

Dialogové systémy

Luděk Bártek

Fakulta informatiku
Masarykova univerzita

podzim 2013

Obsah

- 1 Speech Recognition Grammar Specification
- 2 Semantic Intepretation for Speech Recognition
- 3 Ostatní standardy W3C VoiceBrowser Activity

SRGS

- W3C specifikace jazyka pro zápis bezkontextových gramatik pro podporu rozpoznávání řeči.
- Aktuální verze 1.0 (březen 2004).
- Nahradil původně používaný standard JSGF.
- Dvě varianty zápisu gramatiky:
 - XML
 - Augmented Backus-Naur Form (ABNF).
- Liší se pouze zápis nikoliv vyjadřovací síla.
- Možnost použití způsobu zápisu závisí na použité platformě.
 - Vetší podpora pro XML formát.

Gramatiky (Teorie Formálních Jazyků)

- Gramatika $G = (N, \Sigma, P, S)$
 - N - množina neterminálních symbolů
 - Σ - abeceda
 - P - množina pravidel
 - S - kořenový neterminál.
- Bezkontextová gramatika
 - gramatika $G = (N, \Sigma, P, S)$
 - pravidla ve tvaru: $N \rightarrow \{N \cup \Sigma\}^*$.

XML Formát Gramatiky

- XML prolog.
- Kořenový element - grammar.
- Atributy:
 - root - pravidlo odpovídající kořenovému neterminálnímu symbolu
 - xml:lang - jazyk gramatiky
 - version - použitá verze SRGS (aktuálně pouze 1.0)
 - mode
 - dtmf
 - voice - implicitní hodnota
 - ...
- Element grammar - obsahuje množinu pravidel (elementů rule).

ABNF Gramatika

- ABNF hlavička
 - `#ABNF verze [kódování]`
`#ABNF 1.0 ISO-8859-2`
 - `root $ jméno pravidla;` - kořenový neterminální symbol
 - `language jazyk;`
 - `mode voice|dtmf;`
`#ABNF 1.0 UTF-8 root $pozdrav;`
`language cs-CZ;`
`mode voice;`

Zápis pravidla

- Levá strana pravidla:
 - XML formát
 - element rule
 - atribut id - jednoznačný identifikátor pravidla
 - obsah - pravá strana pravidla

```
<rule id="pozdrav"> ahoj </rule>
```
 - ABNF
 - `<id pravidla>`
 - `$pozdrav = ahoj;`

Zápis pravidla

- Pravá strana pravidla

- může obsahovat terminální a neterminální symboly:

- sekvenci
- varianty

- XML formát

- tělo elementu rule

```
<rule id="vstup">  
  Proved <ruleref uri="#prikazy"/>  
  s parametry <ruleref uri="#parametry"/>.  
</rule>
```

- ABNF

- $\$ \langle \text{neterminál} \rangle = \langle \text{pravá strana} \rangle$

$\$ \text{vstup} =$

Proveď $\$ \langle \text{http://www.nekde.cz/grammar.gram\#prikazy} \rangle$

s parametry $\$ \langle \text{http://www.nekde.cz/grammar.gram\#parametry} \rangle$

nebo

$\$ \text{vstup} = \text{Proved } \$ \text{prikazy s parametry } \$ \text{parametry}$

Sekvence

- Posloupnost terminálních a neterminálních symbolů.
 - $X \rightarrow YZa$
- Lze ji rozdělit na logické části.
- XML zápis:
 - zapsat přímo

```
<rule id="spojeni">
  Chci jet z <ruleref uri="#misto"/>.
</rule>
```
 - dělení na logické části
 - využitelnost
 - počet opakování dané části (atribut repeat)
 - sémantická interpretace

Sekvence

Ukázky

- XML Formát:

```
<rule id="spojeni">  
  Chci jet <item>z <ruleref uri="#misto"/> </item>  
  <item> do <ruleref uri="#misto"/> </item>  
  <item> <ruleref uri="#druh"/></item>  
  <item> <ruleref uri="#datum"/></item>  
  <item> v <ruleref uri="#cas"/></item>  
</rule>
```

- ABNF zápis:

```
$spojeni = Chci jet z $misto do $misto $druh  
          $datum v $cas.
```

Varianty

- umožňují uživateli zadat jeden z možných vstupů
 - $X \rightarrow Y|Z|a$

- XML zápis:

```
<rule id="barvy">
  <one-of>
    <item>cervena</item>
    <item>zelena</item>
    <item>modra</item>
  </one-of>
</rule>
```

- ABNF zápis

```
$barvy = (cervena|zelena|modra)
```

Opakování

- Umožňuje specifikaci:
 - nepovinných částí promluvy
 - opakovaných částí promluvy
- XML zápis
 - pomocí atributu repeat u elementu item

```
<rule id="adresa">  
  www  
  <item repeat="1-2">  
    tečka <ruleref uri="#castAdresy"/>  
  </item>  
  tečka <ruleref uri="#tld"/>  
</rule>
```

Opakování

- ABNF zápis
 - za prvek uvedeme počet opakování uzavřený do $\langle \rangle$
\$adresa = www \$castAdresy $\langle 1-2 \rangle$ \$tld
\$castAdresy = tecka \$text
- počet opakování
 - číslo - *číslo* krát
 - číslo1- číslo2 - *číslo1* - *číslo2* krát
 - číslo- - *číslo* - ∞ krát

