

SIN04: Řečová interakce a sociální sítě

Luděk Bártek

Fakulta informatiky
Masarykova univerzita
Brno

podzim 2013

Obsah

- 1 Rozpoznávání plynulé řeči
- 2 Dialogová komunikace
- 3 Dialogové systémy

Rozpoznávání plynulé řeči

- Hlavní rozdíly oproti rozpoznávání izolovaných slov:
 - nelze vytvořit databázi vzorů
 - nutno brát zřetel na prozodické faktory
 - nutno určovat hranice mezi slovy
 - nutno vypořádat se s výplňkovými zvuky a chybami řeči.
- Řešení – statistický přístup:
 - jazykový model – popis promluv daného jazyka včetně jejich četností.
 - model uživatele – popis stylu vyjadřování daného uživatele.
- Příklad: HMM vrátí stejnou pravděpodobnost pro slova máma a nána
 - nejspíše se použije máma – je častější.

Rozpoznávání plynulé řeči

Úskalí

- Problém – úspěšnost obecného rozpoznávání může klesnout až k cca 50
- Metody pro zvýšení úspěšnosti:
 - znalost domény (money vs. honey).
 - specifikováním rozpoznávaných promluv.
 - např. pomocí gramatiky pro rozpoznávání řeči (JSGF, SRGS, ...) – více u dialogových systémů.

Dialog - základní pojmy

- *Dialog* - z řeckého dialogos (od slovesa dialogemai - rozvažovat, rozmlouvat).
- *Dialog* - komunikace dvou účastníků (člověk- člověk, člověk - počítač, ...), skládá se z *promluv*.
- *Promluva* - souvislé sdělení jednoho z účastníků dialogu.
- *Obrat* - promluva a reakce druhé strany na ní.
- Průběh dialogu každého účastníka se řídí jeho *dialogovou strategií*.
- *Hodnotící funkce* $E(L)$ - přiřazuje dialogu reálné číslo - ohodnocení dialogu daným účastníkem.
- Uspořádanou čtveřici $M = (S_1, S_2, E_1, E_2)$, kde S_i jsou dialogové strategie a E_i hodnotící funkce účastníků nazveme *dialogová komunikace*.

Dialogová strategie

- Zobrazení $U \times Q \rightarrow R \times Q$
 - U - množina vstupních promluv
 - R - množina výstupních promluv
 - Q - množina stavů.
- Určuje následující krok dialogu v závislosti na vstupní promluvě a stavu dialogu (jeho předchozím průběhu).
- Hodnotící funkce dialogu E přiřazuje danému dialogu reálné číslo popisující jeho úspěšnost z pohledu daného účastníka dialogu.

Klasifikace dialogových strategií

- Dělení dialogu z pohledu hodnotící funkce:
 - d je kooperativní dialog pokud $E_1(d) = E_2(d)$
 - d je nekooperativní dialog pokud $E_1(d) \neq E_2(d)$
 - d je dialog s nulovým součtem pokud $E_1(d) = -E_2(d)$
- Další krok dialogu je vždy určen dialogovou strategií jedné z komunikujících stran.
 - Např. jedna strana se ptá, druhá odpovídá.
- U dialogové komunikace člověk (uživatel) - počítač rozlišujeme:
 - strategii s iniciativou systému - systém se ptá, uživatel odpovídá
 - strategii s iniciativou uživatele - uživatel se ptá, systém odpovídá
 - strategii se smíšenou iniciativou - strany se v určování průběhu dialogu střídají. Většinou je z počátku iniciativa ponechána na uživateli. Systém se ujímá iniciativy pokud uživatel nezadal korektně všechny potřebné údaje.

Pravidla pro vedení kooperativního dialogu (H. P. Grice)

- Autorem britský jazykovědec Herbert Paul Grice.
- Pokud chtějí dva lidé vést kooperativní dialog měli by se držet následujících aspektů:
 - 1 Aspekt informativnosti:
 - Buď přiměřeně informativní - nesděluj druhé straně ani méně ani více informací, než je pro daný dialog potřeba.
 - 2 Aspekt přesvědčivosti:
 - Neuváděj nepravdivé informace.
 - Neuváděj informace, které nelze dokázat nebo doložit.
 - 3 Aspekt způsobu:
 - Informace v replice by měla být co nejvíce explicitní.
 - Vyhýbejte se nejednoznačností.
 - Usilujte o stručnost.
 - Buďte disciplinovaní, udržujte v dialogu pořádek, např. neměňte bezdůvodně cíl/téma dialogu.

