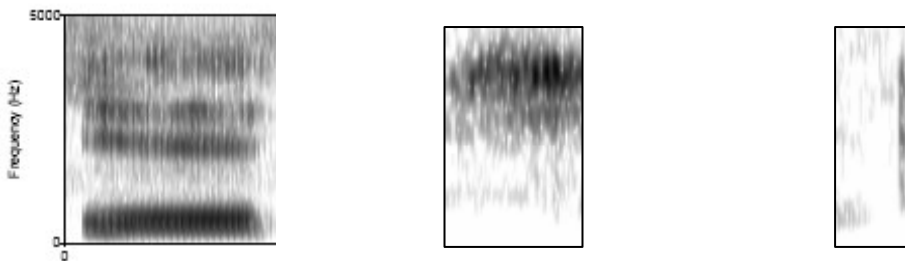


**Tentamen Taal- en Spraaktechnologie**  
 31 augustus 2011 13.00-15.00, Drift 23, zaal 1.07

1. [1 punt] Wat voor type klank (klinker, nasaal, stemhebbend/stemloze fricatief of plofklank, etc) staan in de onderstaande drie spectrogrammen, en waarom?



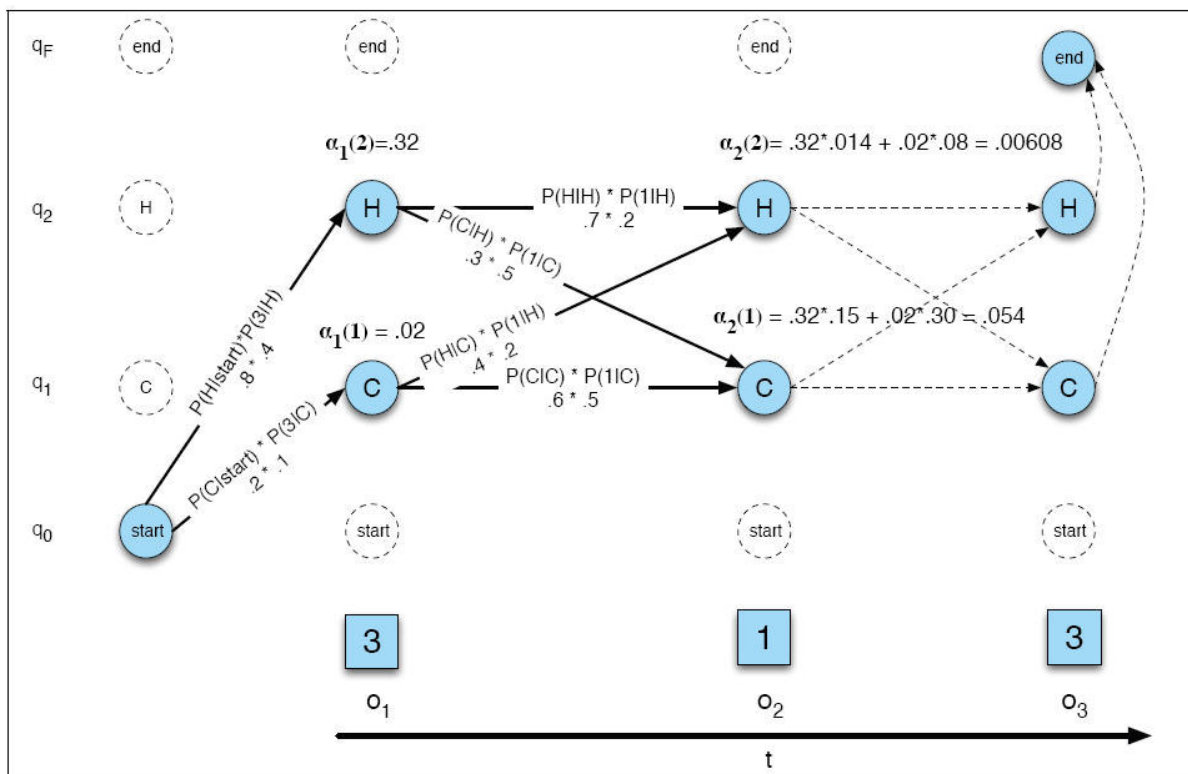
2. [1 punt] Leg op hoofdlijnen de werking van unit selection voor spraaksynthese uit.

