

Para inglês ver:
Questões fundamentais e uma abordagem nova

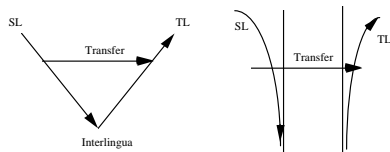
Diana Santos

Linguateca
www.linguateca.pt

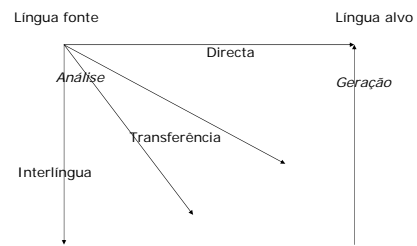
Reforçando algumas opiniões

- O meu funil como alternativa ao triângulo de Vauquois
- O problema da falta de conhecimento da língua dos intervenientes, e da relação com a tradução humana
- O problema do BLEU e do seu mau uso (“entra lixo, sai lixo”)
- A ligação com a localização: ou o que é aceitável numa língua é inaceitável noutra

Triângulo ou funil? (Santos, 1998)



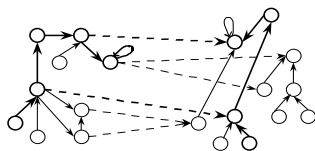
Uma visão gráfica



Primeira Escola de Verão da Linguateca

O modelo da rede de tradução

- Descrição independente das duas línguas
- Pontes entre diferentes categorias



Primeira Escola de Verão da Linguateca

A grande maioria das pessoas que fazem TA

- não têm respeito pelos tradutores
- não sabem nem querem saber das diferenças entre as línguas
- muitas vezes ficam influenciadas pela “má tradução” que produzem, embotando o seu sentido crítico
- muitas vezes imaginam que a tradução está correcta mas estão a traduzir de uma língua fonte inventada/imaginada por eles
- são demasiado transigentes com o sistema
 - Já faz tanta coisa bem!
 - Têm duas bitolas (para TA, isto chega)

A localização é tanto ou mais difícil que a tradução

- Toda a gente pensa que a sabe e pode fazer, e o resultado são coisas como esta:
- Recensões “automáticas”: tradução de números para inglês, mas inaceitável em Portugal
 - Literature awareness Ignorant of a major relevant piece of work
 - Scientific/technical soundness Major mistakes
 - Significance of results No one will ever read this paper
 - Clarity of presentation Some problems with grammar or structure
 - Overall rating Reject
 - Reviewer confidence I'm very familiar with this area

Formas de ver a TA (Ward & Jurafsky, 2000)

- Interessadas no processo: como fazer a tradução (directa, transf, interl)
- Interessadas no resultado: quais as medidas a maximizar, dados os requisitos (estatística)
 - $\text{argmax}_T (\text{fluência } (T) \times \text{fidelidade } (S,T))$
- Maior respeito pela tradução humana?

Primeira Escola de Verão da Linguateca

O factor humano na TA

- As pessoas não gostam de ajudar o computador
 - tradução interactiva não funciona
- As pessoas gostam de corrigir o computador
 - pós-edição e conseqüente responsabilidade final

Primeira Escola de Verão da Linguateca

BLEU (Bilingual Evaluation Understudy) (Papineni et al, 2001)

- using n -gram similarity of a candidate to a set of reference translations (sentence based)
- modified precision of a candidate translation:
 - number of clipped words (n-grams) that occur in any reference transl. / number of total words (n-grams) in the candidate
 - sum of clipped n-grams in all sentences / sum of candidate n-grams
- word-weighted average of sentence-level modified precisions, rather than a sentence-weighted average
- combination of the modified precisions of 1 to 4 grams
- sentence-brevity penalty

ESLLI 2007

Example from Papineni et al

Candidate 1: It is a guide to action which ensures that the military always obeys the commands of the party.

Candidate 2: It is to insure the troops forever hearing the activity guidebook that party direct.

P1=17/18
P2=5/18

Reference 1: It is a guide to action that ensures that the military will forever heed Party commands.

Reference 2: It is the guiding principle which guarantees the military forces always being under the command of the Party.

