

Zadanie č. 3

Pomocou nástroja FLEX zostrojíte lexikálny analyzátor pre lexikálne elementy programovacieho jazyka popísané nižšie. Symboly programovacieho jazyka používajú štandardnú anglickú abecedu 26 písmen (malé aj veľké písmená), t.j. žiadne slovenské písmená s diakritikou!

Kódy príslušných lexém si zvolíte podľa uváženia!

- znaky `\t`, `\n`, `\r`, `\f`, medzera sa ignorujú
- **Kľúčové slová:** `begin`, `end`, `if`, `then`, `while`, `do`, `procedure`, `function`, `Integer`, `Real`, `Boolean`, `type`, `var`, `array`, `of`, `record`
- **Operátory:** `+`, `-`, `*`, `/`, `mod`, `and`, `or`, `not`, `<`, `<=`, `>`, `>=`, `=`, `<>`, `:=`
(t.j. slová *mod*, *and*, *or*, *not* predstavujú operátory, niektoré operátory pozostávajú z 2 znakov, napr. operátor priradenia `:=`)
- **Separátory:** `(`, `)`, `[`, `]`, `...`, `..`, `..`, `;`, `:`
(t.j. zátvorky, hranaté zátvorky, dve bodky, bodka, čiarka, bodkočiarka, dvojbodka)
- **Identifikátory.** Prvý znak identifikátora môže byť **len** malé alebo veľké písmeno. Ďalšie znaky identifikátora môžu byť malé písmená, veľké písmená alebo číslice 0,1,2,3,4,5,6,7,8,9. Veľkosť identifikátora je aspoň 1 znak.
- **Konštanty**
 - *Celočíselná konštanta* - pozostáva z postupnosti čísl 0,1,2,3,4,5,6,7,8,9. **môže/nemusí** začínať znamienkom `+` (plus) alebo `-` (mínus). **Môže** obsahovať ľubovoľný počet núl na začiatku - t.j. jedná sa **len** o ľubovoľnú postupnosť čísl, pred ktorou môže byť znamienko alebo plus, alebo mínus.
 - *Reálna konštanta* - Postupnosť čísl 0 až 9 nasledovaná bodkou nasledovaná postupnosťou čísl 0 až 9. Pred bodkou aj za bodkou musí byť aspoň 1 číslica. Na začiatku oboch postupností čísl môže byť ľubovoľný počet núl. Konštanta môže začínať znamienkom plus (`+`) alebo mínus (`-`).
 - *Boolean konštanta* - `true` / `false`. (t.j. slová *true* alebo *false*, ako napr. v C++).

Váš program/zdrojový súbor pre FLEX ďalej **musí** obsahovať testovaciu **main** funkciu, ktorá bude vypisovať kódy jednotlivých lexém rozpoznaných na vstupe, spolu s textovou hodnotou samotnej lexémy. Navyše v prípade, že pre lexému **konštantu** máte pridelený len jeden kód, musí program informovať, o akú konštantu sa jedná (celočíselná, reálna, boolean), napr. pomocou nejakej ďalšej premennej **typ** - tak, ako sme to robili na cvičení.

Čo odovzdať do AIS?

Do AIS odovzdajte **zdrojový súbor pre program FLEX**, t.j. súbor s koncovkou **.l**, napr. **analyzator.l**

Neodovzdávajte zdrojové kódy **lex.yy.c** ani výsledný program **lex.yy.exe**. Taktiež **nemusíte** odovzdávať žiadnu dokumentáciu.

Zadanie je hodnotené 10 bodmi.

DEADLINE zadania je 15. máj 2016, 24:00, t.j. polnoc z 15.5. na 16.5.. Hodnotím len tie zadania, ktoré si viem stiahnuť z AIS. Preto neodovzdanie zadania / **neskoré odovzdanie zadania** napr. mailom po deadline, sa hodnotí nula bodmi. Taktiež sa uistite, že ste zadanie riadne odovzdali, nie len nahrali a zabudli potvrdiť jeho odovzdanie (čaročrasy AIS sú žiaľ mimo môj dosah).

Príklad identifikátorov, konštánt

- Platné identifikátory: **ab**, **a0**, **A1**, **aB1a**
- Neplatné identifikátory: **1a** (začína číslom), **A+1** (obsahuje iný znak ako písmeno/číslicu), **Mačka1** (obsahuje znak mimo anglickej abecedy)
- Platné celočíselné konštanty: 1, 01, 00005, 5000, +5, +0, -0, -001, +05
- Neplatné celočíselné konštanty: 5- (znamienko nie je na prvom mieste)
- Platné reálne konštanty: 5.0, +5.0, -1.0, +10.00, +10.01
- Neplatné reálne konštanty: 5. (chýbajú číslice za bodkou), 5,1 (použitá čiarka miesto bodky), .5 (chýbajú číslice pred bodkou)