

Dialogové systémy

Luděk Bártek

Laboratoř vyhledávání a dialogu, Fakulta Informatiky Masarykovy Univerzity,
Brno

jaro 2012

- 1876 – udělen patent na telefon A. G. Bellovi
- WWW
 - 1989 – po CERN koluje článek HyperText and Cern (Tim Burnes Lee)
 - Vánoce 1990 – demonstrován řádkový webový prohlížeč a editor.
 - 1991 – všeobecná dostupnost WWW na počítačích v CERN.
 - 1994 – první setkání W3 konsorcia.

W3C Voice Browser Activity

Dialogové
systémy

Luděk Bártek

W3C Voice
Browser
Activity

SRGS

XML Formát
SRGS

ABNF formát
SISR

- 1999 – založena W3C Voice Browser Working Group.
- Cíl – návrh standardů umožňujících přístup k Webu pomocí hlasu a telefonu.
- Členové:
 - HP
 - Nuance Communications
 - Lucent Technologies
 - Motorola
 - ScanSoft
 - IBM
 - Tellme Networks
 - Vocalocity
 - ...

Standardy W3C Voice Browser Activity

Dialogové
systémy

Luděk Bártek

W3C Voice
Browser
Activity

SRGS

XML Formát
SRGS

ABNF formát
SISR

- VoiceXML – jazyk pro popis dialogových strategií.
- Speech Recognition Grammar Specification – jazyk pro zápis gramatik pro podporu rozpoznávání řeči.
- Semantic Interpretation for Speech Recognition – jazyk pro podporu sémantické interpretace.
- Speech Synthesis Markup Language – jazyk pro popis prozodických charakteristik pro syntézu řeči.

Standardy W3C Voice Browser Activity

pokračování

Dialogové
systémy

Luděk Bártek

W3C Voice
Browser
Activity

SRGS

XML Formát
SRGS

ABNF formát
SISR

- Pronunciation Lexicon Specification – popis výslovnosti pro rozpoznávání a syntézu řeči.
- Call Control XML – jazyk pro popis řízení telefonního spojení uživatele a systému.
- State Chart XML – jazyk pro popis obecně použitelných stavových automatů.

Standardy W3C Voice Browser Activity

Zpracování

Dialogové
systémy

Luděk Bártek

W3C Voice
Browser
Activity

SRGS

XML Formát
SRGS

ABNF formát
SISR

- Standardy jsou značkovací jazyky – nutná interpretace
- Existuje řada platforem:
 - Volně dostupné desktopové– JVoiceXML, PublicVoiceXML, ...
 - Komerční desktopové – Optimtalk – dříve existovala volně dostupná verze; laboratoř LSD má zakoupenou licenci na laboratorní stroje.
 - Volně dostupné on-online - Asterisk+VoiceGlue resp. OpenVXI, ...
 - komerční on-line – Voxeo Prophecy, BevoCal Cafe – lze vyzkoušet a omezeně používat on-line (max. 2 paralelní hovory).

Speech Recognition Grammar Specification

Dialogové
systémy

Luděk Bártek

W3C Voice

Browser

Activity

SRGS

XML Formát

SRGS

ABNF formát

SISR

- W3C specifikace jazyka pro zápis bezkontextových gramatik pro podporu rozpoznávání řeči.
- Aktuální verze 1.0.
- Nahradil původně používaný standard JSGF
- Existují dvě varianty zápisu:
 - XML
 - Augmented Backus-Naur Form (ABNF).
- Liší se pouze způsob zápisu nikoliv vyjadřovací síla.
- Možnost použitého zápisu závisí na použité platformě.
 - širší podpora pro XML formát

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

- Začíná XML prologem
 - `<?xml version="1.0" encoding="..."?>`.
- Kořenový element – *grammar*; obsahuje množinu pravidel (elementů *rule*).
- Atributy:
 - *version* – použitá verze standardu SRGS (aktuálně 1.0).
 - *xml:lang* – kód jazyka gramatiky.
 - *root* – id pravidla odpovídajícího kořenovému neterminálu.
 - *mode* – pro jaký způsob komunikace je gramatika určena:
 - *dtmf* – pomocí DTMF kódů
 - *voice* – hlasově; implicitní hodnota.

- Element *rule*:
 - atributy:
 - *id* – identifikátor pravidla (odpovídá neterminálnímu symbolu na levé straně pravidla).
 - Obsah – pravá strana pravidla:
 - textový obsah – posloupnost terminálních symbolů
 - element *ruleref* – neterminální symbol; odkazovaný pomocí atributu *uri*.
 - element *one-of* – varianty (operátor |).
 - element *item* – logické členění sekvence; umožňuje např. uvést počet opakování dané části promluvy.

