

Dialogové systémy

Luděk Bártek

Fakulta informatiky
Masarykova univerzita

podzim 2014

Obsah

- 1 Implementace dialogových systémů
- 2 Používané standardy
- 3 Nástroje pro tvorbu dialogových rozhraní

Přístup k dialogovému rozhraní

- **Telefonní:**
 - PSTN (Public Switched Telephone Network) - veřejná telefonní síť.
 - Nutnost digitalizace uživatelského vstupu a připojení počítače k PSTN:
 - voice-modem
 - ISDN modem
 - výstup lze již přímo zpracovat pomocí ASR.
 - VoIP - protokol pro přenos hlasu přes IP.
 - Většinou jako rozšíření ústředny o dialogový manažer (platformu).
 - Např. Asterisk + VoiceGlue, ...
 - komunikace prostřednictvím VoIP protokolů
 - Možnost využití DTMF (Dual Tone Multi-Frequency).

Přístup k dialogovému rozhraní

- Textové:
 - Odpadá nutnost digitalizace uživatelského vstupu.
 - Velmi vhodný pro ladění rozhraní:
 - není nutno řešit chyby ASR, ...
 - Lze využít i IM
 - Asterisk + XMPP (Jabber).

Části dialogových systémů

- Rozpoznávání řeči:
 - ASR - nejlépe s podporou rozpoznávání plynulé řeči
 - Pro zvýšení úspěšnosti jsou použity bezkontextové gramatiky - občas slouží jako základ i pro sémantickou interpretaci.
 - Lze částečně nahradit pomocí DTMF.
- Sémantická interpretace:
 - Atributy se sémantickou interpretací u bezkontextových gramatik.
 - Občas se používá keyword spotting.
- Dialogový manažer:
 - logické programování
 - různá řešení pomocí procedurální programovacích jazyků
 - proprietární řešení
 - otevřená řešení.

Části dialogových systémů

- Generátor promluv:
 - pracuje s výstupem z dialogového manažeru
 - bývá součástí dialogového manažeru
 - generuje textovou verzi promluvy.
- Hlasový syntetizér:
 - TTS.

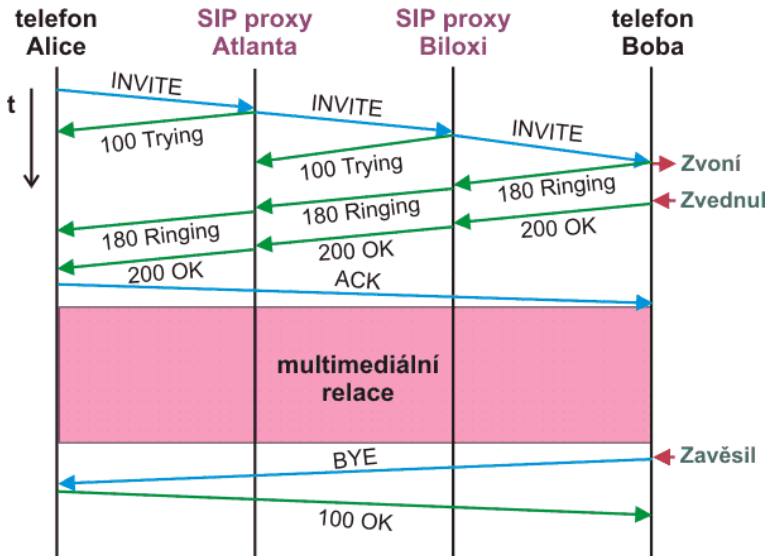
VoIP

- Rodina protokolů pro přenos hlasu přes internet (IP síť).
- Využívá se např. pro IP telefonii.
- Na transportní vrstvě - UDP.
- Na relační vrstvě - Real-Time Protocol (RTP).
- Řada implementací
 - Liší se použitými standardy:
 - H.323 - na ústupu
 - SIP
 - firemní - Skinny (Cisco), HFA - Siemens.
 - Službami a signalizací:
 - podpora video hovorů, zasílání textových zpráv, ...
- Více viz <http://www.protocols.com/pbook/VoIP.htm>

- Session Initiation Protocol (viz SIP RFCs and Drafts, resp. IETF, sbírka Internetových standardů na FI, ...)
- Protokol na aplikační vrstvě.
- Protokol určený pro přenos signalizace v internetové telefonii.
- Na transportní vrstvě využívá UDP.
- Vznikl jako reakce na H.323 - příliš komplexní (obtížně implementovatelný).
 - Je jednodušší.

