

Descrição do trabalho para o sistema AnELL

Isabel Marcelino
26 de Novembro de 2004

O propósito deste relatório é descrever os objectivos do meu trabalho, a forma de os atingir, e o processo prático (*workflow*). Esta versão corresponde ao resultado do meu período de formação em Oslo, em que me inteirei do que a Linguateca espera de mim. Ainda terá de ser aprovada e/ou reformulada pelo LabEL.

Introdução (Motivação do meu trabalho)

- Do ponto de vista linguístico, é uma necessidade rever os textos anotados automaticamente pelo AnELL por causa de:
 - os problemas da ambiguidade.
 - os erros de análise.
 - as palavras não analisadas.

Exemplo duma análise automática de verbos antes de ser revista:

A estrada era famosa pelas curvas e pela sombra das árvores. O Rubro admirava a paisagem que se espraiava para o fundo, enquanto o Rocha mudava de calças por detrás do carro estacionado. Era a hora do equinócio e o calor ia alagando as carnes de suor naquele começo de Agosto. Os dois rapazes esperavam a passagem da camioneta que transportaria mais três companheiros de Braga. Por terem atestado o carro com o material de campismo, não couberam todos.

(*Estão em azul, todas as palavras susceptíveis de serem verbos.*)

Com este exemplo, vemos que há uma real necessidade de rever os textos anotados automaticamente, porque senão, neste caso, todas as palavras em azul seriam analisadas como verbos, mesmo aquelas que no contexto em que ocorrem não o são.

- Do ponto de vista da melhoria do sistema:
 - simplificação da interface.
 - interacção com o utilizador.

I. Principais objectivos/metasp do meu trabalho

Os principais objectivos do meu trabalho são:

1. Fazer um novo modo : o «*Modo supervisionado simples*»

O objectivo desse modo, é ser muito mais simples do que os outros modos:

- *Não anota o texto todo*, por exemplo, pode anotar :. só nomes,
 - . só entidades mencionadas,

- . só verbos,
- . só palavras compostas,
- . só siglas, ...

- O *output* vai ser no *próprio texto* e também numa **lista**.

Exemplos:

- Para os nomes :

<p align="center">Texto anotado <u>sem</u> revisão (todas as palavras que podem ser nomes)</p>	<p align="center">Texto anotado <u>com</u> revisão (só os nomes)</p>
<p>A <u>estrada</u> <u>era</u> famosa pelas <u>curvas</u> e pela <u>sombra</u> das <u>árvores</u>. O <u>Rubro</u> admirava a <u>paisagem</u> que se espalhava para o <u>fundo</u>, enquanto o <u>Rocha</u> mudava de <u>calças</u> por detrás do <u>carro</u> estacionado. <u>Era a hora</u> do equinócio e o <u>calor</u> ia alagando as <u>carnes</u> de suor naquele <u>começo</u> de <u>Agosto</u>. Os dois <u>rapazes</u> esperavam a <u>passagem</u> da <u>camioneta</u> que transportaria <u>mais três</u> <u>companheiros</u> de <u>Braga</u>. Por terem <u>atestado</u> o <u>carro</u> com o <u>material</u> de <u>campismo</u>, <u>não</u> couberam todos.</p>	<p>A <u>estrada</u> era famosa pelas <u>curvas</u> e pela <u>sombra</u> das <u>árvores</u>. O <u>Rubro</u> admirava a <u>paisagem</u> que se espalhava para o <u>fundo</u>, enquanto o <u>Rocha</u> mudava de <u>calças</u> por detrás do <u>carro</u> estacionado. Era a <u>hora</u> do equinócio e o <u>calor</u> ia alagando as <u>carnes</u> de suor naquele <u>começo</u> de <u>Agosto</u>. Os dois <u>rapazes</u> esperavam a <u>passagem</u> da <u>camioneta</u> que transportaria <u>mais três</u> <u>companheiros</u> de <u>Braga</u>. Por terem <u>atestado</u> o <u>carro</u> com o <u>material</u> de <u>campismo</u>, não couberam todos.</p>

Lista dos nomes:

