Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: August 24, 2009

S. McGlashan
Hewlett-Packard
T. Melanchuk
Rain Willow Communications
C. Boulton
NS-Technologies
February 20, 2009

An Interactive Voice Response (IVR) Control Package for the Media Control Channel Framework
draft-ietf-mediactrl-ivr-control-package-05

Status of this Memo

Copyright Notice

Abstract

This document defines a Media Control Channel Framework Package for Interactive Voice Response (IVR) dialog interaction on media connections and conferences. The package defines dialog management request elements for preparing, starting and terminating dialog interactions, as well as associated responses and notifications. Dialog interactions are specified in a dialog language. This package defines a lightweight IVR dialog language (supporting prompt playback, runtime controls, DTMF collect and media recording) and allows other dialog languages to be used. The package also defines elements for auditing package capabilities and IVR dialogs.

Table of Contents

1. Introduction

The Media Control Channel Framework ([I-D.ietf-mediactrl-sip-control-framework]) provides a generic approach for establishment and reporting capabilities of remotely initiated commands. The Control Framework utilizes many functions provided by the Session Initiation Protocol [RFC3261] (SIP) for the rendezvous and establishment of a reliable channel for control interactions. The Control Framework also introduces the concept of a Control Package. A Control Package is an explicit usage of the Control Framework for a particular interaction set. This document defines a Control Package for Interactive Voice Response (IVR) dialogs on media connections and conferences. The term 'dialog' in this document refers to an IVR dialog and is completely unrelated to the notion of a SIP dialog. The term 'IVR' is used in its inclusive sense, allowing media other than voice for dialog interaction.

The package defines dialog management request elements for preparing, starting and terminating dialog interactions, as well as associated responses and notifications. Dialog interactions are specified using a dialog language where the language specifies a well-defined syntax and semantics for permitted operations (play a prompt, record input

from the user, etc). This package defines a lightweight IVR dialog language (supporting prompt playback, runtime controls, DTMF collect and media recording) and allows other dialog languages to be used. These dialog languages are specified inside dialog management elements for preparing and starting dialog interactions. The package also defines elements for auditing package capabilities and IVR dialogs.

This package has been designed to satisfy IVR requirements documented in the Media Server Control Protocol Requirements document ([RFC5167]); more specifically REQ-MCP-28, REQ-MCP-29 and REQ-MCP-30. It achieves this by building upon two major approaches to IVR dialog design. These approaches address a wide range of IVR use cases and are used in many applications which are extensively deployed today.

First, the package is designed to provide the major IVR functionality of SIP Media Server languages such as netann ([RFC4240]), MSCML ([RFC5022]) and MSML ([MSML]) which themselves build upon more traditional non-SIP languages ([H.248.9], [RFC2897]). A key differentiator is that this package provides IVR functionality using the Media Control Channel Framework.

Second, its design is aligned with key concepts of the web model as defined in W3C Voice Browser languages. The key dialog management mechanism is closely aligned with CCXML ([CCXML10]). The dialog functionality defined in this package can be largely seen as a subset

of VoiceXML ([VXML20], [VXML21]): where possible, basic prompting, DTMF collection and media recording features are incorporated, but not any advanced VoiceXML constructs (such as <form>, its interpretation algorithm, or a dynamic data model). As W3C develops VoiceXML 3.0 ([VXML30]), we expect to see further alignment, especially in providing a set of basic independent primitive elements (such as prompt, collect, record and runtime controls) which can be re-used in different dialog languages.

By reusing and building upon design patterns from these approaches to IVR languages, this package is intended to provide a foundation which is familiar to current IVR developers and sufficient for most IVR applications, as well as a path to other languages which address more advanced applications.

This control package defines a lightweight IVR dialog language. The scope of this dialog language is the following IVR functionality:

- playing one or more media resources as a prompt to the user
- runtime controls (including VCR controls like speed and volume)
- collecting DTMF input from the user according to a grammar
- recording user media input

Out of scope for this dialog language are more advanced functions including ASR (Automatic Speech Recognition), TTS (Text-to-Speech), fax, automatic prompt recovery ('media fallback') and media transformation. Such functionality can be addressed by other dialog languages (such as VoiceXML) used with this package, extensions to this package (addition of foreign elements or attributes from another namespace) or other control packages.

The functionality of this package is defined by messages, containing XML [XML] elements, transported using the Media Control Channel Framework. The XML elements can be divided into three types: dialog management elements; a dialog element which defines a lightweight IVR dialog language used with dialog management elements; and finally, elements for auditing package capabilities as well as dialogs managed by the package.

Dialog management elements are designed to manage the general lifecycle of a dialog. Elements are provided for preparing a dialog, starting the dialog on a conference or connection, and terminating execution of a dialog. Each of these elements is contained in a Media Control Channel Framework CONTROL message sent to the media server. When the appropriate action has been executed, the media

server sends a REPORT message (or a 200 response to the CONTROL if it can execute in time) with a response element indicating whether the operation was successful or not (e.g. if the dialog cannot be started, then the error is reported in this response). Once a dialog has been successfully started, the media server can send further event notifications in a framework CONTROL message. This package defines two event notifications: a DTMF event indicating the DTMF activity; and a dialogexit event indicating that the dialog has

exited. If the dialog has executed successful, the dialogexit event includes information collected during the dialog. If an error occurs during execution (e.g. a media resource failed to play, no recording resource available, etc), then error information is reported in the dialogexit event. Once a dialogexit event is sent, the dialog lifecycle is terminated.

The dialog management elements for preparing and starting a dialog specify the dialog using a dialog language. A dialog language has well-defined syntax and semantics for defined dialog operations. Typically dialog languages are written in XML where the root element has a designated XML namespace and, when used as standalone documents, have an associated MIME media type. For example, VoiceXML is an XML dialog language with the root element <vxml> with the designated namespace 'http://www.w3.org/2001/vxml' and standalone documents are associated with the MIME media type 'application/ vxml+xml' ([RFC4267]).

This control package defines its own lightweight IVR dialog language. The language has a root element (<dialog>) with the same designated namespace as used for other elements defined in this package (see Section 8.2). The root element contains child elements for playing prompts to the user, specifying runtime controls, collecting DTMF input from the user and recording media input from the user. The child elements can co-occur so as to provide 'play announcement', 'prompt and collect' as well as 'prompt and record' functionality.

The dialog management elements for preparing and starting a dialog can specify the dialog language either by including inline a fragment with the root element or by referencing an external dialog document. The dialog language defined in this package is specified inline. Other dialog languages, such as VoiceXML, can be used by referencing an external dialog document.

The document is organized as follows. Section 3 describes how this control package fulfills the requirements for a Media Control Channel Framework control package. Section 4 describes the syntax and semantics of defined elements, including dialog management (Section 4.2), the IVR dialog element (Section 4.3) and audit elements (Section 4.4). Section 5 describes an XML schema for these

elements and provides extensibility by allowing attributes and

elements from other namespaces. Section 6 provides examples of package usage. Section 7 describes important security considerations for use of this control package. Section 8 provides information on IANA registration of this control package, including its name, XML namespace and MIME media type. Finally, Section 12 provides additional information on using VoiceXML when supported as an external dialog language.

2. Conventions and Terminology

In this document, BCP 14 [RFC2119] defines the key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL". In addition, BCP 15 indicates requirement levels for compliant implementations.

The following additional terms are defined for use in this document:

Dialog: A dialog performs media interaction with a user following the concept of an IVR (Interactive Voice Response) dialog (this sense of 'dialog' is completely unrelated to a SIP dialog). A dialog is specified as inline XML, or via a URI reference to an external dialog document. Traditional IVR dialogs typically feature capabilities such as playing audio prompts, collecting DTMF input and recording audio input from the user. More inclusive definitions include support for other media types, runtime controls, synthesized speech, recording and playback of video, recognition of spoken input, and mixed initiative conversations.

Application server: A SIP [RFC3261] application server (AS) hosts and executes services such as interactive media and conferencing in an operator's network. An AS influences and impacts the SIP session, in particular by terminating SIP sessions on a media server, which is under its control.

Media Server: A media server (MS) processes media streams on behalf of an AS by offering functionality such as interactive media, conferencing, and transcoding to the end user. Interactive media functionality is realized by way of dialogs which are initiated by the application server.

3. Control Package Definition

This section fulfills the mandatory requirement for information that MUST be specified during the definition of a Control Framework Package, as detailed in Section 8 of [I-D.ietf-mediactrl-sip-control-framework].

3.1. Control Package Name

The Control Framework requires a Control Package to specify and register a unique name and version.

The name and version of this Control Package is "msc-ivr/1.0" (Media Server Control - Interactive Voice Response - version 1.0). Its IANA registration is specified in Section 8.1.

Since this is the initial ("1.0") version of the control package, there are no backwards compatibility issues to address.

3.2. Framework Message Usage

The Control Framework requires a Control Package to explicitly detail the control messages that can be used as well as provide an indication of directionality between entities. This will include which role type is allowed to initiate a request type.

This package specifies CONTROL and response messages in terms of XML elements defined in Section 4, where the message bodies have the MIME media type defined in Section 8.3. These elements describe requests, response and notifications and all are contained within a root <mscivr> element (Section 4.1).

In this package, the MS operates as a Control Server in receiving requests from, and sending responses to, the AS (operating as Control Client). Dialog management requests and responses are defined in Section 4.2. Audit requests and responses are defined in Section 4.4. dialog management and audit responses are carried in a framework 200 response or REPORT message bodies. This package's response codes are defined in Section 4.5.

Note that package responses are different from framework response codes. Framework error response codes (see Section 8 of [I-D.ietf-mediactrl-sip-control-framework]) are used when the request or event notification is invalid; for example, a request is invalid XML (400), or not understood (500).

The MS also operates as a Control Client in sending event notification to the AS (Control Server). Event notifications

(Section 4.2.5) are carried in CONTROL message bodies. The AS MUST respond with a Control Framework 200 response.

3.3. Common XML Support

The Control Framework requires a Control Package definition to specify if the attributes for media dialog or conference references are required.

This package requires that the XML Schema in Section 17.1 of [I-D.ietf-mediactrl-sip-control-framework] MUST be supported for media dialogs and conferences.

The package uses "connectionid" and "conferenceid" attributes for various element definitions (Section 4). The XML schema (Section 5) imports the definitions of these attributes from the framework schema.

3.4. CONTROL Message Body

The Control Framework requires a Control Package to define the control body that can be contained within a CONTROL command request and to indicate the location of detailed syntax definitions and semantics for the appropriate body types.

When operating as Control Server, the MS receives CONTROL messages body with the MIME media type defined in Section 8.3 and containing an <mscivr> element (Section 4.1) with either a dialog management or audit request child element.

The following dialog management request elements are carried in CONTROL message bodies to MS: <dialogprepare> (Section 4.2.1),

<dialogstart> (Section 4.2.2) and <dialogterminate> (Section 4.2.3)elements.

The <audit> request element (Section 4.4.1) is also carried in CONTROL message bodies.

When operating as Control Client, the MS sends CONTROL messages with the MIME media type defined in Section 8.3 and a body containing an <mscivr> element (Section 4.1) with a notification <event> child element (Section 4.2.5).

3.5. REPORT Message Body

The Control Framework requires a control package definition to define the REPORT body that can be contained within a REPORT command request, or that no report package body is required. This section

indicates the location of detailed syntax definitions and semantics for the appropriate body types.

When operating as Control Server, the MS sends REPORT bodies with the MIME media type defined in Section 8.3 and containing a <mscivr> element (Section 4.1) with a response child element. The response element for dialog management requests is a <response> element (Section 4.2.4). The response element for an audit request is a <auditresponse> element (Section 4.4.2).

3.6. Audit

The Control Framework encourages Control Packages to specify whether auditing is available, how it is triggered as well as the query/ response formats.

This Control Packages supports auditing of package capabilities and dialogs on the MS. An audit request is carried in a CONTROL message (see Section 3.4) and an audit response in a REPORT message (or a 200 response to the CONTROL if it can execute the audit in time) (see Section 3.5).

The syntax and semantics of audit request and response elements is defined in Section 4.4.

3.7. Examples

The Control Framework recommends Control Packages to provide a range of message flows that represent common flows using the package and this framework document.

This Control Package provides examples of such message flows in Section 6.

4. Element Definitions

This section defines the XML elements for this package. The elements are defined in the XML namespace specified in Section 8.2.

The root element is <mscivr> (Section 4.1). All other XML elements (requests, responses and notification elements) are contained within it. Child elements describe dialog management (Section 4.2) and audit (Section 4.4) functionality. The IVR dialog element (contained within dialog management elements) is defined in Section 4.3. Response status codes are defined in Section 4.5 and type definitions in Section 4.6.

Implementation of this control package MUST address the Security Considerations described in Section 7).

Implementation of this control package MUST adhere to the syntax and semantics of XML elements defined in this section and the schema

(Section 5). Since XML Schema is unable to support some types of syntactic constraints (such as attribute and element co-occurrence), some elements in this package specify additional syntactic constraints in their textual definition. If there is a difference in constraints between the XML schema and the textual description of elements in this section, the textual definition takes priority.

The XML schema supports extensibility by allowing attributes and elements from other namespaces. Implementations MAY support additional capabilities by means of attributes and elements from other (foreign) namespaces. Attributes and elements from foreign namespaces are not described in this section.

Some elements in this control package contain attributes whose value is a URI. These elements include: <dialogprepare> (Section 4.2.1), <dialogstart> (Section 4.2.2), <media> (Section 4.3.1.5), <grammar> (Section 4.3.1.3.1), and <record> (Section 4.3.1.4). The MS MUST support one or more schemes using communication protocols suitable for fetching resources (e.g. HTTP).

Usage examples are provided in Section 6.

4.1. <mscivr>

The <mscivr> element has the following attributes (in addition to standard XML namespace attributes such as xmlns):

version: a string specifying the mscivr package version. The value is fixed as '1.0' for this version of the package. The attribute is mandatory.

The <mscivr> element has the following defined child elements, only one of which can occur:

1. dialog management elements defined in Section 4.2:

   <dialogprepare> prepare a dialog. See Section 4.2.1

<dialogstart> start a dialog. See Section 4.2.2

<dialogterminate> terminate a dialog. See Section 4.2.3

<response> response to a dialog request. See Section 4.2.4

<event> dialog or subscription notification. See Section 4.2.5

2. audit elements defined in Section 4.4:

<audit> audit package capabilities and managed dialogs. See Section 4.4.1

<auditresponse> response to an audit request. See Section 4.4.2

For example, a request to the MS to start an IVR dialog playing a prompt:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <dialogstart connectionid="ssd3r3~sds345b">
  <dialog>
   <prompt>
    <media loc="http://www.example.com/welcome.wav"/>
   </prompt>
  </dialog>
 </dialogstart>
</mscivr>
```

and a response from the MS that the dialog started successfully:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <response status="200" dialogid="d1"/>
</mscivr>
```

and finally a notification from the MS indicating that the dialog exited upon completion of playing the prompt:


```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <event dialogid="d1">
  <dialogexit status="1">
   <promptinfo termmode="completed"/>
```

```
    </dialogexit>
   </event>
  </mscivr>
```

4.2. Dialog Management Elements

This section defines the dialog management XML elements for this control package. These elements are divided into requests, responses and notifications.

Request elements are sent to the MS to request a specific dialog operation to be executed. The following request elements are defined:

<dialogprepare>: prepare a dialog for later execution

<dialogstart>: start a (prepared) dialog on a connection or conference

<dialogterminate>: terminate a dialog

Responses from the MS describe the status of the requested operation. Responses are specified in a <response> element (Section 4.2.4) which includes a mandatory attribute describing the status in terms of a numeric code. Response status codes are defined in Section 4.5. The MS MUST respond to a request message with a response message. If the MS is not able to process the request and carry out the dialog operation, the request has failed and the MS MUST indicate the class of failure using an appropriate 4xx response code. Unless an error response code is specified for a class of error within this section, implementations follow Section 4.5 in determining the appropriate status code for the response.

Notifications are sent from the MS to provide updates on the status of a dialog or operations defined within the dialog. Notifications are specified in an <event> element (Section 4.2.5).

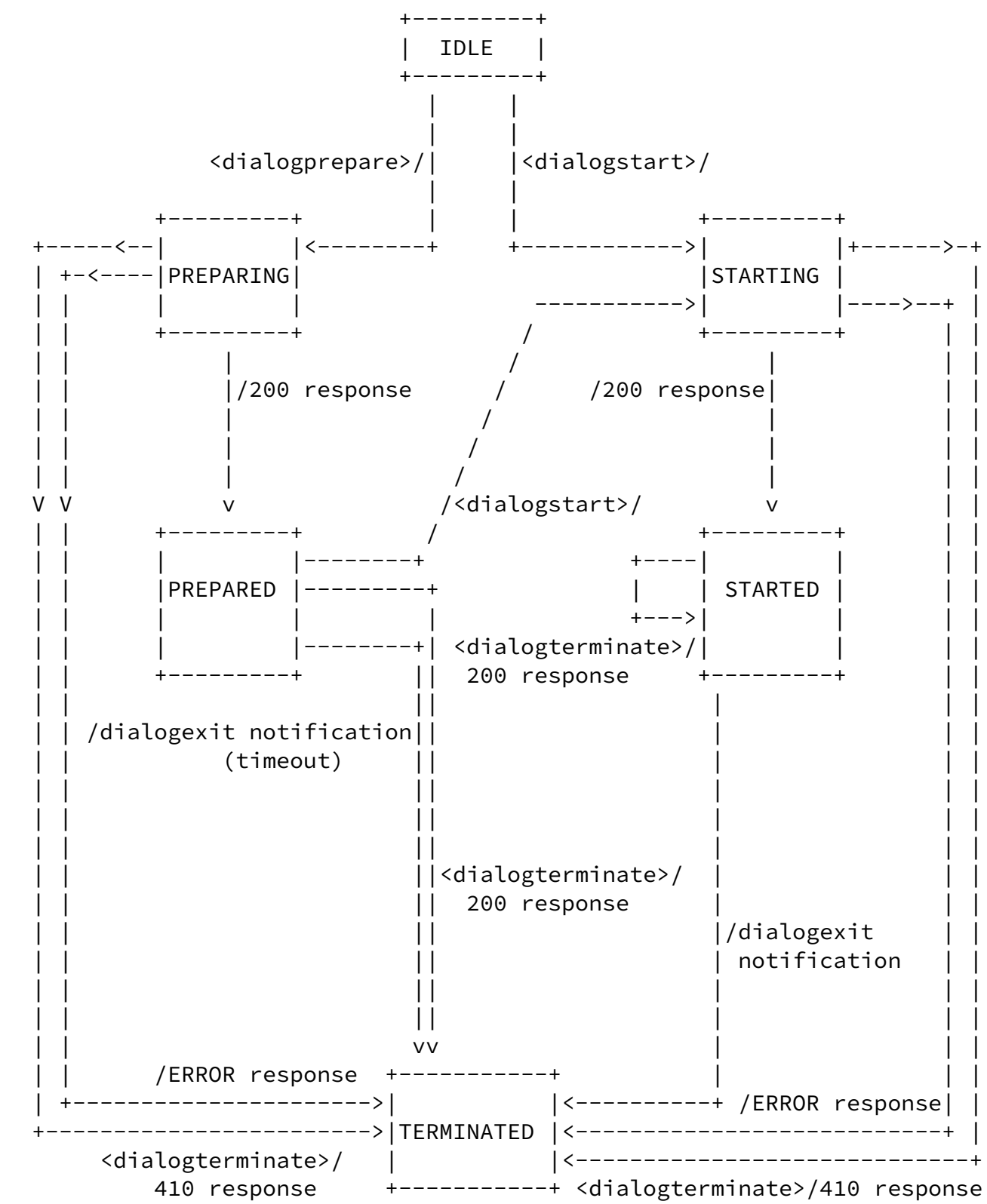


Figure 1: Dialog Lifecycle

The MS implementation MUST adhere to the dialog lifecycle shown in

Figure 1, where each dialog has the following states:

IDLE: the dialog is uninitialized.

PREPARING: the dialog is being prepared. The dialog is assigned a valid dialog identifier (see below). If an error occurs the dialog transitions to the TERMINATED state and the MS MUST send a response indicating the error. If the dialog is terminated before preparation is complete, the dialog transitions to the TERMINATED state and the MS MUST send a 410 response (Section 4.5) for the prepare request.

PREPARED: the dialog has been successfully prepared and the MS MUST send a 200 response indicating the prepare operation was successful. If the dialog is then terminated, the dialog transitions to the TERMINATED state. If the duration the dialog remains in the PREPARED state exceeds the maximum preparation duration, the dialog transitions to the TERMINATED state and the MS MUST send a dialogexit notification with the appropriate error status code (see Section 4.2.5.1). A maximum preparation duration of 30s is RECOMMENDED.

STARTING: the dialog is being started. If the dialog has not already been prepared, it is first prepared and assigned a valid dialog identifier (see below). If an error occurs the dialog transitions to the TERMINATED state and the MS MUST send a response indicating the error. If the dialog is terminated, the dialog transitions to the TERMINATED state and the MS MUST send a 410 response (Section 4.5) for the start request.

STARTED: the dialog has been successfully started and is now active. The MS MUST send a 200 response indicating the start operation was successful. If any dialog events occurs which were subscribed to, the MS MUST send a notifications when the dialog event occurs. When the dialog exits (due to normal termination, an error or a terminate request), the MS MUST send a dialogexit notification event (see Section 4.2.5.1) and the dialog transitions to the TERMINATED state.

TERMINATED: the dialog is terminated and its dialog identifier is no longer valid. Dialog notifications MUST NOT be sent for this dialog.

Each dialog has a valid identifier until it transitions to a TERMINATED state. The dialog identifier is assigned by the MS unless the <dialogprepare> or <dialogstart> request already specifies a identifier (dialogid) which is not associated with any other dialog on the MS. Once a dialog is in a TERMINATED state, its dialog identifier is no longer valid and can be reused for another dialog.

The identifier is used to reference the dialog in subsequent requests, responses and notifications. In a <dialogstart> request, the dialog identifier can be specified in the prepareddialogid attribute indicating the prepared dialog to start. In <dialogterminate> and <audit> requests, the dialog identifier is specified in the dialogid attribute, indicating which dialog is to be terminated or audited respectively. If these requests specify a dialog identifier already associated with another dialog on the MS, the MS sends a response with a 405 status code (see Section 4.5) and the same dialogid as in the request. The MS MUST specify a dialog identifier in notifications associated with the dialog. The MS MUST specify a dialog identifier in responses unless it is a response to a syntactically invalid request.

For a given dialog, the <dialogprepare> or <dialogstart> request elements specify the dialog content to execute either by including inline a <dialog> element (the dialog language defined in this package, see Section 4.3) or by referencing an external dialog document (a dialog language defined outside this package). When referencing an external dialog document, the request element contains a URI reference to the remote document (specifying the dialog definition) and, optionally, a type attribute indicating the MIME media type associated with the dialog document. Consequently, the dialog language associated with a dialog on the MS is identified either inline by a <dialog> child element or by a src attribute referencing a document containing the dialog language. The MS MUST support inline the IVR dialog language defined in Section 4.3. The MS MAY support other dialog languages by reference.

4.2.1. <dialogprepare>

The <dialogprepare> request is sent to the MS to request preparation of a dialog. Dialog preparation consists of (a) retrieving external

dialog document and resources (if required), and (b) validating the dialog document syntactically and semantically.

A prepared dialog is executed when the MS receives a <dialogstart> request referencing the prepared dialog identifier (see Section 4.2.2).

The <dialogprepare> element has the following attributes:

src: specifies the location of an external dialog document to prepare. A valid value is a URI (see Section 4.6.9). If the URI scheme is unsupported, the MS sends a <response> with a 420 status code (Section 4.5). If the document cannot be retrieved within the timeout interval, the MS sends a <response> with a 409 status code. If the document contains a type of dialog language which

the MS does not supported, the MS sends a <response> with a 421 status code. The attribute is optional. There is no default value.

type: specifies the type of the external dialog document indicated in the 'src' attribute. A valid value is a MIME media type (see Section 4.6.10). If the URI scheme used in the src attribute defines a mechanism for establishing the authoratitive MIME media type of the media resource, the value returned by that mechanism takes precedence over this attribute. The attribute is optional. There is no default value.

fetchtimeout: the maximum timeout interval to wait when fetching an external dialog document. A valid value is a Time Designation (see Section 4.6.7). The attribute is optional. The default value is 30s.

dialogid: string indicating a unique name for the dialog. If a dialog with the same name already exists on the MS, the MS sends a <response> with a 405 status code (Section 4.5). If this attribute is not specified, the MS MUST create a unique name for the dialog (see Section 4.2 for dialog identifier assignment). The attribute is optional. There is no default value.

The <dialogprepare> element has one optional child element:

<dialog> an IVR dialog (Section 4.3) to prepare. The element is optional.

The dialog to prepare can either be specified inline with a <dialog> child element or externally (for dialog languages defined outside this specification) using the src attribute. It is a syntax error if both an inline <dialog> element element and a src attribute are specified and the MS sends a <response> with a 400 status code (see Section 4.5). The type and fetchtimeout attributes are only relevant when a dialog is specified as an external document.

For example, a <dialogprepare> request to prepare an inline IVR dialog with a single prompt:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <dialogprepare>
  <dialog>
   <prompt>
    <media loc="http://www.example.com/welcome.wav"/>
   </prompt>
  </dialog>
 </dialogprepare>
```



```
</mscivr>
```

In this example, a request with a specified dialogid to prepare a VoiceXML dialog document located externally:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <dialogprepare dialogid="d2" type="application/voicexml+xml"
                src="http://www.example.com/mydialog.vxml"
                fetchtimeout="15s"/>
</mscivr>
```

Since MS support for dialog languages other than the IVR dialog language defined in this package is optional, if the MS does not support the dialog language it would send a response with the status code 409 (Section 4.5). Further information on using VoiceXML can be found in Section 12.

