

## Semestrální zkouška ZRE, řádný termín, 14.5.2014, skupina $\xi$

Login: ..... Příjmení a jméno: ..... Podpis: .....  
(prosím čitelně!)

1. Co se ve zpracování řeči dá dělat s filtrem  $\frac{1}{A(z)}$ , jehož koeficienty  $a_i$  jsou určeny pomocí lineární predikce ?

- 
2. Jaký je princip odvození koeficientů lineární predikce (nepište plné rovnice, spíše postup).

- 
3. Koeficienty prediktoru LPC 2. řádu jsou  $a_2^{(2)} = -0.25$  a  $a_1^{(2)} = -0.25$ . Určete pomocí algoritmu Levinsona-Durbina koeficienty prediktoru 3. řádu, pokud víte, že  $k_3 = -0.2$ . Pokud Vám to pomůže, původní signál měl délku  $N = 5$  a vzorky  $1 \quad 0 \quad 1 \quad 1 \quad 0$ .

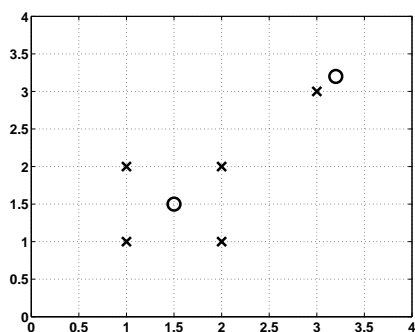
$$a_3^{(3)} = \dots\dots\dots$$

$$a_2^{(3)} = \dots\dots\dots$$

$$a_1^{(3)} = \dots\dots\dots$$

- 
4. Nakreslete schéma libovolné detekce základního tónu. Předpokládejte, že už máme k dispozici jeden rámeček řeši (neřešte předzpracování a dělení na rámce).

- 
5. Pracujeme s dvourozměrnými vektory. Pro trénování kódové knihy vektorového kvantování (VQ) o velikosti 2 máme k dispozici 5 trénovacích vektorů (křížky na obrázku). Kódové vektory jsou inicialisovány na hodnoty označené kolečky. Uveďte, jaké budou výsledné hodnoty kódových vektorů.



6. Tabulka uvádí hodnoty původního a kódovaného/dekódovaného signálu. Určete poměr signálu k šumu (SNR) v deciBellech.

$x[n]$	100	100	100	-100	-100
$\hat{x}[n]$	99	101	101	-99	-101

SNR = ..... dB

---

7. Jaký je v kódování řeči typu CELP (codebook excited linear prediction) rozdíl mezi open-loop pitch analysis a closed-loop pitch analysis ?
- 

8. Jaký je význam adaptivní a fixní kódové knihy v kódování řeči typu CELP (codebook excited linear prediction) ?
- 

9. Jak technika DTW řeší

- variabilitu v časování ?
  - variabilitu v prostoru parametrů ?
- 

10. Je dána matice (mřížka) lokálních vzdáleností DTW (reference svisle, test vodorovně, postup zleva nahoru). Vypočtete DTW vzdálenost (nezapomeňte na normalisaci!) a vyznačte optimální DTW cestu.

9	5
6	3
1	2

---

11. Jak se pro skrytý Markovův model a promluvu  $\mathbf{O}$  určí pravděpodobnost bytí ve stavu  $j$  v čase  $t$ :  $\gamma_j(t)$  ? Nemusíte psát plné rovnice, stačí ukázat graficky nebo popsat slovně. Pokud použijete hodnoty  $\alpha_j(t)$  a  $\beta_j(t)$ , napište, co znamenají.

12. Je dán skrytý Markovův model  $M$  se 4mi stavy (z toho 2 jsou vysílací), s přechodovými pravděpodobnostmi  $a_{12} = 1$ ,  $a_{22} = 0.6$ ,  $a_{23} = 0.4$ ,  $a_{33} = 0.7$ ,  $a_{34} = 0.3$ . Na vstupu je sekvence 5-ti vektorů. Hodnoty funkcí hustoty rozdělení pravděpodobnosti stavů 2 a 3 pro tyto vektory jsou:

	$\mathbf{o}(1)$	$\mathbf{o}(2)$	$\mathbf{o}(3)$	$\mathbf{o}(4)$	$\mathbf{o}(5)$
$b_2[\mathbf{o}(t)]$	0.4	0.3	0.1	0.1	0.1
$b_3[\mathbf{o}(t)]$	0.1	0.1	0.2	0.6	0.7

Určete hodnotu věrohodnosti (likelihood) této sekvence ohodnocené modelem po stavové sekvenci  $X = [1 \ 2 \ 2 \ 3 \ 3 \ 3 \ 4]$ . Stačí napsat všechny činitele, násobit nemusíte.

$$P(O, X|M) = \dots\dots\dots$$

13. Nakreslete rozpoznávací síť pro rozpoznávání sekvencí slov ANO nebo NE, s tím že slova jsou poskládána z modelů fonémů a předpokládáme, že mluvčí bude spíše opakovat to samé slovo — pravděpodobnosti bigramů  $P(\text{současné slovo}|\text{předcházející slovo})$  jsou:  
 $P(\text{ANO}|\text{ANO}) = 0.9$ ,  $P(\text{ANO}|\text{NE}) = 0.1$ ,  $P(\text{NE}|\text{ANO}) = 0.1$ ,  $P(\text{NE}|\text{NE}) = 0.9$ .

14. Vysvětlete, z čeho se kompiluje síť pro rozpoznávání řeči s velkým slovníkem.  
 Pomůcka:  $HCLG = H \circ C \circ L \circ G$ .

15. Co v rozpoznávání řeči označuje pojem “lattice” ?

16. Jaké charakteristiky hlasu podle Vás **lidé** používají pro určení mluvčího ? Ohodnoťte, jak snadno jdou technicky určit od 1 (“difficult”) po 10 (“easy”).

---

17. Co je to MAP-adaptace GMM a jak se používá v rozpoznávání mluvčího ?

---

18. Nakreslete DET křivku systému pro verifikaci mluvčího a vyznačte na ní, který pracovní bod byste doporučili pro aplikaci “hlasová autentizace přístupu k bankovnímu účtu”. Své doporučení zdůvodněte.

---

19. Který přístup pro identifikaci jazyka byste zvolili pro klasifikaci “Britská angličtina” vs. “Indická angličtina” a proč ?

---

20. Popište, na čem jsou trénovány různé části fonotaktického systému pro identifikaci jazyka (PRLM - Phone Recognizer followed by Language Model)

---