<u>estrada</u>	<u>calor</u>
<u>curvas</u>	<u>carnes</u>
<u>sombra</u>	<u>suor</u>
<u>árvores</u>	<u>começo</u>
<u>Rubro</u>	<u>rapazes</u>
<u>paisagem</u>	<u>passagem</u>
<u>fundo</u>	<u>camioneta</u>
<u>Rocha</u>	<u>companheiros</u>
<u>calças</u>	<u>Braga</u>
<u>carro</u>	<u>carro</u>
<u>hora</u>	<u>material</u>
<u>equinócio</u>	<u>campismo</u>

- Para os verbos :

Podemos seleccionar só os verbos sem dar nenhuma informação:

Texto anotado <u>sem</u> revisão (todas as palavras que podem ser verbos)	Texto anotado <u>com</u> revisão (só os verbos)
A estrada <u>era</u> famosa <u>pelas curvas</u> e <u>pela</u> sombra das árvores. O Rubro <u>admirava</u> a paisagem que se <u>espraiava</u> <u>para</u> o <u>fundo</u> , enquanto o Rocha <u>mudava</u> de <u>calças</u> por detrás do carro <u>estacionado</u> . <u>Era</u> a <u>hora</u> do equinócio e o calor <u>ia alagando</u> as <u>carnes</u> de suor naquele <u>começo</u> de Agosto. Os dois rapazes <u>esperavam</u> a passagem da camioneta que <u>transportaria</u> mais três companheiros de Braga. Por <u>terem</u> <u>atestado</u> o carro com o material de campismo, não <u>couberam</u> todos.	A estrada <u>era</u> famosa pelas curvas e pela sombra das árvores. O Rubro <u>admirava</u> a paisagem que se <u>espraiava</u> para o fundo, enquanto o Rocha <u>mudava</u> de calças por detrás do carro estacionado. <u>Era</u> a hora do equinócio e o calor <u>ia alagando</u> as carnes de suor naquele começo de Agosto. Os dois rapazes <u>esperavam</u> a passagem da camioneta que <u>transportaria</u> mais três companheiros de Braga. Por <u>terem</u> <u>atestado</u> o carro com o material de campismo, não <u>couberam</u> todos.

Lista dos verbos:

<u>era</u> <u>admirava</u> <u>espraiava</u> <u>mudava</u> <u>ia</u> <u>alagando</u> <u>esperavam</u> <u>transportaria</u> <u>terem</u> <u>atestado</u> <u>couberam</u>
--

Podemos seleccionar só os verbos e dar o lema de cada verbo:

<p>Texto anotado <u>sem</u> revisão (todas as palavras que podem ser verbos)</p>	<p>Texto anotado <u>com</u> revisão (os verbos com os lemas)</p>
<p>A <u>estrada</u> <u>era</u> famosa <u>pelas</u> <u>curvas</u> e <u>pela</u> sombra das árvores. O Rubro <u>admirava</u> a paisagem que se <u>espraiava</u> <u>para</u> o <u>fundo</u>, enquanto o Rocha <u>mudava</u> de <u>calças</u> por detrás do carro <u>estacionado</u>. <u>Era</u> a <u>hora</u> do equinócio e o calor <u>ia</u> <u>alagando</u> as <u>carnes</u> de suor naquele <u>começo</u> de Agosto. Os dois rapazes <u>esperavam</u> a passagem da camioneta que <u>transportaria</u> mais três companheiros de Braga. Por <u>terem</u> <u>atestado</u> o carro com o material de campismo, não <u>couberam</u> todos.</p>	<p>A estrada <u>era</u> (ser) famosa pelas curvas e pela sombra das árvores. O Rubro <u>admirava</u> (admirar) a paisagem que se <u>espraiava</u> (espraiar) para o fundo, enquanto o Rocha <u>mudava</u> (mudar) de calças por detrás do carro estacionado. <u>Era</u> (ser) a hora do equinócio e o calor <u>ia</u> (ir) <u>alagando</u> (alagar) as carnes de suor naquele começo de Agosto. Os dois rapazes <u>esperavam</u> (esperar) a passagem da camioneta que <u>transportaria</u> (transportar) mais três companheiros de Braga. Por <u>terem</u> (ter) <u>atestado</u> (atestar) o carro com o material de campismo, não <u>couberam</u> (caber) todos.</p>

Lista dos verbos com os lemas :

