

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 10, 2009

S. McGlashan
Hewlett-Packard
T. Melanchuk
Rain Willow Communications
C. Boulton
Avaya
October 7, 2008

An Interactive Voice Response (IVR) Control Package for the Media
Control Channel Framework
draft-ietf-mediactrl-ivr-control-package-01

Status of this Memo

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with [Section 6 of BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at
<http://www.ietf.org/ietf/1id-abstracts.txt>.

The list of Internet-Draft Shadow Directories can be accessed at
<http://www.ietf.org/shadow.html>.

This Internet-Draft will expire on April 10, 2009.

Internet-Draft

IVR Control Package

October 2008

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	5
2.	Conventions and Terminology	9
3.	Control Package Definition	10
3.1.	Control Package Name	10
3.2.	Framework Message Usage	10
3.3.	Common XML Support	11
3.4.	CONTROL Message Body	11
3.5.	REPORT Message Body	11
3.6.	Audit	12
4.	Element Definitions	13
4.1.	<mscivr>	13
4.2.	Dialog Management Elements	15
4.2.1.	<dialogprepare>	18
4.2.2.	<dialogstart>	20
4.2.2.1.	<subscribe>	23
4.2.2.1.1.	<dtmfsub>	24
4.2.2.2.	<stream>	25
4.2.2.2.1.	<region>	26
4.2.3.	<dialogterminate>	26
4.2.4.	<response>	27
4.2.5.	<event>	29
4.2.5.1.	<dialogexit>	29
4.2.5.2.	<dtmfnotify>	30
4.2.6.	<params>	31
4.2.6.1.	<param>	31
4.3.	IVR Dialog Elements	32
4.3.1.	<dialog>	34

4.3.1.1.	<prompt>	36
4.3.1.1.1.	<media>	37
4.3.1.1.2.	<variable>	38
4.3.1.1.3.	<dtmf>	40
4.3.1.2.	<control>	41

4.3.1.3.	<collect>	43
4.3.1.3.1.	<grammar>	45
4.3.1.4.	<record>	47
4.3.2.	Exit Information	51
4.3.2.1.	<promptinfo>	51
4.3.2.2.	<controlinfo>	51
4.3.2.2.1.	<controlmatch>	51
4.3.2.3.	<collectinfo>	51
4.3.2.4.	<recordinfo>	52
4.4.	Audit Elements	52
4.4.1.	<audit>	53
4.4.2.	<auditresponse>	54
4.4.2.1.	<codecs>	56
4.4.2.1.1.	<codec>	56
4.4.2.2.	<capabilities>	57
4.4.2.2.1.	<dialoglanguages>	59
4.4.2.2.2.	<grammartypes>	60
4.4.2.2.3.	<recordtypes>	60
4.4.2.2.4.	<mediatypes>	60
4.4.2.2.5.	<variables>	61
4.4.2.2.5.1.	<variabletype>	61
4.4.2.2.6.	<maxpreparedduration>	62
4.4.2.2.7.	<maxrecordduration>	62
4.4.2.3.	<dialogs>	62
4.4.2.3.1.	<dialogaudit>	62
4.5.	Response Status Codes	63
4.6.	Type Definitions	65
5.	Formal Syntax	67
6.	Examples	91
6.1.	AS-MS Dialog Interaction Examples	91
6.1.1.	Starting an IVR dialog	91
6.1.2.	IVR dialog fails to start	92
6.1.3.	Preparing and starting an IVR dialog	92
6.1.4.	Terminating a dialog	93
6.2.	IVR Dialog Examples	94
6.2.1.	Playing announcements	94

6.2.2.	Prompt and collect	95
6.2.3.	Prompt and record	97
6.2.4.	Runtime controls	98
6.2.5.	Subscriptions and notifications	98
6.3.	Other Dialog Languages	99
6.4.	Foreign Namespace Attributes and Elements	100
7.	Security Considerations	102
8.	IANA Considerations	103
8.1.	Control Package Registration	103
8.2.	URN Sub-Namespace Registration	103
8.3.	Mime Type Registration	103
9.	Change Summary	104

10.	Contributors	112
11.	Acknowledgments	113
12.	References	114
12.1.	Normative References	114
12.2.	Informative References	115
	Authors' Addresses	117
	Intellectual Property and Copyright Statements	118

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 own 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 the IETF MediaCtrl requirements ([[RFC5167](#)]) 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 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, 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:

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

Out of scope for this dialog language are more advanced functions including ASR (Automatic Speech Recognition), TTS (Text-to-Speech), VoiceXML, fax and media transformation. Such functionality may 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 may 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 successfully, 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

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 may 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.

Internet-Draft

IVR Control Package

October 2008

[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](#).

[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](#). 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 Framework Server in receiving requests from, and sending responses to, the AS (operating as Control Framework 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 Framework Client in sending event notification to the AS (Control Framework 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 Framework Server, the MS receives CONTROL messages body containing an <mscivr> element 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 Framework Client, the MS sends CONTROL messages with a body containing a notification <event> 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 should indicate the location of detailed syntax definitions and semantics for the appropriate body types.

When operating as Control Framework Server, the MS sends REPORT bodies containing a <mscivr> element 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 messages and an audit response in a REPORT message (or a 200 response to the CONTROL if it can execute the audit in time).

The syntax and semantics of audit request and response elements is defined in [Section 4.4](#).

[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 adhere to the syntax and semantics of XML elements defined in this section and the schema ([Section 5](#)). The XML schema supports extensibility by allowing attributes and elements from other namespaces. Implementations MAY support attributes and elements from other (foreign) namespaces. If

an MS implementation receives a <mscivr> element containing attributes or elements from another namespace which it does not support, the MS MUST send a 426 response ([Section 4.5](#)).

Attributes and elements from foreign namespaces are not described in this section. In all other cases where there is a difference in constraints between the XML schema and the textual description of elements in this section, the textual definition takes priority.

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 4.3.1.1.1](#)), <grammar> ([Section 4.3.1.3.1](#)), and <record> ([Section 4.3.1.4](#)). While this package is agnostic to the URI schemes supported by the MS, it is RECOMMENDED that the MS 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>
    <dialog>
      <prompt>
        <media src="http://www.example.com/welcome.wav"/>
      </prompt>
    </dialog>
  </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](#)). 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, it is an error and the MS MUST indicate the error in the status code of the response. Both general error codes (e.g. syntax errors such as missing elements or attributes) as well as request-specific status codes (e.g. invalid connectionid) are defined in [Section 4.5](#)

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](#)).