4.2.2. <dialogstart>

The <dialogstart> element is sent to the MS to start a dialog. If the dialog has not been prepared, the dialog is prepared (retrieving an external document and resources if necessary, and the dialog document validated syntactically and semantically). Media processors (e.g. DTMF and prompt queue) are activated and associated with the specified connection or conference.

The <dialogstart> element has the following attributes:

src: specifies the location of an external dialog document to start. A valid value is a URI (see Section 4.6.9). If the URI scheme is unsupported, the MS sends a <response> with a 420 status code (Section 4.5). If the document cannot be retrieved with the timeout interval, the MS sends a <response> with a 409 status code. If the document contains a type of dialog language which the MS does not supported, the MS sends a <response> with a 421 status code. The attribute is optional. There is no default value.

type: specifies the type of the external dialog document indicated in the 'src' attribute. A valid value is a MIME media type (see Section 4.6.10). If the URI scheme used in the src attribute defines a mechanism for establishing the authoratitive MIME media type of the media resource, the value returned by that mechanism takes precedence over this attribute. The attribute is optional. There is no default value.

fetchtimeout: the maximum timeout interval to wait when fetching an external dialog document. A valid value is a Time Designation (see Section 4.6.7). The attribute is optional. The default value is 30s.

dialogid: string indicating a unique name for the dialog. If a dialog with the same name already exists on the MS, the MS sends a <response> with a 405 status code (Section 4.5). If neither the dialogid attribute nor the prepareddialogid attribute is specified, the MS MUST create a unique name for the dialog (see Section 4.2 for dialog identifier assignment). The attribute is optional. There is no default value.

prepareddialogid: string identifying a dialog previously prepared using a dialogprepare (Section 4.2.1) request. If neither the dialogid attribute nor the prepareddialogid attribute is specified, the MS MUST create a unique name for the dialog (see Section 4.2 for dialog identifier assignment). The attribute is optional. There is no default value.

connectionid: string identifying the SIP dialog connection on which this dialog is to be started (see Section 17.1 of [I-D.ietf-mediactrl-sip-control-framework]). The attribute is optional. There is no default value.

conferenceid: string identifying the conference on which this dialog is to be started (see Section 17.1 of [I-D.ietf-mediactrl-sip-control-framework]). The attribute is optional. There is no default value.

Exactly one of the connectionid or conferenceid attributes MUST be specified. If both connectionid and conferenceid attributes are specified or neither are specified, it is a syntax error and the MS sends a <response> with a 400 status code (Section 4.5).

It is an error if the connection or conference referenced by a specific connectionid or conferenceid attribute is not available on the MS at the time the <dialogstart> request is executed. If an invalid connectionid is specified, the MS sends a <response> with a 407 status code (Section 4.5). If an invalid conferenceid is specified, the MS sends a <response> with a 408 status code.

The <dialogstart> element has the following sequence of child elements:

<dialog>: specifies an IVR dialog (Section 4.3) to execute. The element is optional.

<subscribe>: specifies subscriptions to dialog events (Section 4.2.2.1). The element is optional.

<params>: specifies input parameters (Section 4.2.6) for a dialog languages defined outside this specification. The element is optional. If a parameter is not supported by the MS for the external dialog language, the MS sends a <response> with a 427 status code (Section 4.5).

<stream>: determines the media stream(s) associated with the connection or conference on which the dialog is executed (Section 4.2.2.2). The <stream> element is optional. Multiple <stream> elements can be specified.

The dialog to start can be specified either (a) inline with a <dialog> child element, or (b) externally using the src attribute (for dialog languages defined outside this specification), or (c) reference a previously prepared dialog using the prepareddialogid attribute. If exactly one of the src attribute, the prepareddialogid or a <dialog> child element is not specified, it is a syntax error and the MS sends a <response> with a 400 status code (Section 4.5). If the prepareddialogid and dialogid attributes are specified, it is also a syntax error and the MS sends a <response> with a 400 status code. The type and fetchtimeout attributes are only relevant when a dialog is specified as an external document.

The <stream> element provides explicit control over which media streams on the connection or conference are used during dialog execution. For example, if a connection supports both audio and video streams, a <stream> element could be used to indicate that only the audio stream is used in receive mode. In cases where there are multiple media streams of the same type for a dialog, the AS MUST use <stream> elements to explicitly specify the configuration. If no <stream> elements are specified, then the default media configuration is that defined for the connection or conference.

If a <stream> element is in conflict with (a) another <stream> element, (b) with specified connection or conference media capabilities, (c) with a SDP label value as part of the connectionid (see Section 17.1 of [I-D.ietf-mediactrl-sip-control-framework]) then the MS sends a <response> with a 411 status code (Section 4.5). If the media stream configuration is not supported by the MS, then the MS sends a <response> with a 428 status code (Section 4.5).

The MS MAY support multiple, simultaneous dialogs being started on

the same connection or conference. For example, the same connection can receive different media streams (e.g. audio and video) from different dialogs, or receive (and implicitly mix where appropriate) the same type of media streams from different dialogs. If the MS does not support starting another dialog on the same connection or conference, it sends a <response> with a 432 status code (Section 4.5) when it receives the second (or subsequent) dialog request.

For example, a request to start an ivr dialog on a connection subscribing to DTMF notifications:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <dialogstart connectionid="connection1">
  <dialog>
   <prompt>
    <media loc="http://www.example.com/getpin.wav"/>
   </prompt>
   <collect maxdigits="2"/>
  </dialog>
 <subscribe>
  <dmtfsub matchmode="all"/>
 </subscribe>
 </dialogstart>
</mscivr>
```

In this example, the dialog is started on a conference where only audio media stream is received:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <dialogstart conferenceid="conference1">
  <dialog>
   <record maxtime="384000s"/>
  </dialog>
  <stream media="audio" direction="recvonly"/>
</mscivr>
```

4.2.2.1. <subscribe>

The <subscribe> element allows the AS to subscribe to, and be notified of, specific events which occur during execution of the dialog. Notifications of dialog events are delivered using the <event> element (see Section 4.2.5).

The <subscribe> element has no attributes.

The <subscribe> element has the following sequence of child elements (0 or more occurrences):

<dtmfsub>: Subscription to DTMF input during the dialog (Section 4.2.2.1.1). The element is optional.

The MS MUST support <dtmfsub> subscription for the IVR dialog language defined in this specification (Section 4.3). It MAY support other dialog subscriptions (specified using attributes and child elements from a foreign namespace). If the MS does not support a subscription specified in a foreign namespace, the MS sends a response with a 431 status code (see Section 4.5).

4.2.2.1.1. <dtmfsub>

The <dtmfsub> element has the following attributes:

matchmode: controls which DTMF input are subscribed to. Valid values are: "all" - notify all DTMF key presses received during the dialog; "collect" - notify only DTMF input matched by the collect operation (Section 4.3.1.3); and "control" - notify only DTMF input matched by the runtime control operation (Section 4.3.1.2). The attribute is optional. The default value is "all".

The <dtmfsub> element has no child elements.

DTMF notifications are delivered in the <dtmfnotify> element (Section 4.2.5.2).

For example, the AS wishes to subscribe to DTMF key press matching a runtime control:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <dialogstart dialogid="d3" connectionid="connection1">
  <dialog>
   <prompt>
    <media loc="http://www.example.com/getpin.wav"/>
   </prompt>
   <control ffkey="2" rewkey="3"/>
  </dialog>
 <subscribe>
  <dmtfnotify matchmode="control"/>
 </subscribe>
 </dialogstart>
```

```
</mscivr>
```

Each time a '2' or '3' DTMF input is received, the MS sends a notification event:


```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <event dialogid="d3">
   <dtmfnotify matchmode="collect" dtmf="2"
       timestamp="2008-05-12T12:13:14Z"/>
 /event>
</mscivr>
```

4.2.2.2. <stream>

The <stream> element has the following attributes:

media: a string indicating the type of media associated with the stream. The following values MUST be used for common types of media: "audio" for audio media, and "video" for video media. The attribute is mandatory.

label: a string indicating the SDP label associated with a media stream ([RFC4574]). The attribute is optional.

direction: a string indicating the direction of the media flow between a dialog and its end point conference or connection. Defined values are: "sendrecv" (the dialog can send and receive media), "sendonly" (the dialog can only send media), "recvonly" (the dialog can only receive media) and "inactive" (stream is not to be used). The default value is "sendrecv". The attribute is optional.

The <stream> element has the following sequence of child elements:

<region>: an element to specify the region within a mixer video layout where a media stream is displayed (Section 4.2.2.2.1). The element is optional.

<priority>: an element to configure priority associated with the stream in the conference mix (Section 4.2.2.2.2). The element is

optional.

If conferenceid is not specified or if the "media" attribute does not have the value of "video", then the MS ignores the <region> and <priority> elements.

For example, assume a user agent connection with multiple audio and video streams associated with the user and a separate web camera. In this case, the dialog could be started to record only the audio and video streams associated with the user:


```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <dialogstart connectionid="connection1">
  <dialog>
   <record maxtime="384000s"/>
  </dialog>
  <stream media="audio" label="camaudio" direction="inactive"/>
  <stream media="video" label="camvideo" direction="inactive"/>
  <stream media="audio" label="useraudio" direction="recvonly"/>
  <stream media="video" label="uservideo" direction="recvonly"/>
</mscivr>
```

Using the <region> element, the dialog can be started on a conference mixer so that the video output from the dialog is directed to a specific region within a video layout. For example:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <dialogstart conferenceid="conference1">
  <dialog>
   <prompt>
     <media loc="http://www.example.com/presentation.3gp"/>
   </prompt>
  </dialog>
  <stream media="video" direction="sendonly">
    <region>r1</region>
  </stream>
</mscivr>
```

4.2.2.2.1. <region>

The <region> element is used to specify the region within a video layout where a video media stream is displayed.

The <region> element has no attributes and its content model specifies the name of the region layout.

If the region name is invalid, then the MS reports a 416 status code (Section 4.5) in the response to the request element containing the <region> element.

4.2.2.2.2. <priority>

The <priority> element is used to explicitly specify the priority of the dialog for presentation in a conference mix.

The <priority> element has no attributes and its content model specifies a positive integer (see Section 4.6.5). The lower the value, the higher the priority.

4.2.3. <dialogterminate>

A dialog can be terminated by sending a <dialogterminate> request element to the MS.

The <dialogterminate> element has the following attributes:

dialogid: string identifying the dialog to terminate. If the specified dialog identifier is invalid, the MS sends a response with a 405 status code (Section 4.5). The attribute is mandatory.

immediate: indicates whether a dialog in the STARTED state is to be terminated immediately or not (in other states, termination is always immediate). A valid value is a boolean (see Section 4.6.1). A value of true indicates that the dialog is terminated immediately and the MS MUST send a dialogexit notification (Section 4.2.5.1) without report information. A value of false indicates that the dialog terminates after the current iteration and the MS MUST send a dialogexit notification with report information. The attribute is optional. The default value is false.

The MS MUST reply to <dialogterminate> request with a <response> element (Section 4.2.4), reporting whether the dialog was terminated successful or not.

For example, immediately terminating a STARTED dialog with dialogid "d4":

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <dialogterminate dialogid="d4" immediate="true"/>
</mscivr>
```

If the dialog is terminated successfully, then the response to the dialogterminate request would be:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <response status="200" dialogid="d4"/>
</mscivr>
```

4.2.4. <response>

Responses to dialog management requests are specified with a <response> element.

The <response> element has following attributes:

status: numeric code indicating the response status. Valid values are defined in Section 4.5. The attribute is mandatory.

reason: string specifying a reason for the response status. The attribute is optional. There is no default value.

dialogid: string identifying the dialog. If the request specifies a dialogid, then that value is used. Otherwise, with <dialogprepare> and <dialogstart> requests, the dialogid generated by the MS is used. If there is no available dialogid because the request is syntactically invalid (e.g. a <dialogterminate> request with no dialogid attribute specified), then the value is the empty string. The attribute is mandatory.

connectionid: string identifying the SIP dialog connection associated with the dialog (see Section 17.1 of [I-D.ietf-mediactrl-sip-control-framework]). The attribute is optional. There is no default value.

conferenceid: string identifying the conference associated with the dialog (see Section 17.1 of [I-D.ietf-mediactrl-sip-control-framework]). The attribute is optional. There is no default value.

For example, a response when a dialog was prepared successfully:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <response status="200" dialogid="d5"/>
</mscivr>
```

The response if dialog preparation failed due to an unsupported dialog language:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <response status="421" dialogid="d5"
    reason="Unsupported dialog language: application/voicexml+xml"/>
</mscivr>
```

In this example a <dialogterminate> request does not specify a dialogid:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <dialogterminate/>
</mscivr>
```

The response status indicates a 400 (Syntax error) status code and dialogid attribute has an empty string value:


```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <response status="400" dialogid=""
    reason="Attribute required: dialogid"/>
</mscivr>
```

4.2.5. <event>

When a dialog generates a notification event, the MS sends the event using an <event> element.

The <event> element has the following attributes:

dialogid: string identifying the dialog which generated the event. The attribute is mandatory.

The <event> element has the following child elements, only one of which can occur:

<dialogexit>: indicates that the dialog has exited (Section 4.2.5.1).

<dtmfnotify>: indicates that a DTMF key press occurred (Section 4.2.5.2).

4.2.5.1. <dialogexit>

The <dialogexit> event indicates that a prepared or active dialog has exited because it is complete, has been terminated, or because an error occurred during execution (for example, a media resource cannot be played). This event MUST be sent by the MS when the dialog exits.

The <dialogexit> element has the following attributes:

status: a status code indicating the status of the dialog when it exits. A valid value is a non-negative integer (see Section 4.6.4). The MS MUST support the following values:

0 indicates the dialog has been terminated by a <dialogterminate> request.

1 indicates successful completion of the dialog.

2 indicates the dialog terminated because the connection or conference associated with the dialog has terminated.

3 indicates the dialog terminated due to exceeding its maximum duration.

4 indicates the dialog terminated due to an execution error.

5 Reserved for future use.

6 Reserved for future use.

7 Reserved for future use.

8 Reserved for future use.

9 Reserved for future use.

The MS MAY define other values. The AS MUST treat any status code it does not recognize as being equivalent to 4 (dialog execution error). The attribute is mandatory.

reason: a textual description which the MS SHOULD use to provide a reason for the status code; e.g. details about an error. A valid value is a string (see Section 4.6.6). The attribute is optional. There is no default value.

The <dialogexit> element has the following sequence of child elements:

<promptinfo>: report information (Section 4.3.2.1) about the prompt execution in an IVR <dialog>. The element is optional.

<controlinfo>: reports information (Section 4.3.2.2) about the control execution in an IVR <dialog>. The element is optional.

<collectinfo>: reports information (Section 4.3.2.3) about the collect execution in an IVR <dialog>. The element is optional.

<recordinfo>: reports information (Section 4.3.2.4) about the record execution in an IVR <dialog>. The element is optional.

<params>: reports exit parameters (Section 4.2.6) for a dialog language defined outside this specification. The element is optional.

For example, an active <dialog> exits normally the MS sends a dialogexit <event> reporting information:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <event dialogid="d6"/>
  <dialogexit status="1">
   <collectinfo dtmf="1234"  termmode="match"/>
  </dialogexit>
 </event>
</mscivr>
```

4.2.5.2. <dtmfnotify>

The <dtmfnotify> element provide a notification of DTMF input received during the active dialog as requested by a <dtmfsub> subscription (Section 4.2.2.1).

The <dtmfnotify> element has the following attributes:

matchmode: indicates the matching mode specified in the subscription request. Valid values are: "all" - all DTMF key presses notified individually; "collect" - only DTMF input matched by the collect operation notified; and "control" - only DTMF input matched by the control operation notified. The attribute is optional. The default value is "all".

dtmf: DTMF key presses received according to the matchmode. A valid value is a DTMF string (see Section 4.6.3) with no space between characters. The attribute is mandatory.

timestamp: indicates the time (on the MS) at which the last key press occurred according to the matchmode. A valid value is a dateTime expression (Section 4.6.12). The attribute is mandatory.

For example, a notification of DTMF input matched during the collect operation:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <event dialogid="d3">
   <dtmfnotify matchmode="collect" dtmf="3123"
     timestamp="2008-05-12T12:13:14Z"/>
 /event>
</mscivr>
```

4.2.6. <params>

The <params> element is a container for <param> elements (Section 4.2.6.1).

The <params> element has no attributes, but the following child elements are defined (0 or more):

<param>: specifies a parameter name and value (Section 4.2.6.1).

For example, usage with a dialog language defined outside this specification to send additional parameters into the dialog:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
<dialogstart type="application/x-dialog"
             src="nfs://nas01/dialog4" connectionid="c1">
  <params>
   <param name="mode">playannouncement</param>
   <param name="prompt1">nfs://nas01/media1.3gp</param>
   <param name="prompt2">nfs://nas01/media2.3gp</param>
  </params>
 </dialogstart>
</mscivr>
```

4.2.6.1. <param>

The <param> element describes a parameter name and value.

The <param> element has the following attributes:

name: a string indicating the name of the parameter. The attribute is mandatory.

type: specifies a type indicating how the inline value of the parameter is to be interpreted. A valid value is a MIME media type (see Section 4.6.10). The attribute is optional. The default value is "text/plain".

The <param> element content model is the value of the parameter. Note that a value which contains XML characters (e.g. "<") needs to be escaped following standard XML conventions.

For example, usage with a dialog language defined outside this specification to receive parameters from the dialog when it exits:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <event dialogid="d6"/>
  <dialogexit status="1">
  <params>
   <param name="mode">recording</param>
   <param name="recording1" type="audio/x-wav">

   <![CDATA[
     R0lGODlhZABqALMAAFrMYr/BvlKOVJKOg2xZUKmenMfDw8tgWJpV
     ]]>

   </param>
  </params>
  </dialogexit>
 </event>
</mscivr>
```

4.3. IVR Dialog Elements

This section describes the IVR dialog language defined as part of this specification. The MS MUST support this dialog language.

The <dialog> element is an execution container for operations of playing prompts (Section 4.3.1.1), runtime controls (Section 4.3.1.2), collecting DTMF (Section 4.3.1.3),and recording user input (Section 4.3.1.4. Results of the dialog execution (Section 4.3.2) are reported in a dialogexit notification event.

Using these elements, three common dialog models are supported:

playannouncements: only a <prompt> element is specified in the container. The prompt media resources are played in sequence.

promptandcollect: a <collect> element is specified and, optionally, a <prompt> element. If a <prompt> element is specified and bargein is enabled, playing of the prompt is terminated when bargein occurs, and DTMF collection is initiated; otherwise, the prompt is played to completion before DTMF collection is initiated. If no prompt element is specified, DTMF collection is initiated immediately.

promptandrecord: a <record> element is specified and, optionally, a <prompt> element. If a <prompt> element is specified and bargein is enabled, playing of the prompt is terminated when bargein occurs, and recording is initiated; otherwise, the prompt is played to completion before recording is initiated. If no prompt element is specified, recording is initiated immediately.

In addition, this dialog language supports runtime ('VCR') controls enabling a user to control prompt playback using DTMF.

Each of the core elements - <prompt>, <control>, <collect> and <record> - are specified so that their execution and reporting is largely self-contained. This facilitates their re-use in other dialog container elements. Note that DTMF and bargein behavior affects multiple elements and is addressed in the relevant element definitions.

Execution results are reported in the <dialogexit> notification event with child elements defined in Section 4.3.2. If the dialog terminated normally (i.e. not due to an error or to a <dialogterminate> request), then the MS MUST report the results for the operations specified in the dialog:

<prompt>: <promptinfo> (see Section 4.3.2.1) with at least the termmode attribute specified.

<control>: <controlinfo> (see Section 4.3.2.2) if any runtime controls are matched.

<collect>: <collectinfo> (see Section 4.3.2.3) with the dtmf and termmode attributes specified.

<record>: <recordinfo> (see Section 4.3.2.4) with at least the termmode attribute and one <mediainfo> element specified.

The media format requirements for IVR dialogs are undefined. This package is agnostic to the media types and codecs for media resources and recording which need to be supported by an implementation. For example, a MS implementation might only support audio and in particular the 'audio/basic' codec for media playback and recording. However, when executing a dialog, if an MS encounters a media type or codec which it cannot process, the MS MUST stop further processing and report the error using the dialogexit notification.

4.3.1. <dialog>

An IVR dialog to play prompts to the user, allow runtime controls, collect DTMF or record input. The dialog is specified using a <dialog> element.

A <dialog> element has the following attributes:

repeatCount: number of times the dialog is to be executed. A valid value is a non-negative integer (see Section 4.6.4). A value of 0 indicates that the dialog is repeated until halted by other means. The attribute is optional. The default value is 1.

repeatDur: maximum duration for dialog execution. A valid value is a Time Designation (see Section 4.6.7). If no value is specified, then there is no limit on the duration of the dialog. The attribute is optional. There is no default value.

The repeatDur attribute takes priority over the repeatCount attribute in determining maximum duration of the dialog. See 'repeatCount' and 'repeatDur' in SMIL ([W3C.REC-SMIL2-20051213]) for further information. In the situation where a dialog is repeated more than once, only the results of operations in the last dialog iteration are

reported.

The <dialog> element has the following sequence of child elements (at least one, any order):

<prompt>: defines media resources to play in sequence (see Section 4.3.1.1). The element is optional.

<control>: defines how DTMF is used for runtime controls (see Section 4.3.1.2). The element is optional.

<collect>: defines how DTMF is collected (see Section 4.3.1.3). The element is optional.

<record>: defines how recording takes place (see Section 4.3.1.4). The element is optional.

Although the behavior when both <collect> and <record> elements are specified in a request is not defined in this control package, the MS MAY support this configuration. If the MS does not support this configuration, the MS sends a <response> with a 433 status code.

The MS has the following execution model for the IVR dialog after initialization (initialization errors are reported by the MS in the response):

1. If an error occurs during execution, then the MS terminates the dialog and reports the error in the <dialogexit> event by setting the status attribute (see Section 4.3.2). Details about the error are specified in the reason attribute.

2. The MS initializes a counter to 0.

3. The MS starts a duration timer for the value of the repeatDur attribute. If the timer expires before the dialog is complete, then the MS terminates the dialog and sends a dialogexit whose status attribute is set to 3 (see Section 4.2.5.1). The MS MAY report information in the dialogexit gathered in the last execution cycle (if any).

4. The MS initiates a dialog execution cycle. Each cycle executes

the operations associated with the child elements of the dialog. If a <prompt> element is specified, then execute the element's prompt playing operation and activate any controls (if the <control> element is specified). If no <prompt> is specified or when a specified <prompt> terminates, then start the collect operation or the record operation if the <collect> or <record> elements respectively are specified. If subscriptions are specified for the dialog, then the MS sends a notification event when the specified event occurs. If execution of a child element results in an error, the MS terminates dialog execution (and stops other child element operations) and the MS sends a dialogexit status event, reporting any information gathered.

5. If the dialog execution cycle completes successfully, then the MS increments the counter by one. If the value of the repeatCount attribute is greater than zero and the counter is equal to the value of the repeatCount attribute, then the MS terminates dialog execution and the sends a dialogexit (with a status of 1) reporting operation information collected in the last dialog execution cycle only. Otherwise, another dialog execution cycle is initiated.

4.3.1.1. <prompt>

The <prompt> element specifies a sequence of media resources to play back in document order.

A <prompt> element has the following attributes:

xml:base: A string declaring the base URI from which relative URIs in child elements are resolved prior to fetching. A valid value is a URI (see Section 4.6.9). The attribute is optional. There is no default value.

bargein: Indicates whether user input stops prompt playback unless the input is associated with a specified runtime <control> operation (input matching control operations never interrupts prompt playback). A valid value is a boolean (see Section 4.6.1). A value of true indicates that bargein is permitted and prompt playback is stopped. A value of false indicates that bargein is

not permitted: user input does not terminate prompt playback. The

attribute is optional. The default value is true.

The <prompt> element has the following child elements (at least one, any order, multiple occurrences of elements permitted):

<media>: specifies a media resource (see Section 4.3.1.5) to play. The element is optional.

<variable>: specifies a variable media announcement (see Section 4.3.1.1.1) to play. The element is optional.

<dtmf>: generates one or more DTMF tones (see Section 4.3.1.1.2) to play. The element is optional.

<par>: specifies media resources to play in parallel (see Section 4.3.1.1.3). The element is optional.

If the MS does not support the configuration required for prompt playback to the output media streams and a more specific error code is not defined for its child elements, the MS sends a <response> with a 429 status code (Section 4.5). The MS MAY support transcoding between the media resource format and the output stream format.

The MS has the following execution model for prompt playing after initialization:

1. The MS initiates prompt playback playing its child elements (<media>, <variable>, <dtmf> and <par>) one after another in document order.

2. If any error (including fetching and rendering errors) occurs during prompt execution, then the MS terminates playback and reports its error status to the dialog container (see Section 4.3) with a <promptinfo> (see Section 4.3.2.1) where the termmode attribute is set to stopped and any additional information is set.

3. If DTMF input is received and the value of the bargein attribute is true, then the MS terminates prompt playback and reports its execution status to the dialog container (see Section 4.3) with a <promptinfo> (see Section 4.3.2.1) where the termmode attribute is set to bargein and any additional information is set.

4. If prompt playback is stopped by the dialog container, then the MS reports its execution status to the dialog container (see Section 4.3) with a <promptinfo> (see Section 4.3.2.1) where the termmode attribute is set to stopped and any additional

information is set.

5. If prompt playback completes successfully, then the MS reports its execution status to the dialog container (see Section 4.3) with a <promptinfo> (see Section 4.3.2.1) where the termmode attribute is set to completed and any additional information is set.

4.3.1.1.1. <variable>

The <variable> element specifies variable announcements using predefined media resources. Each variable has at least a type (e.g. date) and a value (e.g. 2008-02-25). The value is rendered according to the variable type (e.g. 25th February 2008) as well as other defined attributes. The precise mechanism for generating variable announcements (including the location of associated media resources) is implementation specific.

A <variable> element has the following attributes:

value: specifies the string to be rendered. A valid value is a string (see Section 4.6.6). The attribute is mandatory.

type: specifies the type to use for rendering. A valid value is a string (see Section 4.6.6). The attribute is mandatory.

format: specifies format information to use in conjunction with the type for the rendering. A valid value is a string (see Section 4.6.6). The attribute is optional. There is no default value.

gender: specifies the gender to use when rendering the variable. Valid values are "male" or "female". The attribute is optional. There is no default value.

xml:lang: specifies the language to use when rendering the variable. A valid value is a language identifier (see Section 4.6.11). The attribute is optional. There is no default value.

The <variable> element has no children.

This package is agnostic to which <variable> values, types and formats are supported by an implementation. If a <variable> element configuration specified in a request is not supported by the MS, the MS sends a <response> with a 425 status code (Section 4.5).

For example, the MS could support <variable> type/format combinations such as:

type=date Supported formats: "mdy" (month day year), "ymd" (year month day), "dym" (day month year), "dm" (day month). The value attribute has the format YYYY-MM-DD (4 digit year, 2 digit month, 2 digit day).

type=time Supported formats: "t12" (12 hour format with am/pm), "t24" (24 hour format). The value attribute has the format HH:MM (2 digit hours, 2 digit minutes).

type=digits Supported formats: "gen" (general digit string), "crn" (cardinal), "ord" (ordinal). The value attribute has the format of "D+" (one or more digits).

This specification is agnostic to the type and codec of media resources into which variable are rendered as well as the rendering mechanism itself. For example, an MS implementation supporting audio rendering could map the <variable> into one or more audio media resources.

Depending on the specific implementation of the <variable> rendering on the MS, execution of this element can be seen as conversion of a <variable> into a list of <media> elements. For example,

