

Úvod do počítačového zpracování řeči

Luděk Bártek

Fakulta informatiky
Masarykova univerzita

podzim 2014

Obsah

- 1 Dialogové systémy
- 2 Struktura dialogového systému
- 3 Dialogová komunikace

Motivace

- Dialogový systém - informační systém s dialogovým (hlasovým/textovým) rozhraním.
- Přirozenější způsob komunikace pro většinu uživatelů než GUI.
- Poskytují nové způsoby komunikace s aplikacemi:
 - telefon
 - hlasová komunikace prostřednictvím počítače s náhlavní soupravou (mikrofonem, reproduktory).
- Možnost komunikace bez použití končetin.
- Zlepšují přístupnost pro uživatele s různými druhy postižení:
 - zrakově postižení
 - imobilní uživatelé
 - ...
- Při dobře navrženém dialogovém rozhraní může být komunikace podobně efektivní jako GUI.
 - grafická komunikace - paralelní
 - hlasová komunikace - sekvenční

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

Historie dialogových systémů

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
- 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 auto společného jednoho z
 - 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 (<http://lsd.fi.muni.cz/>) – Laboratoř vyhledávání a dialogu
 - vedoucí – doc. Kopeček, prof. Zezula
 - zaměření – vyhledávání, dialogové systémy, zpracování zvuku, item asistivní technologie, ...
 - NLP (<http://nlp.fi.muni.cz>) – 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

Ve světě

- World Wide Web Consortium Voice Browser Working Group (<http://www.w3.org/Voice/>)
 - 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, ...
- MIT
- Carnegie Mellon University (CMU)
- OGI
- EPF Lausanne
- ...

Současné oblasti použití

- Dialogové informační systémy o dopravních spojeních.
- Dialogové ovládání některých systémů v automobilech:
 - navigace
 - telefon
 - ...
- Dialogové systémy pro zdravotně postižené -
- Vojenské aplikace.
- ...

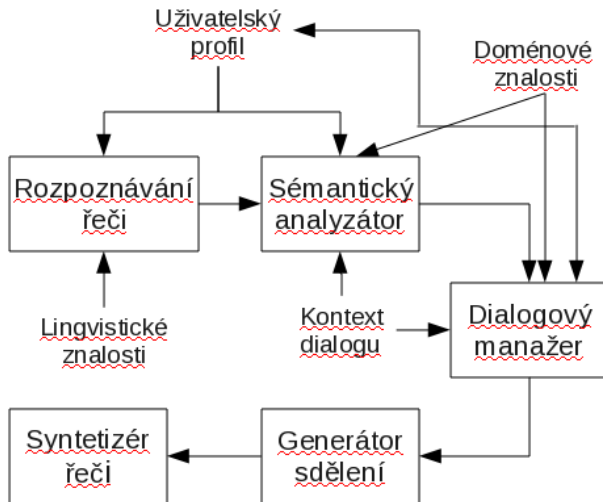
Základní pojmy

- Dialogový systém - informační systém disponující dialogovým rozhraním.
- Dialogové rozhraní - UI, které umožňuje uživateli komunikaci s aplikací prostřednictvím dialogu.
- Dialog - komunikace dvou účastníků (pro nás člověk - počítač).
- Promluva - souvislé sdělení jednoho účastníka dialogu.
- Obrat - promluva a reakce druhého účastníka na ni.
- Dialogová strategie - určuje ke každé promluvě následníka.
- Hodnotící funkce $E(L)$ - přiřazuje dialogu reálné číslo (jeho ohodnocení).
- Dialogová komunikace - uspořádaná čtveřice
 $M = (S_1, S_2, E_1, E_2)$.

Charakteristické rysy dialogových systémů

- Převládající řečová komunikace.
 - Problémy s rozpoznáváním řeči - řešení:
 - omezení problémové domény
 - gramatiky pro rozpoznávání řeči.
 - Bývá doplněna o vstup pomocí:
 - DTMF - telefonní aplikace
 - klávesnice - webové aplikace komunikující dialogem v přirozené řeči.
- Tendence ke komunikaci přirozenou řečí s co nejmenšími omezeními.
 - Vede na rozpoznávání plynulé řeči.
 - Řešení např. omezením množiny akceptovaných promluv pomocí gramatik.
- Snaha o co největší efektivitu a optimalitu komunikace.
 - Změny dialogové strategie v závislosti na zkušenostech uživatele - lze odhadnout z průběhu komunikace:
 - zkušený uživatel - průběh dialogu určuje spíše uživatel
 - nezkušený uživatel - DS se snaží uživatele co nejvíce vést

Dialogový systém



Komponenty dialogového systému

- Rozpoznávání řeči - rozpoznání promluvy
 - rozpoznání promluvy
 - ke zvýšení úspěšnosti využívá
 - lingvistické znalosti
 - uživatelský profil
- Sémantický analyzátor
 - zjištění významu promluvy
 - využívá:
 - uživatelský profil
 - doménové znalosti
 - kontext dialogu
- Dialogový manažer
 - na základě známých faktů rozhoduje o dalším kroku dialogu ze strany systému.
- Generátor sdělení
 - generuje sdělení podle požadavků dialogového manažeru.

