

# Záv re ná zpráva - Robert Ši ška

P i d lení úkol pro projekt *MathMLCan: Kanonizace MathML* na m p ípadlo na ítání konfigurace, práce na jednom z modul - konkrétn `ElementMinimizer`, interface pro p íkazovou ádku a grafické uživatelské prost edí.

## Na ítání konfigurace

Konfigurace využívá dva mechanismy. Nejd íve je na ten properties soubor, který obsahuje seznam modul a další globální parametry jako cesty k DTD a Schema soubor m. Ze seznamu modul v property *modules* jsou moduly v daném po adí na ítány, každý modul si p i inicializaci nahraje sv j vlastní properties soubor.

T ída *MathMLCanonicalizer* m ž e být vytvo ena s argumentem typu *InputStream*, která obsahuje XML konfiguraci. V tomto p ípad je nejd ív tato konfigurace validována proti XML Schema definici pomocí validátoru z balíku *javax.xml.validation* a poté jsou z ní pomocí *javax.xml.stream* API na ítány globální properties a konfigurace modul .

Stejn jako v prvním p ípad , parametry jsou nejd íve na teny z properties souboru. XML konfigurace umož ůje properties pouze p episovat, nikoliv vytvá et nové. Krom elementu *property* m ž e obsahovat ko enový element *config* také element *module*, který dále m ž e obsahovat *property* elementy. Po adí *module* element rozhoduje o po adí aplikace modul na vstupní proud s MathML kódem.

## Práce na modulu *ElementMinimizer*

*ElementMinimizer* je modul používající *javax.xml.stream* API. Jeho ú el je odstran ní element a atribut , které jsou nepodstatné pro význam matematického výrazu. Jsou to hlavn elementy z Presentation layer MathML specifikace popisující p evážn vzhled výrazu. Tento modul také odstra ůje XML komentá e.

Na vstupním proudu modulu je o ekáván MathML kód, který je ten od ko ene do hloubky a každý element je odstran n v etn jeho potomk , pokud je uveden v property *remove\_all*, nebo je odstran n pouze element samotný jeho d ti jsou posunuty na jeho místo, pokud je uveden v property *remove*.

Protože jen malá ást atribut v MathML má vliv na sémantiku výrazu, byl na atributy použit opa ný postup. V konfiguraci modulu specifikujeme, které atributy mají být ponechány, a všechny ostatní jsou smazány. Property *keepAttributes* obsahuje atributy odd lené mezerou, které mají být ponechány ve všech elementech. Dále je možné vytvo it whitelist pro kterýkoliv konkrétní element pomocí property *keepAttributes.jmenoElementu*.

Atributy je možno uvést dv ma zp soby. Bu pouze jménem atributu, nebo výrazem *jmenoAtributu=hodnota*, který má za následek neodstran ní atributu pouze za spln ní podmínky.

P i sestavování výchozí konfigurace modulu jsem vycházel z doporu ení W3C MathML Version 2.0 (Druhá edice) kapitoly 3 - Presentation markup and kapitoly 4.3 - Content Element Attributes.

## GUI a CLI

Pro definici grafického uživatelského rozhraní jsem se rozhodl použít definici layoutu pomocí XML deskriptoru. Tuto metodu umož ůje hned n kolik projekt od malých nadstavb *Swing* knihovny až po velmi rozsáhlé projekty jako je *XUL*. Protože GUI slouží v našem projektu jen pro jednoduchou demonstraci, rozhodl jsem se použít jednoduchou knihovnu *SwiXml*, která generuje Swing objekty z XML souboru.

Pro implementaci interface z p íkazové ádky jsem použil POSIX-like parser z knihovny Apache Commons CLI. Všechny soubory dané jako parametry jsou p e teny a kanonizovány. Lze specifikovat XML konfiguraci pomocí parametru `-c``. Pokud není specifikován žádný soubor, je spušt no GUI.