```
<variable value="2008-02-25" type="date" format="dmy"
xml:lang="en" gender="male"/>
```

could be transformed into audio saying "twenty-fifth of February two thousand and eight" using a list of <media> resources:

```
<media loc="nfs://voicebase/en/male/25th.wav"/>
<media loc="nfs://voicebase/en/male/of.wav"/>
<media loc="nfs://voicebase/en/male/february.wav"/>
<media loc="nfs://voicebase/en/male/2000.wav"/>
<media loc="nfs://voicebase/en/male/and.wav"/>
<media loc="nfs://voicebase/en/male/8.wav"/>
```

4.3.1.1.2. <dtmf>

The <dtmf> element specifies a sequence of DTMF tones for output.

DTMF tones could be generated using <media> resources where the output is transported as RTP audio packets. However, <media> resources are not sufficient for cases where DTMF tones are to be transported as DTMF RTP ([RFC4733]) or in event packages.

A <dtmf> element has the following attributes:

digits: specifies the DTMF sequence to output. A valid value is a DTMF string (see Section 4.6.3). The attribute is mandatory.

level: used to define the power level for which the DTMF tones will be generated. Values are expressed in dBm0. A valid value is an integer in the range of 0 to -96 (dBm0). Larger negative values express lower power levels. Note that values lower than -55 dBm0 will be rejected by most receivers (TR-TSY-000181, ITU-T Q.24A). The attribute is optional. The default value is -6 (dBm0).

duration: specifies the duration for which each DTMF tone is generated. A valid value is a time designation (see Section 4.6.7). The MS MAY round the value if it only supports discrete durations. The attribute is optional. The default value is 100ms.

interval: specifies the duration of a silence interval following each generated DTMF tone. A valid value is a time designation (see Section 4.6.7). The MS MAY round the value if it only supports discrete durations. The attribute is optional. The default value is 100ms.

The <dtmf> element has no children.

If a <dtmf> element configuration is not supported, the MS sends a <response> with a 426 status code (Section 4.5).

4.3.1.1.3. <par>

The <par> element allows media resources to be played in parallel. Each of its child elements specify a media resource (or a sequence of media resources using the <seq> element). When playback of the <par>

element is initiated, the MS begins playback of all its child elements at the same time. This element is modeled after the <par> element in SMIL ([W3C.REC-SMIL2-20051213]).

The <par> element has the following attributes:

endsync: indicates when playback of the element is complete. Valid values are: "first" - indicates that the element is complete when any child element reports that it is complete; "last" - indicates it is complete when every child elements are complete. The attribute is optional. The default value is "last".

If the value is "first", then playback of other child element is stopped when one child element reports it is complete.

The <par> element has the following child elements (at least one, any

order, multiple occurrences of each element permitted):

<seq>: specifies a sequence of media resources to play in a parallel with other <par> child elements (see Section 4.3.1.1.3.1). The element is optional.

<media>: specifies a media resource (see Section 4.3.1.5) to play. The element is optional.

<variable>: specifies a variable media announcement (see Section 4.3.1.1.1) to play. The element is optional.

<dtmf>: generates one or more DTMF tones (see Section 4.3.1.1.2) to play. The element is optional.

If a <par> element configuration is not supported, the MS sends a <response> with a 435 status code (Section 4.5).

Runtime <control>s (Section 4.3.1.2) apply to each child element playing parallel. For example, pause and resume controls cause all child elements to be paused and resumed respectively.

If the <par> element is stopped by the prompt container (e.g. bargein or dialog termination), then playback of all child elements is stopped. The playback duration (Section 4.3.2.1) reported for the

<par> element is the duration of parallel playback, not the cumulative duration of each child element played in parallel.

For example, a request to playback audio and video media in parallel:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
<dialogstart connectionid="c1">
  <dialog>
   <prompt>
    <par>
     <media type="audio/x-wav"
            loc="http://www.example.com/media/comments.wav"/>
     <media type="video/3gpp;codecs='s263'"
            loc="http://www.example.com/media/camera.3gp"/
    </par>
   </prompt>
  </dialog>
 </dialogstart>
</mscivr>
```

When the <prompt> element is executed, it begins playback of its child element in document order sequence. In this case, there is only one child element, a <par> element itself containing audio and

video <media> child element. Consequently playback of both audio and video media resources is initiated at the same time. Since the endsync attribute is not specified, the default value "last" applies. The <par> element playback is complete when the media resource with the longest duration is complete.

4.3.1.1.3.1. <seq>

The <seq> element specifies media resources to be played back in sequence. This allows a sequence of media resources to be played at the same time as other children of a <par> element are played in parallel. For example, a sequence of audio resources while a video resource is played in parallel. This element is modeled after the <seq> element in SMIL ([W3C.REC-SMIL2-20051213]).

The <seq> element has no attributes.

The <seq> element has the following child elements (at least one, any

order, multiple occurrences of each element permitted):

<media>: specifies a media resource (see Section 4.3.1.5) to play. The element is optional.

<variable>: specifies a variable media announcement (see Section 4.3.1.1.1) to play. The element is optional.

<dtmf>: generates one or more DTMF tones (see Section 4.3.1.1.2) to play. The element is optional.

Playback of a <seq> element is complete when all child elements in the sequence are complete. If the <seq> element is stopped by the <par> container, then playback of the current child element is stopped (remaining child elements in the sequence are not played).

For example, a request to play a sequence of audio resources in parallel with a video media:


```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
<dialogstart connectionid="c1">
  <dialog>
   <prompt>
    <par endsync="first">
     <seq>
       <media type="audio/x-wav"
            loc="http://www.example.com/media/date.wav"/>
       <media type="audio/x-wav"
            loc="http://www.example.com/media/intro.wav"/>
       <media type="audio/x-wav"
```

```
              loc="http://www.example.com/media/main.wav"/>
          <media type="audio/x-wav"
              loc="http://www.example.com/media/end.wav"/>
        </seq>
        <media type="video/3gpp;codecs='s263'"
              loc="rtsp://www.example.com/media/camera.3gp"/
      </par>
     </prompt>
    </dialog>
   </dialogstart>
  </mscivr>
```

When the <prompt> element is executed, it begins playback of the <par> element containing a <seq> element and a video <media> element. The <seq> element itself contains a sequence of audio <media> elements. Consequently playback of the video media resource is initiated at the same time as playback of the sequence of the audio media resources is initiated. Each audio resource is played back after the previous one completes. Since the endsync attribute is set to "first", the <par> element playback is complete when either all the audio resources in <seq> have been played to completion or the video <media> is complete, whichever occurs first.

4.3.1.2. <control>

The <control> element defines how DTMF input is mapped to runtime controls, including prompt playback controls.

DTMF input matching these controls MUST NOT cause prompt playback to interrupted (i.e. no prompt bargein), but causes the appropriate operation to be applied; for examples, speeding up prompt playback.

DTMF input matching these controls has priority over <collect> input for the duration of prompt playback. If incoming DTMF matches a specified runtime control, then the DTMF is not available to the <collect> operation, including its digit buffer. Once prompt playback is complete, runtime controls are no longer active.

The <control> element has the following attributes:

gotostartkey: maps a DTMF key to skip directly to the start of the prompt. A valid value is a DTMF Character (see Section 4.6.2).

The attribute is optional. There is no default value.

gotoendkey: maps a DTMF key to skip directly to the end of the prompt. A valid value is a DTMF Character (see Section 4.6.2). The attribute is optional. There is no default value.

skipinterval: indicates how far a MS skips backwards or forwards through prompt playback when the rewind (rwkey) of fast forward key (ffkey) is pressed. A valid value is a Time Designation (see Section 4.6.7). The attribute is optional. The default value is 6s.

ffkey: maps a DTMF key to a fast forward operation equal to the value of 'skipinterval'. A valid value is a DTMF Character (see Section 4.6.2). The attribute is optional. There is no default value.

rwkey: maps a DTMF key to a rewind operation equal to the value of 'skipinterval'. A valid value is a DTMF Character (see Section 4.6.2). The attribute is optional. There is no default value.

pauseinterval: indicates how long a MS pauses prompt playback when the pausekey is pressed. A valid value is a Time Designation (see Section 4.6.7). The attribute is optional. The default value is 10s.

pausekey: maps a DTMF key to a pause operation equal to the value of 'pauseinterval'. A valid value is a DTMF Character (see Section 4.6.2). The attribute is optional. There is no default value.

resumekey: maps a DTMF key to a resume operation. A valid value is a DTMF Character (see Section 4.6.2). The attribute is optional. There is no default value.

volumeinterval: indicates the increase or decrease in playback volume (relative to the current volume) when the volupkey or voldnkey is pressed. A valid value is a percentage (see Section 4.6.8). The attribute is optional. The default value is 10%.

volupkey: maps a DTMF key to a volume increase operation equal to the value of 'volumeinterval'. A valid value is a DTMF Character (see Section 4.6.2). The attribute is optional. There is no default value.

voldnkey: maps a DTMF key to a volume decrease operation equal to the value of 'volumeinterval'. A valid value is a DTMF Character (see Section 4.6.2). The attribute is optional. There is no default value.

speedinterval: indicates the increase or decrease in playback speed (relative to the current speed) when the speedupkey or speeddnkey is pressed. A valid value is a percentage (see Section 4.6.8). The attribute is optional. The default value is 10%.

speedupkey: maps a DTMF key to a speed increase operation equal to the value of the speedinterval attribute. A valid value is a DTMF Character (see Section 4.6.2). The attribute is optional. There is no default value.

speeddnkey: maps a DTMF key to a speed decrease operation equal to the value of the speedinterval attribute. A valid value is a DTMF Character (see Section 4.6.2). The attribute is optional. There is no default value.

external: allows one or more DTMF keys to be declared as external controls (for example: video camera controls); the MS can send notifications when a matching key is activated using <dtmfnotify> (Section 4.2.5.2). A valid value is a DTMF String (see Section 4.6.3). The attribute is optional. There is no default value.

If the same DTMF is specified in more than one DTMF key control attribute - except the pausekey and resumekey attributes - the MS sends a <response> with a 413 status code (Section 4.5).

The MS has the following execution model for runtime control after initialization:

1. If an error occurs during execution, then the MS terminates runtime control and the error is reported to the dialog container. The MS MAY report controls executed successfully before the error in <controlinfo> (see Section 4.3.2.2).

2. Runtime controls are active only during prompt playback (if no <prompt> element is specified, then runtime controls are ignored). If DTMF input matches any specified keys (for example

the ffkey), then the MS applies the appropriate operation

immediately. If a seek operation (ffkey, rwkey) attempts to go beyond the beginning or end of the prompt queue, then the MS automatically truncates it to the prompt queue beginning or end respectively. If a volume operation (voldnkey, volupkey) attempts to go beyond the minimum or maximum volume supported by the platform, then the MS automatically limits the operation to minimum or maximum supported volume respectively. If a speed operation (speeddnkey, speedupkey) attempts to go beyond the minimum or maximum playback speed supported by the platform, then the MS automatically limits the operation to minimum or maximum supported speed respectively. If the pause operation attempts to pause output when it is already paused, then the operation is ignored. If the resume operation attempts to resume when the prompts are not paused, then the operation is ignored. If a seek, volume or speed operation is applied when output is paused, then the MS also resumes output automatically.

3. If DTMF control subscription has been specified for the dialog, then each DTMF match of a control operation is reported in a <dtmfnotify> notification event (Section 4.2.5.2).

4. When the dialog exits, all control matches are reported in a <controlinfo> element (Section 4.3.2.2).

4.3.1.3. <collect>

The <collect> element defines how DTMF input is collected.

The <collect> element has the following attributes:

cleardigitbuffer: indicates whether the digit buffer is to be cleared. A valid value is a boolean (see Section 4.6.1). A value of true indicates that the digit buffer is to be cleared. A value of false indicates that the digit buffer is not to be cleared. The attribute is optional. The default value is true.

timeout: indicates the maximum time to wait for user input to begin. A valid value is a Time Designation (see Section 4.6.7). The attribute is optional. The default value is 5s.

interdigittimeout: indicates inter-digit timeout value to use when recognizing DTMF input. A valid value is a Time Designation (see Section 4.6.7). The attribute is optional. The default value is 2s.

termtimeout: indicates the terminating timeout value to use when recognizing DTMF input. A valid value is a Time Designation (see Section 4.6.7). The attribute is optional. The default value is 0s (no delay).

escapekey: specifies a DTMF key that indicates the DTMF collection is to be re-initiated. A valid value is a DTMF Character (see Section 4.6.2). The attribute is optional. There is no default value.

termchar: specifies a DTMF character for terminating DTMF input collection using the internal grammar. A valid value is a DTMF character (see Section 4.6.2). To disable termination by a conventional DTMF character, set the parameter to an unconventional character like 'A'. The attribute is optional. The default value is '#'.

maxdigits: The maximum number of digits to collect using an internal digits (0-9 only) grammar. A valid value is a positive integer (see Section 4.6.5). The attribute is optional. The default value is 5.

The <collect> element has the following child elements:

<grammar>: indicates a custom grammar format (see Section 4.3.1.3.1). The element is optional.

The custom grammar takes priority over the internal grammar. If a <grammar> element is specified, the MS MUST use it for DTMF collection.

The MS has the following execution model for DTMF collection after initialization:

1. The DTMF collection buffer MUST NOT receive DTMF input matching <control> operations (see Section 4.3.1.2).

2. If an error occurs during execution, then the MS terminates collection and reports the error to the dialog container (see Section 4.3). The MS MAY report DTMF collected before the error in <collectinfo> (see Section 4.3.2.3).

3. The MS clears the digit buffer if the value of the cleardigitbuffer attribute is true.

4. The MS activates a timer with the duration of the value of the timeout attribute. If the timer expires before DTMF input collection begins, then collection execution terminates, the

<collectinfo> (see Section 4.3.2.3) has the termmode attribute set to noinput and the execution status reported to the dialog container.

5. If DTMF collect input matches the value of the escapekey attribute, then the MS cancels the timer and re-initializes DTMF collection.

6. Other DTMF collect input is matched to the grammar. Valid DTMF patterns are either a simple digit string where the maximum length is determined by the maxdigits attribute and which can be terminated by the character in the termchar attribute; or a custom DTMF grammar specified with the <grammar> element. The attributes interdigittimeout and termtimeout control interdigit timeout and the terminating timeout respectively.

7. If the collect input completely matches the grammar, the timer is canceled, the MS terminates collection execution and reports execution status to the dialog container with <collectinfo> (see Section 4.3.2.3) where the termmode attribute set to match.

8. If the collect input does not match the grammar, the MS cancels the timer, terminates collection execution and reports execution status to the dialog container with a <collectinfo> (see Section 4.3.2.3) where the termmode attribute set to nomatch.

4.3.1.3.1. <grammar>

The <grammar> element allows a custom grammar, inline or external, to be specified. Custom grammars permit the full range of DTMF characters including '*' and '#' to be specified for DTMF pattern matching.

The <grammar> element has the following attributes:

src: specifies the location of an external grammar document. A valid value is a URI (see Section 4.6.9). If the URI scheme is unsupported, the MS sends a <response> with a 420 status code (Section 4.5). If the resource cannot be retrieved within the timeout interval, the MS sends a <response> with a 409 status code. If the grammar format is not supported, the MS sends a <response> with a 424 status code. The attribute is optional. There is no default value.

type: identifies the preferred type of the grammar document identified by the src attribute. A valid value is a MIME media type (see Section 4.6.10). If the URI scheme used in the src attribute defines a mechanism for establishing the authoratitive

MIME media type of the media resource, the value returned by that mechanism takes precedence over this attribute. The attribute is optional. There is no default value.

fetchtimeout: the maximum interval to wait when fetching a grammar resource. A valid value is a Time Designation (see Section 4.6.7). The attribute is optional. The default value is 30s.

The <grammar> element allows inline grammars to be specified. XML grammar formats MUST use a namespace other than the one used in this specification. Non-XML grammar formats MAY use a CDATA section.

The MS MUST support the [SRGS] XML grammar format ("application/srgs+xml") and MS MAY support KPML ([RFC4730]) or other grammar formats. If the grammar format is not supported by the MS, then the MS sends a <response> with a 424 status code (Section 4.5).

For example, the following fragment shows DTMF collection with an

inline SRGS grammar:


```
<collect cleardigitbuffer="false" timeout="20s"
        interdigittimeout="1s">
        <grammar>
                <grammar xmlns="http://www.w3.org/2001/06/grammar"
                                version="1.0" mode="dtmf">
                  <rule id="digit">
                   <one-of>
                        <item>0</item>
                        <item>1</item>
                        <item>2</item>
                        <item>3</item>
                        <item>4</item>
```

```
                        <item>5</item>
                        <item>6</item>
                        <item>7</item>
                        <item>8</item>
                        <item>9</item>
                    </one-of>
                </rule>

                <rule id="pin" scope="public">
                 <one-of>
                        <item>
                         <item repeat="4">
                          <ruleref uri="#digit"/>
                                </item>#</item>
                                <item>* 9</item>
                 </one-of>
                </rule>

          </grammar>
            </grammar>
           </collect>
```

The same grammar could also be referenced externally (and take advantage of HTTP caching):

```
<collect cleardigitbuffer="false" timeout="20s">
   <grammar type=""application/srgs+xml"
            src="http://example.org/pin.grxml"/>
</collect>
```

4.3.1.4. <record>

The <record> element specifies how media input is recorded.

The <record> element has the following attributes:

timeout: indicates the time to wait for user input to begin. A valid value is a Time Designation (see Section 4.6.7). The attribute is optional. The default value is 5s.

vadinitial: Control whether voice activity detection (VAD) is used

to initiate the recording operation. A valid value is a boolean (see Section 4.6.1). A value of true indicates the MS MUST initiate recording if the VAD detects voice on the configured inbound audio streams. A value of false indicates that the MS MUST NOT initiate recording using VAD. The attribute is optional. The default value is false.

vadfinal: Control whether voice activity detection (VAD) is used to terminate the recording operation. A valid value is a boolean (see Section 4.6.1). A value of true indicates the MS MUST terminate recording if the VAD detects a period of silence (whose duration is specified by the finalsilence attribute) on configured inbound audio streams. A value of false indicates that the MS MUST NOT terminate recording using VAD. The attribute is optional. The default value is false.

dtmfterm: Indicates whether the recording operation is terminated by DTMF input. A valid value is a boolean (see Section 4.6.1). A value of true indicates that recording is terminated by DTMF input. A value of false indicates that recording is not terminated by DTMF input. The attribute is optional. The default value is true.

maxtime: indicates The maximum duration of the recording. A valid value is a Time Designation (see Section 4.6.7). The attribute is optional. The default value is 15s.

beep: indicates whether a 'beep' is to be played immediately prior to initiation of the recording operation. A valid value is a boolean (see Section 4.6.1). The attribute is optional. The default value is false.

finalsilence: indicates the interval of silence that indicates the end of voice input. This interval is not part of the recording itself. This parameter is ignored if the vadfinal attribute has the value false. A valid value is a Time Designation (see Section 4.6.7). The attribute is optional. The default value is 5s.

append: indicates whether recorded data is appended or not to a recording location if a resource already exists. A valid value is a boolean (see Section 4.6.1). A value of true indicates that recorded data is appended to the existing resource at a recording

location. A value of false indicates that recorded data is to overwrite the existing resource. The attribute is optional. The default value is false.

If either the vadinitial or vadfinal attribute is set to true and the MS does not support VAD, the MS sends a <response> with a 434 status code (Section 4.5).

The <record> element has the following child element (0 or more occurrences):

<media>: specifies the location and type of the media resource for uploading recorded data (see Section 4.3.1.5). The MS uploads recorded data to this resource as soon as possible after recording is complete. The element is optional.

If multiple <media> elements are specified, then media input is to be recorded in parallel to multiple resource locations.

If no <media> child element is specified, the MS MUST provide a recording location where the recording format is implementation-specific. The recording location and format are reported in <recordinfo> (Section 4.3.2.4) when the dialog terminates. The recording MUST be available from this location until the connection or conference associated with the dialog on the MS terminates.

If the MS does not support the configuration required for recording from the input media streams to one or more <media> elements and a more specific error code is not defined for its child elements, the MS sends a <response> with a 423 status code (Section 4.5).

Note that an MS MAY support uploading recorded data to recording locations at the same time the recording operation takes place. Such implementations need to be aware of the requirements of certain recording formats (e.g. WAV) for metadata at the beginning of the uploaded file, that the finalsilence interval is not part of the recording and how these requirements interact with the URI scheme.

The MS has the following execution model for recording after initialization:

1. If an error occurs during execution (e.g. authentication or communication error when trying to upload to a recording location), then the MS terminates record execution and reports the error to the dialog container (see Section 4.3). The MS MAY report data recorded before the error in <recordinfo> (see Section 4.3.2.4).

2. If DTMF input (not matching a <control> operation) is received during prompt playback and the prompt bargein attribute is set to true, then the MS activates the record execution. Otherwise, the MS activates it after the completion of prompt playback.

3. If a beep attribute with the value of true is specified, then the MS plays a beep tone.

4. The MS activates a timer with the duration of the value of the timeout attribute. If the timer expires before the recording operation begins, then the MS terminates the recording execution and reports the status to dialog container with <recordinfo> (see Section 4.3.2.4) where the termmode attribute is set to noinput.

5. Initiation of the recording operation depends on the value of the vadinitial attribute. If vadinitial has the value false, then the recording operation is initiated immediately. Otherwise, the recording operation is initiated when voice activity is detected.

6. When the recording operation is initiated, a timer is started for the value of the maxtime attribute (maximum duration of the recording). If the timer expires before the recording operation is complete, then the MS terminates recording execution and reports the execution status to the dialog container with <recordinfo> (see Section 4.3.2.4) where the termmode attribute set to maxtime.

7. During the record operation input media streams are recording to a location and format specified in one or more <media> child elements. If no <media> child element is specified, the MS records input to an implementation-specific location and format.

8. If the dtmfterm attribute has the value true and DTMF input is detected during the record operation, then the MS terminates recording and its status is reported to the dialog container with a <recordinfo> (see Section 4.3.2.4) where the termmode attribute is set to dtmf.

9. If vadfinal attribute has the value true, then the MS terminates the recording operation when a period of silence, with the duration specified by the value of the finalsilence attribute, is detected. This period of silence is not part of the final

recording. The status is reported to the dialog container with a <recordinfo> (see Section 4.3.2.4) where the termmode attribute is set to finalsilence.

For example, a request to record audio and video input to separate locations:


```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
<dialogstart connectionid="c1">
  <dialog>
   <record maxtime="30s" vadinitial="false" vadfinal="false">
    <media type="audio/x-wav"
        loc="http://www.example.com/upload/audio.wav"/>
    <media type="video/3gpp;codecs='s263'"
        loc="http://www.example.com/upload/video.3gp"/>
   </record>
  </dialog>
 </dialogstart>
</mscivr>
```

When the <record> element is executed, it immediately begins recording of the audio and video (since vadinitial is false) where the destination locations are specified in the <media> child elements. Recording is completed when the duration reaches 30s or the connection is terminated.

4.3.1.5. <media>

The <media> element specifies a media resource to playback from (see Section 4.3.1.1) or record to (see Section 4.3.1.4). In the playback case, the resource is retrieved and in the recording case, recording data is uploaded to the resource location.