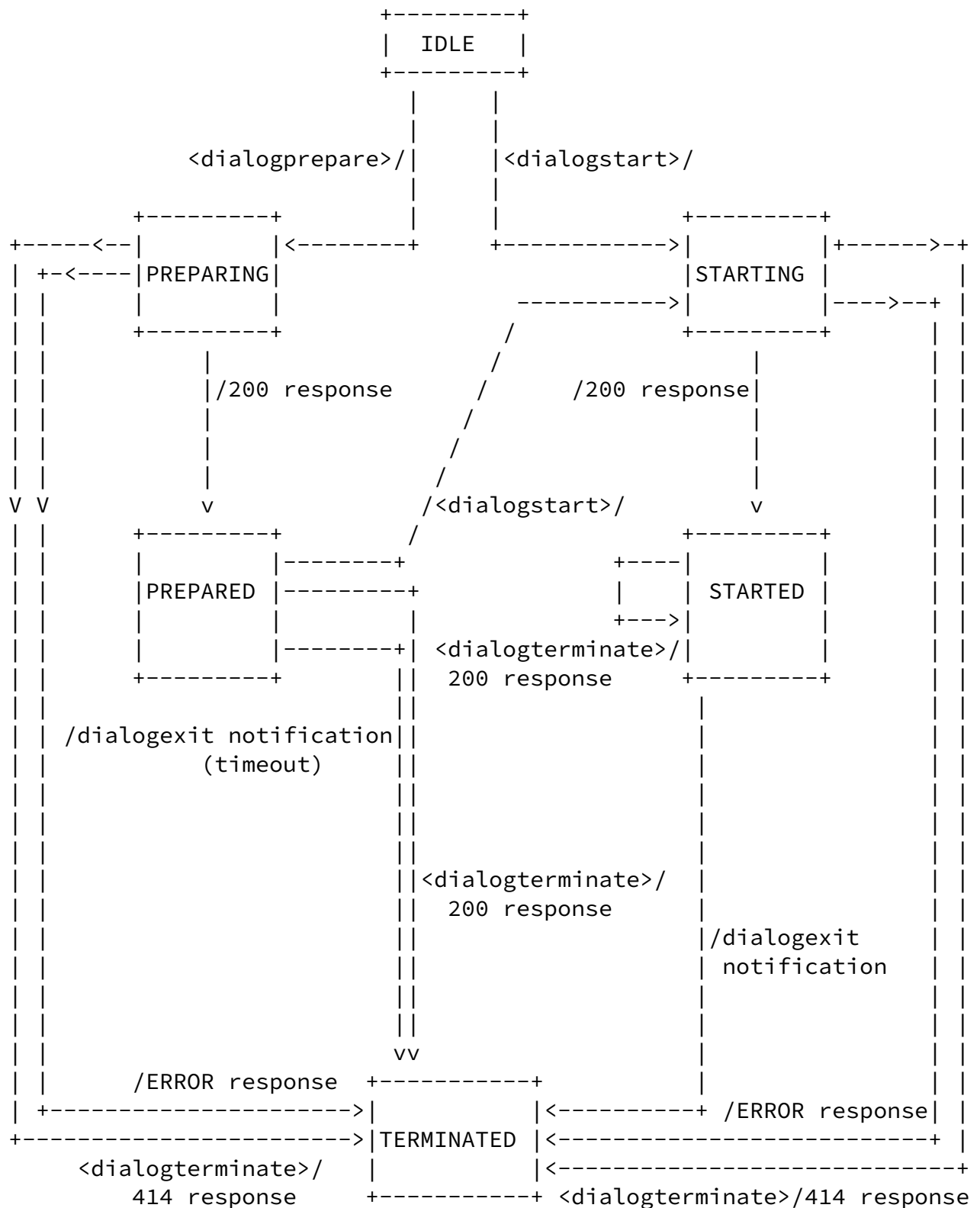


Figure 1: Dialog Lifecycle

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 type. If the dialog is terminated before preparation is complete, the dialog transitions to the TERMINATED state and the MS MUST send a 414 response ([Section 4.5](#)) for the prepare request.

PREPARED: the dialog has been successfully prepared and the MS MUST send a 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 an error status code ([Section 4.5](#)). 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 type. If the dialog is terminated, the dialog transitions to the TERMINATED state and the MS MUST send a 414 response ([Section 4.5](#)) for the start request.

STARTED: the dialog has been successfully started and is now active. The MS MUST send a 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 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, the MS MUST send a response with a 401 status code ([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 <dialogterminate> request without any dialog identifier specified.

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 unknown or unsupported, the MS MUST send a <response> with a 415 status code ([Section 4.5](#)). If the document cannot be retrieved, the MS MUST send a <response> with a 410 status code. If the document contains a type of dialog which the MS does not supported, the MS MUST send a <response> with a 409 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](#)). The MS MAY use the value to assist the remote source in selecting the appropriate resource type (e.g. with HTTP 'accept' header) and to determine how the document is to be processed if the protocol does not provide an authoritative MIME media type for the returned resource. The attribute is optional.

fetchtimeout: the maximum interval to wait when fetching an external dialog document. A valid value is a Time Designation (see [Section 4.6.7](#)). If the external document cannot be fetched within the maximum interval, the MS MUST send a <response> with a 417 status code ([Section 4.5](#)). 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 MUST send a <response> with a 401 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 and a src attribute are specified and the MS MUST send a <response> with a 418 status code ([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 src="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](#)).

[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 unknown or unsupported, the MS MUST send a <response> with a 415 status code ([Section 4.5](#)). If the document cannot be retrieved, the MS MUST send a <response> with a 410 status code. If the document contains a type of dialog which the MS does not supported, the MS MUST send a <response> with a 409 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](#)). The MS MAY use the value to assist the remote source in selecting the appropriate resource type (e.g. with HTTP 'accept' header) and to determine how the document is to be processed if the protocol does not provide an authoritative MIME media type for the returned resource. The attribute is optional.

fetchtimeout: the maximum interval to wait when fetching an external dialog document. A valid value is a Time Designation (see [Section 4.6.7](#)). If the external document cannot be fetched within the maximum interval, the MS MUST send a <response> with a 417 status code ([Section 4.5](#)). 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 MUST send a <response> with a 401 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 MUST send a <response> with a 418 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 MUST send a <response> with a 403 status code ([Section 4.5](#)). If an invalid conferenceid is specified, the MS MUST send a <response> with a 404 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 cannot be understood or processed by the MS for the external dialog language, the MS MUST send a <response> with a 425 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 may 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 MUST send a <response> with a 418 status code ([Section 4.5](#)). If the prepareddialogid and dialogid attributes are specified, it is also a syntax error and the MS MUST send a <response> with a 418 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, it is RECOMMENDED that the configuration is explicitly specified using <stream> elements. 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](#)]) or (d) if the media stream configuration is not supported by the MS, then the MS MUST send a <response> with a 413 status code ([Section 4.5](#)).