3. [1 punt]. Met het Forward-Backward algoritme wordt op basis van observaties en een gegeven HMM een betere versie van de HMM (A en B) berekend. Daarvoor moet je onder andere het verwachte aantal overgangen van state i naar state j berekenen. Hoe doe je dat?

4. [2 punten] Hieronder staat het Viterbi netwerk voor ijsjes eten op warme en koude dagen. Er worden altijd 1, 2 of 3 ijsjes gegeten.

a) Wat zijn de gebruikte A en B matrices?

b) Wat is de meest waarschijnlijk volgorde van koude en warme dagen geweest?



5. [1 punt] Bereken de “cosine similarity” tussen *house* en *office* gegeven de volgende informatie:

	chair	table	roof
house	5	1	1
office	6	1	0

waar *chair*, *table*, *roof* verschillende dimensies zijn. Zijn *house* en *office* gelijk?

6. [1 punt] Welke knoop is de laagste gemeenschappelijke “subsumer” (klasse) voor de knopen Cbf3a en Cbf3c?

Cbf3 contains three proteins, Cbf3a, Cbf3b and Cbf3c.

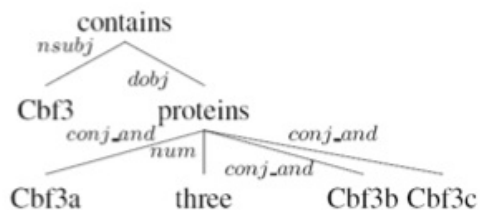


Figure 1: Stanford parser output

Representation: dependency paths				
Cbf3	$\xrightarrow{nsubj}$	contains	$\xleftarrow{dobj}$	proteins $\xrightarrow{conj\_and}$ Cbf3a
Cbf3	$\xrightarrow{nsubj}$	contains	$\xleftarrow{dobj}$	proteins $\xrightarrow{conj\_and}$ Cbf3b
Cbf3	$\xrightarrow{nsubj}$	contains	$\xleftarrow{dobj}$	proteins $\xrightarrow{conj\_and}$ Cbf3c

8. [1 punt] Hieronder staan enkele regulier expressies. Teken het overgangsdigram voor de bijhorende automaat: geef de toestanden, overgangen, begintoestand (aangeduid met een binnenkomend pijltje) en eindtoestand(en) (omcirkeld).

- $[[a,b]^+, b^*]$
- $\square$  (de taal die uitsluitend het lege rijtje bevat)
- $\{\}$  (de lege taal)

9. [2 punten] In deze opgave gebruik je de notatie van FSA Utilities die je ook in het practicum hebt gebruikt.

Maak een transducer die voor NL bijvoeglijke naamwoorden de (regelmatige) vergrotende trap vormt. Voorbeeld: *moeilijk* / *moeilijker*

Zorg dat je transducer zeker werkt voor de gegeven voorbeelden en de volgende invoer: *suf*, *wijs*, *boos*, *dor*, *zuur*, *blij*, *vers*, *lief*, *druk*, *ziek*, *vreemd*, *week*, *vlug*, *snel*, *slecht*, *rot*, *rood*, *raar*, *net*, *nat*, *echt*, *klein*, *solide*

#### Opmerkingen

- het regelmatige geval is toevoeging van de uitgang -er aan de onverbogen vorm van het adjectief
- je moet rekening houden met het onderscheid gesloten vs open lettergreep (*dom* / *dommer* vs *sloom* / *slomer*) en met de wisseling stemloos/stemhebbend (*doof* / *dover* vs *dof* / *doffer*)
- adjectieven die op -r uitgaan krijgen een ingevoegde -d- (*duur* / *duurder*)
- adjectieven die op een doffe -e uitgaan krijgen de uitgang -r (*stupid* / *stupider*)