

Informàtica Aplicada a la Traducció  
Examen de prova  
14 de desembre de 2002  
Segona part (1 hora i 40 min.)

**Instruccions:**

- En aquesta part de l'examen podeu consultar qualsevol material que hàgeu portat però no podeu compartir material ni parlar amb ningú.
- Si teniu algun dubte, alceu la mà i aniré a atendre-us.

**Qüestions:**

1. Indiqueu quina classe d'ambigüitat presenten aquestes frases (justifiqueu breument la vostra resposta):

- (a) "El mecánico revisó la suspensión del auto de Garzón" (1: Aquest mecànic és un expert en legislació i s'ha llegit la resolució judicial sencera; 2: Els amortidors del cotxe de Garzón ja necessitaven una revisió).
- (b) "Mi vecino aseguró que había bajado la bolsa" (1: La bossa d'escombraires ja és al contenidor; 2: Les meues accions de borsa no deixen de donar-me disgustos).
- (c) "A qui has dit que telefonarien?" (1: Per a qui seria la telefonada que has dit que farien? 2: A qui has dit que es produiria una telefonada?)
- (d) "A pesar de haber sido soldado, salió despedido del avión" (1: El sistema fotogràfic estava fortament fixat al fuselatge però es va soltar de l'aparell quan l'avió va girar en ple vol; 2: Malgrat el seu passat militar gloriós, el president el va destituir abans d'arribar a l'aeroport de destinació).
- (e) "Els lladres van ser atrapats a una fàbrica incendiada per un policia" (1: Els lladres van ser capturats pel comissari a una fàbrica abandonada; 2: La fàbrica on van ser capturats va ser l'objectiu d'un agent piròman).

- (f) "Coto privado de caza" (1: Aquesta àrea de caça no és pública; 2: Aquesta és una àrea sense caça).
- (g) "Vull ballar i cantar cançons de bressol" (1: Ballarem durant una estona i després et cantaré perquè et dormes; 2: M'estime tant ballar cançons de bressol com cantar-ne).
- (h) "El manifestant es refà de la pallissa que li van donar en l'hospital" (1: La manifestació va ser una mica violenta i algunes persones han hagut de ser traslladades a l'hospital; 2: Quina pallissa va donar-li l'infermera al quiròfan!).

2. Els buscadors d'internet permeten buscar documents que complisquen criteris més complexos que el fet d'incloure uns mots concrets. Així, alguns permeten indicar condicions de búsqueda com ara  $m_1$  near  $m_2$ , que busquen aquelles pàgines en què el mot  $m_1$  apareix a prop del mot  $m_2$ .

També n'hi ha que permeten restringir les adreces sobre les quals es farà la búsqueda (per exemple, afegint la condició `site:www.ua.es` indiquem que només volem buscar documents inclosos en el servidor de la Universitat d'Alacant).

Raoneu com es poden combinar aquestes possibilitats per facilitar la construcció d'un diccionari monolingüe de cites en què cada accepció s'acompanya amb un exemple d'ús pres de l'obra d'algun escriptor important. Com exemple inicial, podeu inspirar-vos en un terme (sense cites) polisèmic com:

**banco.** *m.* Asiento, con respaldo o sin él, en que pueden sentarse varias personas. || 2. Conjunto de peces que van juntos en gran número. || 3. Establecimiento público de crédito, constituido en sociedad por acciones.

Considereu també el cas dels termes homògrafs. Expliqueu-ho tot amb exemples en la llengua que vulgueu.

3. Imagineu que teniu un sistema de traducció automàtica que treballa amb dues llengües,

diguem-ne  $A$  i  $B$ , en els dos sentits de traducció:  $A \rightarrow B$  i  $B \rightarrow A$ , que traduïm un text origen  $T$  en llengua  $A$  a la llengua  $B$  mitjançant aquest traductor automàtic, generant un text  $T'$ , i que després usem aquest mateix sistema per a traduir  $T'$  de nou a la llengua  $A$ ; anomenarem  $T''$  el nou text en llengua  $A$ .

$$T \xrightarrow{A \rightarrow B} T' \xrightarrow{B \rightarrow A} T'' \quad (1)$$

El text  $T''$  serà previsiblement diferent del text  $T'$ . Trieu dues llengües  $A$  i  $B$  i indiqueu quins canvis són previsibles, classificant-los segons la naturalesa lingüística dels fenòmens que han causat els canvis, explicant la raó del resultat si cal amb un exemple. Heu d'indicar *tres tipus diferents* de canvi.

4. Imagineu que sou part d'un equip de desenvolupament d'un sistema de traducció automàtica de l'anglès al català basat en l'estratègia que hem anomenat a classe transferència morfològica avançada. Els informàtics del projecte us demanen consell sobre les regles de reordenament del sistema, ja que, per motius tècnics, només poden afegir-n'hi tres.

Indiqueu quines serien les 3 regles que proposaríeu, tenint en compte que han de produir, com a mínim, tres oracions ben traduïdes en el corpus d'oracions següents (la traducció ideal s'indica entre parèntesi, tot i que no sempre podrà ser aconseguida):

- (a) A dark autumn night (Una nit fosca de tardor)
- (b) A high tide (Una marea alta)
- (c) A magic dark silhouette (Una silueta fosca màgica)
- (d) An autumn tide (Una marea de tardor)
- (e) A dark magic silhouette (Una silueta màgica fosca)
- (f) A dark autumn high tide (Una marea alta de tardor fosca)
- (g) A dark night (Una nit fosca)

Deixeu de banda la concordança i centreu-vos només en els reordenaments. Assenyaleu quina seria la traducció del sistema per a totes les oracions anteriors usant el conjunt de regles que heu proposat.

5. Hi ha ambigüitats de tipus lèxic que poden ser sempre correctament resoltes després de fer una anàlisi morfològica. No obstant això, hi ha d'altres que només poden ser tractades si es fa una anàlisi sintàctica (tot i que l'ambigüitat siga de tipus lèxic) i fins i tot n'hi ha que requeririen una anàlisi semàntica per a resoldre-les.

Elegiu una llengua origen i una llengua meta (francès, anglès, alemany, català o espanyol) i poseu un exemple d'oració per a cada un dels tres casos anteriors, on siga necessari un determinat nivell d'anàlisi per tal de resoldre una ambigüitat i produir-ne la traducció correcta. Expliqueu quina informació usa el sistema en cada cas per a prendre una decisió.