

Semetrální zkouška CZR, 21.5.2004, skupina B

Login:

Podpis:

-
1. Jazyk UFONů má pouze 4 fonémy: G,R,F,M s těmito apriorními pravděpodobnostmi: $p(G) = 0.1$, $p(R) = 0.2$, $p(F) = 0.3$, $p(M) = 0.4$. Vypočtěte průměrnou míru informace za 1 vteřinu, víte-li, že UFOn za vteřinu vysloví průměrně 123 fonémů.

- A. 246 bitů.
- B. 227 bitů.
- C. 256 bitů.
- D. 492 bitů.

-
2. Signál před vzorkováním obsahoval tón o kmitočtu 5 kHz. Byl navzorkován na 8 kHz a *byl* použit antialiasingový filtr. Co bude obsahovat signál po rekonstrukci ze vzorků?

- A. Tón o kmitočtu 5 kHz
- B. Tón o kmitočtu 3 kHz
- C. Tón o kmitočtu 1 kHz
- D. Nic

-
3. Signál navzorkovaný na 8000 Hz má délku 1.34s. Kolik z něj vyrobíme celých rámců, pokud má být jejich délka 20 ms a překrytí 10 ms?

- A: 800
- B: 133
- C: 266
- D: 200

-
4. Ustředňujete signál pomocí on-line odhadu střední hodnoty: $\bar{s}[n] = \gamma\bar{s}[n-1] + (1-\gamma)s[n]$, kde $\gamma = 0.999$, ale je evidentní, že střední hodnota se velice rychle mění a odhad “nestíhá”. Jakou použijete hodnotu γ ?

- A. 0.99
- B. 1.0
- C. 1.1
- D. 0

-
5. LPC Filtr má přenosovou funkci

$$H(z) = \frac{1}{A(z)} = \frac{1}{1 + 1.1z^{-1} + 0.5z^{-2}}$$

Jeho póly leží v bodech:

- A. $-0.55 \pm j0.44$
- B. $0.44 \pm j0.55$
- C. $0.44 + j0.55, \quad 0.55 - j0.44$
- D. $0.44 - j0.55, \quad 0.55 + j0.44$

6. Rámec signálu má délku $N = 12$. Hodnoty signálu pro $n = 0 \dots 11$ jsou 1, 2, 3, 4, 3, 2, 1, 0, -1, -2, -3, -4. Jaká je hodnota prvního (a jediného) koeficientu a_1 prediktoru prvního řádu?

- A. 0.45
- B. -0.45
- C. 0.81
- D. -0.81

7. Při autokorelační metodě určení základního tóny bylo nalezeno maximum autokorelační funkce na indexu $k = 100$. Vzorkovací frekvence řeči byla 22.1 kHz. Jaký je základní tón řečníka?

- A. 100 Hz.
- B. 4.52×10^{-3} Hz
- C. 221 Hz.
- D. 110.5 Hz.

8. “Mřížka” lokálních vzdáleností DTW vypadá následovně (test vodorovně, reference svisle, začátek vlevo dole):

$$\begin{bmatrix} 5 & 3 & 2 \\ 4 & 1 & 1 \\ 1 & 3 & 2 \end{bmatrix}$$

DTW-vzdálenost je:

- A. 1
- B. 1.16
- C. 1.25
- D. 7

9. HMM má 4 stavy a matici přechodových pravděpodobností:

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0.6 & 0.4 & 0 \\ 0 & 0 & 0.7 & 0.3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Rozpoznávaná sekvence \mathbf{O} má 4 vektory, pro něž známe tyto vysílací pravděpodobnosti: $b_2[\mathbf{o}(1)] = 0.1$, $b_2[\mathbf{o}(2)] = 0.1$, $b_3[\mathbf{o}(3)] = 0.2$, $b_3[\mathbf{o}(4)] = 0.2$. Určete pravděpodobnost vyslání sekvence $\mathcal{P}(\mathbf{O}, X|M)$ po stavové sekvenci $X = [1 \ 2 \ 2 \ 3 \ 3 \ 4]$.

- A. 1×10^{-5}
- B. 2×10^{-5}
- C. 3×10^{-5}
- D. 4×10^{-5}

10. Pro stejný model a sekvenci 4 vektorů určete, která stavová sekvence není možná:

- A: $X = [1 \ 1 \ 2 \ 2 \ 3 \ 4]$
- B: $X = [1 \ 2 \ 2 \ 2 \ 3 \ 4]$
- C: $X = [1 \ 2 \ 2 \ 3 \ 3 \ 4]$
- D: $X = [1 \ 2 \ 3 \ 3 \ 3 \ 4]$