This specification allows multiple, simultaneous dialogs to be started on the same connection or conference. It is RECOMMENDED the MS support the following cases:

1. different media streams used in different dialogs; e.g. audio only on one dialog and video only on another dialog
2. the same media stream received by different dialogs
3. use of implicit mixing (where appropriate) when the same type of media stream is sent from different dialogs

If the MS does not support starting another dialog on the same connection or conference it MUST report an error when starting that dialog.

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 src="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"/>
  </dialogstart>
</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.

Internet-Draft

IVR Control Package

October 2008

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 a <dtmfsub> subscription. It MAY support other dialog subscriptions. If the MS does not support a requested subscription, it MUST send a <response> with a 412 status code ([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 src="http://www.example.com/getpin.wav"/>
      </prompt>
      <control ffkey="2" rewkey="3"/>
    </dialog>
    <subscribe>
      <dtmfnotify matchmode="control"/>
    </subscribe>
  </dialogstart>
</mscivr>
```

```
</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. It is strongly RECOMMENDED that the following values are 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" (media can be sent and received), "sendonly" (media can only be sent), "recvonly" (media can only be received) and "inactive" (stream is not to be used). The default value is "sendrecv". The attribute is optional.

The <stream> element has the following child element:

<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.

If conferenceid is not specified or if the "media" attribute does not have the value of "video", then the MS MUST ignore the <region> element.

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 src="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 MUST report a 416 status code ([Section 4.5](#)) in the response to the request element containing the <region> element.

[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. The MS MUST send a response with a 402 status code ([Section 4.5](#)) if the specified dialog identifier is invalid. 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 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 (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="409" 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 409 (attribute required) 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="408" 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 success or failure of the dialog. A valid value is a non-negative integer (see [Section 4.6.4](#)). A value of 0 indicates that the dialog has been terminated by a <dialogterminate> request. A value of 1 indicates success. A value of 2 indicates that the dialog terminated because the connection or conference associated with the dialog has terminated. A value of 3 indicates the dialog terminated due to exceeding its maximum duration. A value of 4 indicates the dialog terminated due to an execution error. Any other value indicates an error defined by the MS. The attribute is mandatory.

reason: a textual description providing 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>
```

[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.

Internet-Draft

IVR Control Package

October 2008

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/BvIkOVJK0g2xZUKmenMfDw8tgWJpV
      ]]>

    </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 recording, type and termmode attributes 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 may choose to support only 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:

<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.

The MS MUST send a <response> with a 418 status code ([Section 4.5](#)) if no child element is specified or a <control> element is specified but not a <prompt> element. Although the behavior when both <collect> and <record> elements are specified is not defined in this control package, the MS MAY support this configuration. If the MS does not

support this configuration, the MS MUST send a <response> with a 418 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.

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 (any order, multiple occurrences of each element permitted):

<media>: media resource (see [Section 4.3.1.1.1](#)) to play. The element is optional.

<variable>: specifies a variable media announcement (see [Section 4.3.1.1.2](#)) to play. The element is optional.

<dtmf>: generates one or more DTMF tones (see [Section 4.3.1.1.3](#)) to play. The element is optional.

If no child element is specified, then the MS MUST send a <response> with a 418 status code ([Section 4.5](#)).

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

1. The MS initiates prompt playback playing each <media>, <variable> and <dtmf> 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](#). <media>

The <media> element specifies a media resource to play.

A <media> element has the following attributes:

src: specifies the location of the media resource. A valid value is a URI (see [Section 4.6.9](#)). If the URI scheme is unknown or unsupported, the MS MUST send a <response> with a 415 status code ([Section 4.5](#)). If the resource cannot be retrieved, the MS MUST send a <response> with a 410 status code. If the format of the media resource is not supported, the MS MUST send a <response> with a 419 status code. The attribute is mandatory.

type: specifies the type of the media resource indicated in the 'src' attribute. The MS MAY use the value to assist the remote source in selecting the appropriate resource type (e.g. with HTTP 'accept' header) and to determine how the resource is to be processed if the protocol does not provide an authoritative MIME media type for the returned resource. The value may include additional parameters for guiding playback; for example, [[RFC4281](#)] defines a 'codec' parameter for 'bucket' media types like video/3gpp. A valid value is a MIME media type (see [Section 4.6.10](#)). 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](#)). If the resource cannot be fetched within the maximum interval, the MS MUST send a <response> with a 417 status code ([Section 4.5](#)). 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 playing. 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 playing. 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 <media> element has no children.

[4.3.1.1.2.](#) <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.

Internet-Draft

IVR Control Package

October 2008

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. However it is RECOMMENDED that an implementation support the following type/format combinations:

type=date Supported formats: "mdy" (month day year), "ymd" (year month day), "dym" (day month year), "dm" (day month)

type=time Supported formats: "t12" (12 hour format with am/pm), "t24" (24 hour format)

type=digits Supported formats: "gen" (general digit string), "crn" (cardinal), "ord" (ordinal)

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 may map the <variable> into one or more audio media resources.

If a <variable> element configuration is not supported by the MS, the

MS MUST send a <response> with a 421 status code ([Section 4.5](#)).

Depending on the specific implementation of the <variable> rendering on the MS, execution of this element may 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 src="nfs://voicebase/en/male/25th.wav"/>
<media src="nfs://voicebase/en/male/of.wav"/>
<media src="nfs://voicebase/en/male/february.wav"/>
<media src="nfs://voicebase/en/male/2000.wav"/>
<media src="nfs://voicebase/en/male/and.wav"/>
<media src="nfs://voicebase/en/male/8.wav"/>
```

[4.3.1.1.3](#). <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](#)). Implementations may round the value if they only support 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](#)). Implementations may round the value if they only support 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 MUST send a <response> with a 422 status code ([Section 4.5](#)).

