

Conexum LITERAL – descrição do protótipo maio de 2005

Definição

O software LITERAL é um programa construído na linguagem de programação C++. Isto lhe permite ser compilado e utilizado em plataformas Microsoft Windows e Linux. Seu objetivo é, dentro da suíte OpenOffice, verificar a gramática do texto escrito.

Implementação

O software foi criado seguindo as normas da orientação a objetos, onde todo o sistema é descrito em diagramas e posteriormente transformado em linguagem de programação. Para tanto, o programa foi dividido em 11 (onze) classes diferentes, cada uma cuidando de uma parte do programa e interagindo entre si. O LITERAL conta com classes para armazenamento de dicionários (léxico, verbos e regências), manipulação de frases, análise de concordância (nominal e verbal) e análise de regência (nominal e verbal).

No momento atual, o LITERAL já possui uma versão beta para a plataforma Windows. A Figura 1, mostra a tela do programa no seu atual estágio.

Conexum LITERAL - v0.5 Beta

Léxico | Divisão de Sintagmas | Concordância Nominal/Verbal | Regência Nominal/Verbal

Digite uma frase (com pontuação final):
Encontraram abandonados o posto e os restaurantes. Máximo: 120 caracteres

Classificação Gramatical:

Palavra	Classe	Modificada	Gênero	Número
encontraram	Verbo	AÇÃO	Ambos	Plural
abandonados	Adjetivo	ADSUJ	Masculino	Plural
o	Artigo	ADSUJ	Masculino	Singular
posto	Substantivo	SUJ	Masculino	Singular
e	Conjunção	ADFRASE	Ambos	Ambos

Complementos:

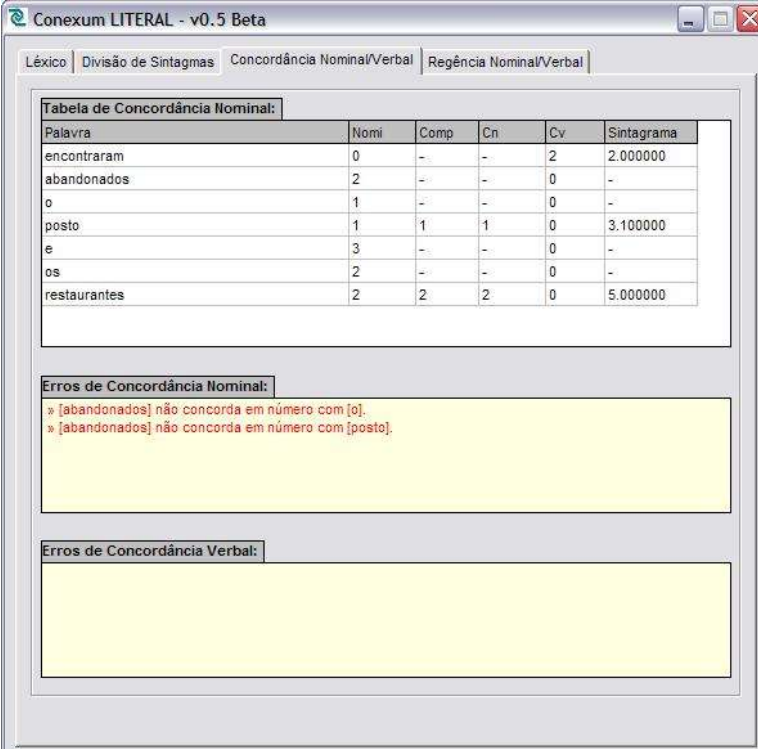
Sujeito	-
Verbo	encontraram
Predicado	encontraram abandonados o posto
Advérbio	-
Complementoaprep	-
Complementoprep	posto
Oracor	0
Orasub	0
Oração	

Ocorrência de erros: verificar erros nas abas correspondentes

>> Erro(s) de Concordância Nominal

Figura 1. Tela da versão beta do LITERAL.

Neste exemplo da Figura 1, foi apresentada a frase “Encontraram abandonados o posto e os restaurantes”. O programa acusou erro concordância nominal, indicando em vermelho na parte inferior da tela. Clicando na aba “Concordância Nominal/Verbal” o programa mostra qual o tipo de erro ocorrido. A Figura 2 ilustra esta tela.



Palavra	Nomi	Comp	Cn	Cv	Sintagma
encontraram	0	-	-	2	2.000000
abandonados	2	-	-	0	-
o	1	-	-	0	-
posto	1	1	1	0	3.100000
e	3	-	-	0	-
os	2	-	-	0	-
restaurantes	2	2	2	0	5.000000

Erros de Concordância Nominal:

- » [abandonados] não concorda em número com [o].
- » [abandonados] não concorda em número com [posto].

Erros de Concordância Verbal:

Figura 2. Tela de indicação de erros de concordância nominal e verbal.

A implementação das regras de concordância verbal e de regência verbal ainda não está completa. Um dicionário com mais de 14.000 verbos está sendo construído para o programa. Atualmente, o LITERAL conta com pouco mais de 150 verbos e um dicionário léxico com mais de 50.000 palavras da língua portuguesa. A conclusão de um módulo que gera hipóteses de correção para os erros encontrados já está em fase de projeto.

O Corretor ortográfico

Juntamente com o LITERAL foi construído um programa, independente, para revisão e correção ortográfica. Assim como o LITERAL, este programa também foi escrito em linguagem C++.

Com uma implementação mais simples do que o módulo gramatical, o ortográfico apenas verifica se uma palavra está escrita corretamente. Caso não esteja, o programa indica uma lista de possibilidades para substituição.

A Figura 3 mostra a tela e um exemplo do funcionamento do corretor ortográfico.



Figura 3. Tela do corretor ortográfico.

Ao ser digitada uma palavra, o programa faz uma varredura em todo o dicionário (o mesmo utilizado na análise léxica do LITERAL) em busca de uma entrada igual. Caso não encontre, o corretor identifica um erro de ortografia. Automaticamente, o programa inicia uma série de cálculos de semelhança entre palavras do dicionário e o que foi digitado. Cada análise gera um valor que indica sua semelhança à entrada. Na tela as palavras semelhantes são mostradas ao usuário na forma de uma lista. Ao lado desta lista o valor da “taxa de semelhança” é apresentado. No final do processamento, o programa recomenda ao usuário (na caixa inferior esquerda da tela) a palavra conhecida que alcançou maior taxa de semelhança. Possivelmente, esta é a palavra correta.

O corretor ortográfico está parcialmente dentro da versão atual do LITERAL. O motivo é o objetivo do LITERAL de ser inserido na suíte OpenOffice. Esta suíte já possui um corretor ortográfico, logo não há necessidade de outro módulo de mesma natureza. Sua função dentro do LITERAL, no momento, é apenas encontrar a classificação gramatical de uma palavra do dicionário. Na versão final (dentro da suíte OpenOffice) este módulo poderá ser descartado.