

Semestrální zkouška CZR/ZRE, řádný termín, 11.5.2006, skupina A

Login:

Podpis:

1. Při kódování řeči máme vektorově zakódovat vektory o velikosti 12 koeficientů, máme k dispozici 24 bitů. Předpokládejme, že srovnání P -prvkového vstupního vektoru s P -prvkovým kódovým vektorem nás stojí P operací. Náročnost kódování se jednou kódovou knihou je tedy $12 \times 2^{24} = 201 \times 10^6$ operací. Kolikrát méně operací budeme provádět, pokud použijeme split-VQ:

- první kódová kniha bude kódovat 6 koeficientů a alokujeme pro ni 12 bitů.
- druhá kódová kniha bude kódovat 4 koeficientů a alokujeme pro ni 10 bitů.
- třetí kódová kniha bude kódovat 2 koeficienty a alokujeme pro ni 2 bity.

- A. asi $11000 \times$ méně.
B. asi $9000 \times$ méně.
C. asi $7000 \times$ méně.
D. asi $4000 \times$ méně.

2. K výpočtu DTW máte k dispozici neúplnou mřížku lokálních vzdáleností \mathbf{D} a neúplnou mřížku částečných kumulovaných vzdáleností \mathbf{G} (bez hodnot ∞ v nultém řádku a nultém sloupci a hodnoty 0 vlevo dole).

		D				G		
reference		10	5	2				
						20	15	12
						16	10	12
						8	13	5
		test				test		

Doplňte poslední řádek mřížky \mathbf{G} :

- A. 30 20 14
B. 30 25 14
C. 30 20 19
D. 30 25 19

3. Je dán signál: cosínusovka o normované frekvenci $\frac{1}{4}$ se stejnosměrnou složkou:

$$x[n] = 0.5 + \cos\left(2\pi \frac{1}{4}n\right)$$

Spočítejte energii takového signálu vztaženou na jeden vzorek.

- A. $E=0.75$ B. $E=1.0$ C. $E = \frac{1}{\sqrt{2}}$ D. $E=4$

4. Při výpočtu koeficientů prediktoru pomocí Levinsona-Durbina vyšly koeficienty prediktoru druhého řádu takto:

$$a_1^{(2)} = -0.0429, \quad a_2^{(2)} = 0.0564$$

Ve třetím kroku vyšel $k_3 = -0.0636$. Vypočtete koeficienty prediktoru řádu 3.

- A. $a_1^{(3)} = -0.0465, \quad a_2^{(3)} = -0.0592, \quad a_3^{(3)} = -0.0636$
B. $a_1^{(3)} = -0.0592, \quad a_2^{(3)} = -0.0465, \quad a_3^{(3)} = -0.0636$
C. $a_1^{(3)} = 0.0592, \quad a_2^{(3)} = -0.0465, \quad a_3^{(3)} = -0.0636$
D. $a_1^{(3)} = -0.0465, \quad a_2^{(3)} = 0.0592, \quad a_3^{(3)} = -0.0636$

5. Při LPC syntéze vyšla energie syntetizovaného rámce 456. Požadovaná energie je ovšem 789 (obě energie jsou vztažené na jeden vzorek). Aby měl syntetizovaný rámec požadovanou energii, je nutné jej:

A: násobit $\frac{456}{789}$ B: násobit $\frac{789}{456}$ C: násobit $\sqrt{\frac{456}{789}}$ D: násobit $\sqrt{\frac{789}{456}}$

6. Filtr druhého řádu

$$\frac{1}{A(z)} = \frac{1}{1 - 0.0429z^{-1} + 0.0564z^{-2}}$$

má na vstupu jednotkový impuls. Jaký je výstupní vzorek $y[2]$ tohoto filtru ?

