

## **Oponentský posudek na disertační práci Mgr. Petra Podveského**

### **„Speech Recognition of Czech Using Finite-State Machines”**

Disertační práce Mgr. Petra Podveského „Speech Recognition of Czech Using Finite-State Machines” se zabývá problematikou z oblasti jazykového modelování v systémech rozpoznávání souvislé řeči. Práce je zaměřena na konstrukci modelů pro mluvenou češtinu, jako reprezentanta skupiny slovanských jazyků. Tyto jazyky vykazují obecně vysokou flexi při tvorbě slov a též relativně volnější strukturu vazeb mezi slovy v jednotlivých promluvách. To má za následek značný nárůst položek ve slovníku a obtíže, jak snížit množství slov mimo slovník, která se mohou objevit při rozpoznávání. Zvyšování rozsahu slovníku znamená i potřebu značného nárůstu dat, ze kterých se budou trénovat jazykový model. Protože se disertační práce zabývá těmito vysoce potřebnými tématy, které souvisí s jazykovými potřebami českého národního prostředí, hodnotím téma disertace jako vysoce aktuální.

I když vlastní cíle disertace jsou popsány velmi stručně - disertant si klade pouze rámcový cíl zaměřený na zlepšení rozpoznávání řeči pro češtinu s využitím morfolgie při jazykovém modelování - konkrétní cíle disertace jsou nicméně směřovány na potlačení nepříznivých vlivů slov mimo slovník (OOV, Out-Of-Vocabulary) při jazykovém modelování a vlastním rozpoznávání. Takové cíle považuji rozhodně za disertabilní.

Úvodní kapitoly disertace se věnují sice stručnému, ale přitom výstižnému shrnutí základních problémů, které je třeba řešit při konstrukci systémů rozpoznávání souvislé řeči. Důraz je přitom kladen zejména na přehled metod jazykového modelování. Jako velmi zajímavou a dobře zpracovanou hodnotím 3. kapitolu, ve které disertant shrnul a popsal nové trendy v modelování jazyka.

Stěžejní část disertační práce je zaměřena na modelování češtiny a zejména na „boj” se slovy mimo slovník. Disertant nejprve na reálných příkladech ukázal důsledky bohaté flexe češtiny a diskutoval známé přístupy odvozené od morfologického zpracování, které se snaží s tímto jevem při jazykovém modelování vhodně zacházet. Výsledkem této diskuse se staly dva zcela nové návrhy, které se snaží zmírnit vliv OOV.

První metoda je založena na shromáždění slov, která jsou mimo původní rozpoznávaný slovník, přičemž tato slova jsou dodána sice jednoduchým nicméně účinným způsobem do navrženého jazykového modelu. Z popisu realizace tohoto způsobu modelování vyplývá stejná váha, se kterou byla všechna slova do modelu vkládána. Zde se nabízí ještě velký prostor pro další experimentování. Například u slov, která byla přidána na základě doplnění slov s jinými pádovými koncovkami, případně slov, která lze řadit do jistých sémantických kategorií (např. jména, číslovky ap.), mohly být pravděpodobnosti odvozeny s jistou znalostí statistik obdobných slov v korpusu, z něž vznikl slovník i n-gramový jazykový model. To by zřejmě znamenalo velké množství dalších experimentů, nicméně to doporučuji k uvážení pro další práce.

Klíčová idea druhé metody spočívá ve dvouprůchodovém rozpoznávání, kdy výstup

prvního průchodu nabídne výsledek, který i v případě přítomných a špatně rozpoznávaných OOV slov bude předpokládat, že řetězec fonémů v rozpoznávaných slovech je foneticky blízký správné posloupnosti slov. Ta je pak ve druhém průchodu zkoušena s využitím rozsáhlého slovníku.

Správnost návrhu obou metod byla otestována při rozpoznávání souvislé řeči. Experimenty byly provedeny s využitím „Korpusu Lidové noviny“, korpusu “Czech Broadcast News”, a korpusu shromážděného při zpracování českých výpovědí svědků holocaustu v projektu MALACH. Tyto experimenty prokázaly správnost navrženého postupu a prokazatelně snížily chybovost rozpoznávání. Protože v obou případech šlo o původní dosud netestovaný způsob úspěšného potlačení vlivu OOV, nemohu než konstatovat úspěšnost splnění cílů disertace. Pro odbornou diskusi mám nicméně dva dotazy:

- Je nějaké zásadní omezení pro rozšiřování slovníku pro případ, že bychom chtěli OOV přiblížit skutečné nule?
- Je možné využít některou z navržených metod snižování OOV (např. po vhodné modifikaci) v real-timeových dekodérech?

Disertační práce je napsána úsporným, nicméně obsahově zcela dostatečným způsobem. V práci jsem nenalezl žádné závažnější nedostatky. Závěrem tedy mohu konstatovat, že předložená disertační práce je velmi dobrým příspěvkem k problematice potlačování negativního vlivu vysokého počtu slov mimo slovník. Za hlavní a disertabilní přínos práce Petra Podveského považuji návrh a široké otestování nových metod, které mají schopnosti využívat i slova mimo původní slovník a dosáhnout tak prokazatelného zlepšení funkce systému rozpoznávání řeči. Cíle disertace považuji za splněné. Mohu proto konstatovat, že byly naplněny podmínky zákona č. 111/98 Sb. a lze tedy doporučit, aby byla disertační práce přijata k obhajobě a po jejím úspěšném vykonání byl disertantovi udělen titul Ph.D.



V Plzni 15.8.2006