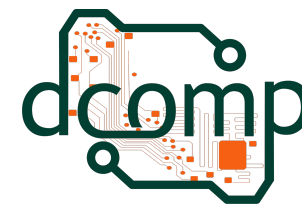




Universidade Federal do Espírito Santo
Centro de Ciências Agrárias – CCA UFES
Departamento de Computação



Processamento de Linguagem Natural (PLN)

Inteligência Artificial

Site: <http://jeiks.net>

E-mail: jacsonrcsilva@gmail.com

Sumário

- Contextualização
- Habilidades e Competências
- História resumida
- Processamento de Linguagem Natural
 - Fonologia;
 - Morfologia;
 - Análise Sintática;
 - Análise Semântica;
 - Análise Pragmática.
- Exemplos de aplicações de PLN
- Conclusões

Contextualização

- O que ocorre quando...
 - ... ligamos para o suporte de uma companhia telefônica?
 - ... falamos um número e o celular faz a ligação?
 - ... digita uma frase com a gramática errada no editor de textos?
 - ... assiste um vídeo em inglês no Youtube e pede para ver as legendas automáticas?

Um pouco de História

- Diversas pesquisas por volta de 1940...
 - Primeira máquina relacionada a Linguagem Natural: Máquina de Tradução (MT).
- Em 1946 máquinas buscavam reconhecer e quebrar códigos na Segunda Guerra Mundial.
- Várias pesquisas seguinte apresentaram falta de análise de ambiguidade léxicas... resultaram em fracassos.
- Em 1957, Chomsky introduziu as ideias de gramática.
- Em 1965, as ideias de Chomsky foram base para outros cientistas trabalharem sobre a representação da gramática.

Um pouco de História

- Nos anos 70, a atenção voltou-se para a semântica, fonemas e planos para a comunicação.
 - McDonald criou um gerador de respostas: MUMMBLE.
- Nos anos 80, revisão da literatura utilizada para endereçar problemas em PLN. E várias soluções começaram a ser pesquisadas.
- Nos anos 90, muitas pesquisas apresentaram sucesso com a identificação da linguagem, tratamento da ambiguidade, etc.
- Na década seguinte e atualmente:
 - Desenvolver e modelar sistemas linguísticos,
 - com sofisticação adequada a determinado objetivo.
 - Conceber o implementar modelos e sistemas de PLN;

O que é PLN?

Um conjunto de teorias de técnicas computacionais para analisar e representar naturalmente textos em um ou mais níveis da análise linguística com propósito de realizar o processamento da linguagem humana para um conjunto de tarefas e aplicações.