Reference 3: It is the practical guide for the army always to heed the directions of the party.

ESLLI 2007

BLEU formulas

$$\text{BLEU} = \text{BP} \cdot \exp\left(\sum_{n=1}^N w_n \log p_n\right) \quad \text{BP} = \begin{cases} 1 & \text{if } c > r \\ e^{(1-r/c)} & \text{if } c \leq r \end{cases}$$

$$\log \text{BLEU} = \min\left(1 - \frac{r}{c}, 0\right) + \sum_{n=1}^N w_n \log p_n$$

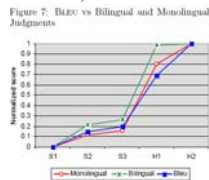
- c, r – length of the candidate or reference translations
- As a baseline, Papineni et al suggest:
 - w_n – uniform weights: $1/N$
 - $N = 4$
- Note that the matches are position independent.

ESLLI 2007

More on BLEU

ESLLI 2007

- Proposed for use in the R&D cycle of machine translation technology
- The more reference translations, the higher the precision
- Even a human translator will hardly score 1 (except if s/he produces a translation equal to one of the reference translations)
- experiments to judge 5 “systems”:
 - 250 Chinese-English sentence pairs = 5 translations of 50 sentences
 - rated by two groups of human judges
 - from 1 (very bad) to 5 (very good)
 - 10 bilinguals and 10 monolinguals
 - linearly normalized by the range



SINTEF

Information and Communication Technologies

13

Crítica do BLEU como é usado agora

- Em vez de várias traduções humanas da MESMA frase...
...várias traduções, por exemplo do EuroParl, que podem não ser boas!
- O BLEU não é uma medida, mas sim uma família de medidas (Lepage & Denoual, 2005). Só com um conjunto de traduções é que o BLEU é uma medida.
- Todos fazem, então há-de estar bem...
- Exemplo da Microsoft no PROPOR 2008: realimentação positiva!
- Stephan Oepen (em conversa): um caso onde a insistência na avaliação do tipo DARPA foi péssima para a comunidade (de TA)

SINTEF

Information and Communication Technologies

14

“O sistema baseado em exemplos mais puro”

- Lepage, Yves & Etienne Denoual. “Purest ever example-based machine translation: Detailed presentation and assessment”, *Machine Translation*, Volume 19, Numbers 3-4 / December, 2005, pp. 251-282.
- Baseado em analogia proporcional
- A:B :: C:D

SINTEF

Information and Communication Technologies

15

Exemplos

- It walks across the street. :: It floats across the river.
: It walked across the street. : It floated across the river.
- I'd like to open these windows. :: I'd like to cash these traveler's checks.
: Could you open a window? : Could you cash a traveler's check?
- Est-ce que ces fenêtres-là, je peux les ouvrir?* :: *Ces chèques de voyage-là, je peux les échanger?*
:Est-ce que vous pouvez m'ouvrir une fenêtre? : *Vous pouvez m'échanger un chèque de voyage?*

SINTEF

Information and Communication Technologies

16

Bibliografia

- Papineni, Kishore, Salim Roukos, Todd Ward & Wei-Jing Zhuw. “BLEU: a Method for Automatic Evaluation of Machine Translation”, Research Report, Computer Science IBM Research Division, T.J.Watson Research Center, RC22176 (W0109-022), 17 September 2001.
- Santos, Diana. “Punctuation and multilinguality: Reflections from a language engineering perspective”. *Working Papers in Applied Linguistics* 4/98, redigert av Jo Terje Ydstie og Anne C. Wollebæk. Oslo: Department of Linguistics, Faculty of Arts, University of Oslo, 1998, pp. 138-60.
- Santos, Diana. “The translation network: A model for the fine-grained description of translations”, in Jean Véronis (ed.), *Parallel Text Processing*, Dordrecht: Kluwer Academic Publishers, 2000, pp.169-186.
- Ward, Nigel & Daniel Jurafsky. “Machine Translation”, in Daniel Jurafsky and James H. Martin (eds.), *Speech and Language Processing*, 1st Edition, Prentice-Hall, 2000, pp. 720–751.

SINTEF

Information and Communication Technologies

17