



SNOBOL4

TAREFA 02 – ESTRUTURA DE LINGUAGENS

Leonardo Sillas
Vinicius Santiago

INTRODUÇÃO

- A linguagem de programação SNOBOL 4 (String Oriented Symbolic Language number 4), originalmente e não oficialmente chamada de SEXI e SCL7, é a quarta e última encarnação de uma série de linguagens de programação específicas destinadas à manipulação de sequências de caracteres (strings).
- É baseada nos princípios de correspondência de padrões (pattern matching) para solucionar problemas de manipulação de strings.
- Foi desenvolvida entre 1962 e 1967 nos Laboratórios Bell, da AT&T, por D. J. Farber, R. E. Griswold, e F. P. Polensky.
- Seu principal diferencial foi um novo tipo de dado chamado “first-class data type”, cujos valores podiam ser manipulados de todas as formas permitidas para todos os tipos diferentes de dados da linguagem, tornando-a dinâmica.
- Suas operações básicas são formação de uma string, busca de padrões e substituição de strings.
- Foi popularmente utilizado nos anos 70 e 80 como linguagem de manipulação de textos e substituído por linguagens novas e mais eficientes como o Awk e Perl.

SNOBOL4 FOI INFLUENCIADO POR:

- **COMIT**, que foi a primeira linguagem específica para processamento de cadeias de caracteres e de correspondência de padrões, projetada para aplicações em tradução de linguagem natural, onde um programa era um conjunto de regras, cada uma com um padrão, uma substituição e outra regra.
- **MARKOV**, um algoritmo de Markov é um sistema de reescrita de sequência de caracteres que usa regras semelhantes a gramática para operar em cadeias de símbolos. Tem poder equivalente ao de uma máquina de Turing e suas cadeias podem representar qualquer expressão matemática a partir da sua notação simples.
- SCL(Symbolic Communication Language), que foi desenvolvida para manipulação de símbolos e inspirou o pattern matching do SNOBOL4.
- Pode ter tido influência também o compilador **SHADOW** que era dirigido por sintaxe. Este compilador antecipou parte da sistemática de padrões em SNOBOL e ICES. Foi utilizado por David J. Farber nos experimentos de padrões iniciais, então pode ter tido influência no SNOBOL.

SNOBOL4 INFLUENCIOU:

- **AWK**, baseada na linguagem C, é utilizada frequentemente por desenvolvedores para processar textos e manipular arquivos. Esta linguagem é considerada por muitos um importante marco para história da programação, tendo tido bastante influência na criação de outras linguagens de programação, como Perl e Lua.
- **SL5**, que é uma linguagem de processamento de listas e strings com sintaxe orientada à expressão. Faz uso de rotinas.
- **ICON**, que é uma linguagem de programação de alto nível, com execução direcionada a objetivos e muitas facilidades para gerenciar strings e padrões de texto.

IMPERATIVA

- Usa declarações que alteram o estado de um programa. Da mesma forma que em linguagens naturais, expressa comandos, um programa imperativo consiste em comandos para o computador executar. A programação imperativa enfoca a descrição de como um programa funciona através de comandos de atribuição baseados em transferência de dados e instruções.
- No SNOBOL4 temos por exemplo:
- $N1 = 43$
- $D = 17$
- $OUTPUT = N1 + D$
- 60

NÃO-ESTRUTURADA

- A programação não estruturada é o paradigma de programação historicamente mais antigo, capaz de criar algoritmos completos de Turing. Muitas vezes, é contrastado com paradigmas de programação estruturada, incluindo programação processual, funcional e orientada a objetos. No SNOBOL4 vemos como exemplo o conceito de GOTO que altera a ordem de execução do programa através de labels, como no exemplo abaixo:
- $N = 0$
- `COPY OUTPUT = INPUT :F(DONE)`
- `N = N + 1 :(COPY)`
- `DONE OUTPUT = 'THERE WERE ' N ' LINES'`
- `END`

INTERPRETADA

- A maioria de suas implementações executam instruções diretamente, sem antes compilar um programa em instruções de linguagem de máquina. O interpretador executa o programa diretamente, traduzindo cada declaração para uma sequência de uma ou mais sub-rotinas já compiladas em código de máquina.

DINÂMICA

- As variáveis não precisam ser pré-declaradas e seus valores podem mudar os tipos com facilidade. No exemplo abaixo, J começa como uma string, mas depois é alterado para que contenha um número inteiro; 42 é impresso.
- `J = "ABC"`
- `A = 20`
- `B = "22"`
- `J = A + B`
- `OUTPUT = J`

PATTERN MATCHING

(CORRESPONDÊNCIA DE PADRÕES)

