Technical Specification Group Services and System Aspects ***TSGS#18(02)0694***
Meeting #18, New Orleans, USA, 9 - 12 December 2002

Source: **TSG-SA WG4**

Title: CRs to TS 26.234 - Corrections (Release 4 and Release 5)

Document for: **Approval**

Agenda Item: **7.4.3**

The following CRs, agreed at the TSG-SA WG4 meeting #23/#24, are presented to TSG SA #18 for approval.

Spec	CR	Rev	Phase	Subject	Cat	Vers	WG	Meeting	S4 doc
26.234	040		Rel-5	Code points for H.263	F	5.2.0	S4	TSG-SA WG4#23	S4-020539
26.234	041	2	Rel-5	File format 3GP based on ISO and not MP4	F	5.2.0	S4	TSG-SA WG4#23	S4-020589
26.234	042	2	Rel-4	Addition regarding IPv6 support in SDP	F	4.4.0	S4	TSG-SA WG4#23	S4-020626
26.234	039	2	Rel-5	Addition regarding IPv6 support in SDP	A	5.2.0	S4	TSG-SA WG4#23	S4-020627
26.234	045	1	Rel-5	Client usage of bandwidth parameter at the media level in SDP	F	5.2.0	S4	TSG-SA WG4#23	S4-020588
26.234	043	2	Rel-4	SMIL authoring instructions	F	4.4.0	S4	TSG-SA WG4#24	S4-020723
26.234	044	1	Rel-5	SMIL authoring instructions	A	5.2.0	S4	TSG-SA WG4#24	S4-020634
26.234	046	1	Rel-4	SMIL Language Profile	F	4.4.0	S4	TSG-SA WG4#24	S4-020699
26.234	047	1	Rel-5	SMIL Language Profile	A	5.2.0	S4	TSG-SA WG4#24	S4-020700
26.234	050	1	Rel-5	Usage of Multiple Media Sample Entries in Media Tracks of 3GP files	F	5.2.0	S4	TSG-SA WG4#24	S4-020724
26.234	051	1	Rel-5	Progressive download of 3GP files	F	5.2.0	S4	TSG-SA WG4#24	S4-020712

3GPP TSG-SA4 Meeting #23 ***Tdoc ⌘S4-020627***
Montreal, Canada, 30 September–4 October 2002

CR-Form-v7

CHANGE REQUEST

⌘ **26.234** **CR** **039** ⌘ **rev** **2** ⌘ Current version: **5.2.0** ⌘

*For **HELP** on using this form, see bottom of this page or look at the pop-up text over the ⌘ symbols.*

Proposed change affects: UICC apps⌘ [ ] ME [X] Radio Access Network [ ] Core Network [ ]

Title:	⌘	Correction regarding undefined IPv6 support in SDP
Source:	⌘	TSG SA WG4
Work item code:	⌘	PSS-E
Date:	⌘	10/12/2002
Category:	⌘	**A**
Release:	⌘	Rel-5

Use one of the following categories:
F *(correction)*
A *(corresponds to a correction in an earlier release)*
B *(addition of feature),*
C *(functional modification of feature)*
D *(editorial modification)*
Detailed explanations of the above categories can be found in 3GPP TR 21.900.

Use one of the following releases:

2	*(GSM Phase 2)*
R96	*(Release 1996)*
R97	*(Release 1997)*
R98	*(Release 1998)*
R99	*(Release 1999)*
Rel-4	*(Release 4)*
Rel-5	*(Release 5)*
Rel-6	*(Release 6)*

Reason for change:	⌘	In June 2002, RFC 3266, "Support for IPv6 in Session Description Protocol" has been published by IETF. This RFC defines how IPv6 addresses shall be handled in an SDP description. To make sure that PSS fully follows IETF standards when used over IPv6 operated network, IPv6 support in SDP must follow the same RFC.
Summary of change:	⌘	- A text added to the SDP section, related to the IPv6 support - A reference to RFC 3266 is included in the References section.
Consequences if not approved:	⌘	IPv6 support in PSS-E SDP descriptions remains undefined. 3GPP PSS-E implementations will not by default follow the standard way defined by IETF – potentially causing interoperability problems.

Clauses affected:	⌘	2 and 5.3.3.1

Other specs affected: ⌘	Y	N	
		N	Other core specifications
		N	Test specifications
		N	O&M Specifications

⌘

Other comments: ⌘

How to create CRs using this form:
Comprehensive information and tips about how to create CRs can be found at http://www.3gpp.org/specs/CR.htm.
Below is a brief summary:

1) Fill out the above form. The symbols above marked ⌘ contain pop-up help information about the field that they are closest to.

2) Obtain the latest version for the release of the specification to which the change is proposed. Use the MS Word "revision marks" feature (also known as "track changes") when making the changes. All 3GPP specifications can be downloaded from the 3GPP server under ftp://ftp.3gpp.org/specs/ For the latest version, look for the directory name with the latest date e.g. 2001-03 contains the specifications resulting from the March 2001 TSG meetings.

3) With "track changes" disabled, paste the entire CR form (use CTRL-A to select it) into the specification just in front of the clause containing the first piece of changed text. Delete those parts of the specification which are not relevant to the change request.

2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

[1] 3GPP TS 22.233: "Transparent End-to-End Packet-switched Streaming Service; Service aspects; Stage 1 "

[2] 3GPP TS 26.233: " Transparent end-to-end packet switched streaming service (PSS); General description".

[3] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications".

[4] IETF RFC 1738: "Uniform Resource Locators (URL)", Berners-Lee, Masinter & McCahill, December 1994.

[5] IETF RFC 2326: "Real Time Streaming Protocol (RTSP)", Schulzrinne H., Rao A. and Lanphier R., April 1998.

[6] IETF RFC 2327: "SDP: Session Description Protocol", Handley M. and Jacobson V., April 1998.

[7] IETF STD 0006: "User Datagram Protocol", Postel J., August 1980.

[8] IETF STD 0007: "Transmission Control Protocol", Postel J., September 1981.

[9] IETF RFC 1889: "RTP: A Transport Protocol for Real-Time Applications", Schulzrinne H. et al., January 1996.

[10] IETF RFC 1890: "RTP Profile for Audio and Video Conferences with Minimal Control", Schulzrinne H. et al., January 1996.

[11] IETF RFC 3267: " RTP payload format and file storage format for the Adaptive Multi-Rate (AMR) Adaptive Multi-Rate Wideband (AMR-WB) audio codecs ", March 2002.

[12] (void)

[13] IETF RFC 3016: "RTP Payload Format for MPEG-4 Audio/Visual Streams", Kikuchi Y. et al., November 2000.

[14] IETF RFC 2429: "RTP Payload Format for the 1998 Version of ITU-T Rec. H.263 Video (H.263+)", Bormann C. et al., October 1998.

[15] IETF RFC 2046: "Multipurpose Internet Mail Extensions (MIME) Part Two: Media Types", N. Freed, N. Borenstein, November 1996.

[16] IETF RFC 3236: " The 'application/xhtml+xml' Media Type ", Baker M. and Stark P., January 2002.

[17] IETF RFC 2616: "Hypertext Transfer Protocol - HTTP/1.1", Fielding R. et al., June 1999.

[18] 3GPP TS 26.071: "Mandatory Speech Codec speech processing functions; AMR Speech Codec; General description".

[19] 3GPP TS 26.101: "Mandatory Speech Codec speech processing functions; AMR Speech Codec; Frame Structure".

[20] 3GPP TS 26.171: "AMR speech codec, wideband; General description".

[21] ISO/IEC 14496-3:2001, "Information technology -- Coding of audio-visual objects -- Part 3: Audio".

[22] ITU-T Recommendation H.263: "Video coding for low bit rate communication".

[23] ITU-T Recommendation H.263: "Annex X, Profiles and levels definition".

[24] ISO/IEC 14496-2:2001, "Information technology -- Coding of audio-visual objects -- Part 2: Visual".

[25] ISO/IEC 14496-2:2001/Amd 2:2002, "Streaming video profile".

[26] ITU-T Recommendation T.81 (1991) | ISO/IEC 10918-1 (1992): "Information technology - Digital compression and coding of continuous-tone still images - Requirements and guidelines.

[27] "JPEG File Interchange Format", Version 1.02, September 1, 1992.

[28] W3C Recommendation: "XHTML Basic", http://www.w3.org/TR/2000/REC-xhtml-basic-20001219, December 2000

[29] ISO/IEC 10646-1 (2000): "Information technology - Universal Multiple-Octet Coded Character Set (UCS) - Part 1: Architecture and Basic Multilingual Plane".

[30] The Unicode Consortium: "The Unicode Standard", Version 3.0 Reading, MA, Addison-Wesley Developers Press, 2000, ISBN 0-201-61633-5.

[31] W3C Recommendation: "Synchronized Multimedia Integration Language (SMIL 2.0)", http://www.w3.org/TR/2001/REC-smil20-20010807/, August 2001.

[32] CompuServe Incorporated: "GIF Graphics Interchange Format: A Standard defining a mechanism for the storage and transmission of raster-based graphics information", Columbus, OH, USA, 1987.

[33] CompuServe Incorporated: "Graphics Interchange Format: Version 89a", Columbus, OH, USA, 1990.

[34] ISO/IEC 14496-1 (2001): "Information technology - Coding of audio-visual objects - Part 1: Systems".

[35] 3GPP TS 26.140: "Multimedia Messaging Service (MMS); Media formats and codecs".

[36] ISO/IEC 15444-1 (2000): "Information technology - JPEG 2000 image coding system: Core coding system; Annex I: The JPEG 2000 file format".

[37] 3GPP TS 26.201: "AMR Wideband Speech Codec; Frame Structure"

[38] IETF RFC 2083: "PNG (Portable Networks Graphics) Specification version 1.0 ", T. Boutell, et. al., March 1997.

[39] W3C Working Draft Recommendation: "CC/PP structure and vocabularies", http://www.w3.org/Mobile/CCPP/Group/Drafts/WD-CCPP-struct-vocab-20010620/, June 2001.

[40] WAP UAProf Specification, http://www1.wapforum.org/tech/documents/WAP-248-UAProf-20011020-a.pdf , October 2001.