- Posloupnost terminálních a neterminálních symbolů.

SAMPLE → Mám rád *TYP* formát SRGS.

```
<rule id="sample">
```

```
  Mám rád <ruleref uri="#typ"/> formát SRGS.
```

```
</rule>
```

- Lze ji rozdělit na logické části:

```
<rule id="address">
```

```
  <item repeat="0-1">
```

```
    <ruleref uri="#server"/>
```

```
  </item>
```

```
  <item repeat="1-5">
```

```
    tečka <ruleref uri="#domena"/>
```

```
  </item>
```

```
</rule>
```

Varianty

Dialogové
systémy

Luděk Bártek

W3C Voice
Browser
Activity

SRGS

XML Formát
SRGS

ABNF formát
SISR

- Element *one-of*.
- Umožňuje specifikovat různé varianty očekávaných vstupů.
- Jednotlivé varianty jsou ohraničeny elementem *item*.
- Příklad:

```
<rule id="barvy">
  <one-of>
    <item>červená</item>
    <item>zelená</item>
    <item>modrá</item>
  </one-of>
</rule>
```

Opakování

Dialogové
systémy

Luděk Bártek

W3C Voice
Browser
Activity

SRGS
XML Formát
SRGS
ABNF formát
SISR

- Umožňuje specifikaci:
 - nepovinných částí promluvy
 - opakujících se částí promluvy
- Zápis – pomocí atributu *repeat* u elementu *item*.
- Možnosti počtů opakování:

- n krát – n :

```
<item repeat="2">opakování</item>
```

- $\langle m, n \rangle$ krát – $m-n$

```
<item repeat="0-1">  
  Chtěl bych  
</item>
```

- $\langle n, \infty \rangle$ krát – $m-$

```
<item repeat="1-">Ahoj</item>
```

Zvláštní pravidla

Dialogové
systémy

Luděk Bártek

W3C Voice
Browser
Activity

SRGS

XML Formát
SRGS

ABNF formát
SISR

- Slouží k zadání:
 - libovolné nespécifikované promluvy – *GARBAGE*
 - nevyslovitelného pravidla (zakázání určité promluvy) – *VOID*
 - vždy platného pravidla (i prázdného) – *NULL*
- Používají se jako zvláštní neterminální symboly:

```
<ruleref special="GARBAGE"/>
```

- Příklad použití:

```
<rule id="spojeni">  
  <ruleref special="GARBAGE"/>  
  z <ruleref uri="#misto"/>  
  do <ruleref uri="#misto"/>  
  <ruleref uri="prostredok"/>  
</rule>
```

ABNF formát SRGS

Dialogové
systémy

Luděk Bártek

W3C Voice

Browser

Activity

SRGS

XML Formát
SRGS

ABNF formát
SISR

- Čistě textový formát gramatiky vycházející z tradičního formátu BNF.

```
<spojeni> ::= Chci jet <kdy> <cim>  
                z <odkud> do <kam>"."
```

```
<kdy> ::= <den> <cas>
```

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

```
...
```

- BNF podobný formát využívá dále např. JSGF

Struktura ABNF zápisu SRGS

Dialogové
systémy

Luděk Bártek

W3C Voice

Browser

Activity

SRGS

XML Formát

SRGS

ABNF formát

SISR

- Hlavička gramatiky – může obsahovat:
 - specifikaci jazyka gramatiky
 - režim gramatiky – voice/dtmf
 - kořenový neterminál
 - ...
- Pravidla gramatiky
 - formát – $\$neterminál = (neterminál|terminál)^*$
 - *neterminál* = identifikátor pravidla u XML formátu.

Struktura hlavičky ABNF zápisu SRGS

Dialogové
systémy

Luděk Bártek

W3C Voice

Browser

Activity

SRGS

XML Formát
SRGS

ABNF formát
SISR

- Začíná identifikací typu dokumentu.
 - #ABNF verze *SRGS kódování_gramatiky*
#ABNF 1.0 ISO-8859-2
- Následuje:
 - specifikace kořenového neterminálu – root \$*neterminál*;
 - jazyk gramatiky – language *kód jazyka*;
language en-US;
 - režim použitelnosti gramatiky – mode (voice|dtmf);

ABNF zápis pravidel gramatiky

Dialogové
systémy

Luděk Bártek

W3C Voice

Browser

Activity

SRGS

XML Formát

SRGS

ABNF formát

SISR

- Sekvence – sekvence terminálních a neterminálních symbolů oddělených mezerou:

```
$pozdrav = dobrý den;
```

```
$datum = $den $mesic $rok;
```