-
11. Určení základního tónu bude nejproblematictější u této hlásky:
- A. 'z'
 - B. 's'
 - C. 'ü'
 - D. 'a'
-
12. Co jsou to centi-sekundové vektory ?
- A. 1 sekunda řeči je parametrizována jedním vektorem obsahujícím 100 MFCC koeficientů.
 - B. Vektory parametrů, jejichž velikost byla z P koeficientů pomocí LDA uměle zvýšena na 100.
 - C. Vektory promluvy o délce 100 sekund.
 - D. Vektory parametrů, kterých je 100 za sekundu.
-
13. Při rozpoznávání metodou DTW dělíme výslednou kumulovanou vzdálenost normalizačním faktorem, který je u základní DTW: $R + T$, kde R je počet vektorů referenční sekvence a T je počet vektorů testovací sekvence. Toto dělení je nutné:
- A. abychom zabránili tomu, že bude testovací sekvence vždy rozpoznána jako jedno nejkratších slov ze slovníku.
 - B. abychom optimalizovali algoritmus DTW na rychlost.
 - C. abychom předešli numerickým chybám.
 - D. abychom optimalizovali paměťové nároky algoritmu DTW.
-
14. V systémech rozpoznávání založených na HMM, které mají pracovat s hlasem libovolného řečníka (speaker-independent), je modelování funkcí hustoty rozložení pravděpodobnosti ve stavech jednou Gaussovou nedostačující. Problém řešíme tak, že:
- A. Povolíme zpětné přechody v modelu (tj. přechodové pravděpodobnosti a_{ij} pod diagonálou nemusejí být nulové).
 - B. Přejdeme od Gaussovek k Newtonovu rozložení hustoty pravděpodobnosti, které má podstatně více parametrů.
 - C. V každém stavu použijeme směs Gaussovek (tzv. Gaussian mixture).
 - D. Omezíme trénovací data na signály pocházející od jednoho mluvčího.
-
15. Jak je nutné upravit algoritmus “token-passing” pro rozpoznávání spojených slov.
- A. Není možné používat logaritmické pravděpodobnosti.
 - B. Každý “token” musí obsahovat strukturu s informacemi o identitě slova, ze kterého byl “vrácen” do společného vstupního stavu všech modelů.
 - C. Namísto selekce jednoho nejlepšího “tokenu” v každém stavu je bezpodmínečně nutné zachovat a propagovat všechny.
 - D. Ve všech modelech je nutné upravit matici přechodových pravděpodobností tak, aby byla diagonální.
-
16. Úloha N -gramových modelů v rozpoznávání spojitě- řeči je:
- A. Generovat pravděpodobnost slova, známe-li parametrizaci.
 - B. Generovat pravděpodobnosti cizích slov.
 - C. Zajistit vynulování přechodových pravděpodobností a_{ij} .
 - D. Generovat apriorní pravděpodobnost slova podmíněného historií jiných slov.
-
17. maxima frekvenčních charakteristik filtrů při Mel-frekvenční analýze:
- A. jsou rozložena lineárně na Melové ose.
 - B. odpovídají pozici každého n -tého koeficientu FFT, kde $n = 5$.
 - C. jsou rozložena lineárně na standardní frekvenční ose.
 - D. jsou rozložena nelineárně na Melové ose.
-

18. Jak se liší lineární diskriminační analýza (LDA) od principal component analysis (PCA):
- A. jednotlivé koeficienty ve vektorech parametrů jsou nejen dekorelovány, ale mají navíc vždy všechny stejnou hodnotu.
 - B. LDA zajišťuje, že jednotlivé koeficienty budou seřazeny podle podílu stejnosměrné složky.
 - C. Na rozdíl od PCA upravuje LDA data pouze pro HMM s úplnými kovariančními maticemi.
 - D. LDA zajišťuje, že jednotlivé koeficienty budou seřazeny podle toho, jak přispívají k separaci tříd, které chceme rozeznávat.
-
19. Který blok syntezátoru řeči je zodpovědný za správný převod čísla 123,56 na “sto dvacet tři celé padesát šest setin”
- A. Normalizace textu.
 - B. Převod grafémy → fonémy.
 - C. Generování prozodie.
 - D. Výběr řečových jednotek z databáze.
-
20. Jaký je princip syntézy signálu metodou OLA (overlap and add)
- A. Z jednotky určené k syntéze se vypočtou LPC koeficienty, ty se sečtou, výsledný zvuk je generován LPC filtrem.
 - B. Jednotky určená k syntéze se rozdělí na rámce vážené oknem, tyto rámce se mohou v čase posunout, pak se sečtou a vznikne výsledný zvuk.
 - C. Jednotky určená k syntéze se rozdělí na rámce, ty se všechny sečtou a přehraje se výsledný zvuk.
 - D. Jednotky určená k syntéze se rozdělí na rámce vážené oknem, v každém rámci se sečtou všechny vzorky, výsledné hodnoty (1 na rámec) se opět sečtou.