[41] W3C Candidate Recommendation: "Resource Description Framework (RDF) Schema Specification 1.0", http://www.w3.org/TR/2000/CR-rdf-schema-20000327, March 2000.

[42] W3C Working Draft Recommendation: "Scalable Vector Graphics (SVG) 1.1 Specification", http://www.w3.org/TR/SVG11 , February 2002.

[43] W3C Working Draft Recommendation: "SVG Mobile Specification", http://www.w3.org/TR/SVGMobile/, February 2002

[44] Scalable Polyphony MIDI Specification Version 1.0, RP-34, MIDI Manufacturers Association, Los Angeles, CA, February 2002.

[45] Scalable Polyphony MIDI Device 5-to-24 Note Profile for 3GPP Version 1.0, RP-35, MIDI Manufacturers Association, Los Angeles, CA, February 2002.

[46] "Standard MIDI Files 1.0", RP-001, in "The Complete MIDI 1.0 Detailed Specification, Document Version 96.1 " The MIDI Manufacturers Association, Los Angeles, CA, USA, February 1996.

[47] WAP Forum Specification: “XHTML Mobile Profile”, http://www1.wapforum.org/tech/terms.asp?doc=WAP-277-XHTMLMP-20011029-a.pdf, October 2001.

[48] "Unicode Standard Annex #13: Unicode Newline Guidelines", by Mark Davis. An integral part of The Unicode Standard, Version 3.1.

[49] IETF RFC 3266: "Support For IPv6 in Session Description Protocol (SDP)", Olson S., Camarillo G. and Roach A. B., June 2002

5.3.3 SDP

5.3.3.1 General

RTSP requires a presentation description. SDP shall be used as the format of the presentation description for both PSS clients and servers. PSS servers shall provide and clients interpret the SDP syntax according to the SDP specification [6] and appendix C of [5]. The SDP delivered to the PSS client shall declare the media types to be used in the session using a codec specific MIME media type for each media. MIME media types to be used in the SDP file are described in clause 5.4 of the present document.

The SDP [6] specification requires certain fields to always be included in an SDP file. Apart from this a PSS server shall always include the following fields in the SDP:

- "a=control:" according to clauses C.1.1, C.2 and C.3 in [5];
- "a=range:" according to clause C.1.5 in [5];
- "a=rtpmap:" according to clause 6 in [6];
- "a=fmtp:" according to clause 6 in [6].

The bandwidth field in SDP is needed by the client in order to properly set up QoS parameters. Therefore, a PSS server shall include the “b=AS:” field at the media level for each media stream in SDP, and a PSS client shall interpret this field. When a client receives SDP, it should ignore the session level “b=AS:” parameter (if present), and instead calculate session bandwidth from the media level bandwidth values of the relevant streams. Note that for RTP based applications , ‘b=AS:’ gives the RTP "session bandwidth" (including UDP/IP overhead) as defined in section 6.2 of [9].

IPv6 addresses in SDP descriptions shall be supported according to RFC 3266[49].

NOTE: The SDP parsers and/or interpreters shall be able to accept NULL values in the 'c=' field (e.g. 0.0.0.0 in IPv4 case). This may happen when the media content does not have a fixed destination address. For more details, see Section C.1.7 of [5] and Section 6 of [6].

3GPP TSG-SA4 Meeting #23 **Tdoc ⌘S4-020539**
Montreal, Canada, 30 September–4 October 2002

CR-Form-v7

CHANGE REQUEST

⌘ **26.234** **CR** **040** ⌘ **rev** **-** ⌘ Current version: **5.2.0** ⌘

*For **HELP** on using this form, see bottom of this page or look at the pop-up text over the ⌘ symbols.*

Proposed change affects: UICC apps⌘ ☐ ME ☒ Radio Access Network ☐ Core Network ☐

Title:	⌘	Code points for H.263
Source:	⌘	TSG SA WG4
Work item code:	⌘	PSS-E
Date:	⌘	10/12/2002
Category:	⌘	**F**
Release:	⌘	Rel-5

Use one of the following categories:
F *(correction)*
A *(corresponds to a correction in an earlier release)*
B *(addition of feature),*
C *(functional modification of feature)*
D *(editorial modification)*
Detailed explanations of the above categories can be found in 3GPP TR 21.900.

Use one of the following releases:
2 *(GSM Phase 2)*
R96 *(Release 1996)*
R97 *(Release 1997)*
R98 *(Release 1998)*
R99 *(Release 1999)*
Rel-4 *(Release 4)*
Rel-5 *(Release 5)*
Rel-6 *(Release 6)*

Reason for change:	⌘	H.263, profile 0 level 10, is a mandatory video decoder for PSS. However, a corresponding bitstream can also be decoded by an MPEG-4 visual decoder (optionally supported) using so-called “short header”. To ensure full support for H.263 it is therefore necessary to signal H.263 as H.263 and not as MPEG-4 visual (in RTP transport and file storage).
Summary of change:	⌘	A recommendation to transport and store H.263 profile 0 bitstreams as H.263, rather than MPEG-4 visual, has been added.
Consequences if not approved:	⌘	If the recommendation is not followed, there is no guarantee that a PSS terminal (supporting only H.263 profile 0) can detect all H.263 profile 0 bitstreams (as such a bitstream could be embedded in an MPEG-4 visual short header stream).

Clauses affected:	⌘	7.4
Other specs affected:	⌘	Y N ☐ N Other core specifications ☐ N Test specifications ☐ N O&M Specifications
	⌘	
Other comments:	⌘	

How to create CRs using this form:
Comprehensive information and tips about how to create CRs can be found at http://www.3gpp.org/specs/CR.htm.
Below is a brief summary:

1) Fill out the above form. The symbols above marked ⌘ contain pop-up help information about the field that they are closest to.

2) Obtain the latest version for the release of the specification to which the change is proposed. Use the MS Word "revision marks" feature (also known as "track changes") when making the changes. All 3GPP specifications can be downloaded from the 3GPP server under ftp://ftp.3gpp.org/specs/ For the latest version, look for the directory name with the latest date e.g. 2001-03 contains the specifications resulting from the March 2001 TSG meetings.

3) With "track changes" disabled, paste the entire CR form (use CTRL-A to select it) into the specification just in front of the clause containing the first piece of changed text. Delete those parts of the specification which are not relevant to the change request.

7.4 Video

ITU-T Recommendation H.263 [22] profile 0 level 10 shall be supported. This is the mandatory video decoder for the PSS. In addition, PSS should support:

- H.263 [23] Profile 3 Level 10 decoder;
- MPEG-4 Visual Simple Profile Level 0 decoder, [24] and [25].

These two video decoders are optional to implement.

An optional video buffer model is given in Annex G of the present document.

NOTE: ~~ITU-T Recommendation H.263 [22] baseline has been mandated to ensure that video-enabled PSS support a minimum baseline video capability and interoperability can be guaranteed (an H.263 [22] baseline bitstream can be decoded by both H.263 [22] and MPEG-4 decoders). It also provides a simple upgrade path for mandating more advanced decoders in the future (from both the ITU-T and ISO MPEG).~~ITU-T Recommendation H.263 profile 0 has been mandated to ensure that video-enabled PSS supports a minimum baseline video capability. Both H.263 and MPEG-4 visual decoders can decode an H.263 profile 0 bitstream. It is strongly recommended, though, that an H.263 profile 0 bitstream is transported and stored as H.263 and not as MPEG-4 visual (short header), as MPEG-4 visual is not mandated by PSS.

3GPP TSG-SA4 Meeting #23
Montreal, Canada, 30 September–4 October 2002

Tdoc ⌘S4-020589

CR-Form-v7

CHANGE REQUEST

⌘ **26.234** **CR** **041** ⌘ **rev** **2** ⌘ Current version: **5.2.0** ⌘

*For **HELP** on using this form, see bottom of this page or look at the pop-up text over the ⌘ symbols.*

Proposed change affects: UICC apps⌘ [ ] ME [X] Radio Access Network [ ] Core Network [ ]

Title:	⌘	3GPP file format based on ISO and not MP4
Source:	⌘	TSG SA WG4
Work item code:	⌘	PSS-E
Date:	⌘	10/12/2002
Category:	⌘	**F**
Release:	⌘	Rel-5

Use one of the following categories:
F *(correction)*
A *(corresponds to a correction in an earlier release)*
B *(addition of feature),*
C *(functional modification of feature)*
D *(editorial modification)*
Detailed explanations of the above categories can be found in 3GPP TR 21.900.

Use one of the following releases:

2	*(GSM Phase 2)*
R96	*(Release 1996)*
R97	*(Release 1997)*
R98	*(Release 1998)*
R99	*(Release 1999)*
Rel-4	*(Release 4)*
Rel-5	*(Release 5)*
Rel-6	*(Release 6)*

Reason for change:	⌘	MPEG has recently reorganized the file format specification upon which the 3GPP file format (3GP) is based. In particular, the file format structure (ISO base media file format) has been separated from the conformance of the MPEG-4 file format (MP4). This reorganization has been made in order to accommodate 3GPP and we should now refer specifically to the ISO file format for the structure of 3GP files, while defining the conformance statement for 3GP files in TS 26.234.
Summary of change:	⌘	The reference to the file format structure has been corrected and the 3GPP file format is now consistently referred to as 3GP (and not MP4). In addition there are some editorial changes in order to conform to the ISO specification as well as the separation into ISO and MP4 specifications. The guidelines for using the old version of MP4 have been replaced by a conformance definition of 3GP.
Consequences if not approved:	⌘	If this CR is not approved, we will refer to an outdated version of the MPEG-4 file format specification addressing both file format structure and MP4 conformance. This has already generated confusion, since it makes 3GP appear as a derivate of MP4, which is not the case. 3GP is a derivate of the ISO base media file format and not of MP4.

Clauses affected:	⌘	2, 3.2, 7.9, 9, D			
Other specs affected:	⌘	Y: Other core specifications	⌘	CR TS 26.140 003	
		N: Test specifications			
		N: O&M Specifications			
Other comments:	⌘				

How to create CRs using this form:

Comprehensive information and tips about how to create CRs can be found at http://www.3gpp.org/specs/CR.htm. Below is a brief summary:

1) Fill out the above form. The symbols above marked ⌘ contain pop-up help information about the field that they are closest to.

