

Dialogové systémy

Luděk Bártek

Fakulta informatiky
Masarykova univerzita

podzim 2013

Obsah

- 1 PLS
- 2 Simulace dialogu metodou Wizard of Oz
- 3 Multimodální dialogové systémy
- 4 Dialogové systémy a emoce

Pronunciation Lexicon Specification (PLS)

- Standard W3C
- Aktuální verze 1.0 (říjen 2008)
- Definuje značkování pro specifikaci slovníků výslovnosti pro podporu syntézy a rozpoznávání řeči.
- Specifikace na stránkách W3C Specifikace na stránkách W3C

Základní elementy PLS

- Kořenový element - lexicon
 - atributy - xmlns - specifikace jmenného prostoru (<http://www.w3.org/2005/01/pronunciation-lexicon>)
 - xml:lang - jazyk dokumentu
 - version - verze dokument (1.0)
 - alphabet - abeceda použitá pro fonetický přepis
- lexeme - obsahuje popis pro jednu lexikální jednotku (slovo, zkratku, . . .
 - musí obsahovat aspoň jeden dceřiný element grapheme
- phoneme - obsahuje fonetický přepis dané lexikální jednotky (většinou se používá IPA).

Ukázka lexikonu

```
<?xml version="1.0" encoding="utf-8"?>
<lexicon
  version ="1.0"
  xmlns="..."
  alphabet="ipa"
  xml:lang="en-US">
<lexeme>
  <grapheme>color</grapheme>
  <phoneme>
k' ʌə
  </phonem>
</lexeme>
</lexicon>
```

Ukázka propojení lexikonu a gramatiky

- XML formát SRGS

```
<grammar xmlns="..." xml:lang="en" version="1.0">  
  <lexicon  
    uri="http://www.example.com/lexicon.file"/>  
  <lexicon  
    uri="http://www.example.com/strange-city-names.file"  
    type="media-type"/>  
  ...  
</grammar>
```

Ukázka propojení lexikonu a gramatiky

- ABNF formát SRGS

```
#ABNF V1.0 ISO-8859-1;  
language en-US;  
lexicon <http://www.example.com/lexicon.file>;  
lexicon <http://www.example.com/strange-city-names.file>  
      <media-type>;  
...
```

Ukázka použití lexikonu v SSML

```
<speak version="1.1" xmlns="..." xml:lang="en-US">  
  <lexicon uri="lexicon.pls" xml:id="pls"/>  
  <lexicon uri="strange-words.file" xml:id="sw"  
    type="media-type"/>  
  <lookup ref="pls"> tokens here are looked up in  
    lexicon.pls  
  <lookup ref="sw"> tokens here are looked up first in  
    strange-words.file and then, if not found, in  
    lexicon.pls  
  </lookup>  
  tokens here are looked up in lexicon.pls  
  </lookup>  
  tokens here are not looked up in lexicon documents ...  
</speak>
```


Metoda Wizard of Oz

- Simulace dialogového rozhraní modelem člověk – člověk.
- Založena na principu popsaném v knize The Wonderful Wizard of Oz (Lyman Frank Baum)
- Princip:
 - Funkce dialogového rozhraní je (skrytě) simulována člověkem.
 - Průběh dialogu je protokolován.
 - Průběh se řídí navrženou dialogovou strategií.
 - Pokud je dostupný prototyp může Wizard pouze modifikovat a předávat komunikaci mezi uživatelem a systémem.
- Občas snaha navodit zdání, že uživatel komunikuje s dialogovým systémem – využívají se různé prostředky:
 - vzdálená komunikace kde osoba simulující dialogové rozhraní komunikuje prostřednictvím TTS
 - použití vokodérů, které změní hlas osoby, která provádí testování, aby zněl jako výstup TTS
 - ...

Tvorba dialogového rozhraní pomocí metody WoZ

- Z korpusu dialogů na dané téma (pro danou doménu) lze vygenerovat dialogové rozhraní následovně:
 - 1 Vytvoříme iniciální korpus metodou WoZ
 - Komunikace pouze čaroděj – uživatel.
 - 2 Odstraní se konflikty a na základě korpusu se vytvoří dialogové rozhraní.
 - 3 Kombinovaně vytvoříme nový korpus.
 - "Čaroděj" se snaží maximálně využívat navržené dialogové rozhraní..
 - 4 Odstranění konfliktů a vygenerování nové verze dialogového rozhraní.
 - 5 Pokud je rozhraní v pořádku, generování končí, jinak se pokračuje krokem 3.

Multimodální dialogová rozhraní

- Mimo mluvenou řeč umožňuje i další způsoby komunikace člověk – počítač:
 - textová komunikace
 - grafická komunikace
 - emoce
 - ...
- Výhody
 - lepší přístupnost.
 - uživatelé s poruchami sluchu,
 - uživatelé s poruchami řeči,
 - ...
 - lepší pochopení pragmatiky projevu
 - možnost pracovat i s emočním rozpoložením – může mít vliv na pragmatiku projevu.