A <media> element has the following attributes:

loc: specifies the location of the media resource. A valid value is a URI (see Section 4.6.9) including authentication information if defined by the URI scheme (e.g. basic access authentication in HTTP). If the URI scheme is not supported by the MS, the MS sends a <response> with a 420 status code (Section 4.5). If the resource is to be retrieved but the MS cannot retrieve it within the timeout interval, the MS sends a <response> with a 409 status

code. If the format of the media resource is not supported, the MS sends a <response> with a 429 status code. The attribute is mandatory.

type: specifies the type of the media resource indicated in the loc attribute. A valid value is a MIME media type (see Section 4.6.10) which, depending on its definition, can include additional parameters (e.g. [RFC4281]). If the URI scheme used in the loc attribute defines a mechanism for establishing the authoratitive MIME media type of the media resource, the value returned by that mechanism takes precedence over this attribute. If additional media parameters are specified, the MS MUST use them to determine media processing. For example, [RFC4281] defines a

'codec' parameter for media types like video/3gpp which would determine which media streams are played or recorded. The attribute is optional. There is no default value.

fetchtimeout: the maximum interval to wait when fetching a media resource. A valid value is a Time Designation (see Section 4.6.7). The attribute is optional. The default value is 30s.

soundLevel: playback soundLevel (volume) for the media resource. A valid value is a percentage (see Section 4.6.4). The value indicates increase or decrease relative to the original recorded volume of the media. A value of 100% (the default) plays the media at its recorded volume, a value of 200% will play the media twice recorded volume, 50% at half its recorded volume, a value of 0% will play the media silently, and so on. See 'soundLevel' in SMIL ([W3C.REC-SMIL2-20051213]) for further information. The attribute is optional. The default value is 100%.

clipBegin: offset from start of media resource to begin playback. A valid value is a Time Designation (see Section 4.6.7). The offset is measured in normal media playback time from the beginning of the media resource. If the clipBegin offset is after the end of media (or the clipEnd offset), no media is played. See 'clipBegin' in SMIL ([W3C.REC-SMIL2-20051213]) for further information. The attribute is optional. The default value is 0s.

clipEnd: offset from start of media resource to end playback. A

valid value is a Time Designation (see Section 4.6.7). The offset is measured in normal media playback time from the beginning of the media resource. If the clipEnd offset is after the end of media, then the media is played to the end. If clipBegin is after clipEnd, then no media is played. See 'clipEnd' in SMIL ([W3C.REC-SMIL2-20051213]) for further information. The attribute is optional. There is no default value.

The fetchtimeout, soundLevel, clipBegin and clipEnd attributes are only relevant in the playback use case. The MS ignores these attributes when using the <media> for recording.

The <media> element has no children.

4.3.2. Exit Information

When the dialog exits, information about the specified operations is reported in a <dialogexit> notification event (Section 4.2.5.1).

4.3.2.1. <promptinfo>

The <promptinfo> element reports the information about prompt execution. It has the following attributes:

duration: indicates the duration of prompt playback in milliseconds. A valid value is a non-negative integer (see Section 4.6.4). The attribute is optional. There is no default value.

termmode: indicates how playback was terminated. Valid values are: 'stopped', 'completed' or 'bargein'. The attribute is mandatory.

The <promptinfo> element has no child elements.

4.3.2.2. <controlinfo>

The <controlinfo> element reports information about control execution.

The <controlinfo> element has no attributes and has 0 or more <controlmatch> child elements each describing an individual runtime

control match.

4.3.2.2.1. <controlmatch>

The <controlmatch> element has the following attributes:

dtmf: DTMF input triggering the runtime control. A valid value is a DTMF string (see Section 4.6.3) with no space between characters. The attribute is mandatory.

timestamp: indicates the time (on the MS) at which the control was triggered. A valid value is an dateTime expression (Section 4.6.12). The attribute is mandatory.

The <controlmatch> element has no child elements.

4.3.2.3. <collectinfo>

The <collectinfo> element reports the information about collect execution.

The <collectinfo> element has the following attributes:

dtmf: DTMF input collected from the user. A valid value is a DTMF string (see Section 4.6.3) with no space between characters. The attribute is optional. There is no default value.

termmode: indicates how collection was terminated. Valid values are: 'stopped', 'match', 'noinput' or 'nomatch'. The attribute is mandatory.

The <collectinfo> element has no child elements.

4.3.2.4. <recordinfo>

The <recordinfo> element reports information about record execution (Section 4.3.1.4).

The <recordinfo> element has the following attributes:

termmode: indicates how recording was terminated. Valid values are:

'stopped', 'noinput', 'dtmf', 'maxtime', and 'finalsilence'. The attribute is mandatory.

duration: indicates the duration of the recording in milliseconds. A valid value is a non-negative integer (see Section 4.6.4). The attribute is optional. There is no default value.

The <recordinfo> element has the following child element (0 or more occurrences):

<mediainfo>: indicates information about a recorded media resource (see Section 4.3.2.4.1). The element is optional.

When the record operation is successful, the MS MUST specify a <mediainfo> element for each recording location. For example, if the <record> element contained three <media> child elements, then the <recordinfo> would contain three <mediainfo> child elements.

4.3.2.4.1. <mediainfo>

The <mediainfo> element reports information about a recorded media resource.

The <mediainfo> element has the following attributes:

loc: indicates the location of the media resource. A valid value is a URI (see Section 4.6.9). The attribute is mandatory.

type: indicates the format of the media resource. A valid value is a MIME media type (see Section 4.6.10). The attribute is mandatory.

size: indicates the size of the media resource in bytes. A valid value is a non-negative integer (see Section 4.6.4). The attribute is optional. There is no default value.

4.4. Audit Elements

The audit elements defined in this section allow the MS to be audited

for package capabilities as well as dialogs managed by the package. Auditing is particularly important for two use cases. First, it enables discovery of package capabilities supported on an MS before an AS starts a dialog on connection or conference. The AS can then use this information to create request elements using supported capabilities and, in the case of codecs, to negotiate an appropriate SDP for a user agent's connection. Second, auditing enables discovery of the existence and status of dialogs currently managed by the package on the MS. This could be used when one AS takes over management of the dialogs if the AS which initiated the dialogs fails or is no longer available (see Security Considerations described in Section 7).

4.4.1. <audit>

The <audit> request element is sent to the MS to request information about the capabilities of, and dialogs currently managed with, this control package. Capabilities include supported dialog languages, grammar formats, record and media types as well as codecs. Dialog information includes the status of managed dialogs as well as codecs.

The <audit> element has the following attributes:

capabilities: indicates whether package capabilities are to be audited. A valid value is a boolean (see Section 4.6.1). A value of true indicates that capability information is to be reported. A value of false indicates that capability information is not to be reported. The attribute is optional. The default value is true.

dialogs: indicates whether dialogs currently managed by the package are to be audited. A valid value is a boolean (see Section 4.6.1). A value of true indicates that dialog information is to be reported. A value of false indicates that dialog information is not to be reported. The attribute is optional. The default value is true.

dialogid: string identifying a specific dialog to audit. The MS sends a response with a 406 status code (Section 4.5) if the specified dialog identifier is invalid. The attribute is optional. There is no default value.

If the dialogs attribute has the value true and dialogid attribute is specified, then only audit information about the specified dialog is reported. If the dialogs attribute has the value false, then no dialog audit information is reported even if a dialogid attribute is specified.

The <audit> element has no child elements.

When the MS receives an <audit> request, it MUST reply with a <auditresponse> element (Section 4.4.2) which includes a mandatory attribute describing the status in terms of a numeric code. Response status codes are defined in Section 4.5. If the request is successful, the <auditresponse> contains (depending on attribute values) a <capabilities> element (Section 4.4.2.2) reporting package capabilities and a <dialogs> element (Section 4.4.2.3) reporting managed dialog information. If the MS is not able to process the request and carry out the audit operation, the audit request has failed and the MS MUST indicate the class of failure using an appropriate 4xx response code. Unless an error response code is specified for a class of error within this section, implementations follow Section 4.5 in determining the appropriate status code for the response.

For example, a request to audit capabilities and dialogs managed by the package:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
  <audit/>
</mscivr>
```

In this example, only capabilities are to be audited:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
  <audit dialogs="false"/>
</mscivr>
```

With this example, only a specific dialog is to be audited:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
  <audit capabilities="false" dialogid="d4"/>
</mscivr>
```

4.4.2. <auditresponse>

The <auditresponse> element describes a response to a <audit> request.

The <auditresponse> element has the following attributes:

status: numeric code indicating the audit response status. The attribute is mandatory. Valid values are defined in Section 4.5.

reason: string specifying a reason for the status. The attribute is optional.

The <auditresponse> element has the following sequence of child elements:

<capabilities> element (Section 4.4.2.2) describing capabilities of the package. The element is optional.

<dialogs> element (Section 4.4.2.3) describing information about managed dialogs. The element is optional.

For example, a successful response to a <audit> request requesting capabilities and dialogs information:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <auditresponse status="200">
  <capabilities>
   <dialoglanguages>
    <mimetype>application/voicexml+xml</mimetype>
   </dialoglanguages>
   <grammartypes/>
   <recordtypes>
    <mimetype>audio/x-wav</mimetype>
    <mimetype>video/3gpp</mimetype>
   </recordtypes>
   <prompttypes>
    <mimetype>audio/x-wav</mimetype>
    <mimetype>video/3gpp</mimetype>
   </prompttypes>
   <variables>
    <variabletype type="date" desc="value formatted as YYYYMMDD">
     <format desc="month year day">mdy</format>
     <format desc="year month day">ymd</format>
     <format desc="day month year">dmy</format>
     <format desc="day month">dm</format>
    </variabletype>
   </variables>
   <maxpreparedduration>600s</maxpreparedduration>
   <maxrecordduration>1800s</maxrecordduration>
```

```
      <codecs>
       <codec>
        <subtype>H.263</subtype>
       </codec>
       <codec>
```



```
        <subtype>H.264</subtype>
       </codec>
       <codec>
        <subtype>PCMU</subtype>
       </codec>
       <codec>
        <subtype>PCMA</subtype>
       </codec>
       <codec>
        <subtype>telephone-event</subtype>
       </codec>
      </codecs>
     </capabilities>
     <dialogs>
      <dialogaudit dialogid="4532" state="preparing"/>
      <dialogaudit dialogid="4599" state="prepared"/>
      <dialogaudit dialogid="1234" state="started" conferenceid="conf1">
       <codecs>
        <codec>
         <subtype>PCMA</subtype>
        </codec>
        <codec>
         <subtype>telephone-event</subtype>
        </codec>
       </codecs>
      </dialogaudit>
     </dialogs>
    </auditresponse>
   </mscivr>
```

4.4.2.1. <codecs>

The <codecs> provides audit information about codecs.

The <codecs> element has no attributes.

The <codecs> element has the following sequence of child elements (0 or more occurrences):

<codec>: audit information for a codec (Section 4.4.2.1.1). The element is optional.

For example, a fragment describing two codecs:


```
<codecs>
  <codec>
   <subtype>PCMA</subtype>
  </codec>
  <codec>
   <subtype>telephone-event</subtype>
  </codec>
</codecs>
```

4.4.2.1.1. <codec>

The <codec> element describes a codec on the MS. The element is modeled on the <codec> element in the XCON conference information data model ([I-D.ietf-xcon-common-data-model]) but allows addition information (e.g. rate, speed, etc) to be specified.

The <codec> element has no attributes.

The <codec> element has the following sequence of child elements:

<subtype>: element describing the codec's name. The possible values of this element are the values of the 'subtype' column of the RTP Payload Format media types per [RFC4855] defined in IANA ([IANA]). The element is mandatory.

<params>: element (Section 4.2.6) describing additional information about the codec. This package is agnostic to the names and values of the codec parameters supported by an implementation. The element is optional.

For example, a fragment with a <codec> element describing the H.263 codec:

```
<codec>
 <subtype>H.263</subtype>
</codec>
```

4.4.2.2. <capabilities>

The <capabilities> element provides audit information about package capabilities.

The <capabilities> element has no attributes.

The <capabilities> element has the following sequence of child elements:

<dialoglanguages>: element (Section 4.4.2.2.1) describing additional dialog languages supported by the MS. The element is mandatory.

<grammartypes>: element (Section 4.4.2.2.2) describing supported <grammar> (Section 4.3.1.3.1) format types. The element is mandatory.

<recordtypes>: element (Section 4.4.2.2.3) describing <media> (Section 4.3.1.5) format types supported for <record> (Section 4.3.1.4). The element is mandatory.

<prompttypes>: element (Section 4.4.2.2.4) describing supported <media> (Section 4.3.1.5) format types for playback within a <prompt> (Section 4.3.1.1). The element is mandatory.

<variables>: element (Section 4.4.2.2.5) describing supported types and formats for the <variable> element (Section 4.4.2.2.5). The element is mandatory.

<maxpreparedduration>: element (Section 4.4.2.2.6) describing the supported maximum duration for a prepared dialog following a <dialogprepare> (Section 4.2.1) request. The element is

mandatory.

<maxrecordduration>: element (Section 4.4.2.2.7) describing the supported maximum duration for a recording <record> Section 4.3.1.4) request. The element is mandatory.

<codecs>: element (Section 4.4.2.1) describing codecs available to the package. The element is mandatory.

For example, a fragment describing capabilities:


```
<capabilities>
 <dialoglanguages>
  <mimetype>application/voicexml+xml</mimetype>
 </dialoglanguages>
 <grammartypes/>
 <recordtypes>
  <mimetype>audio/x-wav</mimetype>
  <mimetype>video/3gpp</mimetype>
 </recordtypes>
 <prompttypes>
  <mimetype>audio/x-wav</mimetype>
  <mimetype>video/3gpp</mimetype>
 </prompttypes>
 <variables/>
 <maxpreparedduration>30s</maxpreparedduration>
```

```
      <maxrecordduration>60s</maxrecordduration>
      <codecs>
       <codec>
        <subtype>H.263</subtype>
       </codec>
       <codec>
        <subtype>H.264</subtype>
       </codec>
       <codec>
        <subtype>PCMU</subtype>
       </codec>
       <codec>
        <subtype>PCMA</subtype>
       </codec>
       <codec>
        <subtype>telephone-event</subtype>
       </codec>
      </codecs>
     </capabilities>
```

4.4.2.2.1. <dialoglanguages>

The <dialoglanguages> element provides information about additional dialog languages supported by the package. Dialog languages are identified by their associated MIME media types. The MS MUST NOT include the mandatory dialog language for this package (Section 4.3).

The <dialoglanguages> element has no attributes.

The <dialoglanguages> element has the following sequence of child elements (0 or more occurrences):

<mimetype>: element whose content model describes a MIME media type (Section 4.6.10) associated with a supported dialog language. The element is optional.

4.4.2.2.2. <grammartypes>

The <grammartypes> element provides information about <grammar> format types supported by the package. The MS MUST NOT include the

mandatory SRGS format type, "application/srgs+xml" (Section 4.3.1.3.1).

The <grammartypes> element has no attributes.

The <grammartypes> element has the following sequence of child elements (0 or more occurrences):

<mimetype>: element whose content model describes a mime type (Section 4.6.10). The element is optional.

4.4.2.2.3. <recordtypes>

The <recordtypes> element provides information about media resource format types of <record> supported by the package (Section 4.3.1.4).

The <recordtypes> element has no attributes.

The <recordtypes> element has the following sequence of child elements (0 or more occurrences):

<mimetype>: element whose content model describes a mime type (Section 4.6.10). The element is optional.

4.4.2.2.4. <prompttypes>

The <prompttypes> element provides information about media resource format types of <prompt> supported by the package (Section 4.3.1.1).

The <prompttypes> element has no attributes.

The <prompttypes> element has the following sequence of child elements (0 or more occurrences):

<mimetype>: element whose content model describes a mime type (Section 4.6.10). The element is optional.

4.4.2.2.5. <variables>

The <variables> element provides information about types and formats for the <variable> element (Section 4.4.2.2.5) supported by the package.

The <variables> element has no attributes.

The <variables> element has the following sequence of child elements (0 or more occurrences):

<variabletype>: element describing the formats support for a given type (Section 4.4.2.2.5.1). The element is optional.

For example, a fragment describing support for <variable> with a "date" type in some common formats.

```
<variables>
 <variabletype type="date" desc="value formatted as YYYYMMDD">
   <format desc="month year day">mdy</format>
   <format desc="year month day">ymd</format>
   <format desc="day month year">dmy</format>
   <format desc="day month">dm</format>
 </variabletype>
</variables>
```

4.4.2.2.5.1. <variabletype>

The <variabletype> element describes the formats supported for <variable> supported type.

The <variabletype> element has the following attributes:

type: indicates a supported value associated with the type attribute of <variable> element.The attribute is manadatory.

desc: a string providing some textual description of the type and format. The attribute is optional.

The <variabletype> element has the following sequence of child elements (0 or more occurrences):

<format>: element with a desc attribute (optional description) and a content model describing a supported format in the <variable> format attribute. The element is optional.

4.4.2.2.6. <maxpreparedduration>

The <maxpreparedduration> element describes the maximum duration for a dialog to remain in the prepared state (Section 4.2) following a <dialogprepare> (Section 4.2.1) request.

The <maxpreparedduration> element has no attributes.

The <maxpreparedduration> element has a content model describing the maximum prepared dialog duration as a time designation (Section 4.6.7).

4.4.2.2.7. <maxrecordduration>

The <maxrecordduration> element describes the maximum recording duration for <record> Section 4.3.1.4) request supported by the MS.

The <maxrecordduration> element has no attributes.

The <maxrecordduration> element has a content model describing the maximum duration of recording as a time designation (Section 4.6.7).

4.4.2.3. <dialogs>

The <dialogs> element provides audit information about dialogs.

The <dialogs> element has no attributes.

The <dialogs> element has the following sequence of child elements (0 or more occurrences):

<dialogaudit>: audit information for a dialog (Section 4.4.2.3.1). The element is optional.

4.4.2.3.1. <dialogaudit>

The <dialogaudit> element has the following attributes:

dialogid: string identifying the dialog. The attribute is mandatory.

state: string indicating the state of the dialog. Valid values are: preparing, prepared, starting, started. The attribute is mandatory.

connectionid: string identifying the SIP dialog connection associated with the dialog (see Section 17.1 of [I-D.ietf-mediactrl-sip-control-framework]). The attribute is optional. There is no default value.

conferenceid: string identifying the conference associated with the dialog (see Section 17.1 of [I-D.ietf-mediactrl-sip-control-framework]). The attribute is optional. There is no default value.

The <dialogaudit> element has the following child element:

<codecs> element describing codecs used in the dialog. See Section 4.4.2.1. The element is optional.

For example, a fragment describing a started dialog which is using PCMU and telephony-event codecs:

```
<dialogaudit dialogid="1234" state="started">
 <codecs>
  <codec>
   <subtype>PCMU</subtype>
  </codec>
  <codec>
   <subtype>telephone-event</subtype>
  </codec>
 </codecs>
</dialogaudit>
```

4.5. Response Status Codes

This section describes the response codes in Table 1 for the status attribute of dialog management <response> (Section 4.2.4) and audit <auditresponse> (Section 4.4.2) responses. The MS MUST support these status response codes. The MS MAY support other response codes. The AS MUST treat any responses it does not recognize as being equivalent to the x00 response code for all classes. For example, if an AS receives an unrecognized response code of 499, it can safely assume that there was something wrong with its request and treat the

response as if it had received a 400 (Syntax error) response code.

4xx responses are definite failure responses from a particular MS. The reason attribute in the response SHOULD identify the failure in more detail, for example, "Mandatory attribute missing: src in media element" for a 400 (Syntax error) response code.

The AS SHOULD NOT retry the same request without modification (for example, correcting a syntax error or changing the connectionid to

use one available on the MS). However, the same request to a different MS might be successful; for example, if another MS supports a capability required in the request.

4xx failure responses can be grouped into three classes: failure due to a syntax error in the request (400); failure due to an error executing the request on the MS (405-419); and failure due to the request requiring a capability not supported by the MS (420-439).

In cases where more than one request code could be reported for a failure, the MS SHOULD use the most specific error code of the failure class for the detected error. For example, if the MS detects that the dialogid in the request is invalid, then it uses a 406 status code. However, if the MS merely detects that an execution error occurred, then 419 is used.

Code	Summary	Description	Informational: AS Possible Recovery Action
200	OK	request has succeeded	
400	Syntax error	request is syntactically invalid: it is not valid with respect to the XML schema specified in Section 5 or it violates a co-occurrence	Change the request so that it is syntactically valid.

		constraint for a request element defined in Section 4.		
401	Reserved for future use			
402	Reserved for future use			
403	Reserved for future use			
404	Reserved for future use			

405	dialogid already exists	request uses a dialogid identifier for a new dialog which is already used by another dialog on the MS (see Section 4.2).	Send an <audit> request (Section 4.4.1) requesting the list of dialog identifiers already used by the MS and then use a dialog identifier which is not listed.	
406	dialogid does not exist	request uses a dialogid identifier for an dialog which does not exist on the MS (see Section 4.2).	Send an <audit> request (Section 4.4.1) requesting the list of dialog identifiers already used by the MS and then use one of the listed dialog identifiers.	
407	connectionid	request uses a	Use another method	

	does not exist	connectionid identifier for a connection which does not exist on the MS.	to determine which connections are available on the MS.	
408	conferenceid does not exist	request uses a conferenceid identifier for a conference which does not exist on the MS.	Use another method to determine which conferences are available on the MS.	
409	Resource cannot be retrieved	request use a URI to reference an external resource (e.g. dialog, media or grammar) which cannot be retrieved within the timeout interval	Check that the resource URI is valid, can be reached from the MS, and that the appropriate authentication is used.	

410	Dialog execution canceled	request to prepare or start a dialog which has been terminated by a <dialogterminate/> request (see Section 4.2)		
411	Incompatible stream configuration	request specifies a media stream configuration which is in conflict with itself, or the connection or conference capabilities (see Section 4.2.2)	Change the media stream configuration to match the capabilities of the connection or conference	

412	Media stream not available	request specifies an operation for which a media stream is not available. For example, playing a video media resource on an connection or conference without video streams.	Check the media stream capability of the connection or conference and use an operation which only uses these capabilities
413	Control keys with same value	the request contains a <control> element (Section 4.3.1.2) where some keys have the same value	Use different keys for the different control operations.
414	Reserved for future use		
415	Reserved for future use		
416	Reserved for future use		
417	Reserved for future use		
418	Reserved for future use		

419	Other execution error	requested operation cannot be executed by the MS.	
420	Unsupported URI scheme	request specifies a URI whose scheme is not supported by the MS	Use a URI scheme which is supported.
421	Unsupported dialog	request references an external dialog	Send an <audit> request

	language	language not supported by the MS	(Section 4.4.1) requesting the MS capabilities and then use one of the listed dialog languages.	
422	Unsupported playback format	request references a media resource for playback whose format is not supported by the MS	Send an <audit> request (Section 4.4.1) requesting the MS capabilities and then use one of the listed playback media formats.	
423	Unsupported record format	request references a media resource for recording whose format is not supported by the MS	Send an <audit> request (Section 4.4.1) requesting the MS capabilities and then use one of the listed record media formats.	
424	Unsupported grammar format	request references a grammar whose format is not supported by the MS	Send an <audit> request (Section 4.4.1) requesting the MS capabilities and then use one of the listed grammar types.	

425	Unsupported variable configuration	request contains a prompt <variable> element (Section 4.3.1.1.1)	Send an <audit> request (Section 4.4.1) requesting the MS	

		not supported by the MS	capabilities and then use one of the listed variable types.
426	Unsupported DTMF configuration	request contains a prompt <dtmf> element (Section 4.3.1.1.2) not supported by the MS	
427	Unsupported parameter	request contains a <param> element (Section 4.2.6.1) not supported by the MS	
428	Unsupported media stream configuration	request contains a <stream> element (Section 4.2.2.2) whose configuration is not supported by the MS.	
429	Unsupported playback configuration	request contains a <prompt> element (Section 4.3.1.1) which the MS is unable to play on the available output media streams	
430	Unsupported record configuration	request contains a <record> element (Section 4.3.1.1) which the MS is unable to record with on the available input media streams	
431	Unsupported foreign namespace attribute or element	the request contains attributes or elements from another namespace which the MS does not support	

432	Unsupported multiple dialog capability	the request tries to start another dialog on the same conference or connection where a dialog is already running	
433	Unsupported collect and record capability	the request contains <collect> and <record> elements and the MS does support these operations simultaneously	
434	Unsupported VAD capability	the request contains a <record> element where Voice Activity Detection (VAD) is required, but the MS does not support VAD.	
435	Unsupported parallel playback	the request contains a prompt <par> element whose configuration is not supported by the MS.	
436	Reserved for future use		
437	Reserved for future use		
438	Reserved for future use		
439	Other unsupported capability	request requires another capability not supported by the MS	

Table 1: status codes

4.6. Type Definitions

This section defines types referenced in attribute and element definitions.

4.6.1. Boolean

The value space of boolean is the set {true, false}.

4.6.2. DTMFChar

A DTMF character. The value space is the set {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, #, *, A, B, C, D}.

4.6.3. DTMFString

A String composed of one or more DTMFChars.

4.6.4. Non-Negative Integer

The value space of non-negative integer is the infinite set {0,1,2,...}.

4.6.5. Positive Integer

The value space of positive integer is the infinite set {1,2,...}.

4.6.6. String

A string in the character encoding associated with the XML element.

4.6.7. Time Designation

A time designation consists of a non-negative real number followed by a time unit identifier.

The time unit identifiers are: "ms" (milliseconds) and "s" (seconds).

Examples include: "3s", "850ms", "0.7s", ".5s" and "+1.5s".

4.6.8. Percentage

A percentage consists of a Positive Integer followed by "%".

Examples include: "100%", "500%" and "10%".

4.6.9. URI

Uniform Resource Indicator as defined in [RFC3986].

4.6.10. MIME Media Type

A string formated as a IANA MIME media type ([MIME.mediatypes]).

4.6.11. Language Identifier

A language identifier labels information content as being of a particular human language variant. Following the XML specification for language identification [XML], a legal language identifier is identified by a RFC4646 ([RFC4646]) and RFC4647 ([RFC4647]) code where the language code is required and a country code or other subtag identifier is optional.

4.6.12. DateTime

A string formated according to the XML schema definition of a dateTime type ([XMLSchema:Part2]).

5. Formal Syntax

This section defines the XML schema for IVR Control Package.

The schema defines datatypes, attributes, dialog management and IVR dialog elements in the urn:ietf:params:xml:ns:msc-ivr namespace. In most elements the order of child elements is significant. The schema is extensible: elements allow attributes and child elements from other namespaces. Elements from outside this package's namespace can occur after elements defined in this package.

