

VoiceXML 在语音自动转接中的应用

胡小勇 汪林林

(重庆邮电学院软件学院 重庆400065)

摘要 VoiceXML(The Voice Extensible Markup Language, VXML)是 XML 语言在语音方面的一个实例,基于 VXML 技术的语音增值业务将有广泛的实用价值和市场前景。本文介绍了 VXML 的常用元素,深入地研究了 VXML 在实际应用中的架构和解析流程,以及 VXML 与 Web 语言的结合。通过语音自动转接实例,探讨了 VXML 在语音业务中的应用,为 VXML 的应用研究提供了有益的参考。

关键词 语音版超文本标记语言,语音自动转接,FIA, VoiceIE, XML

VoiceXML Applied in Voice Active Dial

HU Xiao-Yong WANG Lin-Ling

(Institute of Software, Chongqing University of Post and Telecom, Chongqing400065)

Abstract In the field of voice, as an application of XML language, VXML can combine the Web resources with the phone resources. In this paper we discuss the basic elements of VXML, its' application architecture and interpreting flow and how to use Web language in VXML. In the use of Vxml technology, as well as with the auto voice service offered by the Internet. The software developer can develop auto voice service so as to open up a new network service.

Keywords Voicexml, Vad, FIA, VoiceIE, XML

1 引言

VoiceXML (The Voice Extensible Markup Language, VXML)即语音版可扩展标记语言,是 XML (可扩展的标记语言)在语音方面的一个实例。VXML 能让最终用户通过语音浏览器(voiceIE)同一个基于 Web 的应用程序进行交互操作。在语音自动转接应用中,语音浏览器会通过合成语音或者预先录制的语音文件同用户进行“交谈”。最终用户通过使用自然语言输入需转接的电话人名和电话方式(手机、办公电话等)来进行回应。语音浏览器然后就会接受这个输入,解析或者翻译输入的信息,然后执行 VXML 中的编码动作,为用户自动转接电话。基于 VXML 技术的语音自动转接可代替原有的人工操作,可以节省大量的人力资源。

2 VXML 应用架构和解析流程

2.1 VXML 常用元素

在 VXML 中的常用元素有:

<vxml>vxml 根元素

<form>提交信息和收集数据单元

<field>收集数据

<prompt>提示发音

<filled>收集数据后的操作

<block>可执行代码的容器

<subdialog>调用另一个 vxml 文档,并返回原流程

<return>从 subdialog 中返回

<goto>调用另一个 vxml 文档,不返回原流程

<clear>清除变量值,使其为空

<if><else if></else>条件逻辑判断

<grammar>定义一个语音识别语法

<catch>捕捉事件,如未输入、输入错误等

2.2 VoiceXML 的执行原则

一个 VoiceXML 应用由一个或多个 VoiceXML document(文档)构成。一个 VoiceXML document 或一个 application(应用)中的一系列 document 形成一个交互式的有限状态自动机。用户任何时刻只能处于一个会话状态,即处于某一个 dialog。每个 dialog 决定下一个跃迁到的 dialog,跃迁时指定使用到的 URI(Uniform Resource Indicator, 统一资源指示),即指出了下面那个 document 和 dialog;如果没有指定下一个 dialog,则假定为 document 中的第一个 dialog。如果当前运行完的 dialog 没有指定下一个跳转的目标,或者其中出现明确要求退出的元素,则终止当前的交互。

VoiceXML 的业务逻辑主要在 document 的 dialog 元素范围内实现,而在应用中用到的 dialog 基本上是窗体项即 form 级元素,是在 FIA(form interpretation algorithm, form 的执行规则)主循环中可被访问到的元素。分为两种,一种是域项,包括元素 field、record、transfer、object、subdialog,用于提示用户,收集信息。另一种窗体项就是控制项,控制项包括 block、initial 元素。

Form 元素由内含的 FIA 规则解释,FIA 循环重复地选择一下 form item,然后访问它,这个被选的 form item 是第一个警戒条件不满足的 form item。例如,form item 的默认警戒条件检验该 form item 的变量是否被赋值,这样对于仅有 form item 的 form,用户将被提示依次处理各 field item。

解释一个 form item 通常包括以下步骤:

1. 选择并播放一段或多段提示音。

2. 收集用户输入,或者填充一个或多个 field,或者抛出各种事件(如帮助事件)。

3. 解释最近填充的 field 中的<filled>动作。

在遇到控制状态转移时,如转移到其它 dialog 或者 document 的 goto 元素,或向服务器提交数据的 submit 元素,FIA 将结束.当遇到暗含的 exit 结束标志,如再没有合法的 form item,FIA 也将结束.

一般的 VoiceXML 应用的执行框架也就是基于上述执行模型,对于整个应用以 VoiceXML document 为执行单位,而一个具体的 document 则在其主要的业务逻辑处理元素 form 内按 FIA 算法来执行.

2.3 VXML 应用架构和解析流程

图1是 VXML 应用架构图.

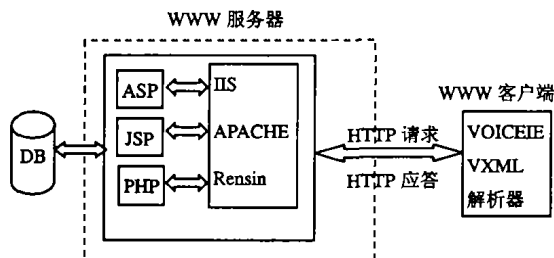


图1 VXML 应用架构图

图1给出了 VXML 的三层应用架构:在客户端语音浏览器通过终端(电话)收集用户的输入,然后根据用户的输入,服务器对后台数据库进行访问,并将结果显示给用户收听.

