

精度 100%の言語処理へ向けて (?)

松崎拓也 (国立情報学研究所)

takuya-matsuzaki@nii.ac.jp

現在の言語処理研究では、テキストに対して人間が付加したアノテーションをどれだけ機械で再現できるか、という尺度で研究の位置づけを行うことが通例となっている。そのような尺度は具体的には色々なものが考えられるが、ここではひとくくりに「精度」と呼んでおく。

さて、「精度」にはふつう上限値があり、であれば 100%の精度を達成するシステムが我々の研究のひとつのゴールである、ということになりそう。100%の精度を目指すことにはどのような意義があるのか、100%の精度が真の目的ではないとしたら、精度による研究の位置づけが近似している「真の評価尺度」とは一体なんなのか、100%の精度を達成するには何が必要で、どのくらい時間がかかるのか、それを一体どうやったら見積もることが出来るのか、といったことを議論したい。