The schema is dependent upon the schema (framework.xsd) defined in Section 17.1 of the Control Framework [I-D.ietf-mediactrl-sip-control-framework]. It is also dependent upon the W3C (xml.xsd) schema for definitions of XML attributes (e.g. xml:base).

```
<?xml version="1.0" encoding="UTF-8"?>
<xsd:schema targetNamespace="urn:ietf:params:xml:ns:msc-ivr"
 elementFormDefault="qualified" blockDefault="#all"
 xmlns="urn:ietf:params:xml:ns:msc-ivr"
 xmlns:fw="urn:ietf:params:xml:ns:control:framework-attributes"
 xmlns:xsd="http://www.w3.org/2001/XMLSchema">

 <xsd:annotation>
  <xsd:documentation>
```

```
   IETF MediaCtrl IVR 1.0 (20090214)

   This is the schema of the IETF MediaCtrl IVR control
   package.

   The schema namespace is urn:ietf:params:xml:ns:msc-ivr

  </xsd:documentation>
 </xsd:annotation>


 <!--
  ############################################################

  SCHEMA IMPORTS

  ############################################################
 -->

 <xsd:import namespace="http://www.w3.org/XML/1998/namespace"
  schemaLocation="xml.xsd">
  <xsd:annotation>
```



```
   <xsd:documentation>
    This import brings in the XML attributes for
    xml:base, xml:lang, etc

    See http://www.w3.org/2001/xml.xsd for latest version
   </xsd:documentation>
  </xsd:annotation>
 </xsd:import>

 <xsd:import
  namespace="urn:ietf:params:xml:ns:control:framework-attributes"
  schemaLocation="framework.xsd">
  <xsd:annotation>
   <xsd:documentation>
    This import brings in the framework attributes for
    conferenceid and connectionid.
   </xsd:documentation>
  </xsd:annotation>
 </xsd:import>
```

```
<!--
 ######################################################

 Extensible core type

 ######################################################
-->


<xsd:complexType name="Tcore">
 <xsd:annotation>
  <xsd:documentation>
   This type is extended by other (non-mixed) component types to
   allow attributes from other namespaces.
  </xsd:documentation>
 </xsd:annotation>
 <xsd:sequence/>
 <xsd:anyAttribute namespace="##other" processContents="lax" />
</xsd:complexType>


<!--
 ######################################################

 TOP LEVEL ELEMENT: mscivr

 ######################################################
```



```
-->

<xsd:complexType name="mscivrType">
 <xsd:complexContent>
  <xsd:extension base="Tcore">
   <xsd:sequence>
    <xsd:choice>
     <xsd:element ref="dialogprepare" />
     <xsd:element ref="dialogstart" />
     <xsd:element ref="dialogterminate" />
     <xsd:element ref="response" />
     <xsd:element ref="event" />
```

```
      <xsd:element ref="audit" />
      <xsd:element ref="auditresponse" />
      <xsd:any namespace="##other" minOccurs="0"
       maxOccurs="unbounded" processContents="lax" />
     </xsd:choice>
    </xsd:sequence>
    <xsd:attribute name="version" type="version.datatype"
     use="required" />
   </xsd:extension>
  </xsd:complexContent>
 </xsd:complexType>

 <xsd:element name="mscivr" type="mscivrType" />


 <!--
  ######################################################

  DIALOG MANAGEMENT TYPES

  ######################################################
 -->

 <!--  dialogprepare -->

 <xsd:complexType name="dialogprepareType">
  <xsd:complexContent>
   <xsd:extension base="Tcore">
    <xsd:sequence>
     <xsd:element ref="dialog" minOccurs="0"
      maxOccurs="1" />
     <xsd:any namespace="##other" minOccurs="0"
      maxOccurs="unbounded" processContents="lax" />
    </xsd:sequence>
    <xsd:attribute name="src" type="xsd:anyURI" />
    <xsd:attribute name="type" type="mime.datatype"/>
```



```
    <xsd:attribute name="fetchtimeout"
     type="timedesignation.datatype" default="30s" />
    <xsd:attribute name="dialogid"
     type="dialogid.datatype" />
   </xsd:extension>
```

```
  </xsd:complexContent>
 </xsd:complexType>

 <xsd:element name="dialogprepare" type="dialogprepareType" />

 <!--  dialogstart -->

 <xsd:complexType name="dialogstartType">
  <xsd:complexContent>
   <xsd:extension base="Tcore">
    <xsd:sequence>
     <xsd:element ref="dialog" minOccurs="0"
      maxOccurs="1" />
     <xsd:element ref="subscribe" minOccurs="0"
      maxOccurs="1" />
     <xsd:element ref="params" minOccurs="0"
      maxOccurs="1" />
     <xsd:element ref="stream" minOccurs="0"
      maxOccurs="unbounded" />
     <xsd:any namespace="##other" minOccurs="0"
      maxOccurs="unbounded" processContents="lax" />
    </xsd:sequence>
    <xsd:attribute name="src" type="xsd:anyURI" />
    <xsd:attribute name="type" type="mime.datatype"/>
    <xsd:attribute name="fetchtimeout"
     type="timedesignation.datatype" default="30s" />
    <xsd:attribute name="dialogid"
     type="dialogid.datatype" />
    <xsd:attribute name="prepareddialogid"
     type="dialogid.datatype" />
    <xsd:attributeGroup ref="fw:framework-attributes" />
   </xsd:extension>
  </xsd:complexContent>
 </xsd:complexType>

 <xsd:element name="dialogstart" type="dialogstartType" />

 <!--  dialogterminate -->

 <xsd:complexType name="dialogterminateType">
  <xsd:complexContent>
   <xsd:extension base="Tcore">
   <xsd:sequence>
```

```
   <xsd:any namespace="##other" minOccurs="0"
     maxOccurs="unbounded" processContents="lax" />
   </xsd:sequence>
   <xsd:attribute name="dialogid"
    type="dialogid.datatype" use="required" />
   <xsd:attribute name="immediate"
    type="boolean.datatype" default="false" />
  </xsd:extension>
 </xsd:complexContent>
</xsd:complexType>

<xsd:element name="dialogterminate" type="dialogterminateType" />

<!--  response -->

<xsd:complexType name="responseType">
 <xsd:complexContent>
  <xsd:extension base="Tcore">
   <xsd:sequence>
   <xsd:any namespace="##other" minOccurs="0"
     maxOccurs="unbounded" processContents="lax" />
   </xsd:sequence>
   <xsd:attribute name="status" type="status.datatype"
    use="required" />
   <xsd:attribute name="reason" type="xsd:string" />
   <xsd:attribute name="dialogid"
    type="dialogid.datatype" use="required" />
   <xsd:attributeGroup ref="fw:framework-attributes" />
  </xsd:extension>
 </xsd:complexContent>
</xsd:complexType>

<xsd:element name="response" type="responseType" />

<!--  event -->

<xsd:complexType name="eventType">
 <xsd:complexContent>
  <xsd:extension base="Tcore">
   <xsd:sequence>
    <xsd:choice>
     <xsd:element ref="dialogexit" minOccurs="0"
      maxOccurs="1" />
     <xsd:element ref="dtmfnotify" minOccurs="0"
      maxOccurs="1" />
     <xsd:any namespace="##other" minOccurs="0"
      maxOccurs="unbounded" processContents="lax" />
    </xsd:choice>
```

```
    </xsd:sequence>
    <xsd:attribute name="dialogid"
     type="dialogid.datatype" use="required" />
   </xsd:extension>
  </xsd:complexContent>
 </xsd:complexType>

 <xsd:element name="event" type="eventType" />

 <!--  dialogexit-->

 <xsd:complexType name="dialogexitType">
  <xsd:complexContent>
   <xsd:extension base="Tcore">
    <xsd:sequence>
     <xsd:element ref="promptinfo" minOccurs="0"
      maxOccurs="1" />
     <xsd:element ref="controlinfo" minOccurs="0"
      maxOccurs="1" />
     <xsd:element ref="collectinfo" minOccurs="0"
      maxOccurs="1" />
     <xsd:element ref="recordinfo" minOccurs="0"
      maxOccurs="1" />
     <xsd:element ref="params" minOccurs="0"
      maxOccurs="1" />
     <xsd:any namespace="##other" minOccurs="0"
      maxOccurs="unbounded" processContents="lax" />
    </xsd:sequence>
    <xsd:attribute name="status"
     type="xsd:nonNegativeInteger" use="required" />
    <xsd:attribute name="reason" type="xsd:string" />
   </xsd:extension>
  </xsd:complexContent>
 </xsd:complexType>

 <xsd:element name="dialogexit" type="dialogexitType" />


 <!--  dtmfnotify-->

 <xsd:complexType name="dtmfnotifyType">
  <xsd:complexContent>
```

```
   <xsd:extension base="Tcore">
   <xsd:sequence>
    <xsd:any namespace="##other" minOccurs="0"
      maxOccurs="unbounded" processContents="lax" />
    </xsd:sequence>
    <xsd:attribute name="matchmode"
```



```
     type="matchmode.datatype" default="all" />
    <xsd:attribute name="dtmf" type="dtmfstring.datatype"
     use="required" />
    <xsd:attribute name="timestamp" type="xsd:dateTime"
     use="required" />
   </xsd:extension>
  </xsd:complexContent>
 </xsd:complexType>

 <xsd:element name="dtmfnotify" type="dtmfnotifyType" />


 <!-- promptinfo -->

 <xsd:complexType name="promptinfoType">
  <xsd:complexContent>
   <xsd:extension base="Tcore">
 <xsd:sequence>
    <xsd:any namespace="##other" minOccurs="0"
      maxOccurs="unbounded" processContents="lax" />
    </xsd:sequence>
    <xsd:attribute name="duration"
     type="xsd:nonNegativeInteger" />
    <xsd:attribute name="termmode"
     type="prompt_termmode.datatype" use="required" />
   </xsd:extension>
  </xsd:complexContent>
 </xsd:complexType>

 <xsd:element name="promptinfo" type="promptinfoType" />



 <!-- controlinfo -->
```

```
<xsd:complexType name="controlinfoType">
 <xsd:complexContent>
  <xsd:extension base="Tcore">
   <xsd:sequence>
    <xsd:element ref="controlmatch" minOccurs="0"
     maxOccurs="unbounded" />
    <xsd:any namespace="##other" minOccurs="0"
     maxOccurs="unbounded" processContents="lax" />
   </xsd:sequence>
  </xsd:extension>
 </xsd:complexContent>
</xsd:complexType>
```



```
<xsd:element name="controlinfo" type="controlinfoType" />


<!-- controlmatch -->

<xsd:complexType name="controlmatchType">
 <xsd:complexContent>
  <xsd:extension base="Tcore">
<xsd:sequence>
   <xsd:any namespace="##other" minOccurs="0"
     maxOccurs="unbounded" processContents="lax" />
   </xsd:sequence>
   <xsd:attribute name="dtmf"
    type="dtmfstring.datatype" />
   <xsd:attribute name="timestamp" type="xsd:dateTime" />
  </xsd:extension>
 </xsd:complexContent>
</xsd:complexType>

<xsd:element name="controlmatch" type="controlmatchType" />


<!--  collectinfo -->

<xsd:complexType name="collectinfoType">
 <xsd:complexContent>
  <xsd:extension base="Tcore">
<xsd:sequence>
```

```
    <xsd:any namespace="##other" minOccurs="0"
      maxOccurs="unbounded" processContents="lax" />
    </xsd:sequence>
    <xsd:attribute name="dtmf"
     type="dtmfstring.datatype" />
    <xsd:attribute name="termmode"
     type="collect_termmode.datatype" use="required" />
   </xsd:extension>
  </xsd:complexContent>
 </xsd:complexType>

 <xsd:element name="collectinfo" type="collectinfoType" />


 <!-- recordinfo -->

 <xsd:complexType name="recordinfoType">
  <xsd:complexContent>
   <xsd:extension base="Tcore">
   <xsd:sequence>
```



```
     <xsd:element ref="mediainfo" minOccurs="0"
      maxOccurs="unbounded" />
     <xsd:any namespace="##other" minOccurs="0"
      maxOccurs="unbounded" processContents="lax" />
    </xsd:sequence>
    <xsd:attribute name="duration"
     type="xsd:nonNegativeInteger" />
    <xsd:attribute name="termmode"
     type="record_termmode.datatype" use="required" />
   </xsd:extension>
  </xsd:complexContent>
 </xsd:complexType>

 <xsd:element name="recordinfo" type="recordinfoType" />

 <!--  mediainfo -->

 <xsd:complexType name="mediainfoType">
  <xsd:complexContent>
   <xsd:extension base="Tcore">
 <xsd:sequence>
```

```
    <xsd:any namespace="##other" minOccurs="0"
      maxOccurs="unbounded" processContents="lax" />
    </xsd:sequence>
    <xsd:attribute name="loc" type="xsd:anyURI"
     use="required" />
    <xsd:attribute name="type" type="mime.datatype"
    use="required"/>
    <xsd:attribute name="size"
     type="xsd:nonNegativeInteger" />
   </xsd:extension>
  </xsd:complexContent>
 </xsd:complexType>

 <xsd:element name="mediainfo" type="mediainfoType" />


 <!--  subscribe -->

 <xsd:complexType name="subscribeType">
  <xsd:complexContent>
   <xsd:extension base="Tcore">
    <xsd:sequence>
     <xsd:element ref="dtmfsub" minOccurs="0"
      maxOccurs="unbounded" />
     <xsd:any namespace="##other" minOccurs="0"
      maxOccurs="unbounded" processContents="lax" />
    </xsd:sequence>
```



```
   </xsd:extension>
  </xsd:complexContent>
 </xsd:complexType>

 <xsd:element name="subscribe" type="subscribeType" />


 <!--  dtmfsub -->

 <xsd:complexType name="dtmfsubType">
  <xsd:complexContent>
   <xsd:extension base="Tcore">
 <xsd:sequence>
    <xsd:any namespace="##other" minOccurs="0"
```

```
      maxOccurs="unbounded" processContents="lax" />
    </xsd:sequence>
    <xsd:attribute name="matchmode"
     type="matchmode.datatype" default="all" />
   </xsd:extension>
  </xsd:complexContent>
 </xsd:complexType>

 <xsd:element name="dtmfsub" type="dtmfsubType" />


 <!-- params -->
 <xsd:complexType name="paramsType">
  <xsd:complexContent>
   <xsd:extension base="Tcore">
    <xsd:sequence>
     <xsd:element ref="param" minOccurs="0"
      maxOccurs="unbounded" />
     <xsd:any namespace="##other" minOccurs="0"
      maxOccurs="unbounded" processContents="lax" />
    </xsd:sequence>
   </xsd:extension>
  </xsd:complexContent>
 </xsd:complexType>

 <xsd:element name="params" type="paramsType" />


 <!--  param -->
 <!--  doesn't extend tCore since its content model is mixed -->
 <xsd:complexType name="paramType" mixed="true">
  <xsd:sequence>
   <xsd:any namespace="##other" minOccurs="0"
    maxOccurs="unbounded" processContents="lax" />
```



```
  </xsd:sequence>
  <xsd:attribute name="name" type="xsd:string" use="required" />
  <xsd:attribute name="type" type="mime.datatype" default="text/plain"/>
  <xsd:anyAttribute namespace="##other" processContents="lax" />
 </xsd:complexType>

 <xsd:element name="param" type="paramType" />
```

```
 <!--  stream -->

 <xsd:complexType name="streamType">
  <xsd:complexContent>
   <xsd:extension base="Tcore">
   <xsd:sequence>
     <xsd:element ref="region" minOccurs="0"
      maxOccurs="1" />
    <xsd:element ref="priority" minOccurs="0"
      maxOccurs="1" />
     <xsd:any namespace="##other" minOccurs="0"
      maxOccurs="unbounded" processContents="lax" />
    </xsd:sequence>
    <xsd:attribute name="media" type="media.datatype"
     use="required" />
    <xsd:attribute name="label" type="label.datatype" />
    <xsd:attribute name="direction"
     type="direction.datatype" default="sendrecv" />
   </xsd:extension>
  </xsd:complexContent>
 </xsd:complexType>

 <xsd:element name="stream" type="streamType" />

<!--  region  -->
<xsd:simpleType name="regionType">
 <xsd:restriction base="xsd:NMTOKEN"/>
</xsd:simpleType>

 <xsd:element name="region" type="regionType" />


 <!--  priority  -->
 <xsd:simpleType name="priorityType">
  <xsd:restriction base="xsd:positiveInteger" />
 </xsd:simpleType>

 <xsd:element name="priority" type="priorityType" />
```

```
<!-- dialog -->

 <xsd:complexType name="dialogType">
  <xsd:complexContent>
   <xsd:extension base="Tcore">
    <xsd:sequence>
     <xsd:element ref="prompt" minOccurs="0"
      maxOccurs="1" />
     <xsd:element ref="control" minOccurs="0"
      maxOccurs="1" />
     <xsd:element ref="collect" minOccurs="0"
      maxOccurs="1" />
     <xsd:element ref="record" minOccurs="0"
      maxOccurs="1" />
     <xsd:any namespace="##other" minOccurs="0"
      maxOccurs="unbounded" processContents="lax" />
    </xsd:sequence>
    <xsd:attribute name="repeatCount"
     type="xsd:nonNegativeInteger" default="1" />
    <xsd:attribute name="repeatDur"
     type="timedesignation.datatype" />
   </xsd:extension>
  </xsd:complexContent>
 </xsd:complexType>

 <xsd:element name="dialog" type="dialogType" />


 <!--  prompt  -->

 <xsd:complexType name="promptType">
  <xsd:complexContent>
   <xsd:extension base="Tcore">
    <xsd:choice minOccurs="1" maxOccurs="unbounded">
     <xsd:element ref="media" />
     <xsd:element ref="variable" />
     <xsd:element ref="dtmf" />
     <xsd:element ref="par" />
     <xsd:any namespace="##other"
      processContents="lax" />
    </xsd:choice>
    <xsd:attribute ref="xml:base" />
    <xsd:attribute name="bargein" type="boolean.datatype"
     default="true" />
   </xsd:extension>
  </xsd:complexContent>
 </xsd:complexType>
```

```
 <xsd:element name="prompt" type="promptType" />


 <!--  media -->

 <xsd:complexType name="mediaType">
  <xsd:complexContent>
   <xsd:extension base="Tcore">
 <xsd:sequence>
    <xsd:any namespace="##other" minOccurs="0"
      maxOccurs="unbounded" processContents="lax" />
    </xsd:sequence>
    <xsd:attribute name="loc" type="xsd:anyURI"
     use="required" />
    <xsd:attribute name="type" type="mime.datatype" />
    <xsd:attribute name="fetchtimeout"
     type="timedesignation.datatype" default="30s" />
    <xsd:attribute name="soundLevel"
     type="percentage.datatype" default="100%" />
    <xsd:attribute name="clipBegin"
     type="timedesignation.datatype" default="0s" />
     <xsd:attribute name="clipEnd"
     type="timedesignation.datatype"/>
   </xsd:extension>
  </xsd:complexContent>
 </xsd:complexType>

 <xsd:element name="media" type="mediaType" />

 <!--  variable -->

 <xsd:complexType name="variableT">
  <xsd:complexContent>
   <xsd:extension base="Tcore">
 <xsd:sequence>
    <xsd:any namespace="##other" minOccurs="0"
      maxOccurs="unbounded" processContents="lax" />
    </xsd:sequence>
    <xsd:attribute name="value" type="xsd:string"
     use="required" />
    <xsd:attribute name="type" type="xsd:string"
     use="required" />
    <xsd:attribute name="format" type="xsd:string" />
```

```
    <xsd:attribute name="gender" type="gender.datatype" />
    <xsd:attribute ref="xml:lang" />
   </xsd:extension>
  </xsd:complexContent>
 </xsd:complexType>
```



```
 <xsd:element name="variable" type="variableT" />

 <!-- dtmf -->

 <xsd:complexType name="dtmfType">
  <xsd:complexContent>
   <xsd:extension base="Tcore">
 <xsd:sequence>
    <xsd:any namespace="##other" minOccurs="0"
      maxOccurs="unbounded" processContents="lax" />
    </xsd:sequence>
    <xsd:attribute name="digits"
     type="dtmfstring.datatype" use="required" />
    <xsd:attribute name="level" type="xsd:integer"
     default="-6" />
    <xsd:attribute name="duration"
     type="timedesignation.datatype" default="100ms" />
    <xsd:attribute name="interval"
     type="timedesignation.datatype" default="100ms" />
   </xsd:extension>
  </xsd:complexContent>
 </xsd:complexType>

 <xsd:element name="dtmf" type="dtmfType" />


 <!--  par  -->

 <xsd:complexType name="parType">
  <xsd:complexContent>
   <xsd:extension base="Tcore">
    <xsd:choice minOccurs="1" maxOccurs="unbounded">
     <xsd:element ref="media" />
     <xsd:element ref="variable" />
     <xsd:element ref="dtmf" />
     <xsd:element ref="seq" />
```

```
    <xsd:any namespace="##other"
     processContents="lax" />
   </xsd:choice>
  </xsd:extension>
 </xsd:complexContent>
</xsd:complexType>

<xsd:element name="par" type="parType" />

<!--  seq  -->

<xsd:complexType name="seqType">
```



```
 <xsd:complexContent>
  <xsd:extension base="Tcore">
   <xsd:choice minOccurs="1" maxOccurs="unbounded">
    <xsd:element ref="media" />
    <xsd:element ref="variable" />
    <xsd:element ref="dtmf" />
    <xsd:any namespace="##other"
     processContents="lax" />
   </xsd:choice>
  </xsd:extension>
 </xsd:complexContent>
</xsd:complexType>

<xsd:element name="seq" type="seqType" />

<!-- control -->

<xsd:complexType name="controlType">
 <xsd:complexContent>
  <xsd:extension base="Tcore">
<xsd:sequence>
   <xsd:any namespace="##other" minOccurs="0"
    maxOccurs="unbounded" processContents="lax" />
   </xsd:sequence>
   <xsd:attribute name="skipinterval"
    type="timedesignation.datatype" default="6s" />
   <xsd:attribute name="ffkey" type="dtmfchar.datatype" />
   <xsd:attribute name="rwkey" type="dtmfchar.datatype" />
   <xsd:attribute name="pauseinterval"
```

```
     type="timedesignation.datatype" default="10s" />
    <xsd:attribute name="pausekey"
     type="dtmfchar.datatype" />
    <xsd:attribute name="resumekey"
     type="dtmfchar.datatype" />
    <xsd:attribute name="volumeinterval"
     type="percentage.datatype" default="10%" />
    <xsd:attribute name="volupkey"
     type="dtmfchar.datatype" />
    <xsd:attribute name="voldnkey"
     type="dtmfchar.datatype" />
    <xsd:attribute name="speedinterval"
     type="percentage.datatype" default="10%" />
    <xsd:attribute name="speedupkey"
     type="dtmfchar.datatype" />
    <xsd:attribute name="speeddnkey"
     type="dtmfchar.datatype" />
    <xsd:attribute name="gotostartkey"
     type="dtmfchar.datatype" />
```



```
    <xsd:attribute name="gotoendkey"
     type="dtmfchar.datatype" />
    <xsd:attribute name="external"
     type="dtmfstring.datatype" />
   </xsd:extension>
  </xsd:complexContent>
 </xsd:complexType>

 <xsd:element name="control" type="controlType" />


 <!-- collect -->

 <xsd:complexType name="collectType">
  <xsd:complexContent>
   <xsd:extension base="Tcore">
    <xsd:sequence>
     <xsd:element ref="grammar" minOccurs="0"
      maxOccurs="1" />
     <xsd:any namespace="##other" minOccurs="0"
      maxOccurs="unbounded" processContents="lax" />
    </xsd:sequence>
```

```
    <xsd:attribute name="cleardigitbuffer"
     type="boolean.datatype" default="true" />
    <xsd:attribute name="timeout"
     type="timedesignation.datatype" default="5s" />
    <xsd:attribute name="interdigittimeout"
     type="timedesignation.datatype" default="2s" />
    <xsd:attribute name="termtimeout"
     type="timedesignation.datatype" default="0s" />
    <xsd:attribute name="escapekey"
     type="dtmfchar.datatype" />
    <xsd:attribute name="termchar"
     type="dtmfchar.datatype" default="#" />
    <xsd:attribute name="maxdigits"
     type="xsd:positiveInteger" default="5" />
   </xsd:extension>
  </xsd:complexContent>
 </xsd:complexType>

 <xsd:element name="collect" type="collectType" />

 <!--  grammar -->
 <!--  doesn't extend tCore since its content model is mixed -->
 <xsd:complexType name="grammarType" mixed="true">
  <xsd:sequence>
   <xsd:any namespace="##other" minOccurs="0"
    maxOccurs="unbounded" processContents="lax" />
```



```
  </xsd:sequence>
  <xsd:attribute name="src" type="xsd:anyURI" />
  <xsd:attribute name="type" type="mime.datatype" />
  <xsd:attribute name="fetchtimeout"
   type="timedesignation.datatype" default="30s" />
  <xsd:anyAttribute namespace="##other" processContents="lax" />
 </xsd:complexType>

 <xsd:element name="grammar" type="grammarType" />

 <!-- record -->

 <xsd:complexType name="recordType">
  <xsd:complexContent>
   <xsd:extension base="Tcore">
```

```
   <xsd:sequence>
     <xsd:element ref="media" minOccurs="0"
      maxOccurs="unbounded" />
     <xsd:any namespace="##other" minOccurs="0"
      maxOccurs="unbounded" processContents="lax" />
    </xsd:sequence>
    <xsd:attribute name="timeout"
     type="timedesignation.datatype" default="5s" />
    <xsd:attribute name="beep" type="boolean.datatype"
     default="false" />
    <xsd:attribute name="vadinitial"
     type="boolean.datatype" default="false" />
    <xsd:attribute name="vadfinal"
     type="boolean.datatype" default="false" />
    <xsd:attribute name="dtmfterm"
     type="boolean.datatype" default="true" />
    <xsd:attribute name="maxtime"
     type="timedesignation.datatype" default="15s" />
    <xsd:attribute name="finalsilence"
     type="timedesignation.datatype" default="5s" />
    <xsd:attribute name="append" type="boolean.datatype"
     default="false" />
   </xsd:extension>
  </xsd:complexContent>
 </xsd:complexType>

 <xsd:element name="record" type="recordType" />

 <!--
  ######################################################

  AUDIT TYPES
```



```
  ######################################################
 -->


 <!-- audit -->

 <xsd:complexType name="auditType">
  <xsd:complexContent>
```

```
  <xsd:extension base="Tcore">
 <xsd:sequence>
    <xsd:any namespace="##other" minOccurs="0"
      maxOccurs="unbounded" processContents="lax" />
    </xsd:sequence>
   <xsd:attribute name="capabilities"
     type="boolean.datatype" default="true" />
   <xsd:attribute name="dialogs"
     type="boolean.datatype" default="true" />
      <xsd:attribute name="dialogid"
     type="dialogid.datatype"/>
   </xsd:extension>
  </xsd:complexContent>
 </xsd:complexType>

 <xsd:element name="audit" type="auditType" />

 <!-- auditresponse -->

 <xsd:complexType name="auditresponseType">
  <xsd:complexContent>
   <xsd:extension base="Tcore">
    <xsd:sequence>
     <xsd:element ref="capabilities" minOccurs="0"
      maxOccurs="1" />
     <xsd:element ref="dialogs" minOccurs="0"
      maxOccurs="1" />
     <xsd:any namespace="##other" minOccurs="0"
      maxOccurs="unbounded" processContents="lax" />
    </xsd:sequence>
    <xsd:attribute name="status" type="status.datatype"
     use="required" />
    <xsd:attribute name="reason" type="xsd:string" />
   </xsd:extension>
  </xsd:complexContent>
 </xsd:complexType>

 <xsd:element name="auditresponse" type="auditresponseType" />

 <!-- codec -->
```



```
 <xsd:complexType name="codecType">
```

```
 <xsd:complexContent>
  <xsd:extension base="Tcore">
   <xsd:sequence>
    <xsd:element ref="subtype" minOccurs="1"
     maxOccurs="1" />
    <xsd:element ref="params" minOccurs="0"
     maxOccurs="1" />
    <xsd:any namespace="##other" minOccurs="0"
     maxOccurs="unbounded" processContents="lax" />
   </xsd:sequence>
  </xsd:extension>
 </xsd:complexContent>
</xsd:complexType>

<xsd:element name="codec" type="codecType" />

<!-- subtype -->

<xsd:simpleType name="subtypeType">
 <xsd:restriction base="xsd:string" />
</xsd:simpleType>

<xsd:element name="subtype" type="subtypeType" />


<!-- codecs -->

<xsd:complexType name="codecsType">
 <xsd:complexContent>
  <xsd:extension base="Tcore">
   <xsd:sequence>
    <xsd:element ref="codec" minOccurs="0"
     maxOccurs="unbounded" />
    <xsd:any namespace="##other" minOccurs="0"
     maxOccurs="unbounded" processContents="lax" />
   </xsd:sequence>
  </xsd:extension>
 </xsd:complexContent>
</xsd:complexType>

<xsd:element name="codecs" type="codecsType" />


<!-- capabilities -->

<xsd:complexType name="capabilitiesType">
 <xsd:complexContent>
```

```
   <xsd:extension base="Tcore">
    <xsd:sequence>
     <xsd:element ref="dialoglanguages" minOccurs="1"
      maxOccurs="1" />
     <xsd:element ref="grammartypes" minOccurs="1"
      maxOccurs="1" />
     <xsd:element ref="recordtypes" minOccurs="1"
      maxOccurs="1" />
     <xsd:element ref="prompttypes" minOccurs="1"
      maxOccurs="1" />
     <xsd:element ref="variables" minOccurs="1"
      maxOccurs="1" />
      <xsd:element ref="maxpreparedduration" minOccurs="1"
      maxOccurs="1" />
      <xsd:element ref="maxrecordduration" minOccurs="1"
      maxOccurs="1" />
     <xsd:element ref="codecs" minOccurs="1"
      maxOccurs="1" />
     <xsd:any namespace="##other" minOccurs="0"
      maxOccurs="unbounded" processContents="lax" />
    </xsd:sequence>
   </xsd:extension>
  </xsd:complexContent>
 </xsd:complexType>

 <xsd:element name="capabilities" type="capabilitiesType" />


 <!-- mimetype -->

 <xsd:element name="mimetype" type="mime.datatype" />


 <!-- dialoglanguages -->

 <xsd:complexType name="dialoglanguagesType">
  <xsd:complexContent>
   <xsd:extension base="Tcore">
    <xsd:sequence>
     <xsd:element ref="mimetype" minOccurs="0"
      maxOccurs="unbounded" />
     <xsd:any namespace="##other" minOccurs="0"
      maxOccurs="unbounded" processContents="lax" />
    </xsd:sequence>
   </xsd:extension>
  </xsd:complexContent>
```

```
 </xsd:complexType>
```



```
 <xsd:element name="dialoglanguages" type="dialoglanguagesType" />

 <!-- grammartypes -->

 <xsd:complexType name="grammartypesType">
  <xsd:complexContent>
   <xsd:extension base="Tcore">
    <xsd:sequence>
     <xsd:element ref="mimetype" minOccurs="1"
      maxOccurs="unbounded" />
     <xsd:any namespace="##other" minOccurs="0"
      maxOccurs="unbounded" processContents="lax" />
    </xsd:sequence>
   </xsd:extension>
  </xsd:complexContent>
 </xsd:complexType>

 <xsd:element name="grammartypes" type="grammartypesType" />


 <!-- recordtypes -->

 <xsd:complexType name="recordtypesType">
  <xsd:complexContent>
   <xsd:extension base="Tcore">
    <xsd:sequence>
     <xsd:element ref="mimetype" minOccurs="0"
      maxOccurs="unbounded" />
     <xsd:any namespace="##other" minOccurs="0"
      maxOccurs="unbounded" processContents="lax" />
    </xsd:sequence>
   </xsd:extension>
  </xsd:complexContent>
 </xsd:complexType>

 <xsd:element name="recordtypes" type="recordtypesType" />


    <!-- prompttypes -->
```

```
 <xsd:complexType name="prompttypesType">
  <xsd:complexContent>
   <xsd:extension base="Tcore">
    <xsd:sequence>
     <xsd:element ref="mimetype" minOccurs="0"
      maxOccurs="unbounded" />
     <xsd:any namespace="##other" minOccurs="0"
      maxOccurs="unbounded" processContents="lax" />
```



```
    </xsd:sequence>
   </xsd:extension>
  </xsd:complexContent>
 </xsd:complexType>

 <xsd:element name="prompttypes" type="prompttypesType" />

<!--  variables -->

 <xsd:complexType name="variablesType">
  <xsd:complexContent>
   <xsd:extension base="Tcore">
    <xsd:sequence>
     <xsd:element ref="variabletype" minOccurs="0"
      maxOccurs="unbounded" />
     <xsd:any namespace="##other" minOccurs="0"
      maxOccurs="unbounded" processContents="lax" />
    </xsd:sequence>
   </xsd:extension>
  </xsd:complexContent>
 </xsd:complexType>

 <xsd:element name="variables" type="variablesType" />


 <xsd:complexType name="variabletypeType">
  <xsd:complexContent>
   <xsd:extension base="Tcore">
    <xsd:sequence>
     <xsd:element ref="format" minOccurs="0"
      maxOccurs="unbounded" />
     <xsd:any namespace="##other" minOccurs="0"
```

```
      maxOccurs="unbounded" processContents="lax" />
    </xsd:sequence>
    <xsd:attribute name="type" type="xsd:string" use="required" />
    <xsd:attribute name="desc" type="xsd:string"/>
   </xsd:extension>
  </xsd:complexContent>
 </xsd:complexType>

 <xsd:element name="variabletype" type="variabletypeType" />

<!--  format  -->
<!--  doesn't extend tCore since its content model is mixed -->
 <xsd:complexType name="formatType" mixed="true">
  <xsd:sequence>
   <xsd:any namespace="##other" minOccurs="0"
    maxOccurs="unbounded" processContents="lax" />
```



```
  </xsd:sequence>
  <xsd:attribute name="desc" type="xsd:string" />
  <xsd:anyAttribute namespace="##other" processContents="lax" />
 </xsd:complexType>

 <xsd:element name="format" type="formatType" />

<!--  maxpreparedduration -->

 <xsd:element name="maxpreparedduration"
type="timedesignation.datatype"/>


<!--  maxrecordduration -->

 <xsd:element name="maxrecordduration"
type="timedesignation.datatype"/>


 <!-- dialogs -->

 <xsd:complexType name="dialogsType">
  <xsd:complexContent>
   <xsd:extension base="Tcore">
    <xsd:sequence>
```

```
    <xsd:element ref="dialogaudit" minOccurs="0"
     maxOccurs="unbounded" />
    <xsd:any namespace="##other" minOccurs="0"
     maxOccurs="unbounded" processContents="lax" />
   </xsd:sequence>
  </xsd:extension>
 </xsd:complexContent>
</xsd:complexType>

<xsd:element name="dialogs" type="dialogsType" />


<!-- dialogaudit -->

<xsd:complexType name="dialogauditType">
 <xsd:complexContent>
  <xsd:extension base="Tcore">
   <xsd:sequence>
    <xsd:element ref="codecs" minOccurs="0"
     maxOccurs="1" />
    <xsd:any namespace="##other" minOccurs="0"
     maxOccurs="unbounded" processContents="lax" />
   </xsd:sequence>
```



```
    <xsd:attribute name="dialogid"
     type="dialogid.datatype" use="required" />
    <xsd:attribute name="state" type="state.datatype"
     use="required" />
    <xsd:attributeGroup ref="fw:framework-attributes" />
  </xsd:extension>
 </xsd:complexContent>
</xsd:complexType>

<xsd:element name="dialogaudit" type="dialogauditType" />


<!--
 ######################################################

 DATATYPES

 ######################################################
```

```
 -->

 <xsd:simpleType name="version.datatype">
  <xsd:restriction base="xsd:NMTOKEN">
   <xsd:enumeration value="1.0" />
  </xsd:restriction>
 </xsd:simpleType>
 <xsd:simpleType name="mime.datatype">
  <xsd:restriction base="xsd:string" />
 </xsd:simpleType>
 <xsd:simpleType name="dialogid.datatype">
  <xsd:restriction base="xsd:string" />
 </xsd:simpleType>
 <xsd:simpleType name="boolean.datatype">
  <xsd:restriction base="xsd:NMTOKEN">
   <xsd:enumeration value="true" />
   <xsd:enumeration value="false" />
  </xsd:restriction>
 </xsd:simpleType>
 <xsd:simpleType name="gender.datatype">
  <xsd:restriction base="xsd:NMTOKEN">
   <xsd:enumeration value="female" />
   <xsd:enumeration value="male" />
  </xsd:restriction>
 </xsd:simpleType>
 <xsd:simpleType name="state.datatype">
  <xsd:restriction base="xsd:NMTOKEN">
   <xsd:enumeration value="preparing" />
   <xsd:enumeration value="prepared" />
```



```
   <xsd:enumeration value="starting" />
   <xsd:enumeration value="started" />
  </xsd:restriction>
 </xsd:simpleType>
 <xsd:simpleType name="status.datatype">
  <xsd:restriction base="xsd:positiveInteger">
   <xsd:pattern value="[0-9][0-9][0-9]" />
  </xsd:restriction>
 </xsd:simpleType>
 <xsd:simpleType name="media.datatype">
  <xsd:restriction base="xsd:string" />
```

```
 </xsd:simpleType>
 <xsd:simpleType name="label.datatype">
  <xsd:restriction base="xsd:string" />
 </xsd:simpleType>
 <xsd:simpleType name="direction.datatype">
  <xsd:restriction base="xsd:NMTOKEN">
   <xsd:enumeration value="sendrecv" />
   <xsd:enumeration value="sendonly" />
   <xsd:enumeration value="recvonly" />
   <xsd:enumeration value="inactive" />
  </xsd:restriction>
 </xsd:simpleType>
 <xsd:simpleType name="timedesignation.datatype">
  <xsd:annotation>
   <xsd:documentation>
    Time designation following Time in CSS2
   </xsd:documentation>
  </xsd:annotation>
  <xsd:restriction base="xsd:string">
   <xsd:pattern value="(\+)?([0-9]*\.)?[0-9]+(ms|s)" />
  </xsd:restriction>
 </xsd:simpleType>
 <xsd:simpleType name="dtmfchar.datatype">
  <xsd:annotation>
   <xsd:documentation>
    DTMF character [0-9#*A-D]
   </xsd:documentation>
  </xsd:annotation>
  <xsd:restriction base="xsd:string">
   <xsd:pattern value="[0-9#*A-D]" />
  </xsd:restriction>
 </xsd:simpleType>
 <xsd:simpleType name="dtmfstring.datatype">
  <xsd:annotation>
   <xsd:documentation>
    DTMF sequence [0-9#*A-D]
   </xsd:documentation>
```



```
  </xsd:annotation>
  <xsd:restriction base="xsd:string">
   <xsd:pattern value="([0-9#*A-D])+" />
  </xsd:restriction>
```

```
 </xsd:simpleType>
 <xsd:simpleType name="percentage.datatype">
  <xsd:annotation>
   <xsd:documentation>
    whole integer followed by '%'
   </xsd:documentation>
  </xsd:annotation>
  <xsd:restriction base="xsd:string">
   <xsd:pattern value="([0-9])+%" />
  </xsd:restriction>
 </xsd:simpleType>

 <xsd:simpleType name="prompt_termmode.datatype">
  <xsd:restriction base="xsd:NMTOKEN">
   <xsd:enumeration value="completed" />
   <xsd:enumeration value="bargein" />
   <xsd:enumeration value="stopped" />
  </xsd:restriction>
 </xsd:simpleType>

 <xsd:simpleType name="collect_termmode.datatype">
  <xsd:restriction base="xsd:NMTOKEN">
   <xsd:enumeration value="match" />
   <xsd:enumeration value="noinput" />
   <xsd:enumeration value="nomatch" />
   <xsd:enumeration value="stopped" />
  </xsd:restriction>
 </xsd:simpleType>

 <xsd:simpleType name="record_termmode.datatype">
  <xsd:restriction base="xsd:NMTOKEN">
   <xsd:enumeration value="noinput" />
   <xsd:enumeration value="dtmf" />
   <xsd:enumeration value="maxtime" />
   <xsd:enumeration value="finalsilence" />
   <xsd:enumeration value="stopped" />
  </xsd:restriction>
 </xsd:simpleType>

 <xsd:simpleType name="matchmode.datatype">
  <xsd:restriction base="xsd:NMTOKEN">
   <xsd:enumeration value="all" />
   <xsd:enumeration value="collect" />
   <xsd:enumeration value="control" />
```

```
  </xsd:restriction>
 </xsd:simpleType>

</xsd:schema>
```

