

TKOM - projekt wstępny

Sebastian Pietras

1 Temat projektu

Stworzenie języka, w którym typem wbudowanym będzie czas oraz interpretera tego języka.

2 Opis języka

Żeby nadać językowi trochę charakteru, będzie on nosił nazwę: **Timon**

Język umożliwia przechowywanie informacji o:

- samej dacie (`date`)
- samej godzinie (`hour`)
- dacie i godzinie razem (`datetime`)
- różnicy czasu (`timedelta`)

Pozwala również na dokonywanie operacji na tych typach danych:

- `date1 - date2 = timedelta`
- `hour1 - hour2 = timedelta`
- `datetime1 - datetime2 = timedelta`
- `datetime - date = timedelta`
- `datetime - hour = timedelta`
- `date1 +- timedelta = date2`
- `hour1 +- timedelta = hour2`
- `datetime1 +- timedelta = datetime2`

- `timedelta1 +/- timedelta2 = timedelta3`
- `timedelta1 */ number = timedelta2`

Operacje takie jak mnożenie dat nie miałyby sensu.

`date` przechowuje informacje o dniu, miesiącu i roku (np. `date = 10.04.2012`), `hour` o godzinie, minucie i sekundzie (np. `hour = 16:34:57`), `datetime` to połączenie `date` i `hour` (np. `datetime = 10.04.2012~16:34:57`), a `timedelta` o latach, miesiącach, tygodniach, dniach, godzinach, minutach i sekundach (np. `timedelta = '10Y 1W 5h'`).

Formaty literałów:

- `date` - dzień.miesiąc.rok
- `hour` - godzina:minuta:sekunda
- `datetime` - dzień.miesiąc.rok~godzina:minuta:sekunda
- `timedelta` - 'lata Y miesiące M tygodnie W dni D
godziny h minuty m sekundy s'

Wszystkie jednostki czasu w `timedelta` są opcjonalne, '' oznacza wartość zerową.

Bardziej formalny opis znajduje się w gramatyce opisanej przez EBNF.

Daty są takie jak w zwykłym kalendarzu gregoriańskim. Uwzględnione są lata przestępne. Wszystkie lata muszą być z naszej ery, dodatkowo istnieje rok 0.

Czas jest 24-godzinny.

Dostępna jest pętla `from`, która pozwala iterować między dwiema wartościami czasowymi podaną jednostką czasu. Na przykład:

```
from 01.01.2020 to 10.01.2020 by days as it
```

wykona się 10 razy i w każdym przebiegu zmienna `it` będzie przechowywać kolejny dzień.

Pozostałe cechy języka:

- dynamiczne typowanie
- typ liczbowy i znakowy
- instrukcja `print`
- instrukcja warunkowa `if else`
- definiowanie funkcji na poziomie globalnym
- deklarowanie zmiennych na każdym poziomie
- operacje logiczne i matematyczne o różnych priorytetach
- obsługa komentarzy wieloliniowych (między znakami `#`)
- dostęp do szczegółowych informacji o czasie (np. `date.days`)

3 Przykłady wykorzystania języka

Przykład 1

```
fun printDaysBetweenDates(d1,d2)
{
    # saving to d3 each consecutive day between d1 and d2 #

    from d1 to d2 by days as d3
    {
        print d3;
    };

    return 0;
};

var dt = 10.04.2018~10:57:00; # moment in time #
var d = 12.04.2018; # date #

printDaysBetweenDates(dt, d);
```

Przykład 2

```
var start_time = 10.06.2020;
var delay = '10W 10D';

var prev_t1 = 25.05.2020;
var prev_t2 = 20.05.2020;

if start_time + (prev_t1 - prev_t2) + delay <= 20.06.2020
{
    print "we have time";
}
else
{
    print "we dont have time";
};
```

Przykład 3

```
var t1 = 15:57:23;
var t2 = 20:45:00;
var td = t2 - t1;
var h = td.hours; # accessing specific timedelta info #
var napis = "hours between " + t1 + " " + t2 + " : ";

print napis + h;
```