Multimodální komunikace počítač – člověk

- Textová:
 - Mimo hlasový výstup je navíc zobrazen i odpovídající textový výstup.
 - Lze využít prostředky pro IM, SMS, ...
- Grafická:
 - Talking Heads – mimo hlasový výstup je navíc zobrazena tvář (hlava, celý člověk, ... jejíž pohyby, zejména úst, odpovídají mluvené řeči.
 - Komunikace znakovou řečí – mluvené slovo je překládáno na znakovou řeč (viz Guimeraes et al. – Structure of the Brazilian Sign Language (Libras) for Computational Tools: Citizenship and Social, in Organizational, Business, and Technological Aspects of the Knowledge Society, CCIS vol. 112, Springer, Heidelberg, 2010, pp. 365 — 370.)
 - Znaková řeč prezentována pomocí rukou nebo avatara.

Multimodální komunikace člověk – počítač

- Široké spektrum možností zadávání vstupu uživatelem jinak než hlasem:
 - klávesnice (počítač, DTMF, SMS, ...)
 - rukou psaný vstup – dotyková obrazovka + pero, ...
 - ústy ovládaná zařízení
 - ovládání pomocí pohybů očí a víček
 - rozpoznávání řeči pomocí sond detekujících činnost svalů a mozku (viz Schultz, T. – Silent and Weak Speech Based on Elektromyography, in Proceedings of 12th International Conference ICCHP 2010 Part 1, Wien, Springer, Heidelberg, pp. 595 – 604, 2010.)
 - rozpoznávání znakové řeči
 - ...
- Často jako doplněk řečového vstupu.

Existující nástroje a standardy

- Proprietární řešení:
 - Součást CSLU Toolkitu.
 - Projekt August.
- Otevřená řešení:
 - Návrhy doporučení W3C týkající se multimodálního přístupu – zatím bez implementace.
 - Využívají a propojují i další standardy W3C (CCXML, XHTML, VoiceXML, SVG, SMIL, ...)
 - Výstup W3C Multimodal Interaction WG

Ukázka multimodální syntetizované řeči

Demonstrační video se syntetizovanou multimodální řečí.

Co jsou to emoce?

- "This is a very tough question, that has produced significant amounts of headaches to scientists in the past ...", "... many researchers have to opted to study systematically phenomena that most consider emotional." (Laval University of Quebec)
- "Only mathematics is certain, so all must be based on mathematics." (R. Descartes)
- Dělení emocí:
 - Primární (základní) – vyskytují se u všech lidí a u části vyšších živočichů.
 - Sekundární (vyšší) – mohou být intelektuální, morální a estetické. Mohou se lišit mezi jednotlivými kulturami.
- Velkých šest:
 - hněv
 - zklamání
 -

Základní emoce

- Velkých šest (R. Descartes):
 - hněv
 - zklamání
 - smutek
 - strach
 - překvapení
- Další autoři:
 - Arnold – hněv, averze, odvaha, sklíčenost, touha, zoufalství, strach, nenávisť, láska, smutek.
 - Ekman, Friesen, Ellsworth – hněv, odpor strach, radost, smutek, překvapení.
 - Frijda — touha, štěstí, zájem, překvapení, údiv, zármutek,
 - ...

Detekce emocí

- Lze provádět pomocí detekce změn různých biometrických vlastností.
 - Změny galvanických vlastností kůže.



- Změny tlaku krve a pulsu.



Detekce emocí

- Použitelné biometrické charakteristiky:
 - změny dýchání



- změny elektrické aktivity mozku



- změny charakteristik řeči
- výraz tváře (Yale Face Database).
- ...

Ukázky z Yale Face Database

- Radost



- Smutek



Ukázky z Yale Face Database

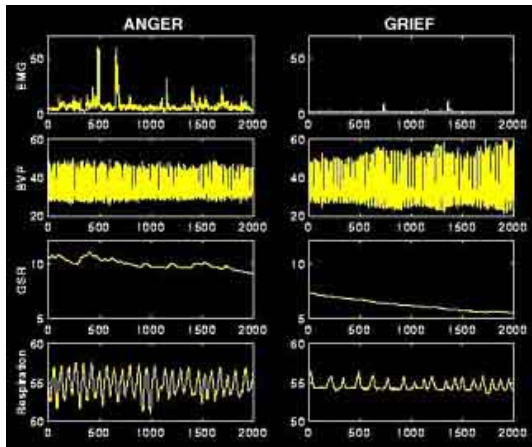
- Ospalost



- Překvapení



Ukázky průběhů charakteristik pro smutek a hněv



Využití emocí v dialogovém rozhraní

- Dialogová rozhraní informačních systémů
 - uzpůsobení dialogové strategie emočnímu stavu uživatele (klid, stres, hněv, ...)
 - přepojení uživatele na lidského operátora.
- Výukové DS:
 - uzpůsobení dialogové strategie koncentraci uživatele.
- ...