- Pro vytvoření a řízení multimediální relace zajišťuje:
 - Lokalizaci účastníka (jednoznačná identifikace - uživatel, adresa SIP serveru).
 - Zjištění stavu účastníka - dostupný, obsazený, přesměrovaný.
 - Zjištění možností účastníka - použitelný kodek, max. přenosová rychlost, použitelnost video hovorů, ...
 - Navázání spojení:
 - využívá se protokol SDP popisující navázané spojení (Session Description Protocol)
 - vlastní hovor využívá protokol RTP (Real-Time Protocol).
 - Řízení probíhajícího spojení:
 - změny parametrů spojení v jeho průběhu
 - ukončení spojení.

Průběh relace přes protokol SIP



Podpora rozpoznávání řeči

- Omezení domény možných vstupů.
- Bezkontextové gramatiky popisující množinu možných vstupů.
- Používané způsoby:
 - prostředky logického programování
 - různá proprietární řešení
 - otevřené standardy
 - JSGF
 - SRGS
 - ...

Java Speech Grammar Specification (JSGF)

- Textový zápis gramatiky nezávislý na platformě a prodeji, sloužící pro podporu rozpoznávání řeči.
- Určen pro použití při rozpoznávání řeči.
- Používá styl a konvence styl a konvence jazyka Java.
- Součást Java Speech API.
- Aktuální verze 1.0 (říjen 1998).
- Využit např. v rozpoznávači Sphinx4J, VoiceXML interpretru VoiceGlue, ...

JSGF - Syntaxe

- Gramatika se skládá z pravidel, které popisují co může být řečeno.
- Syntaxe je case-sensitive.
- Kódování znaků - Unicode.
- Formát hlavičky:

```
#JSGF version [char-encoding [locale]];  
#JSGF V1.0;  
#JSGF V1.0 ISO8859-2;  
#JSGF V1.0 UTF-8 cs_CZ;
```

JSGF - Syntaxe

- Názvy pravidel nesmí obsahovat bílé znaky.
- Neterminální symboly:
 - <názevNeterminálu>, <mesto>, <anone>
- Terminální symboly
 - víceslovní terminály a zvláštní symboly mohou být uzavřeny do uvozovek: "Nové Město na Moravě" "+"
- Zvláštní pravidla
 - <NULL> - pravidlo, které je automaticky použito, aniž by uživatel cokoliv řekl
 - <VOID> - pravidlo, které nemůže být řečeno
- Deklarace gramatiky:
 - grammar *názevBalíku.názevGramatiky*;

JSGF - syntaxe (pokračování)

- Vkládání gramatik - umožňuje používat pravidla nebo gramatiky definované v jiném souboru.
 - `import fullyQualifiedRuleName;`
 - `import fullGrammarName; import <com.sun.speech.app.numbers.one>; import <com.sun.speech.app.numbers.*>;`
- Deklarace gramatiky
 - `grammar názevBalíku.názevGramatiky;`
- Tělo gramatiky
 - *neterminál = pravidlo;*
`<jmeno> = Jan | Jana | ...;`
`<jmeno> = <krestniJmeno> <prijmeni>;`
- Více viz specifikace.