Pravidla pro vedení kooperativního dialogu (H. P. Grice)

Dokončení

- 5 Aspekt zdvořilosti, empatie a etiky:
 - Minimalizujte nároky vůči komunikačnímu partnerovi, maximalizujte výhody pro něj.
 - Minimalizujte nedostatky komunikačního partnera, maximalizujte jeho přednosti.
 - Maximalizujte souhlas s partnerem a minimalizujte nesouhlas.
 - Maximalizujte empatii vůči komunikačnímu partnerovi.

Způsoby vedení dialogu

- Hlasový dialog:
 - tváří v tvář
 - pomocí telekomunikačních prostředků - telefon, vysílačka, ...
 - hlasový IM (Skype, Jabber/Gmail Talk, IP telefonie, ...)
- Textový dialog:
 - pošta - nevýhoda - délka trvání jednoho obratu.
 - elektronická pošta
 - IM
 - různé formy IRC
 - ...

Dialogový systém - základní pojmy

- Dialogový systém - informační systém komunikující s uživatelem formou dialogu v přirozeném jazyce.
 - Většinou se jedná o dialogové rozhraní ke klasickému IS.
- Často komunikace mluvenou řečí.
- Alternativně:
 - komunikace pomocí DTMF
 - textová komunikace přirozenou řečí
 - multimodální komunikace:
 - řeč + obraz (simulace lidské tváře, titulky ve znakové řeči, ...)
 - řeč + text
 - ...

Výhody dialogových rozhraní oproti GUI

- Přírozenější způsob komunikace
 - v ideálním případě by uživatel neměl poznat, že komunikuje s počítačem.
 - umožňuje vedení uživatele procesem zadávání údajů krok za krokem.
- Lepší přístupnost pro lidi s různými druhy postižení:
 - zrakově postižení uživatelé
 - motoricky postižení uživatelé
 - uživatelé mající problémy s ovládním počítače pomocí GUI.
 - zlepšení přístupnosti - multimodální rozhraní.
 - umožňuje přístup ke službám IS bez nutnosti použití vlastnit počítač - stačí telefon.

Nevýhody dialogových rozhraní oproti GUI

- Nevýhody:
 - Rychlost komunikace:
 - sekvenční vnímání zvuku vs. paralelní vnímání obrazu.
 - lze částečně eliminovat pomocí vhodné dialogové strategie.

Historie dialogových systémů

- Eliza
 - počátek 60. let
 - počítačová simulace rozhovoru s psychoterapeutem
 - textové komunikace v přirozeném jazyce
- Parry
 - autor K. M. Colby (1963)
 - simulace paranoidního pacienta - reakce na Elizu
 - v řadě dialogů nebylo možné jednoznačně určit, zda se jedná o simulaci nebo reálného pacienta
- Expertní systémy
 - pracují se znalostní databází vytvořenou experty v dané oblasti
 - znalostní databáze obsahují:
 - fakta
 - inferenční pravidla - pravidla pro odvozování závěrů na základě zjištěných faktů
 - DENDRAL - expertní systém z oblasti organické chemie

Historie dialogových systémů

2.

- Expertní systémy (dokončení):
 - INTERNIST I - expertní systém pro pomoc při diagnostice (1970, University of Pittsburgh Medicine School)
 - MYCIN
 - Stanford University (70. léta)
 - navazuje na INTERNIST I (jeden z autorů společný)
 - obsahoval i pokročilá odvozovací pravidla
 - diagnostika bakteriálních onemocnění
 - ve 3/4 případů shoda s lidským expertem

Aktuální stav oblasti dialogových systémů

FI MU

- Laboratoře:
 - LSD - Laboratoř vyhledávání a dialogu – vedoucí - doc. Kopeček, prof. Zezula
 - zaměření:
 - vyhledávání
 - dialogové systémy
 - zpracování zvuku
 - asistivní technologie
 - ...
 - NLP - Laboratoř zpracování přirozeného jazyka – vedoucí doc. Pala
 - zaměření:
 - textové korpusy
 - slovníky
 - morfologie
 - syntaktická analýza
 - sémantická analýza

