
Strukturovaný posudek bakalářské práce

Tomáš Kubový

Transkripční software pro trénování ASR systémů

1. Informace k zadání

Tato bakalářská práce vychází z potřeb Laboratoře inteligentních komunikačních systémů KIV FAV ZČU. Cílem bylo vyvinout moderní flexibilní modulární aplikaci, která by usnadnila jednak přípravu trénovacích korpusů pro potřeby trénování systémů rozpoznávání řeči, jednak dovolila provádět některé specifické experimenty v oblasti rozpoznávání a klasifikace akustického signálu. Volně dostupné aplikace tohoto typu sice existují, ale buď nesplňují přesně požadavky LIKS, nebo je jejich použití nějakým způsobem omezeno (např. specifickou platformou či požadavky na běhové prostředí). Autor měl tedy za úkol vytvořit aplikaci, která by omezení a nevýhody stávajících řešení překonávala a zároveň představovala solidní základ pro další vývoj.

2. Aktivita během řešení, konzultace, komunikace

..... 7 bodů ^{z 15} max.

Autor práce se na konzultace bohužel dostavoval jen zřídka a intenzita jeho aktivity při řešení úkolu byla velmi kolísavá. V průběhu práce na PRJ5 přišel vlastně jen na začátku, nechal si vysvětlit zadání a pak se dostavil až v lednu pro zápočet. Práce ovšem ani zdaleka nebyla ve stavu, který by udělení zápočtu umožňoval. Lze tedy říci, že autor **občas** pracoval velmi intenzivně. Tento způsob, který vedoucí práce shledává velmi nešťastným, se ale bohužel i přes četná upozornění nijak valně nezměnil.

Na druhou stranu je třeba říci, že autor nakonec práci zvládl dokončit v přijatelné podobě a včas. Konečná podoba aplikace celkem odpovídá představám zadavatele (byť s řadou drobných výhrad). Průvodní dokument byl s vedoucím práce na poslední chvíli částečně konzultován.

3. Splnění požadavků zadání

..... 20 bodů ^{z 25} max.

Předložená bakalářská práce **splňuje všechny body** zadání, některé ovšem s drobnými výhradami uvedenými dále. Autor prokázal obstojné programátorské dovednosti a práci vytvořil podle zavedených zvyklostí a postupů programování v jazyce Java. Aplikoval vhodné technologie, jako např. Java Media Framework, aj., a navrhl aplikaci tak, že bude možné v jejím vývoji pokračovat.

Výsledná podoba uživatelského rozhraní je sice poněkud jiná, než si zadavatel představoval (což je dáno zejména výše zmíněným způsobem vzájemné interakce s autorem), nicméně pro daný úkol použitelná. Aplikace je dostatečně stabilní a výkonná, přijatelně zdokumentovaná a připravená k nasazení. Uživatelská příručka je ovšem dost stručná a představa, že by úspěšně provedla úskalí práce s aplikací uživatele bez průpravy v oblasti IT, je nereálná.

4. Hodnocení formální stránky předložené práce

..... 15 bodů ^{z 25} max.

Dodaný průvodní dokument bakalářské práce je (byť vysázen v MS Wordu) na slušné estetické a typografické úrovni. Práce je napsaná srozumitelně. Autorova technická čeština je místy poněkud těžkopádná, některé konstrukce jsou otrockým překladem anglických výrazů. Vyjadřovací schopnosti autora jsou dobré, stejně jako popisná úroveň většiny dokumentu.

Text práce ovšem není příliš vhodně logicky strukturovaný, např. souvislost částí 2.2 a 2.3 není zjevná. Pokrytí prerekvizitních teoretických informací zejména z oblasti počítačového zpracování zvuku a postupu řešení problému je vyvážené, postrádám hlubší analýzu požadavků na aplikaci (což se ostatně také projevilo výše zmíněným nesouladem představ zadavatele a autora aplikace o její podobě).

Dokument je poměrně stručný, některé důležité části týkající se implementace aplikace jsou dosti povrchní – podle autorova popisu by jistě nebylo možné aplikaci reimplementovat. Text je velmi vhodně doplněn řadou obrázků v dobré kvalitě – rastrové jsou v dostatečném rozlišení, a některé z obrázků jsou dokonce vektorové. Zvýrazňovací řez písma je používán vhodně a žádoucím způsobem přispívá ke zvýšení typografické úrovně i čitelnosti práce.

Překlepy ani gramatické chyby se v práci nevyskytují. Místy lze najít některé ne zcela jasné formulace a ne zcela přesné informace. Horší ovšem je, že autor nemá jasno v použití a významu některých zákl. jednotek, např. na str. 18 uvádí **rychlost zvukového signálu** v jednotkách **kHz/s**, což považuji u studenta technické fakulty za alarmující.

Také převod $176\,400\text{ B/s} = 176,4\text{ kB/s}$ (verbatim, str. 6) je přinejmenším zavádějící. Autor si také plete označení některých formátů s komerčními názvy produktů, a tak na str. 6 hovoří o „CD formátu“ namísto o formátu CD-DA.

Výběr použitých zdrojů odpovídá řešené problematice, rozsah se blíží spíše spodní hranici přijatelnosti.

5. Hodnocení realizačního výstupu

..... **33 bodů** ^{z 35} max.

Software, vyvinutý v rámci řešení této bakalářské práce, je na velmi dobré úrovni. Autor nakonec na poslední chvíli v mohutném vzepjetí sil aplikaci dokončil ke spokojenosti zadavatele. Má všechny požadované a většinu očekávaných funkcí, které bez problémů (pokud lze posoudit) fungují.

Autorem aplikovaný postup přehrávání a označování částí zvukového signálu se ukazuje v praxi jako dobře použitelný a celkově je aplikace až překvapivě rychlá a dostatečně robustní. Architektura aplikace je zvolena vhodně s ohledem na způsob nasazení a budoucí rozšiřování.

Zdrojový kód v jazyce Java je srozumitelný, byť by mohl být více okomentovaný. Dekompozice i objektová analýza je provedená správně. Autor nepoužívá žádné problematické či nestandardní konstrukce, kód je přehledně rozložen do balíků a nelze k němu mít žádné podstatné výhrady.

V aplikaci postrádám (a považuji za důležitou) možnost „překlada“ vkládaných transkripčních značek na srozumitelné výrazy – autorovo řešení funguje tak, že se do transkripce vkládá přímo ten výraz, který je uveden na tlačítku, což není příliš vhodné řešení (navíc naprogramování vhodného by představovalo max. 10 minut práce).

6. Otázky k obhajobě

1. Ve výsledcích na str. 22 a 23 uvádíte grafy závislosti velikosti obsazené paměti a doby výpočtu na velikosti vstupního souboru. Čím si vysvětlujete, že zatímco první je lineární (očekávaně), druhá je exponenciální? Jde o vlastnost nebo chybu interpolace?
2. Proč jste zvolil v panelu vykreslování zvukové vlny symetrickou metodu vykreslování (tj. stejná velikost amplitudy nad i pod osou x), když obecně signál vůbec tuto symetrii vykazovat nemusí a často (v případě stejnosměrného offsetu) ji přímo narušuje?

7. Závěrečné shrnutí

..... **75 bodů** ^{z 100} max.

Práci i přes zmíněné nedostatky **doporučuji k obhajobě.**

V Plzni dne 31. 5. 2012



Ing. Kamil Ekštejn, Ph.D.
KIV FAV ZČU