

**Název práce:** Rozpoznávání koreference pro Universal Dependencies

**Autor:** Ján Faryad

**Ústav:** Ústav formální a aplikované lingvistiky

**Vedoucí bakalářské práce:** Mgr. Michal Novák

**Abstrakt:**

Koreference je důležitým prostředkem pro udržení soudržnosti textu. V univerzálním popisu morfologie a závislostní syntaxe Universal Dependencies (UD) však dosud chyběla možnost jejího vyznačení. Práce představuje způsob, jak koreferenci v projektu UD značit. Součástí práce je i převod dat s anotací koreference z korpusů PDT 3.0 a OntoNotes 5.0 za pomoci nástroje UDPipe pro automatickou analýzu textu ve stylu UD.

Práce se dále zabývá implementací systému pro automatické rozpoznávání koreference zájmen pomocí strojového učení. Nakonec je jednoduchým způsobem vyhodnocena úspěšnost systému. Při návrhu programu je kladen důraz na jazykovou nezávislost a kompatibilitu s rozhraním Udapi používaným pro práci s UD.

**Klíčová slova:** rozpoznávání koreference, koreference, anafora, Universal Dependencies, UD