2) Obtain the latest version for the release of the specification to which the change is proposed. Use the MS Word "revision marks" feature (also known as "track changes") when making the changes. All 3GPP specifications can be downloaded from the 3GPP server under ftp://ftp.3gpp.org/specs/ For the latest version, look for the directory name with the latest date e.g. 2001-03 contains the specifications resulting from the March 2001 TSG meetings.

3) With "track changes" disabled, paste the entire CR form (use CTRL-A to select it) into the specification just in front of the clause containing the first piece of changed text. Delete those parts of the specification which are not relevant to the change request.

2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

[1] 3GPP TS 22.233: "Transparent End-to-End Packet-switched Streaming Service; Service aspects; Stage 1 "

[2] 3GPP TS 26.233: " Transparent end-to-end packet switched streaming service (PSS); General description".

[3] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications".

[4] IETF RFC 1738: "Uniform Resource Locators (URL)", Berners-Lee, Masinter & McCahill, December 1994.

[5] IETF RFC 2326: "Real Time Streaming Protocol (RTSP)", Schulzrinne H., Rao A. and Lanphier R., April 1998.

[6] IETF RFC 2327: "SDP: Session Description Protocol", Handley M. and Jacobson V., April 1998.

[7] IETF STD 0006: "User Datagram Protocol", Postel J., August 1980.

[8] IETF STD 0007: "Transmission Control Protocol", Postel J., September 1981.

[9] IETF RFC 1889: "RTP: A Transport Protocol for Real-Time Applications", Schulzrinne H. et al., January 1996.

[10] IETF RFC 1890: "RTP Profile for Audio and Video Conferences with Minimal Control", Schulzrinne H. et al., January 1996.

[11] IETF RFC 3267: " RTP payload format and file storage format for the Adaptive Multi-Rate (AMR) Adaptive Multi-Rate Wideband (AMR-WB) audio codecs ", March 2002.

[12] (void)

[13] IETF RFC 3016: "RTP Payload Format for MPEG-4 Audio/Visual Streams", Kikuchi Y. et al., November 2000.

[14] IETF RFC 2429: "RTP Payload Format for the 1998 Version of ITU-T Rec. H.263 Video (H.263+)", Bormann C. et al., October 1998.

[15] IETF RFC 2046: "Multipurpose Internet Mail Extensions (MIME) Part Two: Media Types", N. Freed, N. Borenstein, November 1996.

[16] IETF RFC 3236: " The 'application/xhtml+xml' Media Type ", Baker M. and Stark P., January 2002.

[17] IETF RFC 2616: "Hypertext Transfer Protocol - HTTP/1.1", Fielding R. et al., June 1999.

[18] 3GPP TS 26.071: "Mandatory Speech Codec speech processing functions; AMR Speech Codec; General description".

[19] 3GPP TS 26.101: "Mandatory Speech Codec speech processing functions; AMR Speech Codec; Frame Structure".

[20] 3GPP TS 26.171: "AMR speech codec, wideband; General description".

[21] ISO/IEC 14496-3:2001, "Information technology -- Coding of audio-visual objects -- Part 3: Audio".

[22] ITU-T Recommendation H.263: "Video coding for low bit rate communication".

[23] ITU-T Recommendation H.263: "Annex X, Profiles and levels definition".

[24] ISO/IEC 14496-2:2001, "Information technology -- Coding of audio-visual objects -- Part 2: Visual".

[25] ISO/IEC 14496-2:2001/Amd 2:2002, "Streaming video profile".

[26] ITU-T Recommendation T.81 (1991) | ISO/IEC 10918-1 (1992): "Information technology - Digital compression and coding of continuous-tone still images - Requirements and guidelines.

[27] "JPEG File Interchange Format", Version 1.02, September 1, 1992.

[28] W3C Recommendation: "XHTML Basic", http://www.w3.org/TR/2000/REC-xhtml-basic-20001219, December 2000

[29] ISO/IEC 10646-1 (2000): "Information technology - Universal Multiple-Octet Coded Character Set (UCS) - Part 1: Architecture and Basic Multilingual Plane".

[30] The Unicode Consortium: "The Unicode Standard", Version 3.0 Reading, MA, Addison-Wesley Developers Press, 2000, ISBN 0-201-61633-5.

[31] W3C Recommendation: "Synchronized Multimedia Integration Language (SMIL 2.0)", http://www.w3.org/TR/2001/REC-smil20-20010807/, August 2001.

[32] CompuServe Incorporated: "GIF Graphics Interchange Format: A Standard defining a mechanism for the storage and transmission of raster-based graphics information", Columbus, OH, USA, 1987.

[33] CompuServe Incorporated: "Graphics Interchange Format: Version 89a", Columbus, OH, USA, 1990.

[34] ~~ISO/IEC 14496-1 (2001): "Information technology - Coding of audio-visual objects - Part 1: Systems".~~(void)

[35] 3GPP TS 26.140: "Multimedia Messaging Service (MMS); Media formats and codecs".

[36] ~~ISO/IEC 15444-1 (2000): "Information technology - JPEG 2000 image coding system: Core coding system; Annex I: The JPEG 2000 file format".~~(void)

[37] 3GPP TS 26.201: "AMR Wideband Speech Codec; Frame Structure"

[38] IETF RFC 2083: "PNG (Portable Networks Graphics) Specification version 1.0 ", T. Boutell, et. al., March 1997.

[39] W3C Working Draft Recommendation: "CC/PP structure and vocabularies", http://www.w3.org/Mobile/CCPP/Group/Drafts/WD-CCPP-struct-vocab-20010620/, June 2001.

[40] WAP UAProf Specification, http://www1.wapforum.org/tech/documents/WAP-248-UAProf-20011020-a.pdf , October 2001.

[41] W3C Candidate Recommendation: "Resource Description Framework (RDF) Schema Specification 1.0", http://www.w3.org/TR/2000/CR-rdf-schema-20000327, March 2000.

[42] W3C Working Draft Recommendation: "Scalable Vector Graphics (SVG) 1.1 Specification", http://www.w3.org/TR/SVG11 , February 2002.

[43] W3C Working Draft Recommendation: "SVG Mobile Specification", http://www.w3.org/TR/SVGMobile/, February 2002

[44] Scalable Polyphony MIDI Specification Version 1.0, RP-34, MIDI Manufacturers Association, Los Angeles, CA, February 2002.

[45] Scalable Polyphony MIDI Device 5-to-24 Note Profile for 3GPP Version 1.0, RP-35, MIDI Manufacturers Association, Los Angeles, CA, February 2002.

[46] "Standard MIDI Files 1.0", RP-001, in "The Complete MIDI 1.0 Detailed Specification, Document Version 96.1 " The MIDI Manufacturers Association, Los Angeles, CA, USA, February 1996.

[47] WAP Forum Specification: “XHTML Mobile Profile”, http://www1.wapforum.org/tech/terms.asp?doc=WAP-277-XHTMLMP-20011029-a.pdf, October 2001.

[48] "Unicode Standard Annex #13: Unicode Newline Guidelines", by Mark Davis. An integral part of The Unicode Standard, Version 3.1.

[50] ISO/IEC 14496-1:2001/Amd 5, "Information technology - Coding of audio-visual objects - Part 1: Systems - Amendment 5: ISO base media file format".

[51] ISO/IEC 14496-1:2001/Amd 6, "Information technology - Coding of audio-visual objects - Part 1: Systems - Amendment 6: MP4, the MPEG-4 file format".

---------- <text left out> ----------

3.2 Abbreviations

For the purposes of the present document, the abbreviations given in 3GPP TR 21.905 [3] and the following apply.

3GP	3GPP file format
AAC	Advanced Audio Coding
BIFS	Binary Format for Scenes
CC/PP	Composite Capability / Preference Profiles
DCT	Discrete Cosine Transform
GIF	Graphics Interchange Format
HTML	Hyper Text Markup Language
ITU-T	International Telecommunications Union – Telecommunications
JFIF	JPEG File Interchange Format
MIDI	Musical Instrument Digital Interface
MIME	Multipurpose Internet Mail Extensions
MMS	Multimedia Messaging Service
MP4	MPEG-4 file format
PNG	Portable Networks Graphics
PSS	Packet-switched Streaming Service
QCIF	Quarter Common Intermediate Format
RDF	Resource Description Framework
RTCP	RTP Control Protocol
RTP	Real-time Transport Protocol
RTSP	Real-Time Streaming Protocol

SDP	Session Description Protocol
SMIL	Synchronised Multimedia Integration Language
SP-MIDI	Scalable Polyphony MIDI
SVG	Scalable Vector Graphics
UAProf	User Agent Profile
UCS-2	Universal Character Set (the two octet form)
UTF-8	Unicode Transformation Format (the 8-bit form)
UTF-16	Unicode Transformation Format (the 16-bit form)
W3C	WWW Consortium
WML	Wireless Markup Language
XHTML	eXtensible Hyper Text Markup Language
XML	eXtensible Markup Language

---------- <text left out> ----------

7.9 Timed text

If timed text is supported, PSS clients shall support Annex D, clause D.8a, of this specification. There is no support for RTP transport of timed text in this release; 3GP~~P (MP4)~~ files containing timed text may only be downloaded.

NOTE: When a PSS client supports timed text it needs to be able to receive and parse 3GP~~P (MP4)~~ files containing the text streams. This does not imply a requirement on PSS clients to be able to render other continuous media types contained in 3GP~~P (MP4)~~ files, e.g. AMR and H.263, if such media types are included in a presentation together with timed text. Audio and video are instead streamed to the client using RTSP/RTP (see clause 6.2).

---------- <text left out> ----------

9 3GPP file format (i~~I~~nterchange format for MMS)

9.1 General

The 3GPP~~MPEG-4~~ file format (3GP) is based on the ISO base media file format [50] and is defined in this specification. ~~[34]~~It is mandated in [35] to be used for continuous media along the entire delivery chain envisaged by the MMS, independent on whether the final delivery is done by streaming or download, thus enhancing interoperability.