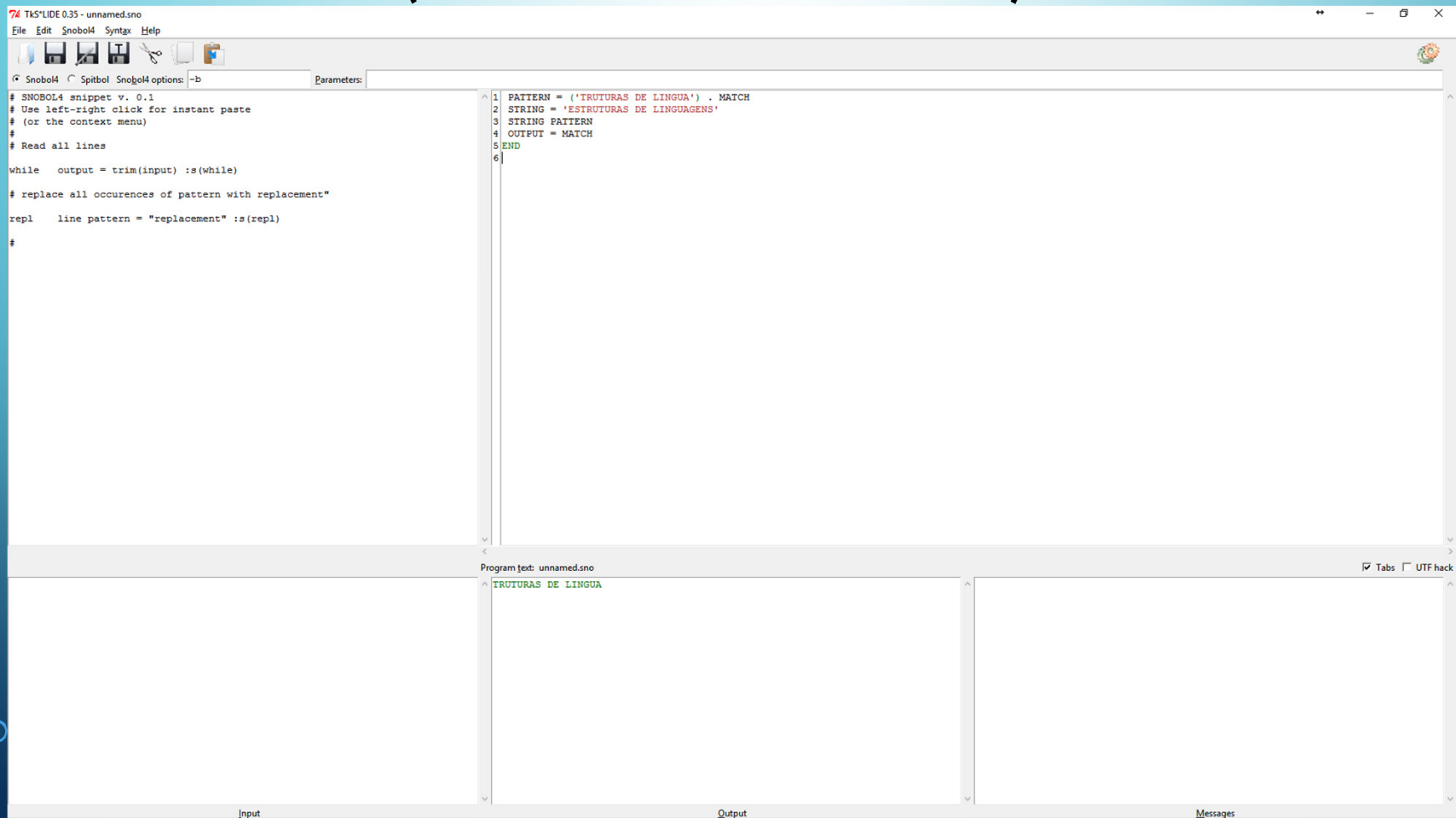
- A correspondência de padrões examina uma sequência de strings para alguma combinação de caracteres, chamado de “padrão”. O processo de correspondência pode ser muito simples ou extremamente complexo. Por exemplo:
- O assunto contém vários nomes de cores. O padrão é a string “BLUE”. A sequência do assunto contém a palavra “BLUE”?
- Um padrão de correspondência requer uma cadeia de strings e um padrão. A string é o primeiro elemento de declaração após o campo de label (se houver). O padrão aparece em seguida, separado da string por espaço em branco. Se STRING é a sequência de caracteres a ser examinada, e PATTERN é o padrão, parece assim:

Label STRING PATTERN

- A correspondência de padrão ocorre se o padrão for encontrado na sequência da string; caso contrário, ele falha. Esse sucesso ou falha pode ser testado no campo GOTO:

Label STRING PATTERN :S(label1) F(label2)

COMPARAÇÃO DE CÓDIGOS REPRESENTATIVOS (PATTERN MATCHING)



The screenshot displays the TkS*LIDE 0.35 Snobol4 editor interface. The main window is titled "TkS*LIDE 0.35 - unnamed.sno" and contains a menu bar (File, Edit, Snobol4, Syntax, Help) and a toolbar. The editor is divided into two panes. The left pane shows a Snobol4 snippet with the following code:

```
# Snobol4 snippet v. 0.1
# Use left-right click for instant paste
# (or the context menu)
#
# Read all lines
while output = trim(input) :s(while)
# replace all occurrences of pattern with replacement"
repl line pattern = "replacement" :s(repl)
#
```

The right pane shows a Snobol4 program with the following code:

```
1 PATTERN = ('TRUTURAS DE LINGUA') . MATCH
2 STRING = 'ESTRUTURAS DE LINGUAGENS'
3 STRING PATTERN
4 OUTPUT = MATCH
5 END
6
```

Below the editor panes, there is a status bar with the text "Program text: unnamed.sno". At the bottom of the window, there are three tabs: "Input", "Output", and "Messages". The "Output" tab is currently selected, showing the text "TRUTURAS DE LINGUA".