6. Examples

This section provides examples of the IVR Control package.

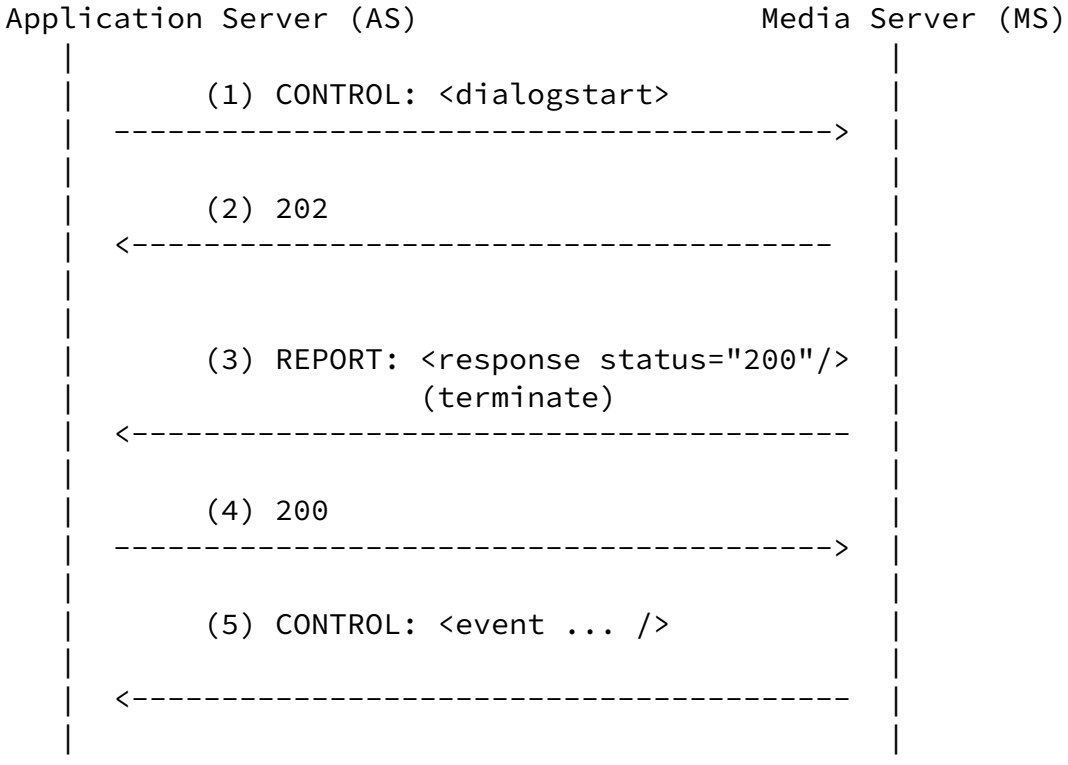
6.1. AS-MS Dialog Interaction Examples

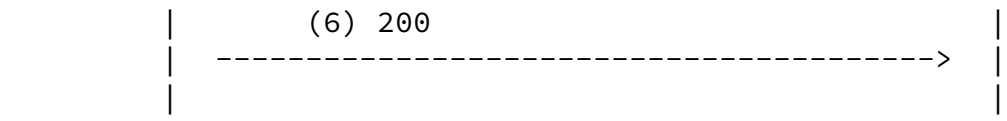
The following example assume a control channel has been established and synced as described in the Media Control Channel Framework ([I-D.ietf-mediactrl-sip-control-framework]).

The XML messages are in angled brackets (with the root <mscivr> omitted); the REPORT status is in round brackets. Other aspects of the protocol are omitted for readability.

6.1.1. Starting an IVR dialog

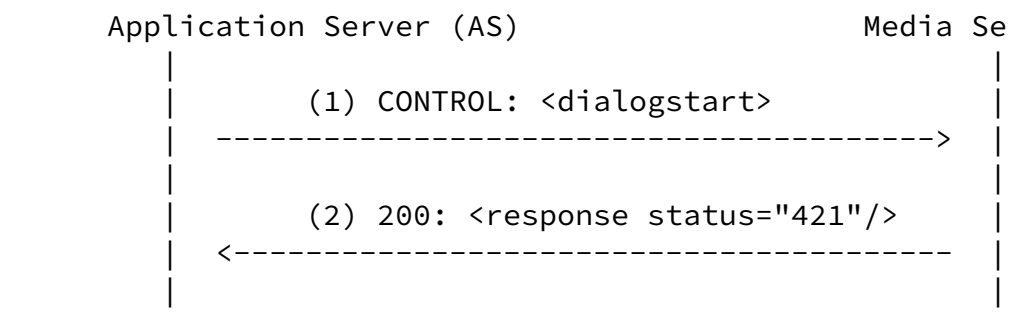
An IVR dialog is started successfully, and dialogexit notification <event> is sent from the MS to the AS when the dialog exits normally.





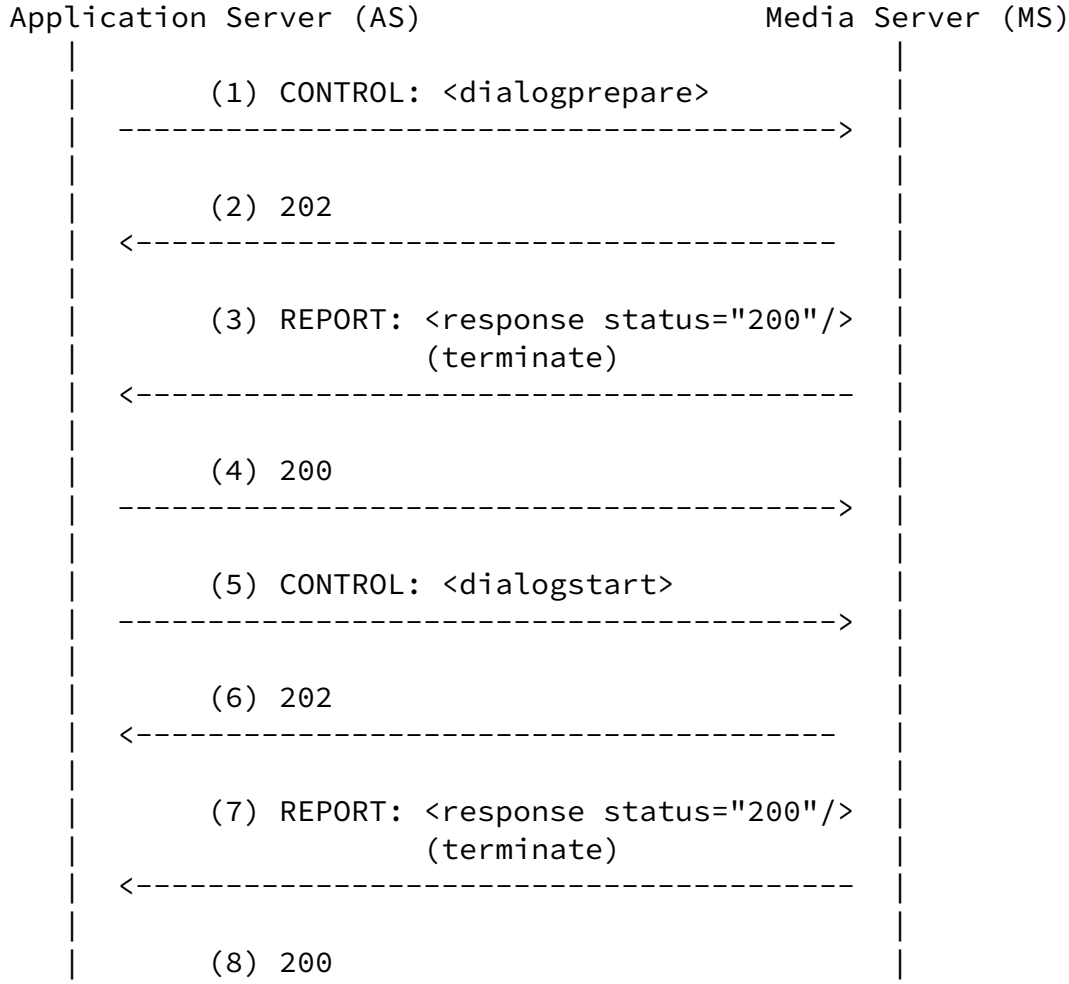
6.1.2. IVR dialog fails to start

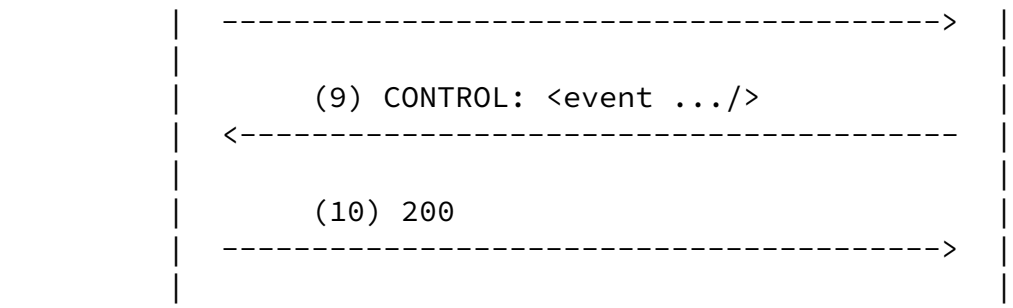
An IVR dialog fails to start due to an unknown dialog language. The <response> is reported in a framework 200 message.



6.1.3. Preparing and starting an IVR dialog

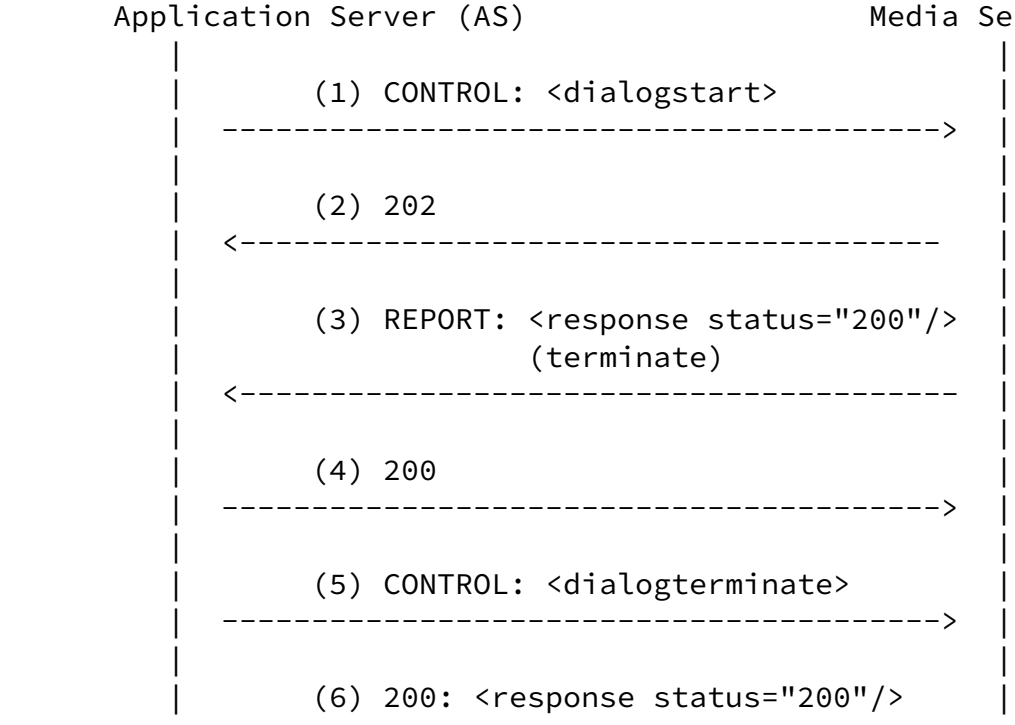
An IVR dialog is prepared and started successfully, and then the dialog exits normally.