In particular, the following stages are considered:

- upload from the originating terminal to the MMS proxy;
- file exchange between MMS servers;
- transfer of the media content to the receiving terminal, either by file download or by streaming. In the first case the self-contained file is transferred, whereas in the second case the content is extracted from the file and streamed according to open payload formats. In this case, no trace of the file format remains in the content that goes on the wire/in the air.

Additionally, the 3GPPMPEG-4 file format should be used for the storage in the servers and the "hint track" mechanism may be used for the preparation for streaming.

The cClause 9.2 of the present document gives the necessary requirements to follow for the 3GPPMPEG-4 file format used in MMS. These requirements will guarantee PSS to interwork with MMS as well as the 3GPPMPEG-4 file format to be used internally within the MMS system. For PSS servers not interworking with MMS there is no requirement to follow these guidelines.

9.2 3GPP fFile format conformanceguidelines

NOTE: The file format used in this specification for timed multimedia (such as video, associated audio and timed text) is structurally based on the MP4 file format as defined in [34]. However, since non-ISO codecs are used here, it is called the 3GPP file format and has its own file extension and MIME type to distinguish these files from MPEG-4 files. When this specification refers to the MP4 file format, it is referring to its structure (ISO file format), not to its conformance definition.The 3GPP file format, used in this specification for timed multimedia (such as video, associated audio and timed text), is structurally based on the ISO base media file format defined in [50]. However, the conformance statement for 3GP files is defined in the present document by addressing the registration of codecs, file identification, file extension and MIME type definition.

NOTE: Codecs or functionalities not conforming to a 3GP file may be ignored.

9.2.1 Registration of non-ISO codecs

MPEG-4 video and AAC audio code streams, as well as How to include the non-ISO code streams AMR narrow-band speech, AMR wideband speech, H.263 encoded video and timed text can be included in 3GPMP4 files as is described in annex D of the present document.

9.2.2 Hint tracks

The hHint tracks are a mechanism that theа server implementation may choose to use in preparation for the streaming of media content contained in 3GPMP4 files. However, it should be observed that the usage of the hint tracks is an internal implementation matter for the server, and it falls outside the scope of the present document.

9.2.3 Self-contained MP4 filesLimitations to the ISO base media file format

All media in the MP4 file shall be self-contained, i.e. there shall not be referencing to external media data from inside the MP4 file. The following limitations to the ISO base media file format [50] shall apply to a 3GP file of this specification:

- there shall be no references to external media outside the file, i.e. a 3GP file shall be self-contained;
- the maximum number of tracks shall be one for video, one for audio and one for text;
- compact sample sizes ('stz2') shall not be used;
- movie fragments shall not be used.

9.2.4 MPEG-4 systems specific elements

For the storage of MPEG-4 media specific information in 3GP files, this specification refers to MP4 [51], which is also based on the ISO base media file format. THowever, tracks relative to MPEG-4 system architectural elements (e.g. BIFS scene description tracks or OD Object descriptors) are optional in 3GP files and shall be ignored. The adoption of the MPEG-4 file format The inclusion of MPEG-4 media does not imply the usage of MPEG-4 systems architecture. The receiving terminal is not required to implement any of the specific MPEG-4 system architectural elements.

9.2.5 Interpretation of ~~MPEG-4~~3GPP file format

All index numbers used in the 3GPP~~MPEG-4~~ file format start with the value one rather than zero, in particular "first-chunk" in Sample to chunk ~~atom~~box, "sample-number" in Sync sample ~~atom~~box and "shadowed-sample-number", "sync-sample-number" in Shadow sync sample ~~atom~~box.

---------- <text left out> ----------

Annex D (normative): ~~Support for non-ISO code streams in MP4~~3GP files – codecs and identification

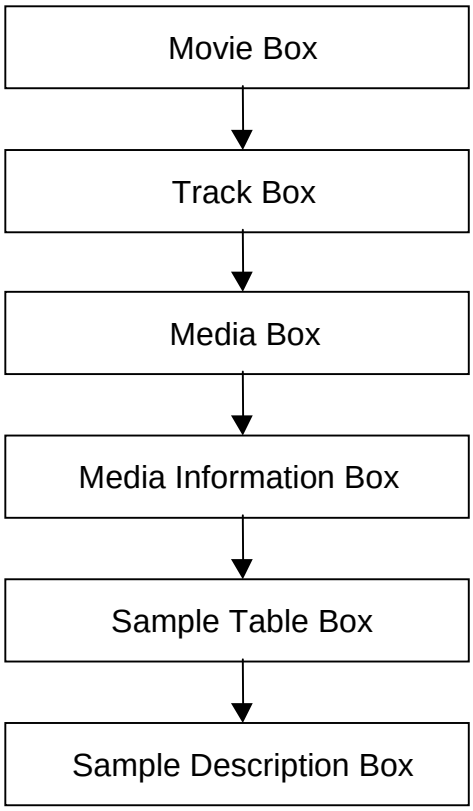
D.1 General

The purpose of this annex is to define the necessary structure for integration of the H.263, MPEG-4 video, AMR~~, and~~ AMR-WB, AAC and timed text media specific information in a~~n~~ 3GP~~MP4~~ file. Clause~~s~~ D.2 ~~to D.4~~ give~~s~~ some background information about the Sample Description ~~atom~~box in the ISO base media file format [50] and clauses D.3 and D.4 about the ~~,~~ MP4VisualSampleEntry ~~atom~~box and the MP4AudioSampleEntry ~~atom~~box in the MPEG-4 file format [51]. T~~hen, t~~he definitions of the SampleEntry ~~atoms~~ boxes for AMR, AMR-WB and H.263 are given in clauses D.5 to D.8. The SampleEntry box for timed text is given in clause D.8a. Finally, the identification of 3GP files is described in clause D.9.

AMR and AMR-WB data is stored in the stream according to the AMR and AMR-WB storage format for single channel header of Annex E [11], without the AMR magic numbers.

D.2 Sample Description ~~atom~~box

In an ISO~~MP4~~ file, Sample Description ~~Atom~~Box gives detailed information about the coding type used, and any initialisation information needed for that coding. The Sample Description ~~Atom~~Box can be found in the ISO file format~~MP4~~ ~~Atom~~Box Structure Hierarchy shown in figure D.1.


Movie Box
Track Box
Media Box
Media Information Box
Sample Table Box
Sample Description Box

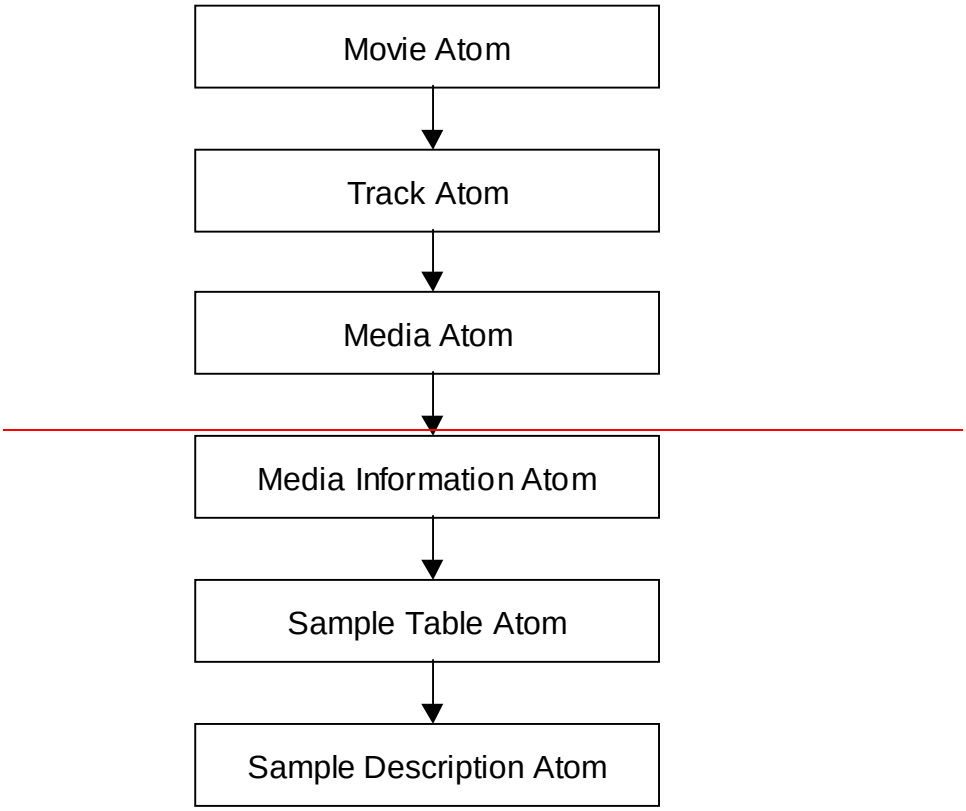

Movie Atom
Track Atom
Media Atom
Media Information Atom
Sample Table Atom
Sample Description Atom

Figure D.1: ISO File FormatMP4 AtomBox Structure Hierarchy

The Sample Description AtomBox can have one or more SampleDescription Entriesy fields. Valid Sample Description Entriesy atoms already defined for ISO and MP4 are include MP4AudioSampleEntry, MP4VisualSampleEntry, and HintSampleEntry and MPEGSampleEntry Atoms. The Sample DescriptionEntriesy Atoms for AMR and AMR-WB shall be AMRSampleEntry, and for H.263 it shall be H263SampleEntry, and for timed text it shall be TextSampleEntryrespectively.

The format of SampleDescriptionEntry and its fields are explained as follows:

SampleDescriptionEntry ::= MP4VisualSampleEntry |

MP4AudioSampleEntry |

HintSampleEntry |

MpegSampleEntryTextSampleEntry |

H263SampleEntry |

AMRSampleEntry

Table D.1: SampleDescriptionEntry fields

Field	Type	Details	Value
MP4VisualSampleEntry		Entry type for visual samples defined in the MPEG-4 specification.	
MP4AudioSampleEntry		Entry type for audio samples defined in the MPEG-4 specification.	
HintSampleEntry		Entry type for hint track samples defined in the MPEG-4ISO specification.	
MpegSampleEntryTextSampleEntry		Entry type for MPEG-related stream samples defined in the MPEG-4 specification.Entry type for timed text samples defined in clause D8a.16 of the present document.	
H263SampleEntry		Entry type for H.263 visual samples defined in clause D.6 of the present document.	
AMRSampleEntry		Entry type for AMR and AMR-WB speech samples defined in clause D.5 of the present document.	

