxquelles le texte fait appel. A cette fin, nous analysons la réponse à la question Q : « Quelle est la cause de l'accident ?» en privilégiant, parmi les nombreuses réponses possibles, les comportements volontaires qui expriment des violations de normes.

L'approche utilisée consiste à déterminer manuellement l'anomalie pour chaque texte, en se basant sur l'analyse des réponses d'un lecteur ordinaire à la question Q, puis à concevoir un système de raisonnement qui reproduit les mêmes réponses

Le système de raisonnement, basée sur la logique des défauts, est organisé en couches. En effet il est très difficile de gérer d'un seul coup, la quantité considérable de connaissances évoquées dans les textes. On répartit alors les prédicats et les règles du système sur plusieurs niveaux. Chaque niveau s'intéresse à un ensemble de phénomènes et produit des résultats directement exploités par les couches inférieures. La couche la plus externe est le résultat d'un analyseur de langue, et la couche la plus interne appelée noyau, contient un nombre limité de prédicats capables d'exprimer toute vraie anomalie. L'approche en couches permet de structurer le système de raisonnement ce qui facilite sa conception.

Pour raisonner afin de détecter les violations des normes, nous avons besoin de les représenter. Nous avons proposé un modèle de représentation permettant de distinguer clairement entre vraie et fausse anomalie d'une part, et d'autre part de tenir compte de l'aspect temporel très pertinent dans notre cas. Nous avons pour cela proposé trois types de modalité qui expriment : 1) devoir faire quelque chose 2) Etre capable de faire quelque chose 3) le déroulement scriptal ou les futurs attendues. Pour exprimer le déroulement temporel des événements de l'accident, nous avons proposé également un modèle de temps basé sur une sémantique d'intervalles et doté de concepts appropriés (les événements, les états, la persistance…). Nous avons ensuite montré comment implémenter les modalités dans le cadre du modèle temporel proposé et en se basant sur une logique de premier ordre réifiée.

Les règles conçues nous ont permis de trouver des résultats prometteurs sur les textes de notre corpus.