

Posudek diplomové práce

Michal Vykypěl: Testovací platforma pro Webové roboty

Cílem práce bylo navrhnout a realizovat prostředí, ve kterém je možné testovat webové roboty. To by mělo umožnit testování bez nutnosti pracovat přímo na reálném webovém prostoru, čímž by došlo i k realizaci tolik potřebného prostředí pro testování různých stahovacích strategií.

Problémy jsou jednak:

- o v rovině technické, například: tvorba prostředí, zpracování velkého množství dat, efektivní práce s potencionálně nekonečnou grafovou strukturou, atp.
- o v rovině teoretické, například: tvorba syntetických simulačních struktur z reálných vzorků částí webu.

Autor navrhl externí datovou strukturu pro uložení grafu, kterou podpořil vlastním alokátozem pro jazyk C++. Tím zajistil vhodné kešování zpracovávaných dat.

Samotné procházení, tedy simulaci práce robota, podpořil příslušným skriptovacím jazykem na bázi jazyka Pascal. Tento prostředek je dostatečně jednoduchý, nabízí několik kontejnerů, podporu pro Page rank, a například i pole front. Poslední jmenovaná struktura umožňuje implementaci dlouhodobého plánování v plánovači robota.

Všechny jmenované prostředky doprovází krátké ukázky použití. Mnohé z nich ale mohly být delší. Stejně tak bylo možné vyjít z nějaké konkrétní implementace robota, a představit jeho „simulační“ protějšek v navrhovaném jazyce.

Transparentní simulátor pro software webového robota konstruoval autor tak, že nejprve našel 1-1 zobrazení z množiny URL na čísla uzlů grafu. Žádosti DNS jsou přesměrováním odbavovány zjednodušeným DNS serverem. Podobně jsou přesměrovány i http požadavky, avšak v jejich případě NATD nenávratně přepíše identifikaci cílového webového serveru (jeho IP adresu). Proto je nutné, aby robot server identifikoval doménové jméno serveru. Bohužel, „jak konkrétně“ už autor neuvádí. Stejně tak práce postrádá návod, jak daný simulační software použít, jak vyrobit syntetické grafové struktury atp.

Protože práce splnila alespoň jeden ze základních cílů, tak ji doporučuji ještě uznat jako práci diplomovou.

V Praze dne 2.9.2008