From the above 6 Sample Entriesatoms, only the MP4VisualSampleEntry, MP4AudioSampleEntry, TextSampleEntry, H263SampleEntry and AMRSampleEntry atoms are taken into consideration, since MPEG specific streams and hint tracks are out of the scope of the present document.

D.3 MP4VisualSampleEntry atombox

The MP4VisualSampleEntry AtomBox is defined as follows:

MP4VisualSampleEntry ::= AtomBoxHeader

Reserved_6

Data-reference-index

Reserved_16

Width

Height

Reserved_4

Reserved_4

Reserved_4

Reserved_2

Reserved_32

Reserved_2

Reserved_2

ESD~~Atom~~Box

Table D.2: MP4VisualSampleEntry fields

Field	Type	Details	Value
~~Atom~~BoxHeader.Size	Unsigned int(32)		
~~Atom~~BoxHeader.Type	Unsigned int(32)		'mp4v'
Reserved_6	Unsigned int(8) [6]		0
Data-reference-index	Unsigned int(16)	Index to a data reference that to use to retrieve the sample data. Data references are stored in data reference ~~Atoms~~boxes.	
Reserved_16	Const unsigned int(32) [4]		0
Width	Unsigned int(16)	Maximum width, in pixels of the stream	
Height	Unsigned int(16)	Maximum height, in pixels of the stream	
Reserved_4	Const unsigned int(32)		0x00480000
Reserved_4	Const unsigned int(32)		0x00480000
Reserved_4	Const unsigned int(32)		0
Reserved_2	Const unsigned int(16)		1
Reserved_32	Const unsigned int(8) [32]		0
Reserved_2	Const unsigned int(16)		24
Reserved_2	Const int(16)		-1
ESD~~Atom~~Box		~~Atom~~Box containing an elementary stream descriptor for this stream.	

The stream type specific information is in the ESD~~Atom~~Box structure, which will be explained later.

This version of the MP4VisualSampleEntry, with explicit width and height, shall be used for MPEG-4 video streams conformant to this specification.

NOTE: width and height parameters together may be used to allocate the necessary memory in the playback device without need to analyse the video stream.

D.4 MP4AudioSampleEntry ~~atom~~box

MP4AudioSampleEntry~~Atom~~Box is defined as follows:

MP4AudioSampleEntry ::= ~~Atom~~BoxHeader

Reserved_6

Data-reference-index

Reserved_8

Reserved_2

Reserved_2

Reserved_4

TimeScale

Reserved_2

ESD~~Atom~~Box

Table D.3: MP4AudioSampleEntry fields

Field	Type	Details	Value
~~Atom~~BoxHeader.Size	Unsigned int(32)		
~~Atom~~BoxHeader.Type	Unsigned int(32)		'mp4a'
Reserved_6	Unsigned int(8) [6]		0
Data-reference-index	Unsigned int(16)	Index to a data reference that to use to retrieve the sample data. Data references are stored in data reference ~~Atoms~~boxes.	
Reserved_8	Const unsigned int(32) [2]		0
Reserved_2	Const unsigned int(16)		2
Reserved_2	Const unsigned int(16)		16
Reserved_4	Const unsigned int(32)		0
TimeScale	Unsigned int(16)	Copied from track	
Reserved_2	Const unsigned int(16)		0
ESD~~Atom~~Box		~~Atom~~Box containing an elementary stream descriptor for this stream.	

The stream type specific information is in the ESD~~Atom~~Box structure, which will be explained later.

D.5 AMRSampleEntry ~~atom~~box

For narrow-band AMR, the ~~atom~~box type of the AMRSampleEntry ~~Atom~~Box shall be 'samr'. For AMR wideband (AMR-WB), the ~~atom~~box type of the AMRSampleEntry ~~Atom~~Box shall be 'sawb'. Each AMR or AMR-WB track shall be associated with a single AMRSampleEntry.

The AMRSampleEntry ~~Atom~~Box is defined as follows:

AMRSampleEntry ::= ~~Atom~~BoxHeader

Reserved_6

Data-reference-index

Reserved_8

Reserved_2

Reserved_2

Reserved_4

TimeScale

Reserved_2

AMRSpecific~~Atom~~Box

Table D.4: AMRSampleEntry fields

Field	Type	Details	Value
~~Atom~~BoxHeader.Size	Unsigned int(32)		
~~Atom~~BoxHeader.Type	Unsigned int(32)		'samr' or 'sawb'
Reserved_6	Unsigned int(8) [6]		0
Data-reference-index	Unsigned int(16)	Index to a data reference that to use to retrieve the sample data. Data references are stored in data reference ~~Atoms~~boxes.	
Reserved_8	Const unsigned int(32) [2]		0
Reserved_2	Const unsigned int(16)		2
Reserved_2	Const unsigned int(16)		16
Reserved_4	Const unsigned int(32)		0
TimeScale	Unsigned int(16)	Copied from media header ~~atom~~box of this media	
Reserved_2	Const unsigned int(16)		0
AMRSpecific~~Atom~~Box		Information specific to the decoder.	

If one compares the MP4AudioSampleEntry ~~Atom~~Box - AMRSampleEntry ~~Atom~~Box the main difference is in the replacement of the ESD~~Atom~~Box, which is specific to MPEG-4 systems, with a~~n~~ ~~atom~~ box suitable for AMR and AMR-WB. The **AMRSpecific~~Atom~~Box** field structure is described in clause D.7.

D.6 H263SampleEntry ~~atom~~box

The ~~atom~~box type of the H263SampleEntry ~~Atom~~Box shall be 's263'.

The H263SampleEntry ~~Atom~~Box is defined as follows:

H263SampleEntry ::= ~~Atom~~BoxHeader

Reserved_6

Data-reference-index

Reserved_16

Width

Height

Reserved_4

Reserved_4

Reserved_4

Reserved_2

Reserved_32

Reserved_2

Reserved_2

H263Specific~~Atom~~Box

Table D.5: H263SampleEntry fields

Field	Type	Details	Value
~~Atom~~Box**Header**.Size	Unsigned int(32)		
~~Atom~~Box**Header**.Type	Unsigned int(32)		's263'
Reserved_6	Unsigned int(8) [6]		0
Data-reference-index	Unsigned int(16)	Index to a data reference that to use to retrieve the sample data. Data references are stored in data reference ~~Atoms~~boxes.	
Reserved_16	Const unsigned int(32) [4]		0
Width	Unsigned int(16)	Maximum width, in pixels of the stream	
Height	Unsigned int(16)	Maximum height, in pixels of the stream	
Reserved_4	Const unsigned int(32)		0x00480000
Reserved_4	Const unsigned int(32)		0x00480000
Reserved_4	Const unsigned int(32)		0
Reserved_2	Const unsigned int(16)		1
Reserved_32	Const unsigned int(8) [32]		0
Reserved_2	Const unsigned int(16)		24
Reserved_2	Const int(16)		-1
H263Specific~~Atom~~Box		Information specific to the H.263 decoder.	

If one compares the MP4VisualSampleEntry – H263SampleEntry ~~Atom~~Box the main difference is in the replacement of the ESD~~Atom~~Box, which is specific to MPEG-4 systems, with a~~n~~ ~~atom~~box suitable for H.263. The **H263Specific~~Atom~~Box** field structure for H.263 is described in clause D.8.

D.7 AMRSpecific~~Atom~~Box field for AMRSampleEntry ~~atom~~box

The AMRSpecific~~Atom~~Box fields for AMR and AMR-WB shall be as defined in table D.6. The AMRSpecific~~Atom~~Box for the AMRSampleEntry ~~Atom~~Box shall always be included if the ~~MP4~~3GP file contains AMR or AMR-WB media.

Table D.6: The AMRSpecific~~Atom~~Box fields for AMRSampleEntry

Field	Type	Details	Value
~~Atom~~BoxHeader.Size	Unsigned int(32)		
~~Atom~~BoxHeader.Type	Unsigned int(32)		'damr'
DecSpecificInfo	AMRDecSpecStruc	Structure which holds the AMR and AMR-WB Specific information	

~~Atom~~Box**Header Size and Type:** indicate the size and type of the AMR decoder-specific ~~atom~~box. The type must be 'damr'.

DecSpecificInfo: the structure where the AMR and AMR-WB stream specific information resides.

The AMRDecSpecStruc is defined as follows:

struct **AMRDecSpecStruc**{

Unsigned int (32)	**vendor**
Unsigned int (8)	**decoder_version**
Unsigned int (16)	**mode_set**
Unsigned int (8)	**mode_change_period**
Unsigned int (8)	**frames_per_sample**

}

The definitions of AMRDecSpecStruc members are as follows:

vendor: four character code of the manufacturer of the codec, e.g. 'VXYZ'. The vendor field gives information about the vendor whose codec is used to create the encoded data. It is an informative field which may be used by the decoding end. If a manufacturer already has a four character code, it is recommended that it uses the same code in this field. Else, it is recommended that the manufacturer creates a four character code which best addresses the manufacturer's name. It can be safely ignored.

decoder_version: version of the vendor's decoder which can decode the encoded stream in the best (i.e. optimal) way. This field is closely tied to the vendor field. It may give advantage to the vendor which has optimal encoder-decoder version pairs. The value is set to 0 if decoder version has no importance for the vendor. It can be safely ignored.

mode_set: the active codec modes. Each bit of the mode_set parameter corresponds to one mode. The bit index of the mode is calculated according to the 4 bit FT field of the AMR or AMR-WB frame structure. The mode_set bit structure is as follows: (B15xxxxxxB8B7xxxxxxB0) where B0 (Least Significant Bit) corresponds to Mode 0, and B8 corresponds to Mode 8.