[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 should skip backwards or forwards through prompt playback when the rewind (rwkey) or 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 should pause 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 MUST send a <response> with a 423 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 DTMF input matches any specified keys (for example the ffkey), then the appropriate operation is applied immediately. If a seek operation (ffkey, rwkey) attempts to go beyond the beginning or end of the prompt queue, then it is automatically truncated to the prompt beginning or end 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 volume operations attempts to go beyond the minimum or maximum volume supported by the platform, then the operation is ignored.
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.

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 may 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 unknown or unsupported, the MS MUST send a `<response>` with a 415 status code ([Section 4.5](#)). If the resource cannot be retrieved, the MS MUST send a `<response>` with a 410 status code. If the grammar format is not supported, the MS MUST send a `<response>`

with a 420 status code. The attribute is optional. There is no

Internet-Draft

IVR Control Package

October 2008

default value.

type: identifies the preferred type of the grammar document identified by the `src` attribute. The MS MAY use the value to assist the remote source in selecting the appropriate resource type (e.g. with HTTP 'accept' header) and to determine how the document is processed if the protocol does not provide an authoritative MIME media type for the returned resource. A valid value is a MIME media type (see [Section 4.6.10](#)). 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](#)). If the resource cannot be fetched within the maximum interval, the MS MUST send a <response> with a 417 status code ([Section 4.5](#)). 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 MUST send a <response> with a 420 status code ([Section 4.5](#)).

For example, the following fragment shows DTMF collection with an inline SRGS grammar:

Internet-Draft

IVR Control Package

October 2008

```
<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 defines how media input is recorded.

The <record> element has the following attributes:

dest: specifies the location where recorded data is to be stored. The MS MUST have uploaded the recorded data to this location as soon as possible after recording is complete. A valid value is a URI (see [Section 4.6.9](#)). If the URI scheme is unknown or unsupported, the MS MUST send a <response> with a 415 status code ([Section 4.5](#)). If the attribute is optional. The MS SHOULD support the inclusion of authentication information in the URI if the URI scheme supports it (e.g. basic access authentication in HTTP). The MS MAY support more advanced authentication mechanisms. If the dest attribute not specified, the MS MUST provide a local recording location (reporting the local URI in the recording attribute of in <recordinfo> [Section 4.3.2.4](#)); the recording MUST be available from this location until the connection or conference associated with the dialog on the MS terminates.

type: specifies the type of the recording format. The type value may include additional parameters for guiding recording; for example, [[RFC4281](#)] defines a 'codec' parameter for 'bucket' media types like video/3gpp. A valid value is a MIME media type (see [Section 4.6.10](#)). If the specified recording format is not supported, the MS MUST send a <response> with a 424 status code ([Section 4.5](#)). The attribute is optional. There is no default value (recording format is MS implementation-dependent if a value is not specified).

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 can be used to initiate the recording operation. A valid value is a boolean (see [Section 4.6.1](#)). A value of true indicates recording may be initiated using voice activity detection. A value of false indicates that recording must not be initiated using voice activity detection. The attribute is optional. The default value is true.

vadfinal: Control whether voice activity detection can be used to terminate the recording operation. A valid value is a boolean (see [Section 4.6.1](#)). A value of true indicates recording may be terminated using voice activity detection. A value of false indicates that recording must not be terminated using voice activity detection. The attribute is optional. The default value is true.

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' should 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 end of speech. 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 should be appended or not to the 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 the 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.

The <record> element has no child elements.

Note that an MS MAY support uploading recorded data to the recording location at the same time the recording operation takes place. Such implementations should be aware of the requirements of certain recording formats (e.g. WAV) for metadata at the beginning of the uploaded file, that the final silence 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 is set to maxtime.
7. During the record operation user media input is recording in the format specified by the value of the type attribute. If the dest attribute is specified, then recorded input is sent to that location. Otherwise, MS uses an internal location.
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 a 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.

[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 'bargain'. 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 the information about record execution.

The <recordinfo> element has the following attributes:

recording: references the location to which media is recorded. A valid value is a URI (see [Section 4.6.9](#)). The attribute is optional. There is no default value.

type: indicates the format of the recording. A valid value is a MIME media type (see [Section 4.6.10](#)). The attribute is optional. There is no default value.

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.

size: indicates the size of the recording in bytes. 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 recording was terminated. Valid values are: 'stopped', 'noinput', 'dtmf', 'maxtime', and 'finalsilence'. The attribute is mandatory.

The <recordinfo> element has no child elements.

[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 may then use this information to create request elements using supported capabilities and, in the case of codecs, to negotiate an appropriate

Internet-Draft

IVR Control Package

October 2008

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 allows one AS to take over management of the dialogs when the AS which initiated the dialogs fails or is no longer available.

[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 MUST send a response with a 402 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 a <audit> request, it MUST reply with a <auditresponse> element ([Section 4.4.2](#)). If the request is successful, <auditresponse> contain (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.

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>
    <mimetype>application/srgs+xml</mimetype>
  </grammartypes>
  <recordtypes>
    <mimetype>application/x-wav</mimetype>
    <mimetype>video/3gpp</mimetype>
  </recordtypes>
  <mediatypes>
    <mimetype>application/x-wav</mimetype>
    <mimetype>video/3gpp</mimetype>
  </mediatypes>
  <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 `<codecs>` element has no attributes.

The `<codecs>` 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 supported <record> ([Section 4.3.1.4](#)) format types. The element is mandatory.

<mediatypes>: element ([Section 4.4.2.2.4](#)) describing supported <media> ([Section 4.3.1.1.1](#)) format types. 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>
    <mimetype>application/srgs+xml</mimetype>
  </grammartypes>
  <recordtypes>
    <mimetype>application/x-wav</mimetype>
    <mimetype>video/3gpp</mimetype>
  </recordtypes>
  <mediatypes>
    <mimetype>application/x-wav</mimetype>
    <mimetype>video/3gpp</mimetype>
  </mediatypes>
  <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):

Internet-Draft

IVR Control Package

October 2008

`<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 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 (1 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 `<record>` format types 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](#). `<mediatypes>`

The `<mediatypes>` element provides information about `<media>` format types supported by the package ([Section 4.3.1.1.1](#)).

The `<mediatypes>` element has no attributes.

The `<mediatypes>` 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 mandatory.

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

The following status codes for dialog management ([Section 4.2.4](#)) and audit ([Section 4.4.2](#)) responses are defined:

code	description
200	OK
401	dialogid already exists
402	dialogid does not exist
403	connectionid does not exist
404	conferenceid does not exist
405	Unknown or unsupported element

406	Element required
407	Unknown or unsupported attribute
408	Attribute required
409	Dialog language not supported
410	Retrieving resource failed
411	Invalid attribute value
412	Subscription not supported
413	Invalid stream configuration

414	Dialog execution canceled
415	Unsupported URI scheme
416	Invalid region identifier
417	Resource cannot be retrieved within timeout interval
418	Syntax constraint violation
419	Media format not supported
420	Grammar format not supported
421	Variable announcement not supported
422	DTMF tone not supported
423	Control keys with same value
424	Recording format not supported
425	Unsupported param parameter
426	Unsupported foreign namespace attribute or element
499	Other error

Table 1: status codes

The MS MAY define other status codes.

[4.6.](#) Type Definitions

This section defines types referenced in attribute 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 formatted 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 formatted 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 (20081003)
```