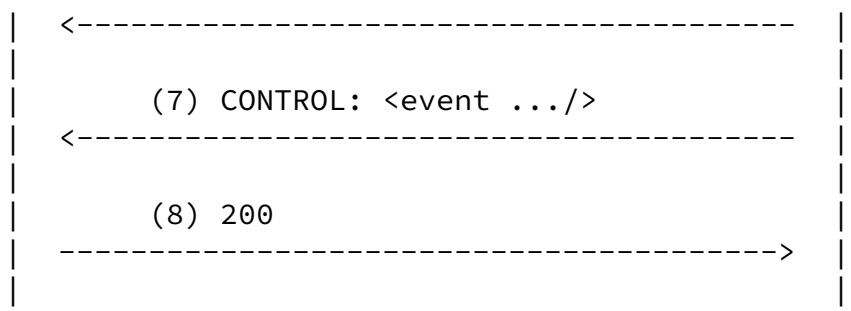

Application Server (AS)
Media Server (MS)
(1) CONTROL: <dialogprepare>
(2) 202
(3) REPORT: <response status="200"/>
(terminate)
(4) 200
(5) CONTROL: <dialogstart>
(6) 202
(7) REPORT: <response status="200"/>
(terminate)
(8) 200



6.1.4. Terminating a dialog

An IVR dialog is started successfully, and then terminated by the AS. The dialogexit event is sent to the AS when the dialog exits.





Note that in (6) the <response> payload to the <dialogterminate/> request is carried on a framework 200 response since it could complete the requested operation before the transaction timeout.

6.2. IVR Dialog Examples

The following examples show how <dialog> is used with <dialogprepare>, <dialogstart> and <event> elements to play prompts, set runtime controls, collect DTMF input and record user input.

The examples do not specify all messages between the AS and MS.

6.2.1. Playing announcements

This example prepares an announcement composed of two prompts where the dialog repeatCount set to 2.


```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
<dialogprepare>
   <dialog repeatCount="2">
    <prompt>
      <media loc="http://www.example.com/media/Number_09.wav"/>
      <media loc="http://www.example.com/media/Number_11.wav"/>
    </prompt>
   </dialog>
</dialogprepare>
</mscivr>
```

If the dialog is prepared successfully, a <response> is returned with status 200 and a dialog identifier assigned by the MS:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
  <response status="200" dialogid="vxi78"/>
</mscivr>
```

The prepared dialog is then started on a conference playing the prompts twice:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <dialogstart prepareddialogid="vxi78" conferenceid="conference11"/>
</mscivr>
```

In the case of a successful dialog, the output is provided in <event>; for example

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <event dialogid="vxi78">
    <dialogexit status="1">
       <promptinfo termmode="completed" duration="24000"/>
    </dialogexit>
 </event>
</mscivr>
```

6.2.2. Prompt and collect

In this example, a prompt is played and then the MS waits for 30s for a two digit sequence:


```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <dialogstart connectionid="7HDY839~HJKSkyHS~HUwkuh7ns">
 <dialog>
  <prompt>
   <media loc="http://www.example.com/prompt1.wav"/>
```

```
  </prompt>
  <collect timeout="30s" maxdigits="2"/>
 </dialog>
 </dialogstart>
</mscivr>
```

If no user input is collected within 30s, then following notification event would be returned:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <event dialogid="vxi81">
    <dialogexit status="1" >
       <promptinfo termmode="completed" duration="4000"/>
       <collectinfo termmode="noinput"/>
    </dialogexit>
 </event>
</mscivr>
```

The collect operation can be specified without a prompt. Here the MS just waits for DTMF input from the user:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <dialogstart connectionid="7HDY839~HJKSkyHS~HUwkuh7ns">
  <dialog>
   <collect/>
  </dialog>
 </dialogstart>
</mscivr>
```

If the dialog is successful, then dialogexit <event> contains the dtmf collected in its result parameter:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <event dialogid="vxi80">
    <dialogexit status="1">
       <collectinfo dtmf="12345" termmode="match"/>
    </dialogexit>
</event>
/mscivr>
```

And finally in this example, one of the input parameters is invalid:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
<dialogstart connectionid="7HDY839~HJKSkyHS~HUwkuh7ns">
 <dialog repeatCount="two">
   <prompt>
     <media loc="http://www.example.com/prompt1.wav"/>
   </prompt>
   <collect cleardigitbuffer="true" bargein="true"
   timeout="4s" interdigittimeout="2s"
   termtimeout="0s" maxdigits="2"/>
 </dialog>
</dialogstart>
</mscivr>
```

The error is reported in the response:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <response status="400" dialogid="vxi82"
  reason="repeatCount attribute value invalid: two"/>
</mscivr>
```

6.2.3. Prompt and record

In this example, the user is prompted, then their input is recorded for a maximum of 30 seconds.

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
<dialogstart connectionid="7HDY839~HJKSkyHS~HUwkuh7ns">
    <dialog>
      <prompt>
       <media loc="http://www.example.com/media/sayname.wav"/>
      </prompt>
      <record dtmfterm="false" maxtime="30s" beep="true"/>
    </dialog>
</dialogstart>
</mscivr>
```

If successful and the recording is terminated by DTMF, the following is returned in a dialogexit <event>:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <event dialogid="vxi83">
  <dialogexit status="1">
   <recordinfo termmode="dtmf">
    <mediainfo type="audio/x-wav"
      loc="http://www.example.com/recording1.wav"/>
   </recordinfo>
  </dialogexit>
 </event>
</mscivr>
```

6.2.4. Runtime controls

In this example, a prompt is played with collect and runtime controls are activated.

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <dialogstart connectionid="7HDY839~HJKSkyHS~HUwkuh7ns">
 <dialog>
  <prompt bargein="true">
   <media loc="http://www.example.com/prompt1.wav"/>
  </prompt>
  <control ffkey=5" rwkey="6" speedupkey="3"
           speeddnkey="4"/>
  <collect maxdigits="2"/>
 </dialog>
 </dialogstart>
</mscivr>
```

Once the dialog is active, the user can press keys 3, 4, 5 and 6 to execute runtime controls on the prompt queue. The keys do not cause bargein to occur. If the user presses any other key, then the prompt is interrupted and DTMF collect begins. Note that runtime controls are not active during the collect operation.

When the dialog is completed successfully, then both control and collect information is reported.

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <event dialogid="vxi81">
    <dialogexit status="1">
       <promptinfo termmode="bargein"/>
       <collectinfo termmode="match" dtmf="14"/>
       <controlinfo>
        <controlmatch dtmf="4" timestamp="2008-05-12T12:13:14Z"/>
        <controlmatch dtmf="3" timestamp="2008-05-12T12:13:15Z"/>
        <controlmatch dtmf="5" timestamp="2008-05-12T12:13:16Z"/>
       <controlinfo>
    </dialogexit>
 </event>
</mscivr>
```

6.2.5. Subscriptions and notifications

In this example, a looped dialog is started with subscription for notifications each time the user input matches the collect grammar:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <dialogstart connectionid="7HDY839~HJKSkyHS">
  <dialog repeatCount="0">
   <collect maxdigits="2"/>
  </dialog>
  <subscribe>
   <dtmfsub matchmode="collect"/>
  </subscribe>
 </dialogstart>
</mscivr>
```

Each time the user input the DTMF matching the grammar, the following notification event would be sent:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <event dialogid="vxi81">
```

```
    <dtmfnotify matchmode="collect" dtmf="12"
     timestamp="2008-05-12T12:13:14Z"/>
  </event>
</mscivr>
```

If no user input was provided, or the input did not match the grammar, the dialog would continue to loop until terminated (or an error occurred).

6.3. Other Dialog Languages

The following example requests that a VoiceXML dialog is started:


```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <dialogstart dialogid="d2"
              connectionid="7HDY839~HJKSkyHS"
              type="application/voicexml+xml"
              src="http://www.example.com/mydialog.vxml"
              fetchtimeout="15s">
  <params>
   <param name="prompt1">nfs://nas01/media1.3gp</param>
   <param name="prompt2">nfs://nas01/media2.3gp</param>
  </params>
 </dialogstart>
</mscivr>
```

If the MS does not support this dialog language, then the response would have the status code 421 (Section 4.5). However, if it does support the VoiceXML dialog language, it would respond with a 200 status, activate the VoiceXML dialog and make the <params> available to the VoiceXML script as described in Section 12.

When the VoiceXML dialog exits, exit namelist parameters are specified using <params> in the dialogexit event:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <event dialogid="d2">
   <dialogexit status="1">
    <params>
     <param name="username">peter</param>
     <param name="pin">1234</param>
```

```
      </params>
     </dialogexit>
    </event>
  </mscivr>
```

6.4. Foreign Namespace Attributes and Elements

An MS can support attributes and elements from foreign namespaces within the <mscivr> element. For example, the MS could support a <listen> element (in a foreign namespace) for speech recognition by analogy to how <collect> support DTMF collection.

In the following example, a prompt and collect request is extended with a <listen> element:


```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr"
 xmlns:ex="http://www.example.com/mediactrl/extensions/1">
 <dialogstart connectionid="7HDY839~HJKSkyHS~HUwkuh7ns">
 <dialog>
  <prompt>
   <media loc="http://www.example.com/prompt1.wav"/>
  </prompt>
  <collect timeout="30s" maxdigits="4"/>
  <ex:listen maxtimeout="30s" >
    <ex:grammar src="http://example.org/pin.grxml"/>
  </ex:listen>
 </dialog>
 </dialogstart>
</mscivr>
```

In the <mscivr> root element, the xmlns:ex attribute declares that "ex" is associated with the foreign namespace URI "http://www.example.com/mediactrl/extensions/1". The <ex:listen>, its attributes and child elements are associated with this namespace. This <listen> could be defined so that it activates an SRGS grammar and listens for user input matching the grammar in a similar manner

to DTMF collection.

If an MS receives this request but does not support the <listen> element, then it would send a 431 response:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <response status="431" dialogid="d560"
  reason="unsupported foreign listen element"/>
</mscivr>
```

If the MS does support this foreign element, it would send a 200 response and start the dialog with speech recognition. When the dialog exits, it provides information about the <listen> execution within <dialogexit>, again using elements in a foreign namespace such as <listeninfo> below:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr"
 xmlns:ex="http://www.example.com/mediactrl/extensions/1">
 <event dialogid="d560">
   <dialogexit status="1">
    <ex:listeninfo speech="1 2 3 4" termmode="match"/>
   </dialogexit>
  </event>
</mscivr>
```

Note that in reply the AS sends a Control Framework 200 response even though the notification event contains an element in a foreign

namespace which it might not understand.

7. Security Considerations

As this control package processes XML markup, implementations MUST address the security considerations of [RFC3023].

Implementations of this control package MUST address security, confidentiality and integrity of messages transported over the

control channel as described in Section 11 of the Media Control channel Framework ([I-D.ietf-mediactrl-sip-control-framework]), including Transport Level Protection, Control Channel Policy Management and Session Establishment. In addition, implementations MUST address security, confidentiality and integrity of User Agent sessions with the MS, both in terms of SIP signaling and associated RTP media flow; see [I-D.ietf-mediactrl-sip-control-framework] for further details on this topic.

Adequate transport protection and authentication are critical, especially when the implementation is deployed in open networks. If the implementation fails to correctly address these issues, it risks exposure to malicious attacks, including (but not limited to):

Denial of Service: An attacker could insert a request message into the transport stream causing specific dialogs on the MS to be terminated immediately. For example, <dialogterminate dialogid="XXXX" immediate="true">, where the value of "XXXX" could be guessed or discovered by auditing active dialogs on the MS using an <audit> request. Likewise, an attacker could impersonate the MS and insert error responses into the transport stream so denying the AS access to package capabilities.

Resource Exhaustion: An attacker could insert into the control channel new request messages (or modify existing ones) with, for instance, <dialogprepare> elements with a very long fetchtimeout attribute and a bogus source URL. At some point this will exhaust the number of connections that the MS is able to make.

Phishing: An attacker with access to the control channel could modify the "loc" attribute of the <media> element in a dialog to point to some other audio file that had different information from the original. This modified file could include a different phone number for people to call if they want more information or need to provide additional information (such as governmental, corporate or financial information).

Data Theft: An attacker could modify a <record> element in the control channel so as to add a new recording location:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <dialogstart>
  <dialog>
   <record>
      <media type="audio/x-wav" loc="(Good URI)"/>
      <media type="audio/x-wav" loc="(Attacker's URI)"/>
   </record>
  </dialog>
 </dialogstart>
</mscivr>
```

The recorded data would be uploaded to two locations indicated by the "{Good URI}" and the "{Attacker's URI}". This allows the attacker to steal the recorded audio (which could include sensitive or confidential information) without the originator of the request necessarily being aware of the theft.

The Media Control Channel Framework permits additional security policy management, including resource access and control channel usage, to be specified at the control package level beyond that specified for the Media Control Channel Framework (see Section 11.3 of [I-D.ietf-mediactrl-sip-control-framework]).

Since creation of IVR dialogs is associated with media processing resources (e.g. DTMF detectors, media playback and recording, etc) on the MS, the security policy for this control package needs to address how such dialogs are securely managed across more than one control channels. Such a security policy is only useful for secure, confidential and integrity protected channels. The identity of control channels is determined by the channel identifier: i.e. the value of the cfw-id attribute in the SDP and Dialog-ID header in the channel protocol (see [I-D.ietf-mediactrl-sip-control-framework]). Channels are the same if they have the same identifier; otherwise, they are different. This control package imposes the following additional security policies:

Responses: The MS MUST only send a response to a dialog management or audit request using the same control channel as the one used to send the request.

Notifications: The MS MUST only send notification events for a dialog using the same control channel as it received the request creating the dialog.

Auditing: The MS MUST only provide audit information about dialogs which have been created on the same control channel as the one upon the <audit> request is sent.

Rejection: The MS SHOULD reject requests to audit or manipulate an existing dialog on the MS if the channel is not the same as the one used when the dialog was created. The MS rejects a request by sending a Control Framework 403 response (see Section 7.4 and Section 11.3 of [I-D.ietf-mediactrl-sip-control-framework]). For example, if a channel with identifier 'cfw1234' has been used to send a request to create a particular dialog and the MS receives on channel 'cfw98969' a request to audit or terminate the dialog, then the MS sends a 403 framework response.

There can be valid reasons why an implementation does not reject an audit or dialog manipulation request on a different channel from the one which created the dialog. For example, a system administrator might require a separate channel to audit dialog resources created by system users and to terminate dialogs consuming excessive system resources. Alternatively, a system monitor or resource broker might require a separate channel to audit dialogs managed by this package on a MS. However, the full implications need to be understood by the implementation and carefully weighted before accepting these reasons as valid. If the reasons are not valid in their particular circumstances, the MS rejects such requests.

There can also be valid reasons for 'channel handover' including high availability support or where one AS needs to take over management of dialogs after the AS which created them has failed. This could be achieved by the control channels using the same channel identifier, one after another. For example, assume a channel is created with the identifier 'cfw1234' and the channel is used to create dialogs on the MS. This channel (and associated SIP dialog) then terminates due to a failure on the AS. As permitted by the Control Framework, the channel identifier 'cfw1234' could then be reused so that another channel is created with the same identifier 'cfw1234', allowing it to 'take over' management of the dialogs on the MS. Again, the implementation needs to understand the full implications and carefully weight them before accepting these reasons as valid. If the reasons are not valid for their particular circumstances, the MS uses the appropriate SIP mechanisms to prevent session establishment when the same channel identifier is used in setting up another control channel (see Section 4 of [I-D.ietf-mediactrl-sip-control-framework]).

8. IANA Considerations

This specification instructs IANA to register a new Media Control Channel Framework Package, a new XML namespace and a new MIME type.

8.1. Control Package Registration

Control Package name: msc-ivr/1.0

8.2. URN Sub-Namespace Registration

XML namespace: urn:ietf:params:xml:ns:msc-ivr

8.3. MIME Type Registration

MIME type: application/msc-ivr+xml

9. Change Summary

Note to RFC Editor: Please remove this whole section.

The following are the changes between the -05 and -04 versions.

- Corrected syntax errors in examples: 4.1, 6.2.2, 6.2.5, 6.3
- 6.3: Corrected error status code for unsupported dialog
- 4.4.2.2.2: <grammartypes>: corrected text to reflect that no additional grammar types need be specified.
- Schema: corrected schema to allow <grammartypes> with no children, and renamed <mediatypes> to <prompttypes> to align with text. Fixed problem with non-deterministic content models.
- 7. Security Considerations: Added requirement that implementations need to secure SIP and RTP sessions with User Agents.

The following are the changes between the -04 and -03 versions.

- 4.2.2.2: Clarified that media stream direction is relative to the dialog (the examples showed this but not the definition).
- 4.3.1.2: Clarified that speed. volume and seek <control> operations beyond the platform's capability are automatically limited to the platform's minimum/maximum. Also clarified that

when the output is paused, then the MS resumes output automatically on speed, volume and seek control operations.

- 6.2.5: Corrected syntax error in example

- 7 Security: Added a denial of service example where the attacker impersonates the MS.

- 7 Security: Clarified that the package security policy for multiple channels is only useful if the channels themselves are secured.

- Added reference for VoiceXML 3.0

- Updated acknowledgements.

- Appendix A: 12.2: Corrected syntax error in example

- Appendix A: 12.2.1: Changed behavior for how VoiceXML session information is populated for a conference; session.conference.name is specified. Added session.connection.originator variable for a connection (MS receives inbound connections but does not create outbound connections).

- Appendix A: 12.2.2: Added session.conference.media for conference media information. Clarified that direction is relative to the dialog.

- Appendix A: 12.2.2: Corrected syntax error in example

- Appendix A: 12.2.3: Changed session parameter mapping to session.values using an associative array.

- Appendix A: 12.5: Changed behavior for handling VoiceXML Call Transfer requests to raising an error.unsupported.transfer event.

The following are the major changes between the -03 and -02 versions.

- Conformance language: Removed unnecessary MUSTs, especially for error codes. Removed lowercase 'should', 'must' and 'may'.

- Introduction: Clarified which MediaCtrl IVR Requirements are satisfied by this package. Added link to Security Considerations Section (also in Section 4.0 and 4.4).

- 4.0: Element definitions. Changed RECOMMENDED to MUST for MS support of communication protocols in URIs.

- 4.2.2:<dialogstart>: Changed RECOMMENDED to MAY for MS support of multiple dialogs on same connection or conference. Changed RECOMMENDED to MUST for using <stream> in cases where connection has multiple streams of the same type.

- 4.2.2.2:<stream>: Changed RECOMMENDED to MUST for use of common media attribute values.

- 4.2.2.1: <subscribe>: Clarified that if the MS does not support a subscription specified in a foreign namespace, then the MS generates a 431 error response.

- 4.2.4: <response>: Clarified that a dialogid with an empty string value is used when the request is syntactically invalid.

- 4.2.5.1: <dialogexit>: Added reserved range of status codes, and tightened up the wording.

- 4.3.1.3:<collect>: Clarified that termtimeout attribute default of 0s meaning no delay.

- 4.3.1.1.1: <variable>: Changed RECOMMENDED to MAY for MS support of date, time and digits <variable>s. Clarified value attribute format for date, time and digits.

- 4.3.1.1.3: <par>: Removed RECOMMENDED for MS support of parallel playback of different media. Added error response code (435) if MS does not support parallel playback configuraton.

- 4.3.1.1.3.1: <seq>: Removed RECOMMENDED for MS support of sequential playback of same media within a <par> (error case already covered by <par> configuration not supported response code).

- 4.3.1.4:<record>: Removed RECOMMENDED for MS support of parallel recording of different media. Clarified wording around uploading recording data to a media resource location.

- 4.3.1.4: <record>: Clarified the definition of vadinitial and vadfinal. Changed the default values to false. Added a response error (434) for when the MS does not support VAD.

- 4.3.1.5: <media>. Removed unnecesaary SHOULD for MS ignoring fetchtimeout, soundLevel, clipBegin and clipEnd when <media> used for recording. Clarified definition of loc and type attributes with stronger conformance language. Similar clarifications of type attributes in <dialogprepare>, <dialogstart> and <grammar>.

- 4.3.2.4.1: <mediainfo>: Clarified usage and strengthen conformance language.

- 4.4.2.22: <grammartypes>: Changed MUST to MUST NOT for inclusion of mandatory SRGS grammar format. Updated examples.

- Updated schema.

- Security Considerations: Major update. Added examples showing malicious attacks when channel security is not correctly addressed. Added more details on multiple channel cases including administrator and monitor channels as well as channel handover.

- Removed affliations in Contributors and Acknowledgements sections.

- Added Appendix A describing how to use VoiceXML with this package if it is supported by the MS.

- Corrected typos and nits.

The following are the major changes between the -02 and -01 versions.

- corrected typos.

- Section 3: Aligned Control Package definitions with requirements in Section 8 of the Control Framework.

- Section 4.2.2.2: Added <priority> child element to <stream> element (alignment with mixer package).

- Following October Interim meeting discussion on response codes, generally clarified usage of error status codes, modified some codes and re-organized the response codes section (Section 4.5) with more guidance and details.

- Section 4.3.1.5: Following October Interim meeting request for parallel playback and record, created a generalized version of <media> used for both playback and record. The 'src' attribute is renamed to 'loc'. Updated <prompt> and <record> definitions as described below.

- Sections 4.3.1.1/4.3.1.1.4: <prompt>: Added <par> child element to allow parallel playback of separate media resources. The <par> element has a <seq> child element to allow a sequence of media resources to be played at the same time as other resources are played in parallel.

- Sections 4.3.1.4/4.3.1.4.1: <record>: Removed 'dest' and 'type' attributes. Added <media> child elements to support parallel recording to separate media resource locations.

- Sections 4.3.2.4/4.3.2.4.1: <recordinfo>: Removed 'dest', 'type' and 'size' attribute. Added <mediainfo> child elements with 'loc', 'type' and 'size' attributes.

- Section 4.4.2.2.4: Renamed <mediatypes> to <prompttypes> to clarify distinction with <recordtypes>.

- Sections 4.3.1.4: <record>: Clarified RFC2119 language around vadinitial and vadfinal behavior.

- Updated schema. Removed some element-specific syntactic constraint statements which are already covered in the schema.

- 4.3.1: occurrence of <control> without a <prompt> no longer treated as a syntax error - instead runtime controls are simply

ignored.

o

The following are the major changes between the -01 and -00 versions.

- 7: Updated security section referencing control framework security and adding policy requirement to address dialog resource management over multiple channels.

- corrected typos and example errors

- 4.2: [IVR-200] Added state machine for dialog lifecycle.

- 4.2: clarified dialog identifier assignment and use, including MS assignment of dialogid in <dialogprepare> and <dialogstart>.

- 4.2/4.2.3: clarified <dialogterminate> behavior when dialog is not in a STARTED state.

- 1/4.2: Clarified concept of dialog language and replaced references to 'dialog types' with dialog languages. Replaced 'dialogtypes' with 'dialoglanguages' in auditing. Clarified that IVR <dialog> is inline and other supported dialog languages are specified by reference. Removed default type values for <dialogprepare> and <dialogstart>.

- 4.4.2.2.1: clarified that the inline dialog language (<dialog>) must not be listed as an additional supported dialog language.

- 4.2.2.2: [IVR-201] Added <region> element to <stream> so that dialog video output can be directed to a specific region a conference video layout.

- 4.3.1.1.2:[IVR-202]: removed ndn format and clarified gen format for digits.

- 4.4.2.1.1:[IVR-203]: added <params> to <codec> to allow additional codec information to be specified.

- 4.5: added error status code for unsupported URI (415), invalid region identifier (416), fetchtimeout exceeded (417), syntactic constraint violation (418), unsupported media format (419), unsupported grammar format (420), unsupported variable announcement (421), unsupported DTMF tone generation (422), conflict with control key values (423), unsupported recording format (424).

- Generally, replaced 'it is an error ...' language with RFC2119 language, making error codes more explicit.

- 4.3.1: Clarified that an MS MAY support <record> and <collect> elements co-occurring in a <dialog> element, but the MS MUST send an error response if they are not supported. Clarified that MS MUST send an error response if <control> is specified without a <prompt> element.

- 4.2.5.2: clarified that the timestamp in <dtmfnotify> is that of the last DTMF in a matching input sequence.

- References: more references now normative.

- 4.3: Replaced passive voice language with active voice language in the description of execution models.

- 1/2: Clarified that the term 'dialog' refers to IVR dialog and is completely unrelated to the term 'SIP dialog'.

- 4: Added clarification that elements with URI attributes are recommended to support one or more communication protocols suitable for fetching resources.

- 4.3.1.4: <record> clarified MS MAY support upload of recording data during recording, and that upload errors (e.g. authentication failures, communication errors, etc) are execution errors. Added 'append' attribute to control behavior when a recorded resource already exists at the recording location.

- 4.2.2: Clarified that an error is reported if <dialogstart> with <params> contains parameters which the MS cannot process for the given dialog language.

- 4.2.6.1: <param> removed 'valuetype' attribute and clarified that the type attribute indicates the MIME media type associated with the inline value.

- 4.3.1.4: Added append attribute to <record> to control whether recordings are appended or not to the recording location resource.

- 4.3.1.1.1: Added clipEnd attribute to <media> to control when playback of the media ends.

- 4.3.1: Clarified that when there are multiple iterations of a dialog (using repeatCount attribute) only the results of the last

dialog iteration are reported.

- 4.4.2.2: Added ability to audit <variable> capability, as well as maximum duration of prepared dialogs and of recordings.
- 4.3.1/4.3.1.1: Clarified the sequence of dialog operations in <dialog> and how <prompt> reports its status.
- 4: Changed handling of unsupported foreign namespace elements and attributes. The MS send a 426 error response if it encounters foreign elements and attributes it does not support.

