

Dialogové systémy

Luděk Bártek

Laboratoř vyhledávání a dialogu, Fakulta Informatiky Masarykovy Univerzity,
Brno

jaro 2022

Multimodální dialogová rozhraní

Dialogové
systémy

Luděk Bártek

Multimodální
dialogová
rozhraní

- Multimodální dialogové rozhraní – mimo mluvenou řeč umožňuje alternativní způsoby komunikace člověk ↔ počítač:
 - textová komunikace
 - grafická komunikace
 - emoce
 - ...
- Multioborové odvětví:
 - počítačová akustika a fonetika
 - počítačová grafika
 - lingvistika
 - psychologie
 - ...

- Výhody:
 - lepší přístupnost – např. pro uživatele s poruchami sluchu/řeči.
 - možnost přesnějšího odvození pragmatiky projevu,
 - běžný způsob komunikace člověk ↔ člověk
 - při komunikaci člověk ↔ člověk se uplatňuje:
 - řečový kanál
 - vizuální kanál
 - emoce
 - ...
 - ...

Multimodální rozhraní

Jednoduché modalitní komponenty

Dialogové
systémy

Luděk Bártek

Multimodální
dialogová
rozhraní

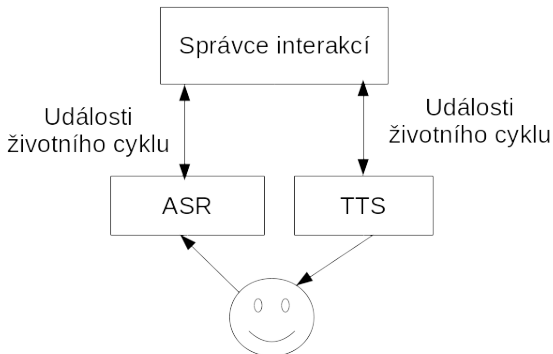
- Atomická modalitní komponenta — nelze ji rozdělit na více jednoduchých komponent.
- Příklady jednoduchých komponent:
 - výstup: TTS, audio výstup, reprezentace GUI (např. HTML), video výstup, ...
 - vstup: nahrávání zvuku, ASR, DTMF, klávesnice, myš, GPS, ...

Schéma komunikace jednoduchých komponent

Dialogové
systémy

Luděk Bártek

Multimodální
dialogová
rozhraní



Obrázek: Ukázka jednoduchých komponent (převzato ze specifikace Multimodal Architecture and Interfaces)

Multimodální rozhraní

Komplexní modalitní komponenty

Dialogové
systémy

Luděk Bártek

Multimodální
dialogová
rozhraní

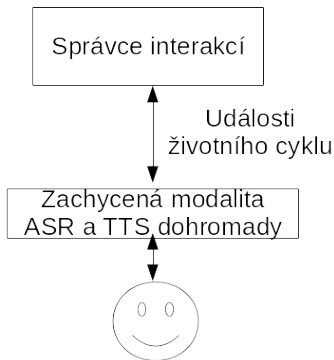
- Může obsahovat funkcionalitu dvou a více jednoduchých komponent.
- Příklady komplexních komponent:
 - GUI – prezentuje informace uživateli a zpracovává události od uživatele
 - VXML – prezentuje informace pomocí TTS a zachytává vstup od uživatele pomocí ASR.
 - ...

Schéma komunikace komplexní komponenty

Dialogové
systémy

Luděk Bártek

Multimodální
dialogová
rozhraní

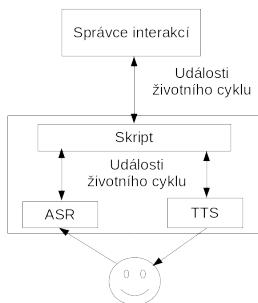


Obrázek: Ukázka komplexní modalitní komponenty (převzato ze specifikace Multimodal Architecture and Interfaces)

Multimodální rozhraní

Zanořené modalitní komponenty

- Zanořená modalitní komponenta = množina modalitních komponent a skript (např. v SCXML), který je spravuje.



Obrázek: Ukázka vnořených modalitních komponent (převzato ze specifikace Multimodal Architecture and Interfaces)

Multimodální komunikace člověk → počítač

Dialogové
systémy

Luděk Bártek

Multimodální
dialogová
rozhraní

- Velmi široké spektrum možností zadávání vstupu jinak než hlasem:
 - klávesnice (počítač, DTMF, SMS, ...)
 - rukou psaný vstup – dotyková obrazovka + pero
 - různá ústy ovládaná zařízení
 - ovládání pomocí pohybů očí a víček
 - rozpoznávání řeči pomocí sond detekujících činnost svalů a mozku [?] – pomocí EEG, EMG, ...
 - rozpoznávání emocí – pomáhá při určení pragmatiky projevu
 - ...
- Často jako doplněk řečového vstupu.

