

Posudek oponenta diplomové práce

Autor/autorka práce: **Ondřej Kupilík**

Název práce: **Kontrola formálních pravidel u dokumentů vybraných textových procesorů**

Obsah práce

Práce má přímočarou strukturu v podstatě kopírující body zadání. Text obsahuje všechny informace potřebné k pochopení vytvářené aplikace, bohužel ale nic více. V teoretické části chybí jakýkoliv přehled o tom, zda podobnou úlohu již někdo někde řešil a jak. Nutno ale poznamenat, že explicitně to v zadání není požadováno. Popis je pochopitelný.

Kvalita řešení a dosažených výsledků

Kvalita řešení mi nepřipadá úplně dobrá, respektive nedokáži si představit, že by tento SW někdo prakticky používal. Zejména ne vyučující z humanitně zaměřených fakult, kteří měli být pravděpodobně cílovou skupinou uživatelů. Myslím si totiž, že je jednodušší otevřít odevzdaný dokument v příslušném textovém procesoru a základní informace (počet slov, počet stran, existenci literatury, existenci stránkování apod.) v něm zjistit několika málo úkony. Je to snazší, než se probírat komplikovaným výsledkovým XML souborem. Práci totiž zcela chybí analýza praktické použitelnosti. Po prvním vyzkoušení by bylo možné vyjmenovat celou řadu dílčích vylepšení zvyšujících použitelnost, ale omezím se jen na jedno – pokud kontrolovaný dokument splňuje všechna stanovená kritéria, mělo by to být z výsledku zřejmé na první pohled, a nikoli až detailní analýzou výstupního XML souboru. Uyučující by se pak mohl soustředit jen na ty dokumenty, které **nesplňují** požadovaná kritéria. V současném stavu musí kontrolovat stejně všechny práce, jen místo „DOC“ musí prohlížet XML soubory. Stačilo jako další krok použít XSLT transformaci. Konkrétní výhrady: `DOCXManager.java` obsahuje metody `loadLiterature()` s 8 zanořeními z toho 5 cyklů, `controlPageNumbers()` pak 10 zanoření. Existence magických čísel (`spacing/240`) nebo řetězců (`pageFormat.equals("PAGE * MERGEFORMAT")`) s bílými znaky! Poskytnutý `rules.xml` nejde externě validovat (viz dále otázku č. 1). Soubor `rules.xml`, který bude uživatel měnit, se nenachází v adresáři `jar` (je zabalen do výsledného `.jar`, kde jej běžný uživatel nemá šanci nalézt) a to v podobě, která se liší od str. 73 (bez komentářů). Nevalidní `rules.xml` způsobí vypsání celé výjimky, což je krajně nepřehledné. Program nenalezne datum bez nevýznamové nuly (např. 29.6.2015 versus 29.06.2015). Výstup informace o použitých fontech a o řádkování, kde jsou vypsány údaje o **všech odstavcích dokumentu**, je pro člověka nepřehledný. Na druhou stranu je ale třeba konstatovat, že program je dobře strukturovaný, čitelný a dostatečně komentovaný. Na CD se nalézá 48 Java souborů o velikosti 140 KB.

Formální úroveň

Práce je pečlivě vysazena a je zřejmé, že před odevzdáním prošla pečlivou kontrolou, protože neobsahuje téměř žádné překlepy, typografické prohřešky a jiné podobné nepravosti. Lze namítnout, že v seznamu zkratk zřejmě DOM nepředstavuje „Ústav pro informace ve vzdělávání“.

Práce s literaturou

Diplomant udává celkem 20 zdrojů, které jsou všechny v práci použity. Uvedené zdroje jsou odpovídající pro řešení práce a jsou relevantní.

Splnění zadání

Diplomant zadání splnil, bohužel však do něj nedal téměř žádnou vlastní invenci – viz výše. Lehkou výhradu mám k bodu č. 4 – viz otázka č. 2.

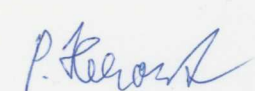
Dotazy k práci

1. Na CD se nachází soubor `rules.xsd`, ovšem všechny dostupné soubory `rules.xml` odkazují na jinou verzi XSD souboru – `rules3.xsd`. Validace podle XSD však probíhá. Jak je to možné?
2. Jak by se aplikace spouštěla pro „sadu dokumentů“ (bod 4. zadání), tj. např. pro 100 různých semestrálních prací odevzdaných studenty v jednom předmětu?

Navrhuji hodnocení známkou **dobře** a práci doporučuji k obhajobě.

V Plzni 29.6.2015

Západočeská univerzita v Plzni
Fakulta aplikovaných věd
katedra informatiky a výpočetní techniky


doc. Ing. Pavel Herout, Ph.D.