

Similarity of XML Data

Předkládaná diplomová práce se zabývá problematikou určování podobnosti XML dat. Vzhledem k tomu, že podobnost je možné analyzovat na různých úrovních, autor práce zvolil problematiku podobnosti na úrovni XML schémat. Zvoleným jazykem pro popis schématu je jazyk DTD, pro určování podobnosti je využito a struktura DTD přizpůsobeno několik ověřených přístupů jako je editační vzdálenost, sémantická podobnost apod. Navržená metoda je experimentálně ověřena na reálných i umělých datech.

Práce je rozdělena do tří částí. V první části (kapitolách 1 a 2) jsou stručně popsány cíle práce a především technologie a teoretické základy, na nichž je založena. Ve druhé části (kapitole 3) je rozebrána problematika podobnosti XML dat z různých úhlů pohledu a nastíněny výhody a nevýhody existujících řešení. Třetí část (kapitoly 4, 5 a 6) zahrnuje vlastní řešení založené na těchto výhodách a nevýhodách – návrh algoritmu, popis prototypové implementace a experimentální výsledky. Kapitola 7 práci shrnuje a uzavírá.

Hlavními klady práce jsou především:

- Kvalitní analýza existujících řešení zadané problematiky.
- Z hlediska stávajících řešení zcela originální využití kombinace editační vzdálenosti a sémantiky názvů elementů/atributů pro určování podobnosti DTD.
- Přizpůsobení využitých algoritmů specifickým vlastnostem jazyka DTD jako jsou složité modely obsahu, operátory nebo omezení počtu výskytů.
- Experimentální analýza navrženého řešení nad umělými i reálnými daty demonstrující jeho vlastnosti.

Práce má kvalitní úpravu a je napsána srozumitelně. V práci je pouze minimální množství překlepů a gramatických chyb. Drobnými nedostatky jsou občasné ne-anglické větné konstrukce a místy přílišná strohost textu. Přestože text práce obsahuje veškeré náležitosti, popis je místy skutečně minimalistický.

I přes výše uvedené drobné připomínky splňuje práce původní zadání i cíl. Student dostatečně pronikl do zadané problematiky a prokázal schopnost zanalyzovat a řešit zadaný problém určování podobnosti XML dat, a to na dobré úrovni teoretické i experimentální. Výsledky práce by jistě stály za publikování ve vhodné odborné publikaci.

Práce Aleše Wojnara podle mého názoru bezpochyby splňuje všechny podmínky na diplomovou práci kladené, a proto ji **doporučuji** jako práci diplomovou **uznat**.

V Praze, 21.4.2008