JSGF - ukázka

JSGF V1.0

```
#import cz.mesta.*;
```

```
#import cz.hodiny.*;
```

```
<koren> = Chci jet <cim>. | Chci jet <cim> z <odkud> do <kar  
        Chci jet <cim> z <odkud> do <kam> v <kdy>;
```

```
<cim> = vlakem | autobusem;
```

```
<odkud> = <czMesto>;
```

```
<kam> = <czMesto>;
```

```
<kdy> = <czCas>;
```


SRGS

- Standard W3C Voice Browser Activity WG.
- Aktuální verze 1.0 (březen 2004).
- Definuje způsob zápisu pravidel a jejich odkazování.
- Dva způsoby zápisu
 - XML
 - ABNF (Augmented BNF).
- Více později při probírání VoiceXML

Sémantická interpretace promluvy

- Většinou řešeno pomocí atributů v gramatice pro rozpoznávání řeči.
- Slouží k určení umístění a hodnoty významných částí uživatelské promluvy.
- JSGF:
 - K pravidlu je přiřazena jeho sémantická interpretace.
 - Zapisuje se:
`{sémantická interpretace};`

- Příklad:

```
<souhlas> = <ano> {ano} | <ne> {ne};
```

```
<ano> = ano | jo | jasně;
```

```
<ne> = ne | ani náhodou;
```

Sémantická interpretace promluvy

Pokračování

- SRGS:
 - K pravidlu je přiřazena jeho sémantická interpretace.
 - Používá se standard Semantic Interpretation for Speech Recognition (SISR)
 - Sémantická interpretace může obsahovat výrazy v jazyce (ECMAScript).
 - Skriptovací jazyk standardizovaný organizací ECMA (European Computer Manufacturer Association).
 - Používá se pro skriptování na straně klienta ve webových stránkách.
 - Implementace - JavaScript, JScript, ActionScript.
 - Jednotlivé implementace přidávají nestandardní knihovny (práce s prohlížečem, ...)
 - Více později při probírání VoiceXML.

Popis dialogových rozhraní

- Popis ve vyšším programovacím jazyce:
 - prostředky logického programování (Prolog)
 - procedurální programovací jazyky (C/C++, Java, ...) - např. projekt AudiC (LSD FI),
- Proprietární řešení.
- Otevřené standardy:
 - VoxML – předchůdce VoiceXML (2. polovina 90. let)
 - VoiceXML – součást W3C Voice Browser Activity
 - CallXML – Voxeo Prophecy

Syntéza řeči

- Tvorba promluvy:
 - 1 Dialogový manažer zvolí rámec pro požadovanou výstupní promluvu.
 - 2 Doplní se do ní hodnoty slotů.
 - 3 Předá se řečovému syntetizéru.
- Značkování prozodických jevů:
 - závislé na použitém TTS
 - Speech Synthesis Markup Language (SSML) - součást standardů W3C VoiceBrowser Activity
 - TTS musí obsahovat podporu pro tento standard.

Online nástroje pro tvorbu dialogových rozhraní

- Webové nástroje
 - Nuance (BeVocal) Café
 - podporované standardy: VoiceXML 2.1, SRGS 1.0
 - obsahuje sadu nástrojů pro ladění dialogových rozhraní
 - Tellme Studio
 - podporované standardy: VoiceXML 2.x, SRGS + SISR
 - Voxeo Prophecy
 - podporované standardy: CalXML, VoiceXML 2.0, SRGS, CCXML
 - další viz např. seznamy na W3C VoiceBrowser Activity nebo na Wikipedii

Desktopové nástroje pro tvorbu dialogových rozhraní

- Trinidkit
 - toolkit pro tvorbu dialogových rozhraní založený na logickém programování (Sicstus Prolog)
- CSLU Toolkit
 - vývojové prostředí pro tvorbu dialogových rozhraní
 - umožňuje snadnou tvorbu i multimodálních dialogových rozhraní
 - pro Win32
- Voxeo Prophecy
 - on-line VoiceXML platforma - bezplatně max. 2 připojení k hostovanému rozhraní
 - možnost instalace serveru na vlastní počítač
 - podobné možnosti jako online verze

Desktopové nástroje pro tvorbu dialogových rozhraní

- Textový editor, nejlépe se zvýrazněním syntaxe + (desktopová) VoiceXML platforma:
 - JVoiceXML
 - PublicVoiceXML
 - OptimTalk
 - ...
- Další viz např. seznamy na W3C Voice Browser Activity nebo na Wikipedii.