The following are the major changes between the -00 of this work group item draft and the individual submission -05 version.

- [IVR01] When the MS sends a notification event in a CONTROL, the AS sends mandatory 200 response (no extended transaction).
- [IVR23] Added a top-level container element, <mscivr>, with version attribute.
- Removed term 'basic' in title, description, elements and IANA registration. Control package name is now 'msc-ivr/1.0'. Namespace is now 'urn:ietf:params:xml:ns:msc-ivr'. Mime type is now 'application/msc-ivr+xml'. Renamed 'basicivr' element to 'dialog' and moved version attribute to mscivr element.
- [IVR15] Updated and simplified XML schema. Ordering of child elements is significant.
- [IVR06] Removed 'volume' and 'offset' from prompt element. Added 'soundLevel' and 'clipBegin' to media element.
- [IVR17]/[IVR06] Removed 'iterations' and 'duration' from prompt. Added 'repeatCount' and 'repeatDur' to dialog element.
- Moved VCR commands from <collect> into separate <control> element. Defined controlinfo element to report runtime control match information.
- [IVR05] Added <subscribe> to <dialogstart> where AS can subscribe

to DTMF key presses (all, control match only, collect match only). Extended <event> to support associated notification.

- Moved definition of <stream> into a separate section.

- [IVR21] Added audit capability: auditing of package capabilities and managed dialogs

- [IVR21] Explicitly stated that an error must be reported if the connection or conference referenced in a <dialogstart> is not available at the time the request is processed on the MS.

- Clarified that the <variable> rendering mechanism is MS implementation specific.

- [IVR09]/[IVR10] Clarified <variable> attribute definitions and added 'gender' attribute.

- [IVR16] Clarified that info must be reported in dialogexit, if the corresponding element is specified in a <dialog>. For example, if <prompt> is specified, then <promptinfo> must be specified if the dialog terminates normally.

- [IVR18] Added 'inactive' value for direction attribute of <stream>.

- [IVR19] Clarified case of <dialogstart> on connection/conference with multiple streams of the same type: recommended to be set explicitly with <stream>s.

- [IVR02] Clarified that multiple dialogs may started simultaneously on the same connection or conference.

- [IVR20] Added maximum duration (10 minutes) for a dialog to remain in the PREPARED state.

- Added <params> in <dialogstart> and <dialogexit> for input/output in other dialog types

- [IVR22] Added fetchtimeout parameter to dialogprepare,

dialogstart, media and grammar elements.

- [IVR04] Added dialogexit status to indicate the connection or conference has been terminated. Added others status errors.

- [IVR08] Clarified that the <control> operation does not interrupt playing prompts and that matched DTMF is not available to <collect> or <record> operations during prompt playback.

- [IVR11] Added runtime controls for speed, goto start/end and external controls.

- Clarified that recordings can be uploaded to dest during or after the recording operation.

- <record>/<collect>: clarified timer handling - timeout refers to waiting time for collect/record to begin.

- Clarified behavior of immediate attribute on <dialogterminate>.

- clarified dialogid lifecycle: dialogids can be re-cycled.

- Clarified error handling.

- Editorial tidy up of sections.

- dialogid attribute on <response> is now mandatory.

- Clarified that the duration specified in finalsilence attribute of <record> is not part of the final recording.

- Clarified that the SRGS XML grammar format is mandatory

The following are the major changes between the -06 of the draft and the -05 version.

- Event notifications are sent in CONTROL messages with the MS acting as Control Framework Client. Compared with the previous approach, this means that a <dialogstart> transaction is now complete when the MS sends a <response>. A new transaction is

initiated by the MS each time the MS sends a notification <event> to the AS.

- Changed conf-id to conferenceid and connection-id to connectionid.

- Clarification of the state model for dialogs

- <dialogprepare>: modified definition of src attribute to allow reference to external dialog documents; added (MIME) type attribute; removed <data> child element.

- <dialogstart>: modified definition of src attribute to allow reference to external dialog documents; added (MIME) type attribute; removed <data> child element;

- <dialogterminate>: modified so that a dialogexit event is always sent for active dialogs (i.e. the dialogexit event is a terminating notification)

- <event> notification simplified and make more extensible. Manual notifications (via <subscribe> element) are removed from the basic package. A <dialogexit> event is defined as <event> child and it can be extended with additional child elements

- <data> element is removed.

- <subscribe> element removed.

- Replaced dialog templates with a general <dialog> element. It has child elements for playing media resource (<prompt>), collecting DTMF (<collect>) and recording (<record>). The functionality is largely unchanged.

- <dialogprepare> and <dialogstart> are extended with <dialog> child element.

- <event> is extended with a <dialogexit> element which contains status and reported information (replacement for output parameters in template dialogs)

- Prompts: now structured as a <prompt> element with <media>, <variable> and <dtmf> children. The <prompt> element has xml:base

attribute, bargein, iterations, duration, volume and offset attributes. The speed attribute is removed. A <media> element has src and type attributes. The maxage and maxstale attributes are removed.

- DTMF input: parameters now specified as attributes of a <collect> element. Custom grammar specified with a <grammar> element as child of <collect> element. Added 'escapekey' to allow the dialog to be retried. Added 'pauseinterval', 'pausekey' and 'resumekey' to allow the prompts to paused/resumed. Added 'volumeinterval', 'volupkey' and voldnkey' to add prompt volume to be increased/ decreased. Moved 'bargein' attribute to prompt.

- Recording: parameters now specified as attributes of <record> element. Added 'dest' and 'beep' attributes.

The following are the major changes between the -05 of the draft and the -04 version.

- Mainly an alignment/evaluation exercise with requirements produced by MEDIACTRL IVR design team.

- playannouncement parameters from Table 7 of '04' version are now reflected in text - schema to be updated.

- Added VCR commands based on MSCML.

The following are the major changes between the -04 of the draft and the -03 version.

- None.

The following are the major changes between the -03 of the draft and the -02 version.

- added "basicivr:" protocol to template dialog types which must be supported as values of the "src" attribute in <dialogprepare> and <dialogstart>. Note alternative: "/basicivr/playannouncement" offered in [RFC4240].

- added "basicivr:" URI schema to IANA considerations

- Added mimetype, vadinitial and vadfinal parameters to 'promptandrecord' dialog type
- updated references

The following are the major changes between the -02 of the draft and the -01 version.

- added version 1.0 to package name
- separate section for element definitions
- dialogterminate treated as request rather than notification
- simplified responses: single element <response>
- removed response elements: <dialogprepared>, <dialogstarted>, <errordialognotprepared>, <errordialognotstarted>
- simplified event notifications to single <event> element carried in a REPORT
- <dialogexit> element replaced with <event name="dialogexit">
- removed <dialoguser> element
- added <stream> element as child of <dialogstart>
- removed 'type' attribute from <dialogprepare> and <dialogstart>
- added dialogid attribute to <dialogprepare> and <dialogstart>
- removed template "Sample Implementation" section
- renamed <namelist> to <data>

- re-organized so that template details after general package framework and element description.

The following are the primary changes between the -01 of the draft

and the -00 version.

- Removed requirement for VoiceXML dialog support
- Added requirement for template dialog support

10. Contributors

Asher Shiratzky provided valuable support and contributions to the early versions of this document.

The authors would like to thank the IVR design team consisting of Roni Even, Lorenzo Miniero, Adnan Saleem, Diego Besprosvan, Mary Barnes and Steve Buko who provided valuable feedback, input and text to this document.

11. Acknowledgments

The authors would like to thank Adnan Saleem, Gene Shtirmer, Dave Burke, Dan York and Steve Buko for expert reviews of this work.

Ben Campbell carried out the RAI expert review on this specification and provided a great deal of invaluable input. Donald Eastlake carried out a thorough security review.

12. Appendix A: Using VoiceXML as a dialog language

The IVR control package allows, but does not require, the MS to support other dialog languages by referencing an external dialog document. This appendix provides MS implementations which support the VoiceXML dialog language ([VXML20], [VXML21], [VXML30]) with additional details about using these dialogs in this package.

This appendix covers preparing (Section 12.1), starting (Section 12.2), terminating (Section 12.3) and exiting (Section 12.4) VoiceXML dialogs as well as handling VoiceXML call transfer (Section 12.5).

12.1. Preparing a VoiceXML dialog

A VoiceXML dialog is prepared by sending the MS a request containing a <dialogprepare> element (Section 4.2.1). The type attribute is set to "application/voicexml+xml" and the src attribute to the URI of the VoiceXML document which is to be prepared by the MS. For example:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <dialogprepare type="application/voicexml+xml"
                src="http://www.example.com/mydialog.vxml"
                fetchtimeout="15s"/>
</mscivr>
```

The VoiceXML dialog environment uses the <dialogprepare> request as an opportunity to fetch and validate the initial document indicated by the src attribute along with any resources referenced in the VoiceXML document marked as prefetchable. Note that the fetchtimeout attribute is not defined in VoiceXML for an initial document but the MS MUST support this attribute in its VoiceXML environment.

The success or failure of the VoiceXML document preparation is reported in the MS response. For example, if the VoiceXML document cannot be retrieved, then a 407 error response is returned. If the document is syntactically invalid according to VoiceXML, then a 400 response is returned. If successful, the response includes a dialogid attribute whose value the AS can use in <dialogstart> element to start the prepared dialog.

12.2. Starting a VoiceXML dialog

A VoiceXML dialog is started by sending the MS a request containing a <dialogstart> element (Section 4.2.2). If a VoiceXML dialog has already been prepared using <dialogprepare>, then the MS starts the dialog indicated by the prepareddialogid attribute. Otherwise, a new VoiceXML dialog can be started by setting the type attribute to

"application/voicexml+xml" and the src attribute to the URI of the VoiceXML document. For example:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <dialogstart connectionid="ssd3r3~sds345b"
              type="application/voicexml+xml"
              src="http://www.example.com/mydialog.vxml"
              fetchtimeout="15s"/>
</mscivr>
```

Note that the fetchtimeout attribute is not defined in VoiceXML for an initial document but the MS MUST support this attribute in its VoiceXML environment. Note also that support for <dtmfsub> subscriptions (Section 4.2.2.1.1) and their associated dialog notification events is not defined in VoiceXML. If such a subscription is specified in a <dialogstart> request, then the MS sends a 439 error response (see Section 4.5).

The success or failure of starting a VoiceXML dialog is reported in the MS response as described in Section 4.2.2.

When the MS starts a VoiceXML dialog, the MS MUST map session information into VoiceXML session variable object. There are 3 types of session information: protocol information (Section 12.2.1), media stream information (Section 12.2.2) and parameter information (Section 12.2.3).

12.2.1. Session protocol information

If the connectionid attribute is specified, the MS assigns protocol information from the SIP dialog associated with the connection to the following session variables in VoiceXML:

session.connection.local.uri Evaluates to the SIP URI specified in the To: header of the initial INVITE

session.connection.remote.uri Evaluates to the SIP URI specified in the From: header of the initial INVITE

session.connection.originator Evaluates to the value of session.connection.remote (MS receives inbound connections but does not create outbound connections)

session.connection.protocol.name Evaluates to "sip". Note that this is intended to reflect the use of SIP in general, and does not distinguish between whether the connection accesses the MS via SIP or SIPS procedures.

session.connection.protocol.version Evaluates to "2.0".

session.connection.redirect This array is populated by information contained in the History-Info ([RFC4244]) header in the initial INVITE or is otherwise undefined. Each entry (hi-entry) in the History-Info header is mapped, in reverse order, into an element of the session.connection.redirect array. Properties of each element of the array are determined as follows:

uri Set to the hi-targeted-to-uri value of the History-Info entry

pi Set to 'true' if hi-targeted-to-uri contains a 'Privacy=history' parameter, or if the INVITE Privacy header includes 'history'; 'false' otherwise

si Set to the value of the 'si' parameter if it exists, undefined otherwise

reason Set verbatim to the value of the 'Reason' parameter of hi-

targeted-to-uri

session.connection.aai Evaluates to the value of a SIP header with the name "aai" if present; otherwise undefined.

session.connection.protocol.sip.requesturi This is an associative array where the array keys and values are formed from the URI parameters on the SIP Request-URI of the initial INVITE. The array key is the URI parameter name. The corresponding array value is obtained by evaluating the URI parameter value as a string. In addition, the array's toString() function returns the full SIP Request-URI.

session.connection.protocol.sip.headers This is an associative array where each key in the array is the non-compact name of a SIP header in the initial INVITE converted to lower-case (note the case conversion does not apply to the header value). If multiple header fields of the same field name are present, the values are combined into a single comma-separated value. Implementations MUST at a minimum include the Call-ID header and MAY include other headers. For example, session.connection.protocol.sip.headers["call-id"] evaluates to the Call-ID of the SIP dialog.

If a conferenceid attribute is specified, then the MS populates the following session variables in VoiceXML:

session.conference.name Evaluates to the value of the conferenceid attribute

12.2.2. Session media stream information

The media streams of the connection or conference to use for the dialog are described in Section 4.2.2, including use of <stream> elements (Section 4.2.2.2) if specified. The MS maps media stream information into the VoiceXML session variable session.connection.protocol.sip.media for a connection, and session.conference.media for a conference. In both variables, the value of the variable is an array where each array element is an

object with the following properties:

type This required property indicates the type of the media associated with the stream (see Section 4.2.2.2 <stream> type attribute definition)

direction This required property indicates the directionality of the media relative to the dialog (see Section 4.2.2.2 <stream> direction attribute definition).

format This property is optional. If defined, the value of the property is an array. Each array element is an object which specifies information about one format of the media stream. The object contains at least one property called name whose value is the subtype of the media format ([RFC4855]). Other properties may be defined with string values; these correspond to required and, if defined, optional parameters of the format.

As a consequence of this definition, when a connectionid is specified there is an array entry in session.connection.protocol.sip.media for each media stream used by the VoiceXML dialog. For an example, consider a connection with bi-directional G.711 mu-law audio sampled at 8kHz where the dialog is started with

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <dialogstart connectionid="ssd3r3~sds345b"
              type="application/voicexml+xml"
              src="http://www.example.com/mydialog.vxml"
              fetchtimeout="15s">
  <stream type="audio" direction="sendonly"/>
 </dialogstart>
</mscivr>
```

In this case, session.connection.protocol.sip.media[0].type evaluates to "audio", session.connection.protocol.sip.media[0].direction to "sendonly" (i.e. this dialog only sends media to the connection - it

does not receive media from the connection), and session.connection.protocol.sip.media[0].format[0].name evaluates to "audio/PCMU" and session.connection.protocol.sip.media[0].format[0].rate evaluates to "8000".

Note that the session variable is updated if the connection or conference media session characteristics for the VoiceXML dialog change (e.g. due to a SIP re-INVITE).

12.2.3. Session parameter information

Parameter information is specified in the <params> child element of <dialogstart>, where each parameter is specified using a <param> element. The MS maps parameter information into VoiceXML session variables as follows:

session.values This is an associative array mapped to the <params> element. It is undefined if a <params> element is not specified. Array keys and values are formed from <param> children of the <params> element. Each array key is the value of the name attribute of a <param> element. If the same name is used in more than one <param> element, then the array key is associated with the last <param> in document order. The corresponding value for each key is an object with two required properties: a "type" property evaluating to the value of the type attribute; and a "content" property evaluating to the content of the <param>. In addition, this object's toString() function returns the value of the "content" property as a string.

For example, a VoiceXML dialog started with one parameter:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
 <dialogstart connectionid="ssd3r3~sds345b"
              type="application/voicexml+xml"
              src="http://www.example.com/mydialog.vxml"
              fetchtimeout="15s">
  <params>
   <param name="mode">playannouncement</param>
  </params>
 </dialogstart>
</mscivr>
```

In this case, session.values would be defined with one item in the array where session.values['mode'].type evaluates to "text/plain" (the default value), session.values['mode'].content evaluates to "playannouncement" and session.values['mode'].toString() also evaluates to "playannouncement".

The MS sends an error response (see Section 4.2.2) if a <param> is not supported by the MS (e.g. the parameter type is not supported).

12.3. Terminating a VoiceXML dialog

When the MS receives a request with a <dialogterminate> element (Section 4.2.3), the MS throws a 'connection.disconnect.hangup' event into the specified VoiceXML dialog. Note that if the immediate attribute has the value true, then the MS MUST NOT return <params> information when the VoiceXML dialog exits (even if the VoiceXML dialog provides such information) - see Section 12.4.

If the connection or conference associated with the VoiceXML dialog terminates, then the MS throws a 'connection.disconnect.hangup' event into the specified VoiceXML dialog.

12.4. Exiting a VoiceXML dialog

The MS sends a <dialogexit> notification event (Section 4.2.5.1) when the VoiceXML dialog is complete, has been terminated or because it exits due to an error. The <dialogexit> status attribute specifies the status of the VoiceXML dialog when it exits and its <params> child element specifies information, if any, returned from the VoiceXML dialog.

A VoiceXML dialog exits when it processes a <disconnect> element, a <exit> element or an implicit exit according to the VoiceXML FIA. If the VoiceXML dialog executes a <disconnect> and then subsequently executes an <exit> with namelist information, the namelist information from the <exit> element is discarded.

The MS reports namelist variables in the <params> element of the <dialogexit>. Each <param> reports on a namelist variable. The MS set the <param> name attribute to the name of the VoiceXML variable. The MS sets the <param> type attribute according to the type of the VoiceXML variable. The MS sets the <param> type to 'text/plain' when the VoiceXML variable is a simple ECMAScript value. If the VoiceXML variable is a recording, the MS sets the <param> type to the MIME media type of the recording and encodes the recorded content as CDATA in the <param> (see Section 4.2.6.1 for an example). If the VoiceXML variable is a complex ECMAScript value (e.g. object, array, etc), the MS sets the <param> type to 'application/json' and converts the variable value to its JSON value equivalent ([RFC4627]. The behavior resulting from specifying an ECMAScript object with circular references is not defined.

If the expr attribute is specified on the VoiceXML <exit> element instead of the namelist attribute, the MS creates a <param> element

with the reserved name '__exit', the type 'text/plain' and the content of the expr attribute. To allow the AS to differentiate between a <dialogexit> notification event resulting from a VoiceXML <disconnect> from one resulting from an <exit>, the MS creates a <param> with the reserved name '__reason', the type 'text/plain', and a value of "disconnect" (without brackets) to reflect the use of VoiceXML's <disconnect> element, and the value of "exit" (without brackets) to an explicit <exit> in the VoiceXML dialog. If the VoiceXML session terminates for other reasons (such as encountering an error), this parameter MAY be omitted or take on platform-specific values prefixed with an underscore.

Table 2 provides some examples of VoiceXML <exit> usage and the corresponding <params> element in the <dialogexit> notification event. It assumes the following VoiceXML variable names and values: userAuthorized=true, pin=1234 and errors=0. The <param> type attributes ('text/plain') are omitted for clarity.

<exit> Usage	<params> Result
<exit>	<params> <param name="__reason">exit</param> </params>
<exit expr="5">	<params> <param name="__reason">exit</param> <param name="__exit">5</param> </params>
<exit expr="'done'">	<params> <param name="__reason">exit</param> <param name="__exit">'done'</param> </params>
<exit expr="userAuthorized">	<params> <param name="__reason">exit</param> <param name="__exit">true</param> </params>
<exit namelist="pin errors">	<params> <param name="__reason">exit</param> <param name="pin">1234</param> <param name="errors">0</param> </params>

Table 2: VoiceXML <exit> mapping examples

12.5. Call Transfer

While VoiceXML is at its core a dialog language, it also provides optional call transfer capability. It is NOT RECOMMENDED to use VoiceXML's call transfer capability in networks involving Application Servers. Rather, the AS itself can provide call routing functionality by taking signaling actions based on the data returned to it, either through VoiceXML's own data submission mechanisms or through the mechanism described in Section 12.4. If the MS encounters a VoiceXML dialog requesting call transfer capability, the MS SHOULD raise an error event in the VoiceXML dialog execution context: an error.unsupported.transfer.blind event if blind transfer is requested, error.unsupported.transfer.bridge if bridge transfer is requested, or error.unsupported.transfer.consultation if consultation transfer is requested.

13. References

13.1. Normative References

[I-D.ietf-mediactrl-sip-control-framework]
Boulton, C., Melanchuk, T., and S. McGlashan, "Media Control Channel Framework", draft-ietf-mediactrl-sip-control-framework-09 (work in progress), February 2009.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC3023] Murata, M., St. Laurent, S., and D. Kohn, "XML Media Types", RFC 3023, January 2001.

[RFC3986] Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier (URI): Generic Syntax", STD 66, RFC 3986, January 2005.

[RFC4574] Levin, O. and G. Camarillo, "The Session Description Protocol (SDP) Label Attribute", RFC 4574, August 2006.

[RFC4646] Phillips, A. and M. Davis, "Tags for Identifying Languages", BCP 47, RFC 4646, September 2006.

[RFC4647] Phillips, A. and M. Davis, "Matching of Language Tags", BCP 47, RFC 4647, September 2006.

[SRGS] Hunt, A. and S. McGlashan, "Speech Recognition Grammar Specification Version 1.0", W3C Recommendation, March 2004.

[W3C.REC-SMIL2-20051213] Bulterman, D., Layaida, N., Jansen, J., Michel, T., Zucker, D., Mullender, S., Koivisto, A., and G. Grassel, "Synchronized Multimedia Integration Language (SMIL 2.1)", World Wide Web Consortium Recommendation REC-SMIL2-20051213, December 2005, <http://www.w3.org/TR/2005/REC-SMIL2-20051213>.

[XML] Bray, T., Paoli, J., Sperberg-McQueen, C M., Maler, E., and F. Yergeau, "Extensible Markup Language (XML) 1.0 (Third Edition)", W3C Recommendation, February 2004.

[XMLSchema:Part2] Biron, P. and A. Malhotra, "XML Schema Part 2: Datatypes Second Edition", W3C Recommendation, October 2004.

13.2. Informative References

[CCXML10] Auburn, R J., "Voice Browser Call Control: CCXML Version 1.0", W3C Working Draft (work in progress), January 2007.

[H.248.9] "Gateway control protocol: Advanced media server packages", ITU-T Recommendation H.248.9.

[I-D.ietf-xcon-common-data-model] Novo, O., Camarillo, G., Morgan, D., Even, R., and J. Urpalainen, "Conference Information Data Model for Centralized Conferencing (XCON)", draft-ietf-xcon-common-data-model-12 (work in progress), October 2008.

[IANA] "IANA registry for RTP Payload Types", <http://www.iana.org/assignments/rtp-parameters>.

[MIME.mediatypes] "IANA registry for MIME Media Types", <http://www.iana.org/assignments/media-types/>.

[MSML] Saleem, A., Xin, Y., and G. Sharratt, "Media Session Markup Language (MSML)", draft-saleem-msml-08 (work in progress), February 2009.

[RFC2897] Cromwell, D., "Proposal for an MGCP Advanced Audio Package", RFC 2897, August 2000.

[RFC3261] Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston, A., Peterson, J., Sparks, R., Handley, M., and E. Schooler, "SIP: Session Initiation Protocol", RFC 3261, June 2002.

[RFC4240] Burger, E., Van Dyke, J., and A. Spitzer, "Basic Network Media Services with SIP", RFC 4240, December 2005.

[RFC4244] Barnes, M., "An Extension to the Session Initiation Protocol (SIP) for Request History Information", RFC 4244, November 2005.

[RFC4267] Froumentin, M., "The W3C Speech Interface Framework Media Types: application/voicexml+xml, application/ssml+xml, application/srgs, application/srgs+xml, application/ccxml+xml, and application/pls+xml", RFC 4267, November 2005.

[RFC4281] Gellens, R., Singer, D., and P. Frojdh, "The Codecs

Parameter for "Bucket" Media Types", RFC 4281, November 2005.

[RFC4627] Crockford, D., "The application/json Media Type for JavaScript Object Notation (JSON)", RFC 4627, July 2006.

[RFC4730] Burger, E. and M. Dolly, "A Session Initiation Protocol (SIP) Event Package for Key Press Stimulus (KPML)", RFC 4730, November 2006.

[RFC4733] Schulzrinne, H. and T. Taylor, "RTP Payload for DTMF Digits, Telephony Tones, and Telephony Signals", RFC 4733, December 2006.

[RFC4855] Casner, S., "Media Type Registration of RTP Payload Formats", RFC 4855, February 2007.

[RFC5022] Van Dyke, J., Burger, E., and A. Spitzer, "Media Server Control Markup Language (MSCML) and Protocol", RFC 5022, September 2007.

[RFC5167] Dolly, M. and R. Even, "Media Server Control Protocol Requirements", RFC 5167, March 2008.

[VXML20] McGlashan, S., Burnett, D., Carter, J., Danielsen, P., Ferrans, J., Hunt, A., Lucas, B., Porter, B., Rehor, K., and S. Tryphonas, "Voice Extensible Markup Language (VoiceXML) Version 2.0", W3C Recommendation, March 2004.

[VXML21] Oshry, M., Auburn, RJ., Baggia, P., Bodell, M., Burke, D., Burnett, D., Candell, E., Carter, J., McGlashan, S., Lee, A., Porter, B., and K. Rehor, "Voice Extensible Markup Language (VoiceXML) Version 2.1", W3C Recommendation, June 2007.

[VXML30] McGlashan, S., Auburn, RJ., Baggia, P., Barnett, J., Bodell, M., Burnett, D., Carter, J., Oshry, M., Rehor, K., Young, M., and R. Hosn, "Voice Extensible Markup Language (VoiceXML) Version 3.0", W3C Working Draft, December 2008.

Authors' Addresses

Scott McGlashan
Hewlett-Packard
Gustav III:s boulevard 36
SE-16985 Stockholm, Sweden

Email: scott.mcglashan@hp.com

Tim Melanchuk
Rain Willow Communications

Email: tim.melanchuk@gmail.com

Chris Boulton
NS-Technologies

Email: chris@ns-technologies.com