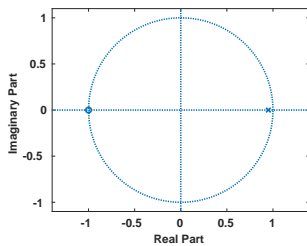


Semestrální zkouška ZRE, řádný termín, 23.5.2017, skupina 333

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

1. Na obrázku je vyznačen nulový bod a pól číslicového filtru $H(z)$. Nakreslete přibližně modul jeho kmitočtové charakteristiky od 0 do poloviny vzorkovací frekvence.



-
2. Uveďte, proč se při použití filtru $A(z)$ k tomu, abychom z řečového signálu $x[n]$ dostali budící signál hlasivek $e[n]$, hovoří o *lineární predikci*. Co se z čeho predikuje ?

-
3. Je dáno 5 vzorků signálu $x[n]$: [16 -8 4 -2 1]. Určete koeficient a_1 prediktoru prvního řádu (nejlépe bez výpočtu, z hlavy!). Pomůcka: $1 - A(z) = -a_1 z^{-1}$.

-
4. Běžný řád LPC filtru používaného v kodérech se vzorkovací frekvencí 8 kHz je $P = 10$. Proč zrovna tolik ?

-
5. Popište princip odhadu frekvence základního tónu (F_0) pomocí autokorelační funkce (ACF). Prosím opravdu jen základy, neuvažujte klipování, úpravu vstupního signálu, vyhlazování odhadu, atd.

6. Proč dochází u algoritmů odhadu frekvence základního tónu (F_0) v některých rámcích k chybě typu “double lag” ?

7. naimplementujte v jazyce C funkci pro výpočet poměru signálu k šumu. Parametr `xin` je vstupní signál, `xout` je signál po kódování a dekódování, `N` je počet vzorků.

```
float snr(float *xin, float *xout, int N) {
```

```
}
```

8. Proč je u metody diferenční pulsní kódové modulace (DPCM) nutné vestavět do kodéru celý dekodér ?

9. Co je u vektorového kvantování *kódový vektor* ?

10. Co je v současných kodérech řeči *analýza syntézou* ? Můžete nakreslit schéma a stručně napsat, co je který funkční blok.

11. V rozpoznávání řeči pomocí dynamického borcení času (DTW) má referenční sekvence R vektorů a testovací sekvence T vektorů. Srovnání probíhá po tzv. cestě o K krocích, která je dána “indexovacími” funkcemi $r[k]$ a $t[k]$. Napište některé podmínky, které pro $r[k]$ a $t[k]$ musí platit.

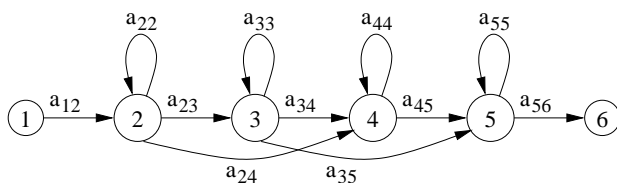
12. Mřížka lokálních vzdáleností u DTW je následující (reference svisle, test vodorovně). Dopačítejte mřížku částečných kumulovaných vzdáleností a určete DTW vzdálenost mezi testem a referencí. Nezapomeňte na normalizaci.

10	10	1
10	1	2
10	1	2
1	10	10

2		

13. Popište princip trénování modelu se směsí Gaussovek (Gaussian Mixture model) popisujícího vektorová data. Pokud chcete psát rovnice, data jsou $\mathbf{x}[1] \dots \mathbf{x}[T]$, ale můžete popsat i slovně.

14. Srytý Markovův model (HMM) na obrázku má reprezentovat promluvu o délce $T = 5$ vektorů. Napište všechny možné stavové sekvence X . Uvědomte si, že v každé sekvenci musí být stav č. 1 na začátku a stav č. 6 na konci. Tyto dva stavy nerepresentují žádný vektor.



15. Je definován levo-pravý HMM se čtyřmi stavy, z toho 2 vysílací, log. přechodové pravděpodobnosti jsou: $\log a_{12} = 0$, $\log a_{22} = -0.51$, $\log a_{23} = -0.92$, $\log a_{33} = -0.36$, $\log a_{34} = -1.2$. Tabulka logaritmu hodnot funkcí hustoty vysílacích pravděpodobností je:

t	...	46	47	48	...
$\log b_2(\mathbf{o}(t))$...	-1	-2	-3	...
$\log b_3(\mathbf{o}(t))$...	-4	-5	-6	...

Provádíme Viterbiho algoritmus pomocí “token passing”. Hodnota tokenu ve stavu 2 v čase 46 je $\Psi_2(46) = -26$. Určete hodnotu tokenu ve stavu 2 v čase 48: $\Psi_2(48) = \dots$

16. Jaké akustické jednotky jsou obvykle modelovány v systémech pro rozpoznávání řeči (ASR) s velkým slovníkem ?

17. Při rozpoznávání řeči pomocí váhovaných konečných stavových převodníků (wFST) je výsledná rozpoznávací síť dána jako $HCLG = H \circ C \circ L \circ G$. Napište nebo nakreslete, jak by měla vypadat gramatika G v případě rozpoznávače pro vytáčení telefonních čísel (předpokládáme, že se bude vytáčet pouze základními číslovkami “nula”, “jedna” ... “devět”).

18. Při rozpoznávání (dekódování) s velkým slovníkem se používá tzv “beam search” - co to je ?

19. Pro vyhodnocení systémů pro ověřování mluvčího (a obecně pro každou detekční úlohu) se používá detection error trade-off (DET) křivka. Uveďte, co je na jejích osách a nakreslete typickou DET křivku.

20. Uveďte, jaké je hlavní použití *neuronových sítí* v systémech pro rozpoznávání řeči (ASR).