

Gedit Syntax Highlighter

Burlacu Emil Claudiu, grupa E1

Universitatea "Alexandru Ioan-Cuza" Iași, Facultatea de Informatica

1 Introducere

Pentru proiectul final la materia **Principii ale limbajelor de programare** am ales dezvoltarea unui Syntax Highlighter pentru editorul text Gedit.

2 Tehnologii utilizate

2.1 XML

Extensible Markup Language (XML) este un limbaj de marcare care definește un set de reguli pentru codificarea documentelor ce poate fi citit atât de programator, cât și de mașina folosind expresii regulate.

Pentru a crea un limbaj pentru editorul text Gedit, regulile acestuia trebuie definite folosind limbajul XML.

3 Detalii de implementare

Codul XML va fi scris într-un fișier cu extensia .lang și va fi pus în folder-ul language-specs din folder-ul de instalare al lui Gedit.

Structura

Documentul XML va avea o celulă principală <language>, în care va fi declarată id-ul și numele limbajului. În această celulă se vor afla alte 3 celule: <metadata>, <styles> și <definitions>.

<metadata>

Celula <metadata> este opțională și oferă o colecție de proprietăți care specifică informație arbitrară despre limbajul pe care vrem să îl definim. O astfel de informație este elementul *globs* care permite editorului să detecteze automat ce sintaxă trebuie folosită. În cazul acestui limbaj, aceasta va apărea astfel:
<property name="globs"*.k</property>

Alte proprietăți utilizate vor fi *line-comment-start*, *block-comment-start* și *block-comment-end*, ce au ca rol definirea simbolurilor pentru a introduce comentarii în sintaxa creată.

<styles>

În celula <styles> vor fi declarate stilurile folosite pentru marcarea sintaxei. Fiecare stil va fi declarat cu ajutorul elementului <style> astfel:

```
<style id="keyword"__name="Keyword"map-to="def:keyword"/>
```

Elementul *id* reprezintă un identificator pentru stilul respectiv, iar elementul *name* este numele stilului. *map-to* este opțional și are rolul de a atribui o formă standard de marcă definită în fișierul *def.lang*.

<definitions>

În această celulă vor fi definite stilurile declarate în celula <styles> prin rezervarea cuvintelor sau cu ajutorul expresiilor regulate.

Prima linie de cod din această celulă este:

```
<context id="k">
```

În cadrul acestei celule vom defini toate stilurile ce aparțin de limbajul K, adică toate stilurile declarate în celula <styles>. Una dintre definițiile acestei sintaxe este:

```
<context id="β_paranthesisβstyle-ref="keyword">
```

```
<start>[</start>
```

```
<end>]</end>
```

```
</context>
```

Această definiție va folosi stilul "keyword" din fișierul *def.lang* pentru a marca parantezele patrate ([]) și tot ce se află între ele.

4 Concluzii

Momentan, am implementat marcarea doar pentru o parte din limbajul K, și anume, comentariile, expresiile Bool, paranteze patrate și acolade, precum și marcarea unelor cuvinte rezervate. Scopul acestui proiect este de a avea un syntax highlighter complet pentru limbajul K în editorul text Gedit, ceea ce include: marcarea tuturor cuvintelor rezervate, ceea ce include operatorii, tipurile de date și funcții, spell-check(marcarea expresiilor care nu sunt scrise corect de către utilizator) și marcarea numelor introduse de utilizator(cu ajutorul expresiilor regulate).