The mapping of existing AMR modes to FT is given in table 1.a in [19]. A value of 0x81FF means all modes and comfort noise frames are possibly present in an AMR stream.

The mapping of existing AMR-WB modes to FT is given in Table 1.a in TS 26.201 [37]. A value of 0x83FF means all modes and comfort noise frames are possibly present in an AMR-WB stream.

As an example, if mode_set = 0000000110010101b, only Modes 0, 2, 4, 7 and 8 are present in the stream.

mode_change_period: defines a number N, which restricts the mode changes only at a multiple of N frames. If no restriction is applied, this value should be set to 0. If mode_change_period is not 0, the following restrictions apply to it according to the frames_per_sample field:

if (mode_change_period < frames_per_sample)

frames_per_sample = k x (mode_change_period)

else if (mode_change_period > frames_per_sample)

mode_change_period = k x (frames_per_sample)

where k : integer [2, …]

If mode_change_period is equal to frames_per_sample, then the mode is the same for all frames inside one sample.

frames_per_sample: defines the number of frames to be considered as 'one sample' inside the ~~MP4~~3GP file. This number shall be greater than 0 and less than 16. A value of 1 means each frame is treated as one sample. A value of 10 means that 10 frames (of duration 20 msec each) are put together and treated as one sample. It must be noted that, in this case, one sample duration is 20 (msec/frame) x 10 (frame) = 200 msec. For the last sample of the stream, the

number of frames can be smaller than frames_per_sample, if the number of remaining frames is smaller than frames_per_sample.

NOTE1: The "hinter", for the creation of the hint tracks, can use the information given by the AMRDecSpecStruc members.

NOTE2: The following AMR MIME parameters are not relevant to PSS: {mode_set, mode_change_period, mode_change_neighbor}. PSS servers should not send these parameters in SDP, and PSS clients shall ignore these parameters if received.

D.8 H263Specific~~Atom~~Box field for H263SampleEntry ~~atom~~box

The H263Specific~~Atom~~Box fields for H. 263 shall be as defined in table D.7. The H263Specific~~Atom~~Box for the H263SampleEntry ~~Atom~~Box shall always be included if the ~~MP4~~3GP file contains H.263 media.

The H263Specific~~Atom~~Box for H263 is composed of the following fields.

Table D.7: The H263Specific~~Atom~~Box fields H263SampleEntry

Field	Type	Details	Value
~~Atom~~BoxHeader.Size	Unsigned int(32)		
~~Atom~~BoxHeader.Type	Unsigned int(32)		'd263'
DecSpecificInfo	H263DecSpecStruc	Structure which holds the H.263 Specific information	
Bitrate~~Atom~~Box		Specific bitrate information (optional)	

~~Atom~~BoxHeader Size and Type: indicate the size and type of the H.263 decoder-specific ~~atom~~box. The type must be 'd263'.

DecSpecificInfo: This is the structure where the H263 stream specific information resides.

H263DecSpecStruc is defined as follows:

struct **H263DecSpecStruc**{

Unsigned int (32)	**vendor**
Unsigned int (8)	**decoder_version**
Unsigned int (8)	**H263_Level**
Unsigned int (8)	**H263_Profile**

}

The definitions of H263DecSpecStruc members are as follows:

vendor: four character code of the manufacturer of the codec, e.g. 'VXYZ'. The vendor field gives information about the vendor whose codec is used to create the encoded data. It is an informative field which may be used by the decoding end. If a manufacturer already has a four character code, it is recommended that it uses the same code in this field. Else, it is recommended that the manufacturer creates a four character code which best addresses the manufacturer's name. It can be safely ignored.

decoder_version: version of the vendor's decoder which can decode the encoded stream in the best (i.e. optimal) way. This field is closely tied to the vendor field. It may give advantage to the vendor which has optimal encoder-decoder version pairs. . The value is set to 0 if decoder version has no importance for the vendor. It can be safely ignored.

H263_Level and H263_Profile: These two parameters define which H263 profile and level is used. These parameters are based on the MIME media type video/H263-2000. The profile and level specifications can be found in [23].

EXAMPLE 1: H.263 Baseline = {H263_Level = 10, H263_Profile = 0}

EXAMPLE 2: H.263 Profile 3 @ Level 10 = {H263_Level = 10 , H263_Profile = 3}

NOTE: The "hinter", for the creation of the hint tracks, can use the information given by the H263DecSpecStruc members.

The Bitrate~~Atom~~Box field shall be as defined in table D.7.1. The Bitrate~~Atom~~Box may be included if the ~~MP4~~3GP file contains H.263 media.

The Bitrate~~Atom~~Box is composed of the following fields.

Table D.7.1: The Bitrate~~Atom~~Box fields

Field	Type	Details	Value
~~Atom~~BoxHeader.Size	Unsigned int(32)		
~~Atom~~BoxHeader.Type	Unsigned int(32)		'bitr'
DecBitrateInfo	DecBitrStruc	Structure which holds the Bitrate information	

~~Atom~~BoxHeader Size and Type: indicate the size and type of the bitrate ~~atom~~box. The type must be 'bitr'.

DecBitrateInfo: This is the structure where the stream bitrate information resides.

DecBitrStruc is defined as follows:

struct **DecBitrStruc**{

Unsigned int (32) **Avg_Bitrate**

Unsigned int (32) **Max_Bitrate**

}

The definitions of DecBitrStruc members are as follows:

Avg_Bitrate: the average bitrate in bits per second of this elementary stream. For streams with variable bitrate this value shall be set to zero.

Max_Bitrate: the maximum bitrate in bits per second of this elementary stream in any time window of one second duration.

D.8a Timed Text Format

This clause defines the format of timed text in downloaded files. In this release, timed text is downloaded, not streamed.

Operators may specify additional rules and restrictions when deploying terminals, in addition to this specification, and behavior that is optional here may be mandatory for particular deployments. In particular, the required character set is almost certainly dependent on the geography of the deployment.

D.8a.1 Unicode Support

Text in this specification uses the Unicode 3.0 [30] standard. Terminals shall correctly decode both UTF-8 and UTF-16 into the required characters. If a terminal receives a Unicode code, which it cannot display, it shall display a predictable result. It shall not treat multi-byte UTF-8 characters as a series of ASCII characters, for example.

Authors should create fully-composed Unicode; terminals are not required to handle decomposed sequences for which there is a fully-composed equivalent.

Terminals shall conform to the conformance statement in Unicode 3.0 section 3.1.

Text strings for display and font names are uniformly coded in UTF-8, or start with a UTF-16 BYTE ORDER MARK (\uFEFF) and by that indicate that the string which starts with the byte order mark is in UTF-16. Terminals shall recognise the byte-order mark in this byte order; they are not required to recognise byte-reversed UTF-16, indicated by a byte-reversed byte-order mark.

D.8a.2 Bytes, Characters, and Glyphs

This clause uses these terms carefully. Since multi-byte characters are permitted (i.e. 16-bit Unicode characters), the number of characters in a string may not be the number of bytes. Also, a byte-order-mark is not a character at all, though it occupies two bytes. So, for example, storage lengths are specified as byte-counts, whereas highlighting is specified using character offsets.

It should also be noted that in some writing systems the number of glyphs rendered might be different again. For example, in English, the characters ‘fi’ are sometimes rendered as a single ligature glyph.

In this specification, the first character is at offset 0 in the string. In records specifying both a start and end offset, the end offset shall be greater than or equal to the start offset. In cases where several offset specifications occur in sequence, the start offset of an element shall be greater than or equal to the end offset of the preceding element.

D.8a.3 Character Set Support

All terminals shall be able to render Unicode characters in these ranges:

a) basic ASCII and Latin-1 (\u0000 to \u00FF), though not all the control characters in this range are needed;

b) the Euro currency symbol (\u20AC)

c) telephone and ballot symbols (\u260E through \u2612)

Support for the following characters is recommended but not required:

a) miscellaneous technical symbols (\u2300 through \u2335)

b) ‘Zapf Dingbats’: locations \u2700 through \u27AF, and the locations where some symbols have been relocated (e.g. \u2605, Black star).

The private use characters \u0091 and \u0092, and the initial range of the private use area \uE000 through \uE0FF are reserved in this specification. For these Unicode values, and for control characters for which there is no defined graphical behaviour, the terminal shall not display any result: neither a glyph is shown nor is the current rendering position changed.

D.8a.4 Font Support

Fonts are specified in this specification by name, size, and style. There are three special names which shall be recognized by the terminal: Serif, Sans-Serif, and Monospace. It is strongly recommended that these be different fonts for the required characters from ASCII and Latin-1. For many other characters, the terminal may have a limited set or only a single font. Terminals requested to render a character where the selected font does not support that character should substitute a suitable font. This ensures that languages with only one font (e.g. Asian languages) or symbols for which there is only one form are rendered.

Fonts are requested by name, in an ordered list. Authors should normally specify one of the special names last in the list.

Terminals shall support a pixel size of 12 (on a 72dpi display, this would be a point size of 12). If a size is requested other than the size(s) supported by the terminal, the next smaller supported size should be used. If the requested size is smaller than the smallest supported size, the terminal should use the smallest supported size.

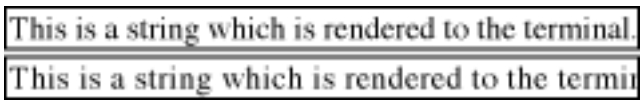
Terminals shall support unstyled text for those characters it supports. It may also support bold, italic (oblique) and bold-italic. If a style is requested which the terminal does not support, it should substitute a supported style; a character shall be rendered if the terminal has that character in any style of any font.

D.8a.5 Fonts and Metrics

Within the sample description, a complete list of the fonts used in the samples is found. This enables the terminal to pre-load them, or to decide on font substitution.

Terminals may use varying versions of the same font. For example, here is the same text rendered on two systems; it was authored on the first, where it just fitted into the text box.

EXAMPLE:



Authors should be aware of this possible variation, and provide text box areas with some 'slack' to allow for rendering variations.