Informace využívané dialogovým systémem

- Lingvistické znalosti - jazykový model, ...
- Uživatelský profil - model řečníka, emoční model, ...
- Doménové znalosti - informace o oblasti DS použitelné pro interpretaci rozpoznané promluvy, pro rozhodování dialogové strategie, ...
 - informace od oblasti dialogového systému
 - použitelné pro:
 - interpretaci rozpoznávané promluvy
 - rozhodování o dalším kroku dialogu (dialogovou strategií)
 - patří sem:
 - např. jaká data je zapotřebí zadat
- Kontext dialogu:
 - uchovává aktuální stav dialogu
 - zadané údaje
 - informace o chybách (nerozpoznané promluvy, chyby v zadávaných údajích)
 - lze použít např. pro použití vhodné gramatiky pro sémantickou interpretaci rozpoznané promluvy.

Dialogová strategie

- Zobrazení $U \times Q \rightarrow U \times Q$.
 - U - promluva
 - Q - stav
- Určuje následující krok dialogu v závislosti na stavu dialogu a vstupní promluvě.
- Hodnotící funkce dialogu E přiřazuje danému dialogu reálné číslo popisující úspěšnost dialogu z pohledu dané strany.
- Dělení dialogů z hlediska hodnotící funkce:
 - d je kooperativní dialog - $E_1(d) = E_2(d)$
 - d je nekooperativní dialog - $E_1(d) \neq E_2(d)$
 - d je dialog s nulovým součtem - $E_1(d) = -E_2(d)$

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

- Autorem Herbert Paul Grice - anglický jazykovědec
- Aspekt informativnosti
 - 1 Buď přiměřeně informativní (ne méně než je potřeba, ne více než je potřeba).
- Aspekt přesvědčivosti
 - 1 Neuváděj nepravdivé informace.
 - 2 Neuváděj informace, které nelze dokázat nebo doložit.
- Aspekt způsobu
 - 1 Informace v replice by měla být co nejvíce explicitní
 - 2 Vyhýbejte se nejednoznačností.
 - 3 Usilujte o stručnost.
 - 4 Buďte disciplinovaní, udržujte v dialogu pořádek.

Pravidla pro vedení kooperativního dialogu

Pokračování

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

Pravidla pro vedení kooperativního dialogu (H. P. Grice) - komunikace člověk počítač

- Aspekt asymetrie
 - 1 Informujte uživatele o všech důležitých charakteristikách, která vybočují z očekávaného normálního průběhu dialogu a která by měl vzít v úvahu k zajištění kooperativity.
 - 2 Zajistěte stručné avšak dostatečné informování uživatele o možnostech systému a jeho omezeních.
 - 3 Informujte srozumitelně a dostatečně uživatele o způsobu interakce ze systémem.

Pravidla pro vedení kooperativního dialogu (H. P. Grice) - komunikace člověk počítač

Pokračování

- Aspekt znalostí a schopností
 - 1 Vezměte v úvahu relevantní znalosti uživatele.
 - 2 Vezměte v úvahu možné uživatelské chybné analogie.
 - 3 Rozlišujte mezi začínajícím a zkušeným uživatelem systému.
 - 4 Vezměte v úvahu legitimní představy uživatele o znalostech a schopnostech systému.
- Aspekt vyjasňování a odstraňování chyb
 - 1 V případě selhání komunikace iniciujte metakomunikaci zajišťující odstranění chyby nebo její vysvětlení.
 - 2 Zajistěte vysvětlující metakomunikaci v případě nekonsistentních nebo nejednoznačných uživatelských vstupních dat.

Aspekty komunikace kooperativního dialogového systému

- aspekt informativnosti
- aspekt přesvědčivosti
- aspekt způsobu
- aspekt zdvořilosti, empatie a etiky
- aspekt asymetrie
- aspekt znalostí a schopností uživatele
- aspekt vyjasňování a odstraňování chyb

Iniciativa v dialogu

- Další krok dialogu je vždy určen dialogovou strategií jedné z komunikujících stran - jedna strana klade dotazy, druhá na ně odpovídá.
- V případě komunikace člověk - počítač lze rozlišit
 - dialog s iniciativou uživatele
 - dialog s iniciativou systému
 - dialog se smíšenou iniciativou.
- V reálném nasazení se používají:
 - dialogy se smíšenou iniciativou
 - dialogy s iniciativou systému.

Iniciativa v dialogu - příklady

- Dialog s iniciativou systému:

System: Zadejte Vaše křestní jméno

Uživatel: Jan

System: Zadejte Vaše příjmení

Uživatel: Novák

- Dialog s iniciativou uživatele:

Uživatel: Chtěl bych bych si rezervovat knihu XY

System: Dobře.

