

理工学専攻修士論文要旨

提出年度：2008年度
提出日：2009年1月27日
専修コース：知能情報コース
学生番号：35607061
学生氏名：川口 純一
研究指導教員：原田 実 教授

(論文題目)

意味グラフベースの言い換えシステム Priest の開発とその有効性の検証

(内容の要旨)

ある言語表現を意味の等しい別表現に置換することを「言い換え」という。例えば、「飛行機を発明したのは誰ですか」を「誰が飛行機を発明しましたか」と変換すると、お互いの文章は言い換え関係にあり、それぞれを合わせて言い換え表現対と呼ぶ。言い換えは、人間が日常使用する言語で頻繁に行われる行為である。原田研究室では、文の言い換えを文の意味構造を表す「意味グラフ」上で行う言い換えエンジン Serpent を 2005 年に開発した。しかし、Serpent による言い換えでは正しく言い換えがされていない不完全な部分が存在している。そこで、言い換え知識の精密化と言い換える精度の更なる向上を目的として言い換えシステム Priest の開発研究を行った。更に、その有用性の検証も同時に行った。

Priest はルールベースのシステムであり、必要に応じて言い換え知識を追加し、それを基に言い換えを行う。言い換え知識は、言い換える対象となる文の意味グラフに対して言い換えできるかを判定する条件となる Head と、Head 照合が成功したときに意味グラフを言い換える情報を記述した Body からなる。Head の中には、言い換え知識を高速に検索するためにインデックスキーの役割となる Key を含めた。Head と Body には、形態素の細部に至るまでの言い換え規則を追加した。

Priest はこの新しい言い換え知識に基づき、意味グラフを言い換えるの行いやすい「言い換え元グラフ」へ変換した後、「言い換え知識の検索」「競合解決」「言い換え知識と言い換え元グラフの照合」「言い換えるの実行」の順に変換処理を行う。言い換え知識の検索では、言い換え知識に記述されたキーと言い換え元グラフに対して生成されたキーを照合し、適した言い換え知識を高速に見つける。競合解決では、検索によって見つかった複数の知識から言い換え後のノード数が最小となるものを選び出す。言い換え知識と言い換え元グラフの照合では、言い換え元グラフの一部が競合解決で選択した言い換え知識の Head を満足するかどうか判定する。言い換えるの実行では、Head と照合した部分グラフを Body で表される部分グラフで置換する。

Priest の評価実験として、奈良先端大の言い換え表現対{(a, b)}232 組を用いて、文 a と文 b 間のグラフ類似度を測定した。文 a を Priest を用いて文 a' に言い換えた結果、言い換え前の(a, b)間の平均類似度は 63.05%、言い換え後の(a', b)間の平均類似度は 84.62%となり、Serpent の同様の結果 75.91%と比較して精度の向上を実現した。また、1 グラフあたりの処理時間は平均 2.23ms となり、Serpent の平均 10.05 ms と比べて向上が見られた。