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">
```

Internet-Draft

IVR Control Package

October 2008

```
<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 component types to
      allow elements and attributes from other namespaces
      to be added.
    </xsd:documentation>
  </xsd:annotation>
```

```

<xsd:sequence>
  <xsd:any namespace="##other" minOccurs="0"
    maxOccurs="unbounded" processContents="lax" />
</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: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: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: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: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: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: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: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:attribute name="recording" type="xsd:anyURI" />
      <xsd:attribute name="type" type="mime.datatype" />
      <xsd:attribute name="duration"
        type="xsd:nonNegativeInteger" />
      <xsd:attribute name="size"
        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" />

<!-- subscribe -->

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

<xsd:sequence>

Internet-Draft

IVR Control Package

October 2008

```
<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: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: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" />

```

```

<!-- 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:any namespace="##other"
    processContents="lax" />
</xsd:choice>
<xsd:attribute ref="xml:base" />
<xsd:attribute name="bargain" 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:attribute name="src" 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="variableType">
  <xsd:complexContent>
    <xsd:extension base="Tcore">
      <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="variableType" />

<!-- dtmf -->

<xsd:complexType name="dtmfType">
  <xsd:complexContent>
    <xsd:extension base="Tcore">
      <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" />

<!-- control -->

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

```

```

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

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

<!-- control -->

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

```

```

<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:attribute name="timeout"
      type="timedesignation.datatype" default="5s" />
    <xsd:attribute name="type" type="mime.datatype" />
    <xsd:attribute name="dest" type="xsd:anyURI" />
    <xsd:attribute name="beep" type="boolean.datatype"
      default="false" />
    <xsd:attribute name="vadinitial"
      type="boolean.datatype" default="true" />
    <xsd:attribute name="vadfinal"
      type="boolean.datatype" default="true" />
    <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: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>

```

Internet-Draft

IVR Control Package

October 2008

```
</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:attribute name="name" type="xsd:string" />
      <xsd:attribute name="policy" type="policy.datatype" />
    </xsd:extension>
  </xsd:complexContent>
</xsd:complexType>
```

```

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

```

```

<xsd:element ref="recordtypes" minOccurs="1"
  maxOccurs="1" />
<xsd:element ref="mediatypes" 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" />

<!-- mediatypes -->

<xsd:complexType name="mediatypesType">

```



```

<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="mediatypes" type="mediatypesType" />

<!-- 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" />

<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: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="policy.datatype">
    <xsd:restriction base="xsd:NMTOKEN">
      <xsd:enumeration value="allowed" />
      <xsd:enumeration value="disallowed" />
    </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">