COMPARAÇÃO DE CÓDIGOS REPRESENTATIVOS (PATTERN MATCHING)

```
C:\Windows\System32\cmd.exe - python
>>> pattern = "TRUTURAS DE LINGUA"
>>> string = "ESTRUTURAS DE LINGUAGENS"
>>> if pattern in string:
...     print(pattern)
...
TRUTURAS DE LINGUA
>>>
```


CONTROLE DE FLUXO

A transferência de controle é realizada pelo comando GOTO. Ele interrompe a execução sequencial normal das instruções, informando qual declaração deve ser executada em seguida. O GOTO aparece no final da instrução, e por ser acompanhado por:

- 1 - dois pontos (:), para desvio incondicional;
- 2 - dois pontos + F (:F), para desvio somente se a condição avaliada retornar false (FAIL);
- 3 - dois pontos + S (:S), para desvio somente se a condição avaliada retornar true (SUCCESS).

```
* PROGRAMA QUE CALCULA A MEDIA E O SOMATÓRIO DE N NUMEROS
  I = 0
  SOMA = 0
  TOPO NUMERO = TRIM(INPUT)
  I = I + 1
  SOMA = SOMA + NUMERO
  FIM MEDIA = SOMA / I
  OUTPUT = 'SOMATORIO =' SOMA
  OUTPUT = 'MEDIA      =' MEDIA
  END
```

1. :F(FIM)

2. : (TOPO)

No exemplo, a transferência “1.” ocorre apenas se houver falha na atribuição da variável ‘NUMERO’, caso contrário, a execução prossegue até o desvio incondicional “2.” onde retorna a execução do programa para a linha referenciada “TOPO”.

COMPARAÇÃO DE CÓDIGOS REPRESENTATIVOS (CONTROLE DE FLUXO)

SNOBOL4

```
TOPO n = input()
      DIFFER(n,'exit') :S(TOPO) F(FIM)
FIM   END
```

PYTHON

- while True:
- n = raw_input()
- if n.strip() == 'exit':
- break

A Variável 'n' está recebendo uma atribuição e enquanto ela não receber a palavra 'exit', ela continuará recebendo atribuições.

AVALIAÇÃO EM COMPARAÇÃO AO PYTHON

READABILITY

Python possui uma melhor Readability pois fica mais transparente à quem estiver lendo o que está sendo feito, pois tanto no primeiro caso onde temos a chamada do método “in” (em) que é mais natural, como se fosse uma simples leitura tanto no segundo, pois a utilização do controle de fluxo e da função DIFFER não são tão naturais e óbvias de se entender quanto um loop com while true/break no python. Reforça-se esse argumento pois em Snobol4 o padrão de STRING PATTERN pode, em um primeiro momento, confundir quem está lendo sobre o que está acontecendo.

WRITABILITY

Nas duas linguagens verificamos a mesma facilidade e simplicidade para atingir o mesmo objetivo, portanto acreditamos que nesse quesito o writability das duas linguagens é o mesmo.

EXPRESSIVIDADE

Se uma linguagem exige muito código para expressar uma abstração, ela é menos expressiva, portanto nesse contexto, por julgarmos que em ambas as linguagens a exigência de código para representar tanto a abstração do match pattern quanto a do controle de fluxo é a mesma, tendo em vista que por ser uma linguagem muito antiga e suas principais características já terem sido incorporadas pelas linguagens mais modernas, a expressividade de ambas é equivalente.

CONCLUSÃO

- Por ser uma linguagem muito antiga, carece de suporte apropriado, seja da própria linguagem ou da comunidade (???). Apesar disso, SNOBOL4 ainda possui vantagens quando o quesito é o tratamento de strings, como por exemplo tipagem dinâmica que fora abordada anteriormente. Foi uma linguagem tão revolucionária na época que influenciou várias linguagens em gerações posteriores e ainda hoje, as linguagens mais modernas possuem algum traço de seu principal diferencial, que foi o pattern matching (correspondência de padrões).

Bibliografia

- Slides: <https://github.com/vinivst/EDL/blob/master/slides/edl-01.pdf>
- Wikipedia:
 - <https://en.wikipedia.org/wiki/SNOBOL>
 - [https://en.wikipedia.org/wiki/Icon_\(programming_language\)](https://en.wikipedia.org/wiki/Icon_(programming_language))
 - <https://pt.wikipedia.org/wiki/AWK>
- Hopl:
 - <http://hopl.info/showlanguage.prx?exp=171>
 - <http://hopl.info/showlanguage.prx?exp=783>
- Site da linguagem:
 - <http://www.snobol4.org/>