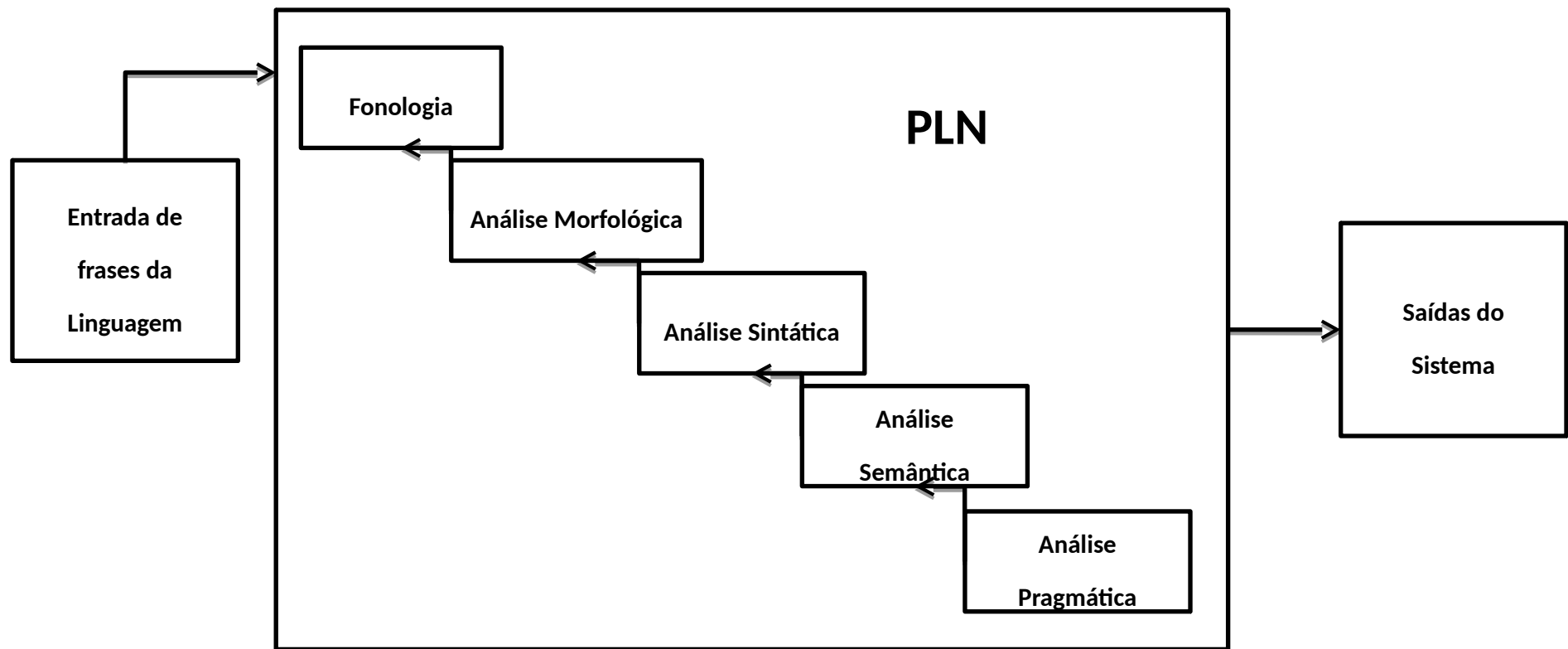
Alguns problemas para tratar linguagens

- Ambiguidade
 - Texto com mais de um significado:
 - A manga ficou amarela.
 - A cachorra da sua sogra esteve no meu terreiro ontem.
- Linguagens diferentes. Ex:
 - Inglês: The white cat .
 - Português: O gato branco.
 - Árabe: القط الابيض
- Identificar qual linguagem está trabalhando.

Processamento de Linguagem Natural

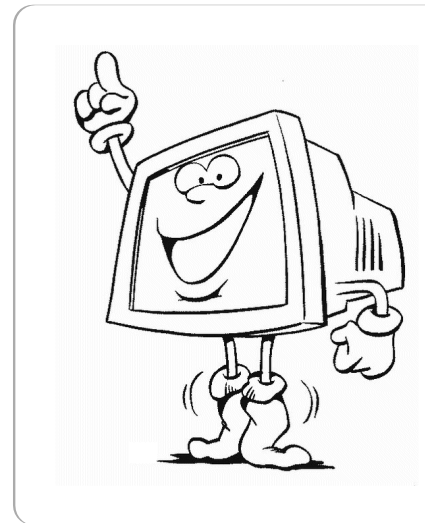
- O sistema precisa ser capaz de processar e manipular a linguagem em diversos níveis.
- E fornecer como saída:
 - Análise sintática, semântica e pragmática;
 - Textos em linguagem natural ou artificial;
 - Diferentes tipos de documentos estruturados.
- A pesquisa da PLN está voltada em três aspectos da língua natural:
 - Som: fonologia;
 - Estrutura: morfologia e sintaxe;
 - Significado: semântica e pragmática.

