Morfolimpíadas: resumo

- Participação gratuita e voluntária de sete grupos diferentes
- Mapa da discordância considerável em termos de opiniões
- Medidas (que nós saibamos, as primeiras a nível mundial) de diferenças de atomização em relação ao português
- Medidas de concordância em termos de forma; forma + PoS; forma + lema; forma + PoS + lema
- Criação de um conjunto de formas revistas com a classificação "certa"
  - mapa dos vários problemas que esta "simples" tarefa envolve
- Criação de um conjunto de programas que transformam a saída de todos os sistemas num mesmo formato comum

Morfolimpíadas: problemas

- As respostas "certas" eram muito duvidosas: fartámo-nos de reescrever a documentação, tirar e adicionar classificações
- Não foi possível pesar as formas da lista dourada em termos de variante (deveria ser equilibrada)
- Idealmente, fenómenos diferentes deviam ser pesados com factores diferentes:falta de um género devia ser menos penalizada do que falta de uma classificação gramatical; proximidade nome/adjectivo mais previsível do que verbo/adjectivo … medidas: perplexidade, entropia
- Medidas todas feitas em termos de comparação entre **análises**
- Formas deviam ser pesadas em termos da sua frequência

Morfolimpíadas: problemas (cont.)

- Não fizemos comparação característica a característica
- Não fizemos comparação ou avaliação de palavras compostas
- Várias coisas a que nos tínhamos proposto não ficaram acabadas a tempo:
  - várias partes da lista dourada (problemas de atomização; hífens; palavras representativas das variantes; lista dourada do ensaio; casos específicos)
  - comparação com a análise certa em contexto (corpora anotados)
  - tratamento da derivação
- A partir de um certo ponto não conseguimos harmonizar (domadores foram mais castradores)
- Demasiada harmonização deita fora exactamente aquilo em que os sistemas podem ser diferentes (e melhores!)

Morfolimpíadas: opções

- Usar precisão como medida única
- Não considerar análises sublexicais
- Não aceitar que formas sistematicamente homónimas pesassem mais
  - reduzir INF, INFP 1 S e INFP 3 S a um
  - reduzir PSI_I 1 S e 3 S a um
  - reduzir PSM_I e PSP_I 3 P a um
  - reduzir número P + número S a um
- Usar verbos com clíticos como UMA forma
- Usar contracções como UMA forma
- Não considerar mais do que uma interpretação gramatical (GRAM)
- Não comparar número e género e caso de palavras gramaticais; considerar o lema como igual à forma

Morfolimpíadas: resultados

- Atomização

  apenas à volta de 80% das unidades são comuns
- Atomização e categoria gramatical

  por volta de 52% de concordância
- Atomização e lema

  por volta de 50% de concordância
- Atomização, categoria gramatical e lema

  por volta de 43% de concordância

Morfolimpíadas: comparação com a dourada

- Relativa ao número de formas da lista dourada também reconhecidas pelo sistema, ou absoluta (todas as formas)?
- Várias opções:
  - tirar INTERJ
  - tirar particípios passados
  - tirar palavras pertencentes à "outra" variante
  - tirar análises raras
- Todos os sistemas entre 86% e 76% de precisão (sem lema nem outros)
- Todos os sistemas entre 94% e 84% de precisão (só PoS)
- Os sistemas variam entre 90% e 78% de precisão (só lema)

Comparação com a dourada

- Tirar análises raras piora o desempenho a <u>todos</u> os sistemas: diminui 4-3% a precisão, aumenta 2-1% a cobertura
- Em termos de variante codificada na lista dourada, alguns sistemas melhoram com uma das variantes (3-2% melhor)
- Um sistema foi consistentemente melhor em todas as categorias -> daí proclamado o vencedor

embora possa haver outras medidas, outros processamentos, que venham a dar outros resultados