```

Internet-Draft

IVR Control Package

October 2008

```
<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="bargain" />
    <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.

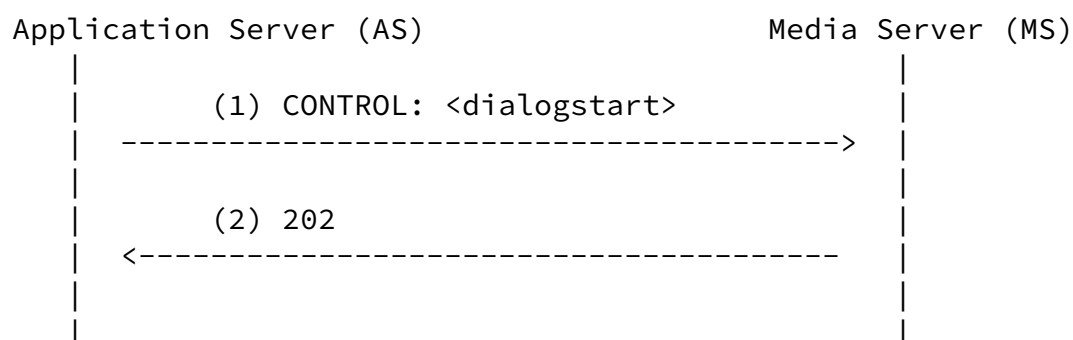
6.1. AS-MS Dialog Interaction Examples

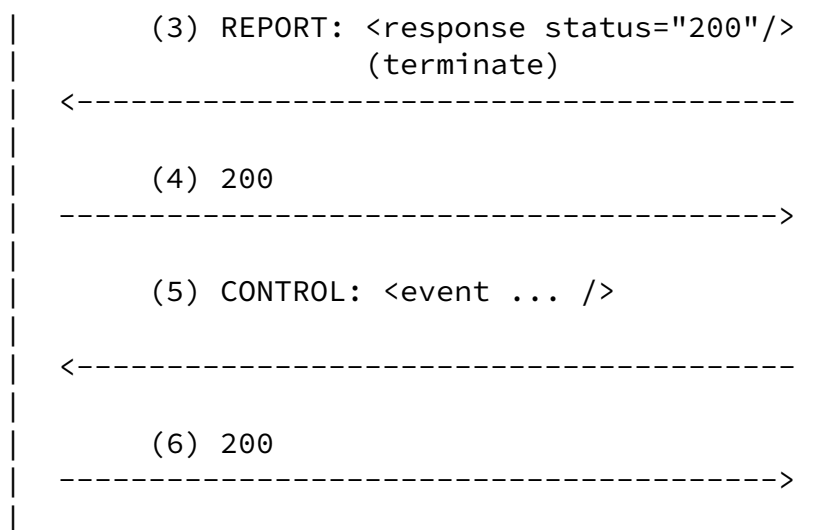
The following example assume a control channel has been established and synced as described in the Media Control Channel Framework ([[I-D.ietf-mediactrl-sip-control-framework](#)]).

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

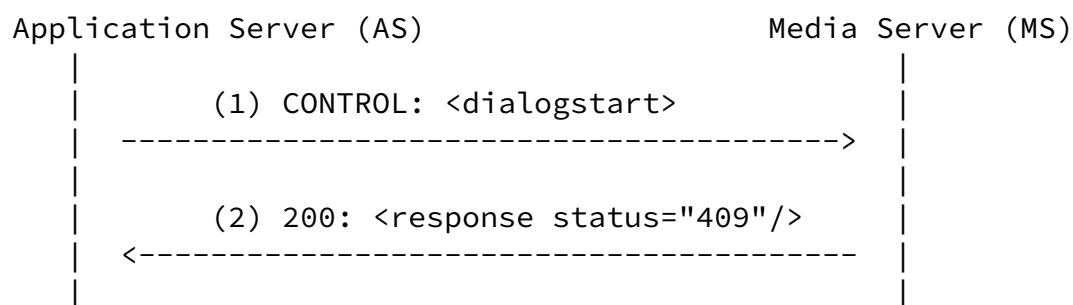