4 Gramatyka

```
program = { nestableStatement | nonNestableStatement } ;

nonNestableStatement = ( functionDefinitionStatement ), ";" ;
nestableStatement = ( functionCallStatement | variableDefinitionStatement |
    variableAssignmentStatement | ifStatement |
    fromStatement | printStatement |
    returnStatement ), ";" ;

functionDefinitionStatement = "fun", identifier, parameters, body ;
functionCallStatement = identifier, parameters ;
variableDefinitionStatement = "var", identifier, [ assignment ] ;
variableAssignmentStatement = identifier, assignment ;
ifStatement = "if", logicExpr, body, [ "else", body ] ;
fromStatement = "from", fromRange, fromStep, fromIterator, body ;
printStatement = "print", mathExpr ;
```

```

returnStatement = "return", mathExpr ;

identifier = nonDigitCharacter, { character } ;
parameters = "(", [ identifier, { ",", identifier } ], ")" ;
body = "{", { nestableStatement }, "}" ;

assignment = assignmentOperator, mathExpr ;

logicExpr = logicAndExpr, { orOperator, logicAndExpr } ;
logicAndExpr = logicEqualExpr, { andOperator, logicEqualExpr } ;
logicEqualExpr = logicRelExpr, [ equalityOperator, logicRelExpr ] ;
logicRelExpr = logicTerm, [ relationOperator, logicTerm ] ;
logicTerm = [ logicNegOperator ], ( mathExpr | parenthesisedLogicExpr ) ;
parenthesisedLogicExpr = "(", logicExpr, ")" ;

fromRange = mathExpr, "to", mathExpr ;
fromStep = "by", timeUnit ;
fromIterator = "as", identifier ;

mathExpr = multMathExpr, { additiveOperator, multMathExpr } ;
multMathExpr = mathTerm, { multiplicativeOperator, mathTerm } ;
mathTerm = [ mathNegOperator ], ( value | parenthesisedMathExpr ) ;
parenthesisedMathExpr = "(", mathExpr, ")" ;

nonDigitCharacter = letter | specialCharacter ;
character = nonDigitCharacter | digit ;

assignmentOperator = "=" ;
orOperator = "|" ;
andOperator = "&" ;
equalityOperator = "==" | "!=" ;
relationOperator = ">" | ">=" | "<" | "<=" ;
logicNegOperator = "!" ;

timeUnit = "years" | "months" | "weeks" | "days" |
           "hours" | "minutes" | "seconds" ;

additiveOperator = "+" | "-" ;
multiplicativeOperator = "*" | "/" ;
mathNegOperator = "-" ;

value = identifier | literal | functionCallStatement | timeInfoAccess ;

literal = numberLiteral | stringLiteral | dateLiteral |
         timeLiteral | datetimeLiteral | timedeltaLiteral ;
timeInfoAccess = identifier, ".", timeUnit ;

numberLiteral = "0" | ( nonZeroDigit, { digit } ) ;
stringLiteral = "'", { stringCharacter }, "'" ;

```

```

dateLiteral = dayNumber, ".", monthNumber, ".", yearNumber ;
timeLiteral = hourNumber, ":", minuteNumber, ":", secondNumber ;
datetimeLiteral = dateLiteral, "~", timeLiteral ;
timedeltaLiteral = "'", [ numberLiteral, "Y" ], [ numberLiteral, "M" ],
                        [ numberLiteral, "W" ], [ numberLiteral, "D" ],
                        [ numberLiteral, "h" ], [ numberLiteral, "m" ],
                        [ numberLiteral, "s" ], "'" ;

letter = "a" | "b" | "c" | "d" | "e" | "f" | "g" | "h" | "i" | "j" |
        "k" | "l" | "m" | "n" | "o" | "p" | "q" | "r" | "s" | "t" |
        "u" | "v" | "w" | "x" | "y" | "z" |
        "A" | "B" | "C" | "D" | "E" | "F" | "G" | "H" | "I" | "J" |
        "K" | "L" | "M" | "N" | "O" | "P" | "Q" | "R" | "S" | "T" |
        "U" | "V" | "W" | "X" | "Y" | "Z" ;
specialCharacter = "_" | "-" ;
digit = "0" | nonZeroDigit ;
nonZeroDigit = "1" | "2" | "3" | "4" | "5" | "6" | "7" | "8" | "9" ;

stringCharacter = ? all visible characters ? - "'" ;

dayNumber = ( "0", nonZeroDigit ) |
            ( ( "1" | "2" ), digit ) |
            ( "3", ( "0" | "1" ) ) ;
monthNumber = ( "0", nonZeroDigit ) | ( "1", ( "0" | "1" | "2" ) ) ;
yearNumber = numberLiteral ;
hourNumber = ( ( "0" | "1" ), digit ) |
            ( "2", ( "0" | "1" | "2" | "3" ) ) ;
minuteNumber = ( "0" | "1" | "2" | "3" | "4" | "5" ), digit ;
secondNumber = minuteNumber ;

```

W gramatyce nie są uwzględnione białe znaki i komentarze, bo mogą występować wszędzie i będą po prostu pomijane przy parsowaniu.

Podstawowe zasady dotyczące dat są ujęte w gramatyce (na przykład to, że nie istnieją dni większe od 31), ale nie zapewnia to całkowitej poprawności. Na poziomie semantycznym trzeba będzie wykonać dodatkowe sprawdzenia (na przykład czy miesiąc ma 30 czy 31 dni albo czy jest to rok przestępny).

5 Zarys struktury interpretera

Programy będzie napisany w języku Python. Na wejście będzie podawana ścieżka do pliku zawierającego skrypt do interpretacji.

Interpreter będzie składał się z kilku etapów:

- analiza leksykalna - zmiana znaków na tokeny
- analiza składniowa - sprawdzenie, czy tokeny są zgodne z gramatyką i stworzenie drzewa składniowego
- analiza semantyczna - sprawdzenie, czy w drzewie znajdują się operacje, które były dozwolone składniowo, ale nie mogą zostać wykonane (np. przypisanie do identyfikatora funkcji zamiast zmiennej)
- uruchomienie - wykonanie po kolei instrukcji z drzewa zgodnie z ich znaczeniem

Analizator leksykalny będzie zbierał znaki w tokeny i zwracały je na żądanie analizatora składniowego, aby jak najszybciej przerwać analizę po wystąpieniu błędu syntaktycznego.

Dodatkowo dostępne będą pomocnicze moduły obsługi plików, błędów i tablicy symboli.