Processamento de Linguagem Natural



Processamento de Linguagem Natural

- Fonologia:
 - Estudo dos sons que compõem as palavras;
 - Utilizado para identificar palavras a partir de sons;
 - Somente se o computador necessitar entender a linguagem falada;



Processamento de Linguagem Natural

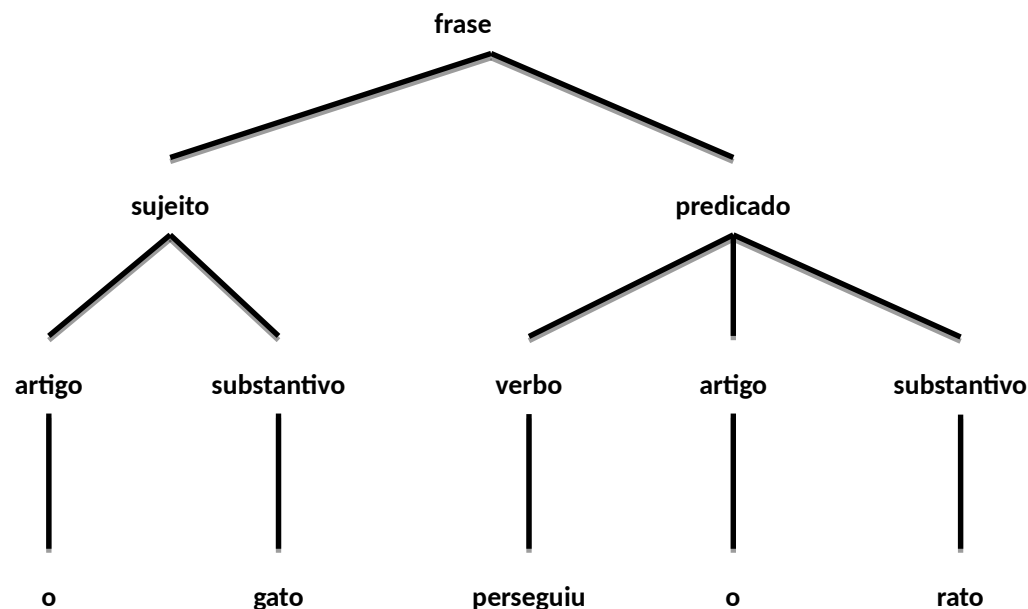
- Análise Morfológica:
 - Primeiro estágio da análise das palavras;
 - Examina os modos que as palavras se desmembram em componentes ; e
 - Como isto afeta o estado gramatical delas:
 - Substantivos;
 - Verbos;
 - Pronomes.
- Nesta fase, pode-se substituir a palavra por outra, mas mantendo o significado.

Análise Morfológica

- Seria ideal ter um dicionário grande com todas as palavras, mas seu tamanho seria enorme.
- Uma abordagem mais prática é:
 - Incluir regras que funcionem com palavras regulares;
 - Adicionar uma lista de palavras irregulares;
 - E tratar domínios específicos bem limitados.
- Exemplo:
 - cantar – cantando – cantor
 - nadar – nadando – nadador

Processamento de Linguagem Natural

- Análise Sintática:
 - Avalia a gramática da linguagem utilizada;
 - Gramática: especificação formal da estruturas das sentenças permitidas na linguagem
 - Gera uma estrutura da sentença analisada:



Análise sintática

- Para realizar a análise da sintática:
 - Definir uma gramática com regras;
 - Regras na Forma Normal de Backus (BNF):
 - Símbolos terminais: que não podem ser derivados;
 - Símbolos não terminais: que podem ser derivados, sendo substituídos por outro(s) símbolo(s);
 - Símbolo inicial: representa uma sentença completa da linguagem;
 - Regras de derivação: definem a estrutura gramática, a utilização dos símbolos.
- A gramática mais utilizada é a Livre de Contexto:
 - somente um símbolo não terminal no lado esquerdo de suas regras de derivação.

Gramática e análise

Gramática	
<i>frase</i>	-> <i>sujeito predicado</i>
<i>sujeito</i>	-> <i>artigo substantivo</i>
<i>predicado</i>	-> <i>verbo artigo substantivo</i>
<i>artigo</i>	-> <i>o</i>
<i>substantivo</i>	-> <i>gato rato</i>
<i>verbo</i>	-> <i>perseguiu</i>

Top-down

- > *frase*
- > *sujeito predicado*
- > *artigo substantivo predicado*
- > *o substantivo predicado*
- > *o gato predicado*
- > *o gato verbo artigo substantivo*
- > *o gato perseguiu artigo substantivo*
- > *o gato perseguiu o substantivo*
- > *o gato perseguiu o rato*