Zvláštní pravidla

- GARBAGE - odpovídá libovolné promluvě až po následující bližší specifikovanou část
- VOID - pravidlo, které nelze vyslovit (zakázání určité promluvy)
- NULL - pravidlo, které je vždy rozpoznáno (může být i prázdné)
- XML formát:
 - ```
<ruleref special="pravidlo" / >
<rule id="spojeni">
 <ruleref special="GARBAGE"/>
 z <ruleref uri="#misto"/> do <ruleref uri="#misto"/>
 <ruleref uri="#prostredek"/>
</rule>
```
- ABNF
  - $\$pravidlo$   
 $\$spojeni = \$GARBAGE z \$misto do \$misto \$prostredek$

# Literatura

- Specifikace W3C.
- Příklady použité na přednášce.

# SISR

- Sémantika – přiřazuje význam slovům a promluvám.
- Sémantika v dialogových systémech – přiřazuje význam promluvám a jejich částem
- SISR - standard W3C pro zpracování sémantiky promluvy.
  - aktuální verze 1.0
  - publikován - duben 2007
  - úzce spjat se standardy
    - ECMAScript
    - SRGS
- Umožňuje přiřazení základních interpretací částem promluvy a vytváření odvozených interpretací pro nadřazená tvrzen
  - přiřazení interpretace částem promluvy
  - odvozování interpretace na základech dílčích interpretací
  - přiřazení interpretace vstupním polím dialogu

## Přřazení interpretace části promluvy

- Sémantická interpretace bývá součástí pravidla SRGS.
- K pravidlu přřazena pomocí elementu/atributu tag.
- XML formát SRGS gramatiky:
  - element tag

```
<item>
 <ruleref uri="souhlas"/><tag>{out='ano'}</tag>
</item>
```
  - atribut tag elementu item

```
<item tag="ano">jo</item>
```
- ABNF tvar:
  - uveden za interpretovanou část promluvy
  - tvar: interpretace

```
$souhlas = jo {ano}
```

## Odvozování interpretace na základě dílčích interpretací

- Zápis pomocí výrazů v jazyce ECMAScript.
- Přiřazeno k pravidlům pomocí elementu tag.
- Interpretace reprezentována pomocí objektů jazyka ECMAScript.
- Stínové proměnné:
  - pro pravidla - objekt *rules*
  - výstup - objekt *out*

# Odvozování interpretace na základě dílčích interpretací

## XML formát SRGS gramatiky

```
<rule id="vlastnictvi">
 <item>Mám
 <item repeat="0-1">
 <ruleref uri="#barva"/>
 </item>
 <ruleref uri="prostredek"/>
 <tag>{out = rules.barva + ';' + rules.prostredek;}</tag>
</item>
</rule>
```

# Odvozování interpretace na základě dílčích interpretací

## ABNF gramatika

```
$vlastnictvi = mam $barva <0-1> $prostredek
 {out = rules.barva + ',' + rules.prostredek;};
$barva = (cervenou {cervena}
 | cervene{cervena}
 | zelenou{zelenou}
 | zelene{zelena});
$prostredek = (auto{auto} | kolobezku{kolobezka});
```

## Přiřazení interpretace vstupním polím VoiceXML formuláře

- Využívají se atributy stínového objektu out
- XML formát:

```
<rule id="vlastnictvi">
 <item>
 Mám <item repeat="0-1"><ruleref uri="#barva"/>
 <ruleref uri="#prostredek"/>
 <tag>
 {
 out.barva = rules.barva;
 out.prostredek = rules.prostredek;
 }
 </tag>
 </item>
</rule>
```

## Přiřazení interpretace vstupním polím VoiceXML formuláře

- ABNF gramatika:

```
$vlastnictvi = mam $barva <0-1> $prostredek
{
 out.barva = rules.barva;
 out.prostredek = rules.prostredek;
};
$barva = (cervenou {cervena}|
 cervene{cervena}|
 zelenou{zelena}|
 zelene{zelena});
$prostredek = (auto{auto} | kolobezku{kolobezka});
```

# Literatura

- Specifikace SISR
- ECMAScript
- Příklady použité na přednášce.

## Další standardy W3C VoiceBrowser Activity

- SSML
- Pronunciation Lexicon Specification
- Call Control XML
- State Chart XML

# Speech Synthesis Markup Language

- Značkovací jazyk pro podporu syntetizované řeči ve webových aplikacích.
- Standard W3C
- Aktuální verze 1.0 (září 2004)
- Vychází z JSGF/JSML (JSpeech Markup Language)
- Cíle:
  - musí umožňovat konzistentní ovládání hlasového výstupu řečovým syntetizérem.
  - musí dovolovat TTS pro co nejširší škálu aplikací a domén
  - musí být internacionalizovaný
  - musí být snadno použitelný pro psaní dokumentů
  - musí být implementovatelný pomocí stávajících technologií
  - JSML dokumenty musí být lidsky čitelné.
- Zbytek viz syntéza řeči.

# 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 W3C

# Call Control XML

- Standard W3C
- Aktuální verze 1.0 (červenec 2011)
- navržen pro ovládání telefonních hovorů z dialogových systémů
- Specifikace W3C
- Umožňuje:
  - sestavení a ovládání konferenčních hovorů
  - přesměrování hovoru
  - ...