

Koreference z mezijazykové perspektivy

Michal Novák

Tématem této práce je studium vlastností koreference s použitím mezijazykových přístupů. Motivací práce je výzkum lingvistické typologie založené na koreferenci. Další motivací je prozkoumání, jestli rozdíly ve způsobech, jak jazyky vyjadřují koreferenci, mohou být využity k natrénování lepších modelů pro rozpoznávání koreference. Navrhujeme dvě mezijazykové metody: rozpoznávání koreference s informací z druhého jazyka a projekci koreference. Výsledky našich experimentů s těmito metodami na česko-anglických datech naznačují, že s ohledem na koreferenci přináší angličtina více informací do češtiny než naopak. Rozpoznávání koreference s informací z druhého jazyka navíc dokázalo při aplikaci na paralelních datech překonat na obou jazycích výsledky jednojazykového systému na rozpoznávání. Při experimentech používáme jednojazykový rozpoznávač koreference a vylepšenou metodu na zarovnání koreferenčních výrazů, které jsme rovněž navrhli v rámci této práce.