Issue52

Title Requests for Clarification
Priority bug   **Status** chatting
Superseder   **Nosy List** eburger, oran, sarvi
Assigned To sarvi   Topics

Created on **2006-02-23.01:14:04** by **sarvi**, last changed **2006-02-23.18:42:09** by **sarvi**.

Messages

msg88 (view)   **Author: sarvi**   **Date: 2006-02-23.01:14:04**

```
Will update the next next draft with the suggestd clarifications.
Adding this thread to the issue tracker list.

See Inline.

    -----Original Message-----
    From: speechsc-bounces@ietf.org
    [mailto:speechsc-bounces@ietf.org] On Behalf Of Andrew Wahbe
    Sent: Wednesday, December 21, 2005 2:03 PM
    To: IETF SPEECHSC (E-mail)
    Subject: [speechsc] Requests for Clarification

    The VoiceXML Forum MRCP Liaison Committee is currently
    evaluating the latest MRCP v2 draft to (a) evaluate the
    compatibility between MRCP v2 and VoiceXML and (b)
    generate test assertions for MRCP v2 based VoiceXML
    browsers and MRCP v2 based media resources. We are
    currently examining the Speech Synthesis portion of the
    specification and have raised issues with the
    specification in prior emails to the SpeechSC list (See
    http://www.ietf.org/mail-archive/web/speechsc/current/msg01605.html
    http://www.ietf.org/mail-archive/web/speechsc/current/msg01606.html
    and
    http://www.ietf.org/mail-archive/web/speechsc/current/msg01
    607.html ).

    These issues (and the responses to them) have been
    discussed by the MRCP Liaison Committee and we would like
    to make the following requests and
    suggestions:

    1) The relationship between the Fetch Hint header and the
    Audio Fetch Hint header should be clarified. More
    specifically, it should be stated that, when specified,
    the Audio Fetch Hint header overrides the Fetch Hint
    header for audio files only.

Sarvi> Sounds good. Will clarify as suggested above.

    2) It should be clarified that  SPEAK completion code 003
    "uri-failure"
    only applies to  fetched SSML files and that  failure to fetch (or
```

process) an audio file will not result in aborting the
SPEAK request.
This does mean, however, that there is no way to
communicate the failure to fetch (or process) the audio
file to the MRCP client. While SSML requires that the
processor "notify the hosting environment" when such a
failure occurs, the members of the committee agree that
logging this event at the MRCP server is sufficient. It
may be advisable for the MRCP specification to suggest
that these events should be logged in some way.
We would also like to suggest that future versions of MRCP
consider adding an event (e.g. "Audio-Exception") to
notify the MRCP client that such a failure has occurred
without aborting the SPEAK request.

Sarvi> Sounds good. Will clarify as above.

3) The definition of the Basic Synthesizer resources is a
bit vague and should be clarified. Its not entirely clear
from the description in the spec how it is supposed to
work. The general consensus in the Committee is that this
resource can be used for audio only prompts. It is
supposed to accept a subset of SSML that only includes
<speak><audio><say-as> and <mark>. What isn't clear, is
how <say-as> is supposed to work in this case and if text
strings are acceptable (you would think no if it wasn't
for <say-as> being allowed). It may also be reasonable to
make <mark> optional; a VoiceXML 2.0 browser certainly
wouldn't need it anyway. We find that clarifications are
needed in order to make any assertions on how a VoiceXML
browser would use a basicsynth resource in an implementation.

Sarvi> With the basic synthesizer it is understood that the rendering capability
would be limited. But I bleieve there is still use for the <sayas> tag. Things
like $200, 1/2 etc can be easily rendered by a basic synthesizer. And as
explained in the SSML specification <sayas> helps tell the processor how to
render things like 1/2 or $200 etc.
So I believe that <sayas> would be usefull within the basic synthesizer, though
in a more limited sense.

Sarvi> Considering <mark> can be very useful to UI implementation and and pretty
simple to implement, I believe we should leave <mark> as is defined today.

A final issue worth noting is that the maxage and maxstale
cache control headers are global in MRCP while VoiceXML
breaks this down by resource type (e.g. audiomaxage,
audiomaxstale, grammarmaxage, grammarmaxstale, etc.). This
may be acceptable because the context of each request
should govern the type of file to which these headers
apply. i.e. in a SPEAK request the control audio file
fetches and in RECOGNIZE requests they control grammar
file fetches. As we continue to evaluate the spec we will
keep our eyes open for scenarios where this does not hold.
Thus, we are not requesting any changes related to this
issue at this time.

Sarvi> Sounds good.

Related to the above issue is the fact that the <audio>
tag in VoiceXML extends the attributes defined in SSML by
adding maxstale, maxage, fetchtimeout, and fetchhint (it

```
     also adds expr but that "evaluates away"
     to src). These fetch-related headers override their
     associated properties. Unfortunately, since MRCP is based
     on SSML, these attributes cannot be included in an MRCP
     request; instead, the associated headers would need to be
     set to control this behavior. This obviously introduces a
     problem if a request contains two <audio> tags that had
     these attributes set differently in the original VoiceXML document.

     It would seem that one way to address this problem is to
     break apart an SSML prompt so that each audio file is sent
     in its own request.
     Unfortunately, Issue (2) from above prevents this solution
     from working.
     Consider a prompt with alternate audio files such as:
     <audio maxstale="A" src="A.wav"><audio maxstale="B"
     src="B.wav"/></audio> where maxstale values A and B are
     not the same. These files can't be sent as part of the
     same request due to their maxstale values. However, if
     they are sent as part of separate requests, the client
     would need to know if A.wav could not be fetched in order
     to decide if it should request for B.wav to be played. But
     as discussed above, there is no way for the client to know
     this. The MRCP Liaison Committee believes that the best
     way for this to be addressed is to make a request to the
     W3C Voice Browser Working Group to add these attributes to
     the audio tag in SSML.
     Again, we are not requesting any changes to MRCP related
     to this issue.

Sarvi> Sounds good.

Thanks,
Sarvi


     Regards,

     Andrew Wahbe
     VoiceXML Forum MRCP Liaison Committee
```

History

Date	User	Action	Args
2006-02-23 18:42:09	sarvi	set	nosy: + oran, eburger
2006-02-23 01:14:04	sarvi	create	