Multimodální komunikace počítač → člověk

Dialogové
systémy

Luděk Bártek

Multimodální
dialogová
rozhraní

- Textová:
 - Mimo hlasový výstup je navíc zobrazen odpovídající textový výstup.
 - Lze využít např. prostředky pro IM, SMS,
- Grafická:
 - Talking Heads – mimo hlasový výstup je navíc zobrazena tvář, jejíž pohyby, zejména úst odpovídají řeči (ukázky[?]).
 - Komunikace znakovou řečí
 - Zobrazeny ruce, které se pohybují synchronně s řečí.[?]
 - Často jako doplněk mluvící hlavy nebo formou avatara.
- Emoce – emoce získány způsobem popsáním dříve.

- Proprietární řešení:
 - součást CSLU Toolkitu [?]
 - projekt August[?]
- Otevřená řešení:
 - Doporučení W3C týkající se multimodálního přístupu – existují platformy, které některé z nich implementují.
 - Využívají a propojují i další standardy W3C (CCXML, XHTML, VoiceXML, SVG, SMIL, ...).
 - Výstup W3C Multimodal Interaction Working Group

Návrhy standardů Multimodal Interaction WG

Dialogové
systémy

Luděk Bártek

Multimodální
dialogová
rozhraní

- Multimodal Architecture and Interfaces – popis architektury a podporovaných rozhraní.
- EMMA (Extensible MultiModal Annotation Markup Language) – standard pro výměnu údajů mezi jednotlivými komponentami multimodálního dialogového systému.
- InkML – standard pro data zadaná pomocí dotykových obrazovek a elektronického pera.
- EmotionML – slouží k anotování dat z pohledu emocí.
- ...

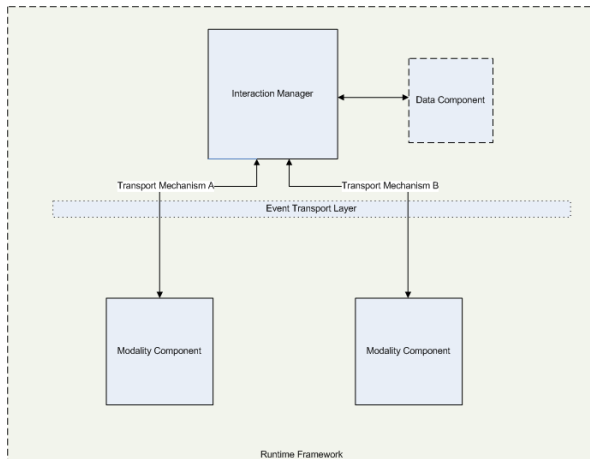
Multimodal Initiative WG

Návrh architektury systému

Dialogové
systémy

Luděk Bártek

Multimodální
dialogová
rozhraní



Obrázek: Návrh architektury multimodálního systému (převzato ze specifikace W3C)

Multimodal Architecture and Interfaces

Dialogové
systémy

Luděk Bártek

Multimodální
dialogová
rozhraní

- Specifikace W3C MultiModal Interaction Working Group
- Specifikace popisuje:
 - architekturu multimodálních rozhraní (viz 13)
 - protokoly pro komunikace mezi komponentami rozhraní
 - události během životního cyklu rozhraní, které musí rozhraní a komponenty podporovat (viz specifikace).

Extensible Multi-Modal Annotations (EMMA)

Dialogové
systémy

Luděk Bártek

Multimodální
dialogová
rozhraní

- Specifikace W3C MultiModal Interaction Working Group
- Specifikace popisuje:
 - jazyk pro reprezentaci interpretaci uživatelských vstupů, anotace hodnocení důvěryhodnosti vstupů, časových razítek, vstupních kanálů, . . .

Ink Markup Language (InkML)

Dialogové
systémy

Luděk Bártek

Multimodální
dialogová
rozhraní

- Specifikace W3C MMI WG
- Specifikace popisuje značkování pro reprezentaci vstupů pomocí elektronických per a stylusů.
 - popisuje vždy posloupnost tahů, každý ve formě posloupnosti souřadnic tahu ve 2D prostoru.

Emotion Markup Language (EmotionML)

Dialogové
systémy

Luděk Bártek

Multimodální
dialogová
rozhraní

- Specifikace W3C MMI WG
- Slouží k:
 - manuálnímu a automatickému popisu materiálů obsahujícímu emoce
 - automatickému rozpoznávání emocí ze senzorů
 - generování emocí v odpovědích systému.

- Jednoduchý jazyk pro popis konverzačních agentů, který by měl být:
 - otevřený
 - snadno naučitelný
 - multilinguální
 - flexibilní a rozšiřitelný
 - ...
- První návrh z 90. let 20. století.
- Správa AIML Foundation
- Interprety:
 - Pandora Bots
 - Bot Libre
 - knihovna pro Python
 - interpret v Node.js
 - referenční implementace v Javě

-  Massaro, D., Cohen, M. M. – Demos From The Perceptual Science Lab, dostupné na adrese <http://mambo.ucsc.edu/demos.html> (květen 2011).
-  Guimeraes, K., Antunes, D. R., Guilhermino, de F. Guilhermino, D., Lopes da Silva, R. A., Garcia, L. S. – 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.
-  Barnard, E. et al – CSLU Toolkit, dostupné na adrese <http://www.cslu.ogi.edu/toolkit/index.html> (květen 2011).
-  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.



the August Home Page,
<http://www.speech.kth.se/august/> (květen 2011).