A: $y[2] = 0.0429$ B: $y[2] = -0.0564$ C: $y[2] = -0.0546$ D: $y[2] = 0.0546$

7. Skrytý Markovův model zpracovává 39-rozměrné vektory. Jeho vysílací funkce hustoty rozdělení pravděpodobnosti je určena jednou 39-rozměrnou Gaussovskou s parametry $\mu = [1, 1, \dots, 1]^T$ a $\sigma = [4, 4, \dots, 4]^T$ (všechny střední hodnoty a všechny směrodatné odchylky jsou tedy stejné, pro úsporu místa jsou vektory zapsány jako řádkové a pak transponovány). Jaká je hodnota této funkce hustoty pro vstupní vektor $\mathbf{o} = [0, 0, \dots, 0]^T$.

A: $b(\mathbf{o}) = 2.66 \times 10^{-39}$ B: $b(\mathbf{o}) = 2.66 \times 10^{-40}$ C: $b(\mathbf{o}) = 2.66 \times 10^{-41}$ D: $b(\mathbf{o}) = 2.66 \times 10^{-42}$

8. Skrytý Markovův model má 4 stavy, z toho 2 vysílací. Jeho matice logaritmičeských přechodových pravděpodobností je:

$$\mathbf{A} = \begin{bmatrix} -\infty & 0 & -\infty & -\infty \\ -\infty & -0.11 & -2.30 & -\infty \\ -\infty & -\infty & -0.22 & -1.60 \\ -\infty & -\infty & -\infty & -\infty \end{bmatrix}$$

Model má na vstupu sekvenci 3 vektorů, jsou pro ně spočítány následující logaritmičeské vysílací pravděpodobnosti:

	$\mathbf{o}(1)$	$\mathbf{o}(2)$	$\mathbf{o}(3)$
$\log b_2$	-3.38	-3.35	-3.22
$\log b_3$	-4.63	-6.91	-5.71

Určete logaritmičeskou Viterbiho pravděpodobnost vyslání sekvence modelem $\log P^*(O|M)$ pomocí algoritmu token passing.

- A: $\log P^*(O|M) = -16.45$
 B: $\log P^*(O|M) = -20.12$
 C: $\log P^*(O|M) = -24.45$
 D: $\log P^*(O|M) = -30.12$

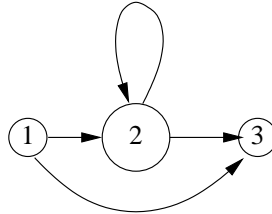
9. Pro de-korelaci příznakových vektorů je použita PCA (Principal component analysis). Pracujeme s dvouprvkovými vektory a PCA byla natrénována na těchto 4 vektorech:

$$\mathbf{o}_1 = \begin{bmatrix} -4 \\ -4 \end{bmatrix}, \quad \mathbf{o}_2 = \begin{bmatrix} -3 \\ -3 \end{bmatrix}, \quad \mathbf{o}_3 = \begin{bmatrix} 3 \\ 3 \end{bmatrix}, \quad \mathbf{o}_4 = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$$

Jaký bude první bázev vektor PCA (směr největší variability) ?

A: $\mathbf{pca}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ B: $\mathbf{pca}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ C: $\mathbf{pca}_1 = \begin{bmatrix} 0.707 \\ 0.707 \end{bmatrix}$ D: $\mathbf{pca}_1 = \begin{bmatrix} -0.707 \\ 0.707 \end{bmatrix}$

10. Model krátkého ticha v rozpoznávání spojitě řeči má následující topografii:



Při zpracování vstupní sekvence vektorů tento model:

- A: vždy zpracuje alespoň jeden vektor.
 - B: zpracuje maximálně 10 vektorů.
 - C: zpracuje libovolný nezáporný počet vektorů včetně 0 vektorů.
 - D: nezpracuje nikdy žádný vektor.
-

11. U dialogového systému je úschova informace o současném stavu dialogu úkolem pro:

- A: dialogový manager.
 - B: operační systém.
 - C: syntezátor řeči.
 - D: modul porozumění přirozenému jazyku
-

12. Při syntéze běžného českého ekonomického textu: *“Americká centrální banka podle očekávání zvýšila své úrokové sazby o 1/4 procentního bodu a to na 5%. Jedná se tak o 16. zvýšení úrokových sazeb v řadě.”* bude největším problémem:

