

## Využití Legendrových polynomů pro opravu prozódie

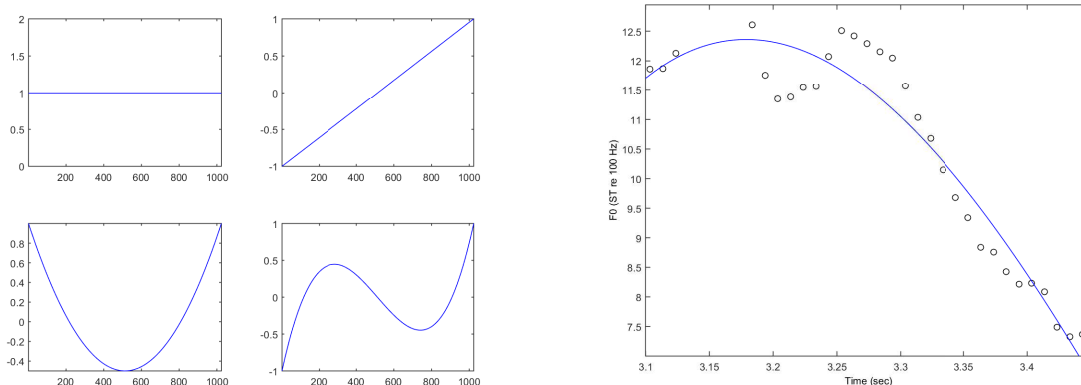
Martin Matura<sup>1</sup>

### 1 Úvod

Přirozenost je velmi důležitým aspektem uměle syntetizované řeči. Nepřirozené změny v prozódii, které vznikají v důsledku neočekávaných výkyvů ve fundamentální frekvenci ( $F_0$ ), působí rušivě a mohou vést k nesrozumitelnosti slova nebo dokonce celé věty. Tato práce se zabývá návrhem postupu, který by opravil popis prozódie v českých korpusech, které se využívají pro syntézu řeči pomocí výběru jednotek.

### 2 Prozódie a Legendrovy polynomy

V korpusech, které používáme pro syntézu řeči, je prozódie popsána abstraktně pomocí takzvaných prozodémů ( $P_0$ ,  $P_1$ ,  $P_2$ ,  $P_3$  a  $P_{0.1}$ ), které přiřazujeme jednotlivým prozodickým slovům. Během procesu syntézy řeči pomocí výběru jednotek dochází k nalezení nejlepší posloupnosti jednotek (v našem případě difonů) požadovaných pro vytvoření určité promluvy. Každá jednotka je popsána určitými parametry a při výběru jednotek se mimo jiné počítá i cena cíle, která by měla zajistit, že jednotka pochází ze správného prozodému. Řečové korpuse, ze kterých jsou však jednotky vybírány, ale někdy obsahují nepřesnosti v prozodickém popisu (převážně prozodémy  $P_0$  nejsou vždy neutrální) a kvůli tomu pak při řetězení jednotek může docházet k neshodě v  $F_0$  a tím pádem k řečovým artefaktům.



(a) První čtyři polynomy - střední hodnota, sklon, tvar údolí a zakřivení.

(b) Interpolace kontury  $F_0$  pro odhad koeficientů Legendrova polynomu.

**Obrázek 1:** Legendrovy polynomy.

Metoda pro opravu českých korpusů, která se zaměřuje na přeznačení neutrálních prozodických slov (označených prozodémem  $P_0$ ), je dvoufázová a využívá Legendrových polynomů

<sup>1</sup> student doktorského studijního programu Aplikované vědy a informatika, obor Kybernetika, e-mail: mate221@kky.zcu.cz

viz Legendre (1785), které jsou definovány Rovnicí 1:

$$P_n(x) = 2^n \sum_{k=0}^n x^k \binom{n}{k} \binom{\frac{n+k-1}{2}}{n}. \quad (1)$$

Během první fáze byl použit *One-Class Support Vector Machine* detektor (OCSVM) pro nalezení anomálií mezi prozoděmy  $P_0$ . Ve druhé fázi je natrénován *Support Vector Machine* klasifikátor (SVC) a nalezené anomálie z prvního kroku jsou klasifikované do 5 tříd –  $P_0, P_1, P_2, P_3$  a  $P_{0.1}$ . Pro natrénování OCSVM i SVC se využívají jednotky popsané prvními čtyřmi Legendrovými polynomy (viz. Obr 1a), k jejichž získání slouží nástroje mPraat pro Matlab od Bořil a Skarnitzl (2016).

### 3 Výsledky

Pomocí popsané dvoufázové metody jsme opravili dva české korpusy – *AJ* a *MR*. V každém z korpusů došlo k nalezení anomálií ve skupině nultých prozoděmů a k jejich následnému přeznačení. Kolik anomálií bylo přeznačeno do jednotlivých tříd ukazuje Tabulka 1.

Korpus	$P_0$ anomálie	$P_0$	$P_1$	$P_2$	$P_3$	$P_{0.1}$
AJ	3,578	1,559 (43.6%)	189 (5.3%)	328 (9.2%)	328 (9.2%)	1,174 (32.8%)
MR	4,174	988 (23.7%)	385 (9.2%)	145 (3.5%)	817 (19.6%)	1,839 (44.1%)

**Tabulka 1:** Přeznačení  $P_0$  anomálií.

K vyhodnocení výsledků byly použity preferenční testy, ve kterých 16 posluchačů porovnávalo kvalitu čtyřiceti párů vět pro každý korpus. Každá dvojice vět se skládala z jedné věty syntetizované z nepřeznačeného korpusu (*TTS-base*) a z jedné věty syntetizované z přeznačeného korpusu (*TTS-new*). Dle výsledků shrnutých v Tabulce 2 můžeme konstatovat, že přeznačení

Korpus	<i>TTS-base</i>	Stejná kvalita	<i>TTS-new</i>	Skóre $s$
AJ	105 (16.4%)	325 (50.8%)	210 (32.8%)	<b>0.164</b>
MR	136 (21.5%)	311 (48.6%)	193 (30.2%)	<b>0.089</b>

**Tabulka 2:** Výsledky poslechových testů.

prozoděmů nezanese do syntézy žádné výrazné zhoršení, kvalita byla zhruba v polovině případů stejná jako před přeznačením a ve zbývajících případech byly spíše preferovány přeznačené korpusy. Výsledky jsou statisticky významné což potvrzuje kladné  $s$  skóre.

#### Poděkování

Příspěvek byl podpořen grantovým projektem SGS-2016-039.

#### Literatura

- Bořil, T. a Skarnitzl, R. (2016) Tools rPraat and mPraat. *Text, Speech, and Dialogue: 19th International Conference*. Brno, Czech Republic, pp. 367–374.
- Legendre, A.-M. (1785) Recherches sur l'attraction des sphéroïdes homogènes. *Mémoires de mathématique et de physique, présentés à l'Académie royale des sciences, par divers sçavans & lus dans ses assemblées*. Paris, pp. 411–435.