- Varianty – příslušné sekvence terminálních a neterminálních symbolů oddělené symbolem '|':

```
$dopravni_prostredok = autobus | vlak;
```

- Opakování:

- volitelné části – uzavřeny do '[']'

- m—n – <m-n>

Ukázka SRGS gramatiky v ABNF notaci

Dialogové
systémy

Luděk Bártek

W3C Voice

Browser

Activity

SRGS

XML Formát

SRGS

ABNF formát

SISR

```
#ABNF 1.0 UTF-8;
```

```
root $url;
```

```
language cs-CZ;
```

```
mode voice;
```

```
$url = [<protokol>][server] tečka
```

```
($domena tečka)<1-3>
```

```
tečka $tld[$cesta];
```

```
$protokol = http | ftp | telnet | gopher | ...;
```

```
$cesta = (/ $adresar) <1-> /[$soubor];
```

```
...
```

Semantic Interpretation for Speech Recognition

Dialogové
systémy

Luděk Bártek

W3C Voice
Browser
Activity
SRGS
XML Formát
SRGS
ABNF formát
SISR

- Sémantika – přiřazuje význam tvrzením.
- Sémantika v dialogových systémech:
 - přiřazuje interpretaci promluvám a jejich částem
 - umožňuje získání relevantních údajů.
- SISR – standard z rodiny W3C Voice Browser Activity
 - slouží k sémantické interpretaci promluv
 - publikován v dubnu 2007
 - aktuální verze 1.0.
 - Je úzce spjat se standardy:
 - ECMA Script – vyhodnocování interpretace používá výrazy jazyka ECMA Script
 - SRGS – vyhodnocování je pomocí atributů přiřazeno gramatice pro rozpoznávání promluvy.
 - JSON – interpretace je vnitřně reprezentována pomocí objektů ve formátu JSON.

Přiřazení interpretace části promluvy

Dialogové
systémy

Luděk Bártek

W3C Voice

Browser

Activity

SRGS

XML Formát

SRGS

ABNF formát

SISR

- Sémantická interpretace bývá součástí pravidel SRGS.
- Přiřazení interpretace k pravidlu – pomocí „tag“:

- XML formát SRGS:

- element tag:

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

- atribut tag:

```
<item tag="ano">jo</item>
```

- ABNF formát SRGS:

- interpretace uvedena za interpretovanou částí promluvy.
- tvar: $\{interpretace\}$

```
$potvrzení = $souhlas {ano} | $nesouhlas {ne}
```

Odvozování interpretace na základě dílčích interpretací

Dialogové
systémy

Luděk Bártek

W3C Voice
Browser
Activity

SRGS

XML Formát
SRGS

ABNF formát
SISR

- Zápis odvození – pomocí výrazů v jazyce ECMAScript.
- Přiřazení pravidel pro odvození k pravidlům gramatiky – pomocí atributu/elementu tag.
- Výsledná interpretace reprezentována pomocí objektů ve formátu JSON.
- Vyhodnocování promluv:
 - přístup k dílčím interpretacím – interpretace neterminálních symbolů na pravé straně:
 - atributy stínové proměnné rules
 - neterminálu N odpovídá atribut N .
 - vrácení výsledné interpretace z pravidla do nadřazeného pravidla – objekt out.
 - vrácení interpretace do dialogu:
 - atributy objektu out
 - vstupnímu poli N odpovídá atribut N .

Vyhodnocování promluv

XML formát

Dialogové
systémy

Luděk Bártek

W3C Voice

Browser

Activity

SRGS

XML Formát

SRGS

ABNF formát

SISR

```
<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>
...
```

Vyhodnocování promluv

ABNF Formát

Dialogové
systémy

Luděk Bártek

W3C Voice
Browser
Activity

SRGS
XML Formát
SRGS
ABNF formát
SISR

```
$vlastnictvi = Mám $barva <0-1> $prostredek  
{  
    out = rules.barva + ";" + rules.prostredek;  
};
```


Přiřazení interpretace vstupním polím

XML Formát

Dialogové
systémy

Luděk Bártek

W3C Voice

Browser

Activity

SRGS

XML Formát

SRGS

ABNF formát

SISR

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

Přiřazení interpretace vstupním polím

ABNF formát

Dialogové
systémy

Luděk Bártek

W3C Voice
Browser
Activity

SRGS

XML Formát
SRGS

ABNF formát
SISR

```
$vlastnictvi = mam $barva <0-1> $prostredek  
{  
  out.barva = rules.barva;  
  out.prostredek = rules.prostredek;  
};
```

- Specifikace SRGS
- Specifikace SISR
- Specifikace ECMAScript
- Specifikace JSON