- A: desambiguace homografů
 - B: prosodická analýza
 - C: výběr vhodných jednotek pro syntézu
 - D: normalizace textu.
-

13. Jedním z nereálných požadavků Lineární diskriminační analýzy, která slouží pro de-korelaci parametrů a omezení jejich počtu, je požadavek stejné kovarianční matice (tedy tvaru rozložení) všech tříd. Která třída zvuků nejvíce porušuje tento požadavek:

- A: všechny znělé hlásky
 - B: ticho
 - C: neznělé sykavky
 - D: dlouhé “r”.
-

14. Proč se při rozpoznávání pomocí kontextově závislých modelů fonémů sdílí stavy ?

- A: kvůli odstranění závislosti na jazyce.
- B: kvůli odstranění závislosti na mluvčím.
- C: kvůli nedostatku dat pro trénování v případě, že by stavy byly samostatné.
- D: kvůli přebytku dat pro trénování každého ze stavů, tento by mohl vést k přetrénování.

15. Při rozpoznávání spontánní řeči (meetingy, telefonní hovory) je největším problémem tvorba jazykového modelu pro spontánní řeč. Tento model je možné natrénovat:

- A: na prepisech spontánní řeči.
 - B: na románech (elektronické verze).
 - C: na sbírce textů (noviny, časopisy, atd.).
 - D: na teletextu.
-

16. Při výpočtu hodnoty funkce hustoty rozdělení pravděpodobnosti pro daný vstupní vektor vystupuje ve výpočtu hodnota

$$\frac{1}{\sqrt{(2\pi)^P}} \prod_{i=1}^P \frac{1}{\sigma_{ji}},$$

kde j je index stavu, P je délka vektoru parametrů a i je počítadlo jednotlivých parametrů.
Během rozpoznávání

- A: je tato hodnota konstantní, takže je možné ji před-počítat a uložit.
 - B: je tato hodnota závislá na vstupním vektoru, takže se nedá před-počítat.
 - C: je tato hodnota dynamicky měněna v závislosti na charakteru vstupního signálu.
 - D: je tato hodnota nevyčíslitelná a zanedbává se.
-

17. Pro generování slovního grafu (word-lattice), který udává i alternativní možnosti k nejlepší Viterbiho cestě, je nutná tato změna Viterbiho dekodéru:

- A: nahrazení Gaussovek konstantami.
 - B: propojení nejen koncových stavů slov zpět na začátek, ale všech stavů.
 - C: umožnění, aby ve stavech přežil více než jeden (nejlepší) token (pivo).
 - D: spuštění dekodéru s několika náhodně změněnými parametry Gaussovek.
-

18. U metody DTW je možné při vyplňování mřížky částečných kumulovaných vzdáleností \mathbf{G} zahájit “back-tracing” pro zjištění optimální srovnávací cesty:

- A: v případě, že při vyplňování dojde k přetečení akumulátoru
 - B: kdykoliv
 - C: nikdy
 - D: teprve až vyplníme celou mřížku
-

19. Jaký je důvod zavedení automatických metod hodnocení kvality kódování, které jsou inspirovány lidským slyšením (PSQM, PESQ):

- A: základní kritéria jako je SNR nebo SEGSNR je velmi složité vypočítat.
 - B: kodéry řeči nejsou primárně určeny pro lidské posluchače
 - C: organizovat poslechové testy je náročné, zdlouhavé a drahé
 - D: V programech MS-PowerPoint a Excel se výsledky PESQ a PSQM dají prezentovat lépe než výsledky poslechových testů.
-

20. Kompresce a expanze a-law a μ -law je ve skutečných implementovaných kodérech řeči technicky realizována

- A: rychlým výpočtem funkcí log a exp.
- B: převodem do spektra pomocí FFT.
- C: aproximací po částech lineární funkcí
- D: vyhledávací tabulkou (look-up table)