

Synteza mowy
Pracownia 3
Zajęcia 7 i 8

Wbrew wcześniejszym zapowiedziom można oddawać zadanie o synteźatorze sylabicznym bez straty na obu zajęciach z tej listy. Uwaga: zarówno synteźator sylabiczny (w małym zakresie), jak i synteźator difoniczny umożliwią zdobycie punktów egzaminacyjnych (informacje na kolejnej liście wraz z dodatkowymi zadaniami z tej tematyki). Dwa wybrane zadania z tej listy można oddać bez straty na pracowni 4 (na pierwszych zajęciach).

Zadanie 1. (3p) Napisz program, który unifikuje głośność plików w danej kartotece. Jego działanie powinno wyglądać tak:

- a) Skanuje wszystkie pliki wav w danej kartotece i wyznacza najgłośniejszy
- b) Zwiększa głośność pozostałych w ten sposób, by wszystkie były mniej więcej tak samo głośne.
- c) Zapisuje podgłośnione pliki w stworzonej do tego celu podkartotece (zachowując nazwy oryginalnych plików)

Zadanie 2. (1p) Przejrzyj i uruchom podany na stronie wykładu program demonstrujący bazę synteźatora difonicznego (program `a_difoniczne.py`). Jeżeli chcesz dalsze zadania z listy robić w innym języku niż Python, powinieneś również skonwertować bazę do wygodnego dla Ciebie formatu. Wskazówkę, jak to zrobić możesz znaleźć w pliku `skonwertuj.py`.

Zadanie 3. (2p) Napisz program, który wymawia wybraną samogłoskę przez t sekund, w którym samogłoska i t są parametrami. Powinieneś korzystać z obu difonów: lewego i prawego.

Zadanie 4. (4p) Napisz program, który jest pełnym synteźatorem difonicznym, ale nie wykonuje żadnych przekształceń na difonach (to znaczy po prostu skleja elementy bazy). Zastanów się, jak najlepiej obsługiwać „difony” takie jak `u_p`, `a_t`, itd (z bezdźwięczną głośką zwartą na drugim miejscu, których nie ma w bazie).

Zadanie 5. (4+1p) Napisz program, który jest pełnym synteźatorem difonicznym, w którym (zależnie od potrzeb) difony są skracane bądź wydłużane. Powinieneś zaimplementować dwie strategie wyznaczania czasów:

- a) fonemy mają czasy równe czasom średnim, nieco wydłużonym w przypadku sylab akcentowanych
- b) każda sylaba nieakcentowana ma ten sam czas (podobnie każda akcentowana), w ramach sylaby fonemy dzielą się czasem proporcjonalnie.

Porównaj naturalność obu podejść. Premia (+1) za wymyślenie i zaimplementowanie trzeciej strategii, będącej połączeniem dwóch powyższych.

Program na wejściu powinien przyjmować znormalizowany tekst w języku polskim. Jakość skracania i wydłużania będzie oceniona w innym zadaniu, do zaliczenia tego zadania wystarczy jedynie unikać grubych błędów (typu jąkanie się, bardzo wyraźne trzaski, gubienie głosek, etc.)

Zadanie 6. (6p) Dodaj możliwość intonacji dla Twojego synteźatora. Możesz założyć, że daną jest fraza, którą powinieneś wymówić jako frazę oznajmującą. Synteźator powinien:

- realizować ciągle opadanie F_0
- zwiększać (w sposób płynny) F_0 w sylabach akcentowanych
- kończyć każdą frazę wyraźnym spadkiem F_0 .

Częstotliwość powinna być interpolowana liniowo (lub wyznaczana za pomocą TD-PSOLA)

Zadanie 7. (4p) Zamień liniową interpolację w poprzednim zadaniu na TD-PSOLA w tym (albo odwrotnie)