D.8a.6 Colour Support

The colour of both text and background are indicated in this specification using RGB values. Terminals are not required to be able to display all colours in the RGB space. Terminals with a limited colour display, with only gray-scale display, and with only black-and-white are permissible. If a terminal has a limited colour capability it should substitute a suitable colour; dithering of text may be used but is not usually appropriate as it results in "fuzzy" display. If colour substitution is performed, the substitution shall be consistent: the same RGB colour shall result consistently in the same displayed colour. If the same colour is chosen for background and text, then the text shall be invisible (unless a style such as highlight changes its colour). If different colours are specified for the background and text, the terminal shall map these to different colours, so that the text is visible.

Colours in this specification also have an alpha or transparency value. In this specification, a transparency value of 0 indicates a fully transparent colour, and a value of 255 indicates fully opaque. Support for partial or full transparency is optional. 'Keying' text (text rendered on a transparent background) is done by using a background colour which is fully transparent. 'Keying' text over video or pictures, and support for transparency in general, can be complex and may require double-buffering, and its support is optional in the terminal. Content authors should beware that if they specify a colour which is not fully opaque, and the content is played on a terminal not supporting it, the affected area (the entire text box for a background colour) will be fully opaque and will obscure visual material behind it. Visual material with transparency is layered closer to the viewer than the material which it partially obscures.

D.8a.7 Text rendering position and composition

Text is rendered within a region (a concept derived from SMIL). There is a text box set within that region. This permits the terminal to position the text within the overall presentation, and also to render the text appropriately given the writing direction. For text written left to right, for example, the first character would be rendered at, or near, the left edge of the box, and with its baseline down from the top of the box by one baseline height (a value derived from the font and font size chosen). Similar considerations apply to the other writing directions.

Within the region, text is rendered within a text box. There is a default text box set, which can be over-ridden by a sample.

The text box is filled with the background colour; after that the text is painted in the text colour. If highlighting is requested one or both of these colours may vary.

Terminals may choose to anti-alias their text, or not.

The text region and layering are defined using structures from the ISO base media file format.

This track header box is used for text track:

```
aligned(8) class TrackHeaderBox
    extends FullBox('tkhd', version, flags){
    if (version==1) {
```

```
        unsigned int(64)    creation_time;
        unsigned int(64)    modification_time;
        unsigned int(32)    track_ID;
        const unsigned int(32)  reserved = 0;
        unsigned int(64)    duration;
    } else { // version==0
        unsigned int(32)    creation_time;
        unsigned int(32)    modification_time;
        unsigned int(32)    track_ID;
        const unsigned int(32)  reserved = 0;
        unsigned int(32)    duration;
    }
    const unsigned int(32)[2]   reserved = 0;
    int(16) layer;
    template int(16) alternate_group = 0;
    template int(16)    volume = 0;
    const unsigned int(16)  reserved = 0;
    template int(32)[9]     matrix=
        { 0x00010000,0,0,0,0x00010000,0,tx,ty,0x40000000 };
        // unity matrix
    unsigned int(32) width;
    unsigned int(32) height;
}
```

Visually composed tracks including video and text are layered using the ‘layer’ value. This compares, for example, to z-index in SMIL. More negative layer values are towards the viewer. (This definition is compatible with that in ISO/MJ2).

The region is defined by the track width and height, and translation offset. This corresponds to the SMIL region. The width and height are stored in the track header fields above. The sample description sets a text box within the region, which can be over-ridden by the samples.

The translation values are stored in the track header matrix in the following positions:

{ 0x00010000,0,0, 0,0x00010000,0, tx, ty, 0x40000000 }

These values are fixed-point 16.16 values, here restricted to be integers (the lower 16 bits of each value shall be zero). The X axis increases from left to right; the Y axis from top to bottom. (This use of the matrix is conformant with ISO/MJ2.)

So, for example, a centered region of size 200x20, positioned below a video of size 320x240, would have track_width set to 200 (widh= 0x00c80000), track_height set to 20 (height= 0x00140000), and tx = (320-200)/2 = 60, and ty=240.

Since matrices are not used on the video tracks, all video tracks are set at the coordinate origin. Figure D.2 provides an overview:

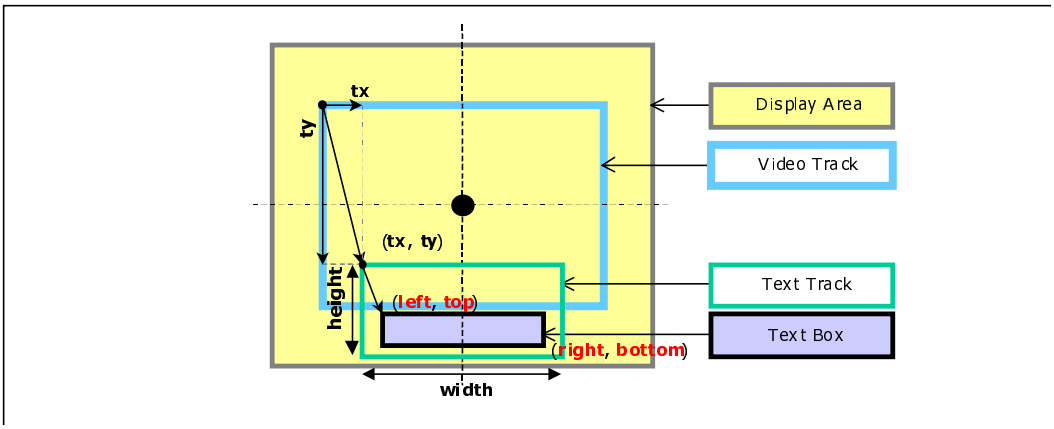


Figure D.2: Illustration of text rendering position and composition

The top and left positions of the text track is determined by the tx and ty, which are the translation values from the coordinate origin (since the video track is at the origin, this is also the offset from the video track). The default text box set in the sample description sets the rendering area unless over-ridden by a 'tbox' in the text sample. The box values are defined as the relative values from the top and left positions of the text track.

It should be noted that this only specifies the relationship of the tracks within a single 3GP ~~(MP4)~~ file. If a SMIL presentation lays up multiple files, their relative position is set by the SMIL regions. Each file is assigned to a region, and then within those regions the spatial relationship of the tracks is defined.

D.8a.8 Marquee Scrolling

Text can be 'marquee' scrolled in this specification (compare this to Internet Explorer's marquee construction). When scrolling is performed, the terminal first calculates the position in which the text would be displayed with no scrolling requested. Then:

a) If scroll-in is requested, the text is initially invisible, just outside the text box, and enters the box in the indicated direction, scrolling until it is in the normal position;

b) If scroll-out is requested, the text scrolls from the normal position, in the indicated direction, until it is completely outside the text box.

The rendered text is clipped to the text box in each display position, as always. This means that it is possible to scroll a string which is longer than can fit into the text box, progressively disclosing it (for example, like a ticker-tape). Note that both scroll in and scroll out may be specified; the text scrolls continuously from its invisible initial position, through the normal position, and out to its final position.

If a scroll-delay is specified, the text stays steady in its normal position (not initial position) for the duration of the delay; so the delay is after a scroll-in but before a scroll-out. This means that the scrolling is not continuous if both are specified. So without a delay, the text is in motion for the duration of the sample. For a scroll in, it reaches its normal position at the end of the sample duration; with a delay, it reaches its normal position before the end of the sample duration, and remains in its normal position for the delay duration, which ends at the end of the sample duration. Similarly for a scroll out, the delay happens in its normal position before scrolling starts. If both scroll in, and scroll out are specified, with a delay, the text scrolls in, stays stationary at the normal position for the delay period, and then scrolls out – all within the sample duration.

The speed of scrolling is calculated so that the complete operation takes place within the duration of the sample. Therefore the scrolling has to occur within the time left after scroll-delay has been subtracted from the sample duration. Note that the time it takes to scroll a string may depend on the rendered length of the actual text string. Authors should consider whether the scrolling speed that results will be exceed that at which text on a wireless terminal could be readable.

Terminals may use simple algorithms to determine the actual scroll speed. For example, the speed may be determined by moving the text an integer number of pixels in every update cycle. Terminals should choose a scroll speed which is as fast or faster than needed so that the scroll operation completes within the sample duration.

Terminals are not required to handle dynamic or stylistic effects such as highlight, dynamic highlight, or href links on scrolled text.

The scrolling direction is set by a two-bit field, with the following possible values:

00b – text is vertically scrolled up ('credits style'), entering from the bottom of the bottom and leaving towards the top.

01b – text is horizontally scrolled ('marquee style'), entering from the right and leaving towards the left.

10b – text is vertically scrolled down, entering from the top and leaving towards the bottom.

11b – text is horizontally scrolled, entering from the left and leaving towards the right.

D.8a.9 Language

The human language used in this stream is declared by the language field of the media-header ~~atom~~box in this track. It is an ISO 639/T 3-letter code. The knowledge of the language used might assist searching, or speaking the text. Rendering is language neutral. Note that the values 'und' (undetermined) and 'mul' (multiple languages) might occur.

D.8a.10Writing direction

Writing direction specifies the way in which the character position changes after each character is rendered. It also will imply a start-point for the rendering within the box.

Terminals shall support the determination of writing direction, for those characters they support, according to the Unicode 3.0 specification. Note that the only required characters can all be rendered using left-right behaviour. A terminal which supports characters with right-left writing direction shall support the right-left composition rules specified in Unicode.

Terminals may also set, or allow the user to set, an overall writing direction, either explicitly or implicitly (e.g. by the language selection). This affects layout. For example, if upper-case letters are left-right, and lower-case right-left, and the Unicode string ABCdefGHI shall be rendered, it would appear as ABCfedGHI on a terminal with overall left-right writing (English, for example) and GHIdefABC on a system with overall right-left (Hebrew, for example).

Terminals are not required to support the bi-directional ordering codes (\u200E, \u200F and \u202A through \u202E).