Bottom-up

- o gato perseguiu o rato*
- > *artigo gato perseguiu o rato*
- > *artigo substantivo perseguiu o rato*
- > *sujeito perseguiu o rato*
- > *sujeito verbo o rato*
- > *sujeito verbo artigo rato*
- > *sujeito verbo artigo substantivo*
- > *sujeito predicado*
- > *frase*

Exemplo prático

- Uma das formas de fazer no computador
 - Prolog (GNU Prolog)
 - SWI-Prolog,
 - utilizando o DCG – Definite Clause Grammar.

- *exemplos...*

Processamento de Linguagem Natural

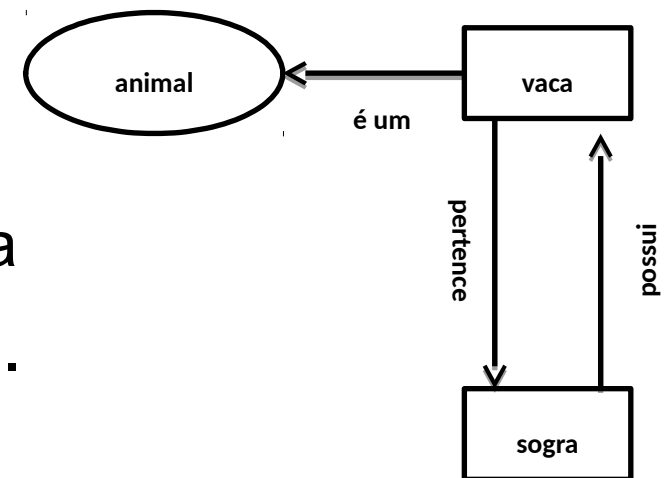
- Análise Semântica

- Associa um significado a estrutura sintática;
- Semântica:
 - considera os significados das palavras;
- Análise semântica:
 - análise para extrair um significado de uma declaração
- Exemplo:
 - Frase: O menino comeu o bolo.
 - Semântica: uma pessoa alimentou/digeriu um alimento.
- Pode-se representá-la por uma rede semântica:



Análise semântica

- Pode ser mais completa,
 - exemplo:
 - abrigando informações sobre natureza do gato: objeto, animal, quadrúpede, mamífero, etc.
 - permitindo dizer que ele gosta de leite.
- Também é útil para:
 - eliminar ambiguidade, ex.:
 - a vaca da minha sogra é branca
 - auxiliar na análise pragmática.