Aktuální stav oblasti dialogových systémů

ČR

- FIT VUT Brno:
 - analýza signálu
 - rozpoznávání řeči
 - systémy pro automatizovaný záznam a zpracování konferencí
 - ...
- ZČU Plzeň
 - rozpoznávání řeči
 - syntéza řeči
 - dialogové systémy
 - ...
- ČVUT Praha
 - syntéza řeči
 - počítačová lingvistika
 - ...

Aktuální stav oblasti dialogových systémů

Zbytek světa

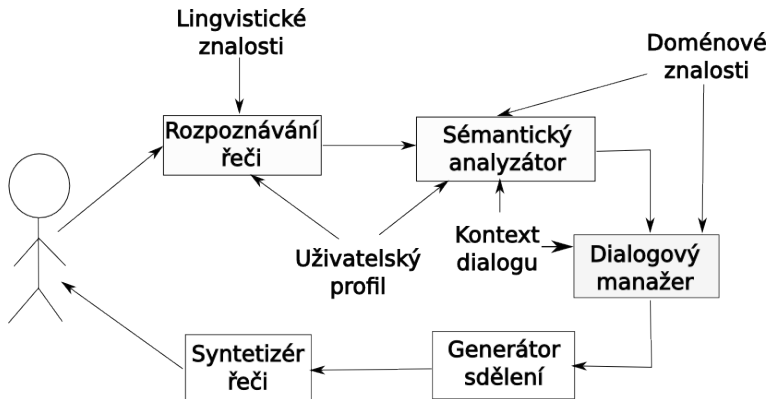
- World Wide Web Consortium Voice Browser Working Group
 - vývoj a správa standardů pro tvorbu dialogových rozhraní
 - vývoj a správa standardů pro tvorbu multimodálních dialogových rozhraní
 - členové:
 - IBM
 - Nuance Communication
 - Lucent Technologies
 - Motorola
 - ScanSoft
 - Tellme Networks
 - Vocalocity
 - ...

Aktuální stav oblasti dialogových systémů

Akademická sféra

- MIT
- Carnegie Mellon University (CMU)
- OGI
- EPF Lausanne
- ...

Struktura dialogového systému



Komponenty dialogového systému (1.)

- Uživatel - koncové zařízení, které zprostředkovává komunikaci uživatel ↔ dialogový systém.
 - telefon
 - IP telefon
 - IM klient
 - ...
- Rozpoznávání řeči:
 - Převádí mluvené slovo na text.
 - Využívá se:
 - rozpoznávání plynulé řeči
 - rozpoznávání izolovaných slov
 - Pro zvýšení úspěšnosti se používají lingvistické znalosti (n-gramy, gramatiky pro rozpoznávání řeči).
 - Používané standardy:
 - JSGF - definována v Java Speech API
 - SRGS - standard W3C Voice Browser WG

Komponenty dialogového systému (2.)

- Sémantický analyzátor:
 - Snaží se získat relevantní informace z rozpoznaného textu.
 - např.. hodnoty vstupních polí
 - Využívá db. lingvistických znalostí (atributové gramatiky, ...).
- Dialogový manažer
 - Řídí průběh komunikace s uživatelem na základě:
 - aktuálního stavu dialogu – informace o dosavadním průběhu (zadané údaje, chybovost uživatele, chybovost rozpoznávání řeči pro daného uživatele, ...)
 - použité dialogové strategie.
 - K implementaci lze využít standard VoiceXML

Komponenty dialogového systému (3.)

- Generátor promluv
 - Na základě sémantických informací od dialogového manažeru generuje smysluplnou promluvu pro uživatele.
 - Lze použít např. rámce a sloty.
 - rámec - předpřipravená promluva
 - slot - proměnlivá část rámce, doplněná informací od dialogového manažeru.
 - Výstup může být doplněn o prozodické informace.
 - Výstup může být popsán pomocí standardů:
 - SABLE
 - SSML
- Syntetizér řeči
 - Na základě promluvy od generátoru provádí syntézu dané promluvy.