Uživatel: A film UV.

System: Dobře.

Uživatel: To je vše.

System: Vaše rezervace knihy XY a filmu UV byla přijata.

Iniciativa v dialogu - příklady

- Dialog se smíšenou iniciativou:

Uživatel: Chtěl bych si zaregistrovat předmět PB123.

Systém: S jakým zakončením?

Uživatel: Zkouška.

Systém: Registruji Vám předmět PB123 se zakončením
zkouškou. Souhlasí?

Uživatel: Ano.

Zpětná vazba v dialogovém systému

- Před tím, než systém předá získané informace k dalšímu zpracování je vhodné provést jejich verifikaci.
 - chyby rozpoznávání řeči
 - chyba uživatele
 - ...
- Způsoby ověření získaných dat
 - sumarizující zpětná vazba
 - zpětná vazba "echo"
 - implicitní zpětná vazba
 - explicitní zpětná vazba.
- V případě nesouhlasné reakce uživatele následuje opravný dialog.

Ukázky zpětné vazby

- Sumarizující zpětná vazba:

Uživatel: Jmenuji se Jan Novak a narodil jsem se
8.8. 1988 v ABC.

System: Bylo zadáno jméno Jan Novák, datum narození
8.8. 1988 a místo narození ABC.

- Zpětná vazba "echo":

System: Jak se jmenujete?

Uživatel: Jan Novák.

System: Jmenujete se Jan Novák. Kdy jste se narodil.

Uživatel: 8.8. 1988

System: Narodil jste se 8.8. 1988. Ve kterém městě?

Uživatel: ABC System: Narodil jste se v městě ABC.

Ukázky zpětné vazby

Dokončení

- Implicitní zpětná vazba:

System: Jméno studenta.

Uživatel: Jan Novák.

System: Ve kterém městě se Jan Novák narodil?

...

- Explicitní zpětná vazba:

System: Zadejte jméno studenta.

Uživatel: Jan Novák

System: Student se jmenuje Jan Novák, je to tak?

Ukázka opravného dialogu

Uživatel: Jmenuji se Jan Horák a narodil jsem se
8.8. 1988 v ABC.

Systém: Bylo zadáno jméno Jan Novák, datum narození
8.8. 1988 a místo narození ABC.

Uživatel: Ne. Jméno je Jan Horák.

Systém: Zadal jste jméno Jan Horňák, datum narození
8.8. 1988 a místo ABC.

Uživatel: Ne, jméno má být Jan Novák.

Systém: Zadal jste jméno Jan Novák, datum narození
8.8. 1988 a místo ABC.

Další aspekty dialogových strategií

- Možnost přerušit systém.
- Možnost iniciovat korekce a metakomunikaci uživatelem. Nerozumím, mohl byste to prosím zopakovat?
- Možnost získání nápovědy. Nevím, co po mě požadujete.
- Detekce znalostí a schopností uživatele.
 - Částečně lze odvodit z předchozího průběhu dialogu.
 - Uzpůsobení iniciativy dialogové strategie - iniciativa systém vs. smíšená iniciativa.

Další aspekty dialogových rozhraní

- Přizpůsobení dialogové strategie - např. uživatel ve spěchu, vystresovaný uživatel, ...
- Detekce emocí uživatele:
 - na základě neverbálních charakteristik hlasu
 - nutná podpora v modulu rozpoznávání řeči.
- Vícejazyčnost (multilingualita):
 - nutnost podpory na straně použité platformy:
 - rozpoznávání řeči
 - sémantická analýza řeči
 - syntéza řeči.
 - Možnost pokusit se o automatickou detekci použitého jazyka.
 - Paralelní rozpoznávání všemi dostupnými rozpoznávači - použije se výstup s maximální pravděpodobností.

Další aspekty dialogových rozhraní

- Multimodalita:
 - Umožňuje paralelní komunikaci více kanály - (obraz, zvuk, hmat).
 - Zlepšuje přístupnost
 - Příklady multimodálních rozhraní:
 - Rozhovor vede avatar (talking head) - vhodné pro uživatele s poruchou slyšení.
 - Ruce/avatar (celé tělo resp. horní polovina) - provádí tlumočení do znakové řeči.
 - Alternativní způsoby vstupu - klávesnice, kamera, snímače aktivity mozku, svalů (krk, obličej, ...), různé joysticky, ...
 - Nutnost synchronizace jednotlivých kanálů.

Další aspekty dialogových rozhraní

Pokračování

- Zdvořilost - viz pravidla vedení kooperativního dialogu
- Prozódie - určení sémantiky a pragmatiky promluvy:
 - určení druhu věty (tázací
(data/masse-dotaz.wav)/oznamovací
(data/masse-ozn.wav)), ...
 - detekce emocí
 - ...
- Učení se z chyb.
 - Zapamatování si nerozpoznané promluvy a pokud uspěje opravný dialog (zpětná vazba), pokus o analýzu původní promluvy a přidání typu promluvy do lingvistických znalostí.