<p><u>era</u> (ser) <u>admirava</u> (admirar) <u>espraiava</u> (espraiar) <u>mudava</u> (mudar) <u>ia</u> (ir) <u>alagando</u> (alagar) <u>esperavam</u> (esperar) <u>transportaria</u> (transportar) <u>terem</u> (ter) <u>atestado</u> (atestar) <u>couberam</u> (caber)</p>
--

Podemos seleccionar só os verbos e dar o tempo e o modo de cada verbo:

<p>Texto anotado <u>sem</u> revisão (todas as palavras que podem ser verbos)</p>	<p>Texto anotado <u>com</u> revisão (os verbos com o tempo e o modo)</p>
<p>A estrada <u>era</u> famosa <u>pelas</u> <u>curvas</u> e <u>pela</u> sombra das árvores. O Rubro <u>admirava</u> a paisagem que se <u>espraiava</u> <u>para</u> o <u>fundo</u>, enquanto o Rocha <u>mudava</u> de <u>calças</u> por detrás do carro <u>estacionado</u>. <u>Era</u> a <u>hora</u> do equinócio e o calor <u>ia</u> <u>alagando</u> as <u>carnes</u> de suor naquele <u>começo</u> de Agosto. Os dois rapazes <u>esperavam</u> a passagem da camioneta que <u>transportaria</u> mais três companheiros de Braga. Por <u>terem</u> <u>atestado</u> o carro com o material de campismo, não <u>couberam</u> todos.</p>	<p>A estrada <u>era</u> (pretérito imperfeito do indicativo) famosa pelas curvas e pela sombra das árvores. O Rubro <u>admirava</u> (pretérito imperfeito do indicativo) a paisagem que se <u>espraiava</u> (pretérito imperfeito do indicativo) para o fundo, enquanto o Rocha <u>mudava</u> (pretérito imperfeito do indicativo) de calças por detrás do carro estacionado. <u>Era</u> (pretérito imperfeito do indicativo) a hora do equinócio e o calor <u>ia</u> (pretérito imperfeito do indicativo) <u>alagando</u> (gerúndio) as carnes de suor naquele começo de Agosto. Os dois rapazes <u>esperavam</u> (pretérito imperfeito do indicativo) a passagem da camioneta que <u>transportaria</u> (presente do condicional) mais três companheiros de Braga. Por <u>terem</u> (pretérito perfeito do indicativo) <u>atestado</u> (particípio passado) o carro com o material de campismo, não <u>couberam</u> (pretérito perfeito do indicativo) todos.</p>

Lista dos verbos com o tempo e o modo de cada verbo:

<u>era</u> (pretérito imperfeito do indicativo)
<u>admirava</u> (pretérito imperfeito do indicativo)
<u>espraiava</u> (pretérito imperfeito do indicativo)
<u>mudava</u> (pretérito imperfeito do indicativo)
<u>ia</u> (pretérito imperfeito do indicativo)
<u>alagando</u> (gerúndio)
<u>esperavam</u> (pretérito imperfeito do indicativo)
<u>transportaria</u> (presente do condicional)
<u>terem</u> (pretérito perfeito do indicativo)
<u>atestado</u> (particípio passado)
<u>couberam</u> (pretérito perfeito do indicativo)

2. Fazer uma interface muito mais simples e muito mais fácil para o utilizador

Esta interface, é um esboço, para ter uma ideia da simplificação :

The image shows a wireframe sketch of a user interface. It includes the following elements:

- An input field at the top left.
- A checkbox labeled "Notificar acerca da fase de processamento" which is checked.
- A file selection input field with a "Browse..." button.
- A list of checkboxes for processing options, all of which are checked:
 - Palavras compostas
 - Siglas
 - Nomes
 - Verbos
 - Verbos (com a informação do lema)
 - Verbos (com a informação do tempo e do modo)
- A dark horizontal bar at the bottom containing the text "Contacto com o utilizador ?".
- Two buttons at the very bottom: "Enviar Pedido" and "Limpar Formulário".

Agradecimento

O AnELL foi (parcialmente) financiado pela Fundação para a Ciência e Tecnologia, co-financiada pelo POSI, através do projecto POSI/PLP/43931/2001 (Linguateca).