在实际运用中必须将 VXML 语言和 Web 语言相结合才能发挥它通过电话终端访问 Web 资源的巨大作用.我们可以将 Web 语言内嵌于 VXML 中,组成新的应用文档.对这个新应用文档的解析如下:VXML 解析器 voiceIE 对 VXML 程序流程进行解析,通过同用户的语音交谈,收集用户的输入.当需要对用户的输入进行数据库操作时,voiceIE 会对 Webserver 发出 HTTP 请求,触发后台 Web 程序进行数据库操作,并将查询结果以 HTTP 应答返回给 voiceIE,这时用户就能听到相应的内容.在例1中,主流程 index.vxml 等待用户输入帐号,调用 confirm.jsp 操作数据库查询该帐号是否存在,并将查询结果返回给主流程.

例1 Indexvxml

```

<?xml version="1.0"?><!--根元素-->
<vxml version="1.0">
<form><!--收集用户语音输入单元-->
  <field name="callerName"><!--收集用户输入,并将用户输入内容存入 callerName-->
    <grammar src=".../grammar/name.gsl"/>
    <prompt>请说出你的帐号</prompt><!--同用户对话,收集用户谈话-->
  </field>
  <!--同过子对话调用后台 jsp 访问数据库-->
  <subdialog name="confirm.jsp" src="http://localhost:8080/test/jsp/confirm.jsp namelist="callerName">
    <filled><!--根据 confirm.jsp 的返回值判断用户是否存在-->
    <if cond="confirm.ret=="1">
      <prompt>用户存在</prompt>
    <elseif />
      <prompt>用户不存在</prompt>
    </if>
    </filled>
  </subdialog>
</form>
</vxml>

confirm.jsp
<%@ page language="java" charset="gb2312">
<jsp:Usebean id="usebean" class="h.useage"/>
<%
//查询数据库
int Ret=0;
String name = usebean.toChn (request.getParameter (" callerName"));
ResultSet rs=usebean.executeQuery("select * from clientInfo where name="+name+"");

```

```

If(rs.next()) Ret=1;
%>
<?xml version="1.0"?>
<vxml version="1.0">
<form>
  <!--将查询结果付值给 vxml 变量-->
  <var name="ret" expr="(%Ret%)/">
  <block>
    <!--将结果返回调用页面-->
    <return namelist="ret" />
  </block>
</form>
</vxml>

```

3 VXML 在语音转接(vad)中的应用

用户在开通 Vad(voice active dial)业务后,如需拨打电话本上的电话时,只需说出被叫号码,系统将为用户只动转接被叫号码.其主流程如图2所示.

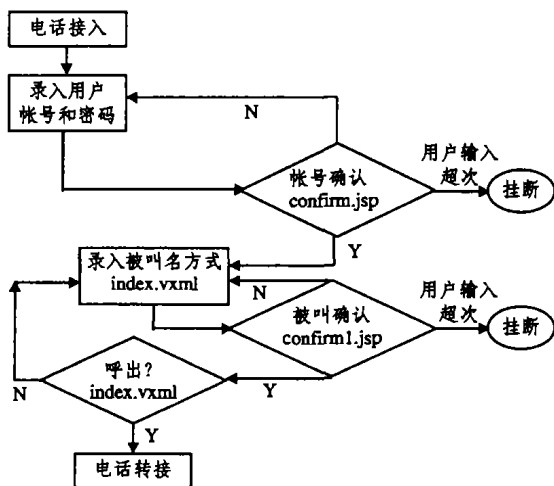


图2

用户拨打接入电话后,进入主流程 index.vxml,voiceIE 开始解析.等待用户说出自己帐号和密码,index.vxml 通过 <field> 录入用户的帐号和密码,通过 <subdialog> 调用 confirm.jsp 对用户的帐号密码进行验证.当验证不通过时返回 index.vxml 提示用户重新录入;用户帐号、密码正确,提示用户说出需转接的人和电话方式(手机,办公室电话,家庭电话),调用 confirm1.jsp 对被叫名和电话方式进行验证,被叫名和电话方式正确时返回 index.vxml 让用户确认是否拨出电话,用户确认后为用户拨出电话,或返回让用户重新录入被叫名和电话方式.在流程中当用户的输入超过三次时,提示用户超次输入后退出流程.

结束语 在本文中我们描述和证明了如何运用 VXML 技术通过电话终端对网络资源进行访问,并通过一个实例对 VXML 在语音领域的应用作了初步的探索.现在我国的语音业务正在大力发展,语音信箱等语音增值业务正在为运营商带来巨大的经济效益,本文所讨论的实例能成为语音业务的下一个增值点.

参考文献

- 1 w3:Voice Extensible Markup Language (Voicexml) Version 2.0 <http://www.w3.org/TR/2004/REC-voicexml20-20040316/>
- 2 Beasley R,等. voicexml 语音应用程序开发(美).机械工业出版社,2002
- 3 Natanya Pitts. XML 技术内幕[M].北京:机械工业出版社,2002
- 4 磨砺你的 VoiceXML 技术. <http://www00.zdnet.com.cn/developer/tech/story/0,2000081602,39156198,00.htm>