An IVR dialog is started successfully, and dialogexit notification <event> is sent from the MS to the AS when the dialog exits normally.





[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.



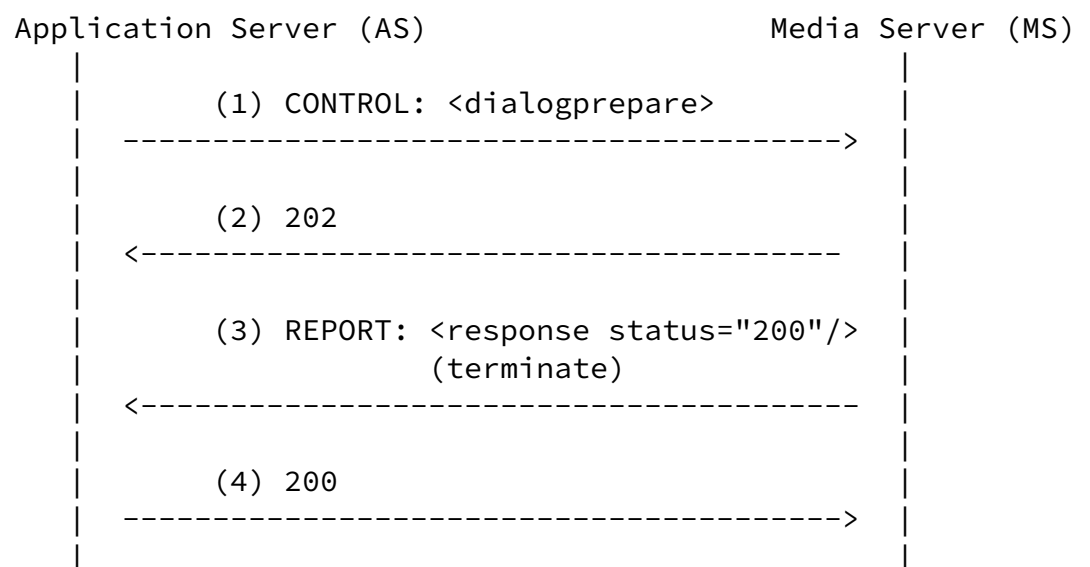
[6.1.3.](#) Preparing and starting an IVR dialog

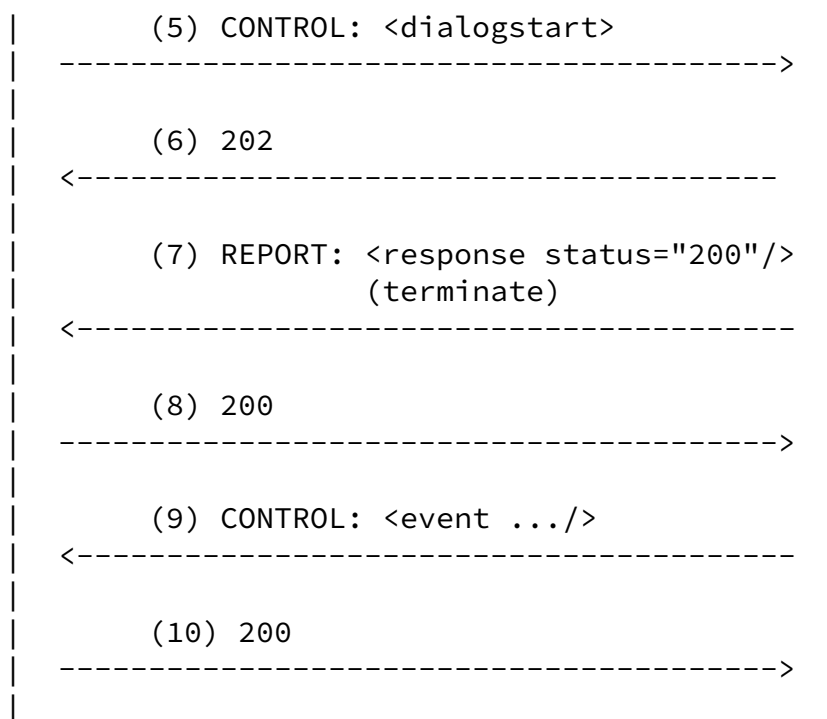
An IVR dialog is prepared and started successfully, and then the dialog exits normally.

Internet-Draft

IVR Control Package

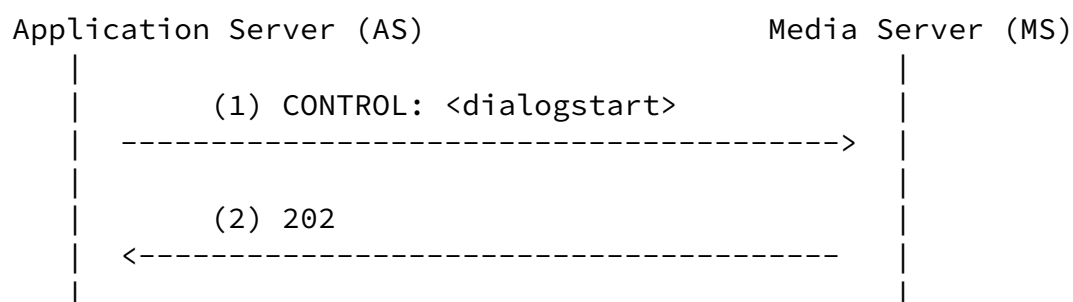
October 2008

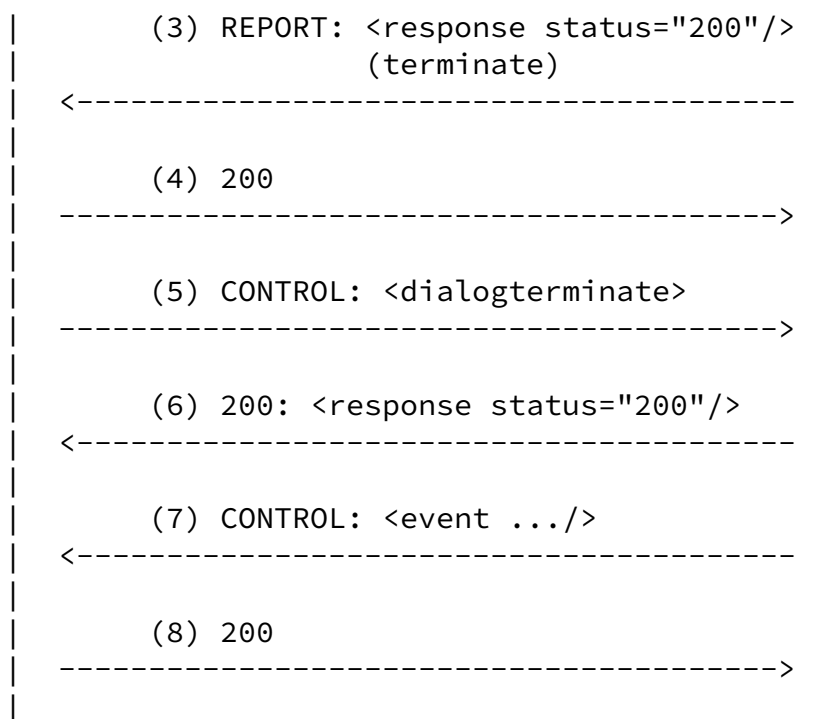




[6.1.4.](#) Terminating a dialog

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