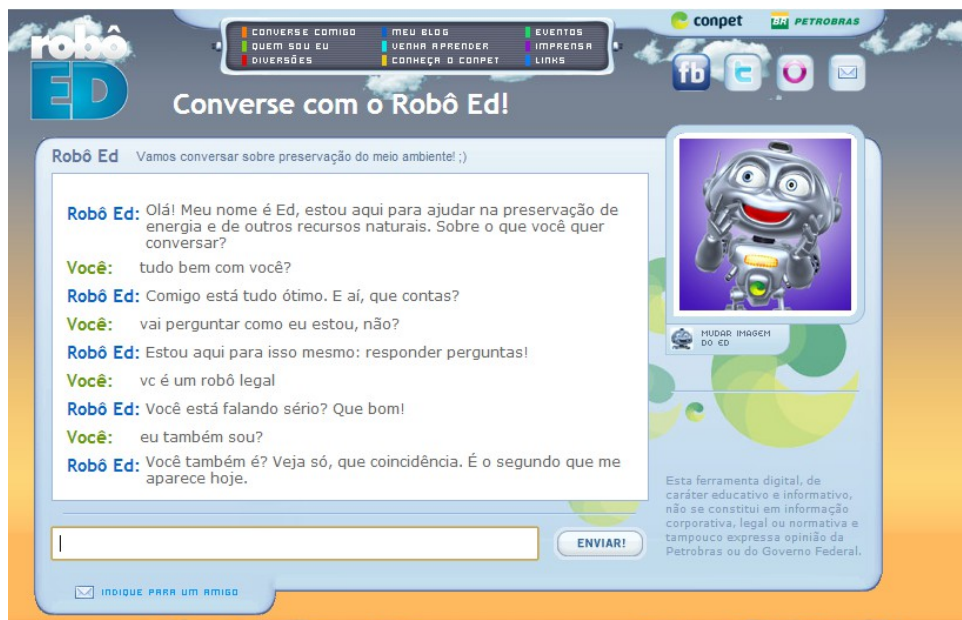


Processamento de Linguagem Natural

- Análise pragmática
- Interpreta o todo e não somente de uma parte:
 - Verifica se o significado da análise semântica está correto;
 - Determina significados que não estejam claros;
- Palavras se associam em:
 - paradigmáticas: através de um significado.
ex.: “nadar” e “água”.
 - sintagmáticas: conectam palavras frequentemente encontradas em um mesmo assunto.
ex.: “água” e “poço”.

Processamento de Linguagem Natural

- Aplicações em consulta de banco de dados com linguagem natural:
 - Denominada Recuperação de Dados.
 - Chatterbots, como “robô Ed” e “Se7e Zoom”



Robô Ed

robô ED

CONVERSE COMIGO | MEU BLOG | EVENTOS
 QUEM SOU EU | VENHA APRENDER | IMPRENSA
 DIVERSÕES | CONHEÇA O CONPET | LINKS

conpet | PETROBRAS

fb | | |

Converse com o Robô Ed!

Robô Ed Vamos conversar sobre preservação do meio ambiente! ;)

Robô Ed: Olá! Meu nome é Ed, estou aqui para ajudar na preservação de energia e de outros recursos naturais. Sobre o que você quer conversar?

Você: tudo bem com você?

Robô Ed: Comigo está tudo ótimo. E aí, que contas?

Você: vai perguntar como eu estou, não?

Robô Ed: Estou aqui para isso mesmo: responder perguntas!

Você: vc é um robô legal

Robô Ed: Você está falando sério? Que bom!

Você: eu também sou?

Robô Ed: Você também é? Veja só, que coincidência. É o segundo que me aparece hoje.

MUDAR IMAGEM DO ED

Esta ferramenta digital, de caráter educativo e informativo, não se constitui em informação corporativa, legal ou normativa e tampouco expressa opinião da Petrobras ou do Governo Federal.

ENVIAR!

INDIQUE PARA UM AMIGO

Se7e Zoom



s e 7 e . z o o m
Zoom

Você: você é mesmo um robô?

SeteZoom: Robô é você.

Você:

Conclusões

- O PLN possui diversas etapas que devem ser seguidas.
- O PLN é utilizado atualmente em diversos sistemas comuns às pessoas.
- O PLN possui diversas pesquisas e meios de implementação.
- O PLN apresenta muito sucesso em diversos sistemas atuais.
- Para a criação de sistemas de reconhecimento e recuperação de dados, é necessária uma pesquisa mais aprofundada na área.

Referências

- Referências Bibliográficas:
 - Coppin, Ben. Inteligência Artificial. Rio de Janeiro: LTC, 2010.
 - Barr, Avron; Feigenbaum, Edward A. The Handbook of Artificial Intelligence, vol. I., Stanford, Calif.: HeurisTech Press; Los Altos, Calif. : William Kaufmann, 1981.
 - LIDDY, E. D., et al. Natural Language Processing. Encyclopedia of Library and Information Science, pages 2126--2136, 2003.
- Prolog:
 - Daniel. The GNU Prolog: <<http://www.gprolog.org>>
 - SWI-Prolog. SWI-Prolog: <<http://www.swi-prolog.org>>
- ChatterBots:
 - CONPET. Robô Ed. Disponível em: <<http://www.ed.conpet.gov.br/br/converse.php>>
 - INSITE. Se7e Zoom. Disponível em: <<http://www.inbot.com.br/sete>>
- Curso de Processamento de Linguagem Natural de Stanford
 - MANNING, Christopher D. <<http://goo.gl/6gluf>>