

# Parsování dat XML

## Semestrální práce k předmětu XD36PJP

ČVUT FEL

obor STM - Softwarové inženýrství, kombinované studium

5. semestr

---

### Gramatika

**1. Terminální symboly** – kromě symbolu „chars“ je každý terminální symbol ve tvaru start a konec

<pobyty></pobyty>	<turnus></turnus>	<lokalita></lokalita>
<zacatek></zacatek>	<konec></konec>	<skupina></skupina>
<cisloSkupiny></cisloSkupiny>	<vedouci></vedouci>	<jmeno></jmeno>
<prijmeni></prijmeni>	<rokNarozeni></rokNarozeni>	<lokalita></lokalita>
<dite></dite>	<pohlavi></pohlavi>	chars

### 2. Neterminální symboly

Pobyty, Turnusy, Turnus, Lokalita, Zacatek, Konec, Datum, Skupina, Skupiny, cisloSkupiny, Vedouci, Osoba, Jmeno, Prijmeni, Rok Narozeni, Deti, Dite, Pohlavi

Jednotlivé **neterminály gramatiky** jsou ve vnitřní paměti reprezentovány objekty stejnojmenných tříd. Neterminální symboly jsou reprezentovány seznamem příslušných objektů.

### 3. Pravidla

Pobyty	→ <pobyty>Turnusy</pobyty>
Turnusy	→ Turnus Turnusy   ε
Turnus	→ <turnus>Lokalita Termin Konec Skupiny</turnus>
Lokalita	→ <lokalita>chars</lokalita>
Zacatek	→ <zacatek>Datum</zacatek>
Konec	→ <konec>Datum</konec>
Datum	→ chars
Skupiny	→ Skupina Skupiny   ε
Skupina	→ <skupina>CisloSkupiny Vedouci Deti</skupina>
CisloSkupiny	→ <cisloSkupiny>chars</cisloSkupiny>
Vedouci	→ <vedouci>Osoba</vedouci>
Osoba	→ Jmeno Prijmeni RokNarozeni
Jmeno	→ <jmeno>chars</jmeno>
Prijmeni	→ <prijmeni>chars</prijmeni>
RokNarozeni	→ <rokNarozeni>chars</rokNarozeni>
Deti	→ Dite Deti   ε
Dite	→ <dite>Osoba Pohlavi</dite>
Pohlavi	→ <pohlavi>chars</pohlavi>

Pro parsování příslušného XML souboru byla použita knihovna StAX, která poskytuje metodu next(), která vrátí jeden element XML souboru - ten je v gramatice uveden jako **terminální symbol**.

Rekurze je použita u těchto pravidel	Turnusy	→ Turnus Turnusy   ε
	Skupiny	→ Skupina Skupiny   ε
	Deti	→ Dite Deti   ε

```
public List<Turnus> parseTurnusy() {
    Turnus t;
    List<Turnus> ty;
    if (isCurrentStartElement(turnus)) {
        t = parseTurnus();
        ty = parseTurnusy();
        ty.add(0, t);
        return ty;
    } else {
        ArrayList<Turnus> turnusy = new ArrayList<Turnus>();
        return turnusy;
    }
}

public List<Skupina> parseSkupiny() {
    Skupina s;
    List<Skupina> sy;
    if (isCurrentStartElement(skupina)) {
        s = parseSkupina();
        sy = parseSkupiny();
        sy.add(0, s);
        return sy;
    } else {
        ArrayList<Skupina> skupiny = new ArrayList<Skupina>();
        return skupiny;
    }
}

public List<Dite> parseDeti() throws SyntaxeError {
    if (isCurrentStartElement(dite)) {
        Dite d = parseDite();
        List<Dite> di = parseDeti();
        di.add(0, d);
        return di;
    } else {
        ArrayList<Dite> deti = new ArrayList<Dite>();
        return deti;
    }
}
```

## Zdroje

[1] PŘEDNÁŠKOVÉ SLAJDY, *předmět XD36PJP* [online]. Poslední aktualizace neuvedeno [cit. 2011-10-04]. Dostupné z WWW: <<https://service.felk.cvut.cz/courses/XD36PJP/>>.

[2] MÜLLER Karel, Ing., *Programovací jazyky* [online]. Poslední aktualizace 2009 [cit. 2011-10-05]. Dostupné z WWW: <[https://service.felk.cvut.cz/courses/XD36PJP/Skripta\\_prednasky.pdf](https://service.felk.cvut.cz/courses/XD36PJP/Skripta_prednasky.pdf)>.

[3] JAVA STAX, Jakob Jenkov [online]. Poslední aktualizace neuvedeno [cit. 2011-10-05]. Dostupné z WWW: <<http://tutorials.jenkov.com/java-xml/stax.html>>.

[4] DANĚČEK Jiří, Ing., *konzultace v průběhu zpracování programu*.

**Zpracovala:** Radoslava Jandová

**V Praze dne 12. 12. 2011**

**Username:** jandora1

**e-mail:** jandora1@fel.cvut.cz