```

<dialogprepare>
  <dialog repeatCount="2">
    <prompt>
      <media src="http://www.example.com/media/Number_09.wav"/>
      <media src="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 src="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" >
      <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:

Internet-Draft

IVR Control Package

October 2008

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
<dialogstart connectionid="7HDY839~HJKSkyHS~HUwkuh7ns">
  <dialog repeatCount="two">
    <prompt>
      <media src="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="411" dialogid="vxi82"
    reason="repeatCount 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 src="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 recording="http://www.example.com/recording1.wav"
        termmode="dtmf"/>
    </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 src="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 repeatCount="0" connectionid="7HDY839~HJKSkyHS~HUwkuh7ns">
    <dialog>
      <collect maxdigits="2"/>
      <subscribe>
        <dtmfsub matchmode="collect"/>
      </subscribe>
    </dialog>
  </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" 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 409 ([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 in the VoiceXML script through the "connection.ccxml.values" object.

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 may support attributes and elements from foreign namespaces within the <mscivr> element. For example, it may 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 src="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 426 response:

```
<mscivr version="1.0" xmlns="urn:ietf:params:xml:ns:msc-ivr">
  <response status="426" 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 may provide 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 must send a Control Framework 200 response even though the notification event contains an element in a foreign namespace which it may not understand.

[7.](#) Security Considerations

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

As a Control Package of the Media Control Channel Framework, security, confidentiality and integrity of messages transported over the control channel MUST be addressed as described in [Section 11](#) of the Media Control channel Framework ([\[I-D.ietf-mediactrl-sip-control-framework\]](#)), including Session Establishment, Transport Level Protection and Control Channel Policy Management.

The Media Control Channel Framework permits additional 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, policy management for this control package MUST address how such dialogs are managed across multiple control channels. This includes which channels are used to deliver dialog event notifications, and whether channels are permitted to originate requests managing a dialog which was not created through that channel (e.g. a dialog has been prepared or started via channel X and a request to terminate the dialog originates from channel Y).

[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

Internet-Draft

IVR Control Package

October 2008

[9.](#) Change Summary

Note to RFC Editor: Please remove this whole section.

The following are the major changes between the -01 and -00 versions.

- o 7: Updated security section referencing control framework security and adding policy requirement to address dialog resource management over multiple channels.
- o corrected typos and example errors
- o 4.2: [IVR-200] Added state machine for dialog lifecycle.
- o 4.2: clarified dialog identifier assignment and use, including MS assignment of dialogid in <dialogprepare> and <dialogstart>.
- o 4.2/4.2.3: clarified <dialogterminate> behavior when dialog is not in a STARTED state.
- o 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>.
- o 4.4.2.2.1: clarified that the inline dialog language (<dialog>) must not be listed as an additional supported dialog language.
- o 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.
- o 4.3.1.1.2:[IVR-202]: removed ndn format and clarified gen format for digits.
- o 4.4.2.1.1:[IVR-203]: added <params> to <codec> to allow additional codec information to be specified.

- o 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).

- o Generally, replaced 'it is an error ...' language with [RFC2119](#) language, making error codes more explicit.
- o 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.
- o 4.2.5.2: clarified that the timestamp in <dtmfnotify> is that of the last DTMF in a matching input sequence.
- o References: more references now normative.
- o 4.3: Replaced passive voice language with active voice language in the description of execution models.
- o 1/2: Clarified that the term 'dialog' refers to IVR dialog and is completely unrelated to the term 'SIP dialog'.
- o 4: Added clarification that elements with URI attributes are recommended to support one or more communication protocols suitable for fetching resources.
- o 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.
- o 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.

- o 4.2.6.1: <param> removed 'valuetype' attribute and clarified that the type attribute indicates the MIME media type associated with the inline value.
- o 4.3.1.4: Added append attribute to <record> to control whether recordings are appended or not to the recording location resource.
- o 4.3.1.1.1: Added clipEnd attribute to <media> to control when playback of the media ends.
- o 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.

- o 4.4.2.2: Added ability to audit <variable> capability, as well as maximum duration of prepared dialogs and of recordings.
- o 4.3.1/4.3.1.1: Clarified the sequence of dialog operations in <dialog> and how <prompt> reports its status.
- o 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.

- o [IVR01] When the MS sends a notification event in a CONTROL, the AS sends mandatory 200 response (no extended transaction).
- o [IVR23] Added a top-level container element, <mscivr>, with version attribute.
- o 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.
- o [IVR15] Updated and simplified XML schema. Ordering of child

elements is significant.

- o [IVR06] Removed 'volume' and 'offset' from prompt element. Added 'soundLevel' and 'clipBegin' to media element.
- o [IVR17]/[IVR06] Removed 'iterations' and 'duration' from prompt. Added 'repeatCount' and 'repeatDur' to dialog element.
- o Moved VCR commands from <collect> into separate <control> element. Defined controlinfo element to report runtime control match information.
- o [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.
- o Moved definition of <stream> into a separate section.
- o [IVR21] Added audit capability: auditing of package capabilities and managed dialogs

- o [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.
- o Clarified that the <variable> rendering mechanism is MS implementation specific.
- o [IVR09]/[IVR10] Clarified <variable> attribute definitions and added 'gender' attribute.
- o [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.
- o [IVR18] Added 'inactive' value for direction attribute of <stream>.
- o [IVR19] Clarified case of <dialogstart> on connection/conference

with multiple streams of the same type: recommended to be set explicitly with <stream>s.

- o [IVR02] Clarified that multiple dialogs may started simultaneously on the same connection or conference.
- o [IVR20] Added maximum duration (10 minutes) for a dialog to remain in the PREPARED state.
- o Added <params> in <dialogstart> and <dialogexit> for input/output in other dialog types
- o [IVR22] Added fetchtimeout parameter to dialogprepare, dialogstart, media and grammar elements.
- o [IVR04] Added dialogexit status to indicate the connection or conference has been terminated. Added others status errors.
- o [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.
- o [IVR11] Added runtime controls for speed, goto start/end and external controls.
- o Clarified that recordings can be uploaded to dest during or after the recording operation.

- o <record>/<collect>: clarified timer handling - timeout refers to waiting time for collect/record to begin.
- o Clarified behavior of immediate attribute on <dialogterminate>.
- o clarified dialogid lifecycle: dialogids can be re-cycled.
- o Clarified error handling.
- o Editorial tidy up of sections.
- o dialogid attribute on <response> is now mandatory.

- o Clarified that the duration specified in finalsilence attribute of <record> is not part of the final recording.
- o 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.

- o 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.
- o Changed conf-id to conferenceid and connection-id to connectionid.
- o Clarification of the state model for dialogs
- o <dialogprepare>: modified definition of src attribute to allow reference to external dialog documents; added (MIME) type attribute; removed <data> child element.
- o <dialogstart>: modified definition of src attribute to allow reference to external dialog documents; added (MIME) type attribute; removed <data> child element;
- o <dialogterminate>: modified so that a dialogexit event is always sent for active dialogs (i.e. the dialogexit event is a terminating notification)
- o <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

- o <data> element is removed.
- o <subscribe> element removed.
- o 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.

- o <dialogprepare> and <dialogstart> are extended with <dialog> child element.
- o <event> is extended with a <dialogexit> element which contains status and reported information (replacement for output parameters in template dialogs)
- o 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.
- o 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 be paused/resumed. Added 'volumeinterval', 'volumekey' and 'voldnkey' to add prompt volume to be increased/decreased. Moved 'bargein' attribute to prompt.
- o 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.

- o Mainly an alignment/evaluation exercise with requirements produced by MEDIACTRL IVR design team.
- o playannouncement parameters from Table 7 of '04' version are now reflected in text – schema to be updated.
- o Added VCR commands based on MSCML.

The following are the major changes between the -04 of the draft and the -03 version.

- o None.

The following are the major changes between the -03 of the draft and the -02 version.

- o 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](#)].
- o added "basicivr:" URI schema to IANA considerations
- o Added mimetype, vadinitial and vadfinal parameters to 'promptandrecord' dialog type
- o updated references

The following are the major changes between the -02 of the draft and the -01 version.

- o added version 1.0 to package name
- o separate section for element definitions
- o dialogterminate treated as request rather than notification
- o simplified responses: single element <response>
- o removed response elements: <dialogprepared>, <dialogstarted>, <errordialognotprepared>, <errordialognotstarted>
- o simplified event notifications to single <event> element carried in a REPORT
- o <dialogexit> element replaced with <event name="dialogexit">
- o removed <dialoguser> element
- o added <stream> element as child of <dialogstart>
- o removed 'type' attribute from <dialogprepare> and <dialogstart>
- o added dialogid attribute to <dialogprepare> and <dialogstart>
- o removed template "Sample Implementation" section
- o renamed <namelist> to <data>

Internet-Draft

IVR Control Package

October 2008

- o 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.

- o Removed requirement for VoiceXML dialog support
- o Added requirement for template dialog support

Internet-Draft

IVR Control Package

October 2008

[10.](#) Contributors

Asher Shiratzky from Radvision 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 and Mary Barnes who provided valuable feedback, input and text to this document.

[11.](#) Acknowledgments

The authors would like to thank Adnan Saleem of Radisys, Gene Shtirmer of Intel, Dave Burke of Google, Dan York of Voxeo and Steve Buko of Dialogic for expert reviews of this work.

[12.](#) References

[12.1.](#) 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-04](#) (work in progress), August 2008.
- [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., Jansen, J., Zucker, D., Layaida, N., Michel, T., 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.

McGlashan, et al. Expires April 10, 2009 [Page 114]

Internet-Draft IVR Control Package October 2008

[12.2](#). 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., and R. Even, "Conference Information Data Model for Centralized Conferencing (XCON)", [draft-ietf-xcon-common-data-model-11](#) (work in progress), June 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-07](#) (work in progress), August 2008.
- [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.
- [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.
- [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, R.J., 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.

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
Avaya
Building 3
Wern Fawr Lane
St Mellons
Cardiff, South Wales CF3 5EA

Email: cboulton@avaya.com

Internet-Draft

IVR Control Package

October 2008

Full Copyright Statement

Copyright (C) The IETF Trust (2008).

This document is subject to the rights, licenses and restrictions contained in [BCP 78](#), and except as set forth therein, the authors retain all their rights.

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE IETF TRUST AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Intellectual Property

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in [BCP 78](#) and [BCP 79](#).

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at <http://www.ietf.org/ipr>.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

