

S rostoucí důležitostí počítačového zpracování přirozeného jazyka narůstá i množství výzkumů na téma automatické analýzy anafory. Příspěvkem k výzkumu této problematiky je rovněž naše diplomová práce, jejímž cílem je vytvořit soubor pravidel pro analýzu anafory v českém jazyce. Vytvořený soubor pravidel obsahuje jak ručně psaná pravidla, tak i pravidla vznikající pomocí systému strojového učení C4.5. K trénování a testování pravidel byla použita anotovaná data z Pražského závislostního korpusu, ve kterém je zachycena zájmenná anafora, kontrola, reciprocita a závislostní vztah doplňků. Právě těmto druhům anafory je věnována naše práce. Vyhodnocení pravidel je provedeno standardními metodami pro hodnocení úplnosti a přesnosti.