

ANALISI SISTEMI DI TRADUZIONE AUTOMATICA "COMMERCIALI"

Esplorazione

- (1) Collegarsi a siti web di news ed iniziare a tradurre alcuni titoli di notizie (scegliete pure le pagine di giornale che volete, ma ricordate di tradurre VERSO la vostra lingua madre! Ad es. gli italiani dovranno collegarsi a siti stranieri e tradurre i titoli dalla lingua in cui sono scritti all'italiano; inserire i dati raccolti nella seguente pagina web: <http://www.ciscl.unisi.it/progetti/news/>)

Cercate di prevedere quali titoli/traduzioni causeranno difficoltà e quali no... riuscite ad intuirne il perchè?

- (2) Collegarsi ai seguenti siti web e ripetere il processo di traduzione automaticamente, quindi confrontare le traduzioni automatiche con le vostre:

<http://www.systransoft.com/index.html> (Systran)
<http://babelfish.altavista.com/> (Babel-fish)
http://www.google.com/language_tools?hl=it (Google)
http://www.worldlingo.com/en/products_services/worldlingo_translator.html (WordLingo)

<http://www.freetranslation.com/> (Free Translation)

<http://www.tranexp.com:2000/Translate/result.shtml> (InterTran)

<http://www-306.ibm.com/software/pervasive/tech/demos/translation.shtml> (IBM WebSphere)
 le pagine web da indicare al traduttore sono:
<http://www.ciscl.unisi.it/progetti/news/it.htm> (per i titoli italiani);
 ... /en.htm (per i titoli inglesi);
 ... /fr.htm (per i titoli francesi);
 ... /sr.htm (per i titoli sardi)

- (3) Le vostre previsioni su quali titoli avrebbero causato difficoltà e quali no erano corrette? Riuscite ad intuirne il perchè?

Valutazione

- (4) **Soggettiva** (per ogni definizione tradotta valutare l'output prodotto sulle tre scale):

	intelligibilità (chiarezza)	informatività (contenuto)	stile (appropriatezza)
Systran	0 1 2 3 4 5 6 7 8 9 10	0 1 2 3 4 5 6 7 8 9 10	0 1 2 3 4 5 6 7 8 9 10
Free Translation	0 1 2 3 4 5 6 7 8 9 10	0 1 2 3 4 5 6 7 8 9 10	0 1 2 3 4 5 6 7 8 9 10
InterTran	0 1 2 3 4 5 6 7 8 9 10	0 1 2 3 4 5 6 7 8 9 10	0 1 2 3 4 5 6 7 8 9 10
IBM WebSphere	0 1 2 3 4 5 6 7 8 9 10	0 1 2 3 4 5 6 7 8 9 10	0 1 2 3 4 5 6 7 8 9 10

- (5) **Oggettiva** (numero di *cancellazioni*, *aggiunte*, *sostituzioni*, *riarrangiamenti* necessari per trasformare un'espressione prodotta in una forma grammaticalmente corretta e semanticamente coerente con l'espressione sorgente)

- (6) **Linguistica** (valutare la presenza/assenza, difficoltà presentate dalle seguenti proprietà):

- ripetizioni (la casa sulla collina non è bella come la villa sul colle)
- espressioni pronominali/referenziali (Luigi chiede a Luisa di fargli vedere una sua foto)
- NP complessi (Il castello di sabbia che era stato costruito dall'accusa)
- dipendenze a distanza (la cartolina che Luca ha scritto è stata consegnata)
- forme irregolari (cadde, cadò...)
- morfologia derivazionale (mangiafuoco, inconsistentemente... aereoplanino)
- espressioni idiomatiche (tirare il calzino, mangiare la foglia...)
- giochi di parole (Amnistia: colpo di grazia)
- ambiguità
 - o lessicali (la vecchia legge la regola)
 - o sintattiche (ho visto il ragazzo nel parco con il cannocchiale)
 - o semantiche (la pesca non è stata fruttuosa)
- frasi dichiarative, interrogative, imperative
- consecutio temporum
- espressioni telegrafiche, contratte

Reverse Engineering

- (7) Tentare di riflettere su come si potrebbero manifestare le seguenti proprietà:
- reversibilità / flessibilità
 - modularità: modulo morfologico, parsing profondo Vs. shallow parsing, analisi semantica; componente di generazione
 - robustezza agli errori (ortografici, sintattici, "semantic")
 - Interlingua Vs. transfer Vs. word-by-word

Riferimenti

Hutchins & Somers (1992) Cap. 9

Gli indirizzi web dei traduttori automatici cambiano frequentemente (e spesso i siti vengono chiusi), verificare con un motore di ricerca se l'URL è corretta.