If vertical text is requested by the content author, characters are laid out vertically from top to bottom. The terminal may choose to render different glyphs for this writing direction (e.g. a horizontal parenthesis), but in general the glyphs should not be rotated. The direction in which lines advance (left-right, as used for European languages, or right-left, as used for Asian languages) is set by the terminal, possibly by a direct or indirect user preference (e.g. a language setting). Terminals shall support vertical writing of the required character set. It is recommended that terminals support vertical writing of text in those languages commonly written vertically (e.g. Asian languages). If vertical text is requested for characters which the terminal cannot render vertically, the terminal may behave as if the characters were not available.

D.8a.11Text wrap

Automatic wrapping of text from line to line is complex, and can require hyphenation rules and other complex language-specific criteria. For these reasons, text is not wrapped in this specification. If a string is too long to be drawn within the box, it is clipped. The terminal may choose whether to clip at the pixel boundary, or to render only whole glyphs.

There may be multiple lines of text in a sample (hard wrap). Terminals shall start a new line for the Unicode characters line separator (\u2028), paragraph separator (\u2029) and line feed (\u000A). It is recommended that terminals follow Unicode Technical Report 13 [48]. Terminals should treat carriage return (\u000D), next line (\u0085) and CR+LF (\u000D\u000A) as new line.

D.8a.12Highlighting, Closed Caption, and Karaoke

Text may be highlighted for emphasis. Since this is a non-interactive system, solely for text display, the utility of this function may be limited.

Dynamic highlighting used for Closed Caption and Karaoke highlighting, is an extension of highlighting. Successive contiguous sub-strings of the text sample are highlighted at the specified times.

D.8a.13Media Handler

A text stream is its own unique stream type. For the 3GPP file format, the handler-type within the ‘hdlr’ ~~atom~~box shall be ‘text’.

D.8a.14Media Handler Header

The 3G text track uses an empty null media header (‘nmhd’), called Mpeg4MediaHeader~~Atom~~Box in the MP4 specification [51], in common with other MPEG streams.

```
aligned(8) class  Mpeg4MediaHeaderAtomBox
    extends FullAtomBox('nmhd', version = 0, flags) {
 }
```

D.8a.15 Style record

Both the sample format and the sample description contain style records, and so it is defined once here for compactness.

```
aligned(8) class StyleRecord {
    unsigned int(16)    startChar;
    unsigned int(16)    endChar;
    unsigned int(16)    font-ID;
    unsigned int(8) face-style-flags;
    unsigned int(8) font-size;
    unsigned int(8) text-color-rgba[4];
}
```

startChar: character offset of the beginning of this style run (always 0 in a sample description)

endChar: first character offset to which this style does not apply (always 0 in a sample description); shall be greater than or equal to startChar. All characters, including line-break characters and any other non-printing characters, are included in the character counts.

font-ID: font identifier from the font table; in a sample description, this is the default font

face style flags: in the absence of any bits set, the text is plain

1 bold

2 italic

4 underline

font-size: font size (nominal pixel size, in essentially the same units as the width and height)

text-color-rgba: rgb colour, 8 bits each of red, green, blue, and an alpha (transparency) value

Terminals shall support plain text, and underlined horizontal text, and may support bold, italic and bold-italic depending on their capabilities and the font selected. If a style is not supported, the text shall still be rendered in the closest style available.

D.8a.16 Sample Description Format

The sample table box ('stbl') contains sample descriptions for the text track. Each entry is a sample entry box of type ‘tx3g’. This name defines the format both of the sample description and the samples associated with that sample description. Terminals shall not attempt to decode or display sample descriptions with unrecognised names, nor the samples attached to those sample descriptions.

It starts with the standard fields (the reserved bytes and the data reference index), and then some text-specific fields. Some fields can be overridden or supplemented by additional boxes within the text sample itself. These are discussed below.

There can be multiple text sample descriptions in the sample table. If the overall text characteristics do not change from one sample to the next, the same sample description is used. Otherwise, a new sample description is added to the table. Not all changes to text characteristics require a new sample description, however. Some characteristics, such as font size, can be overridden on a character-by-character basis. Some, such as dynamic highlighting, are not part of the text sample description and can be changed dynamically.

The TextDescription extends the regular sample entry with the following fields.

```
class FontRecord {
    unsigned int(16)    font-ID;
    unsigned int(8) font-name-length;
    unsigned int(8) font[font-name-length];
}

class FontTableBox() extends Box(‘ftab’) {
    unsigned int(16) entry-count;
    FontRecord  font-entry[entry-count];
}

class BoxRecord {
```

```
    signed int(16)  top;
    signed int(16)  left;
    signed int(16)  bottom;
    signed int(16)  right;
}

class TextSampleEntry() extends SampleEntry ('tx3g') {
    unsigned int(32)    displayFlags;
    signed int(8)       horizontal-justification;
    signed int(8)       vertical-justification;
    unsigned int(8) background-color-rgba[4];
    BoxRecord           default-text-box;
    StyleRecord         default-style;
    FontTableBox        font-table;
}
```

displayFlags:

scroll In	0x00000020	
scroll Out	0x00000040	
scroll direction	0x00000180	/ see above for values
continuous karaoke	0x00000800	
write text vertically	0x00020000	

horizontal and vertical justification: / two eight-bit values from the following list:

left, top	0
centered	1

bottom, right -1

background-color-rgba:
rgb color, 8 bits each of red, green, blue, and an alpha (transparency) value

Default text box: the default text box is set by four values, relative to the text region; it may be over-ridden in samples;

style record of default style: startChar and endChar shall be zero in a sample description

The text box is inset within the region defined by the track translation offset, width, and height. The values in the box are relative to the track region, and are uniformly coded with respect to the pixel grid. So, for example, the default text box for a track at the top left of the track region and 50 pixels high and 100 pixels wide is {0, 0, 50, 100}.

A font table shall follow these fields, to define the complete set of fonts used. The font table is an ~~atom~~box of type 'ftab'. Every font used in the samples is defined here by name. Each entry consists of a 16-bit local font identifier, and a font name, expressed as a string, preceded by an 8-bit field giving the length of the string in bytes. The name is expressed in UTF-8 characters, unless preceded by a UTF-16 byte-order-mark, whereupon the rest of the string is in 16-bit Unicode characters. The string should be a comma separated list of font names to be used as alternative font, in preference order. The special names "Serif", "Sans-serif" and "Monospace" may be used. The terminal should use the first font in the list which it can support; if it cannot support any for a given character, but it has a font which can, it should use that font. Note that this substitution is technically character by character, but terminals are encouraged to keep runs of characters in a consistent font where possible.

D.8a.17 Sample Format

Each sample in the media data consists of a string of text, optionally followed by sample modifier boxes.

For example, if one word in the sample has a different size than the others, a 'styl' box is appended to that sample, specifying a new text style for those characters, and for the remaining characters in the sample. This overrides the style in the sample description. These boxes are present only if they are needed. If all text conforms to the sample description, and no characteristics are applied that the sample description does not cover, no boxes are inserted into the sample data.

```
class TextSampleModifierBox(type) extends Box(type) {
}

class TextSample {
    unsigned int(16)            text-length;
```

```
    unsigned int(8)         text[text-length];
    TextSampleModifierBox   text-modifier[];    // to end of the sample
}
```

The initial string is preceded by a 16-bit count of the number of bytes in the string. There is no need for null termination of the text string. The sample size table provides the complete byte-count of each sample, including the trailing modifier boxes; by comparing the string length and the sample size, you can determine how much space, if any, is left for modifier boxes.

Authors should limit the string in each text sample to not more than 2048 bytes, for maximum terminal interoperability.

Any unrecognised box found in the text sample should be skipped and ignored, and processing continue as if it were not there.

D.8a.17.1 Sample Modifier Boxes

D.8a.17.1.1 Text Style

'styl'

This specifies the style of the text. It consists of a series of style records as defined above, preceded by a 16-bit count of the number of style records. Each record specifies the starting and ending character positions of the text to which it applies. The styles shall be ordered by starting character offset, and the starting offset of one style record shall be greater than or equal to the ending character offset of the preceding record; styles records shall not overlap their character ranges.

```
class TextStyleBox() extends TextSampleModifierBox (‘styl’) {
    unsigned int(16)    entry-count;
    StyleRecord         text-styles[entry-count];
}
```

D.8a.17.1.2 Highlight

'hlit' - Specifies highlighted text: the ~~atom~~box contains two 16-bit integers, the starting character to highlight, and the first character with no highlighting (e.g. values 4, 6 would highlight the two characters 4 and 5). The second value may be the number of characters in the text plus one, to indicate that the last character is highlighted.

```
class TextHighlightBox() extends TextSampleModifierBox (‘hlit’) {
    unsigned int(16)    startcharoffset;
    unsigned int(16)    endcharoffset;
}
class TextHilightColorBox() extends TextSampleModifierBox ('hclr') {
    unsigned int(8)     highlight_color_rgba[4];
}
```

highlight_color_rgb:

rgb color, 8 bits each of red, green, blue, and an alpha (transparency) value

The TextHilightColor Box may be present when the TextHighlightBox or TextKaraokeBox is present in a text sample. It is recommended that terminals use the following rules to determine the displayed effect when highlight is requested:

a) if a highlight colour is not specified, then the text is highlighted using a suitable technique such as inverse video: both the text colour and the background colour change.

b) if a highlight colour is specified, the background colour is set to the highlight colour for the highlighted characters; the text colour does not change.

Terminals do not need to handle text that is both scrolled and either statically or dynamically highlighted. Content authors should avoid specifying both scroll and highlight for the same sample.

D.8a.17.1.3 Dynamic Highlight

'krok' – Karaoke, closed caption, or dynamic highlighting. The number of highlight events is specified, and each event is specified by a starting and ending character offset and an end time for the event. The start time is either the sample start time or the end time of the previous event. The specified characters are highlighted from the previous end-time (initially the beginning of this sample's time), to the end time. The times are all specified relative to the sample's time; that is, a time of 0 represents the beginning of the sample time. The times are measured in the timescale of the track.

The ~~atom~~box starts with the start-time offset of the first highlight event, a 16-bit count of the event count, and then that number of 8-byte records. Each record contains the end-time offset as a 32-bit number, and the text start and end values, each as a 16-bit number. These values are specified as in the highlight record – the offset of the first character to highlight, and the offset of the first character not highlighted. The special case, where the startcharoffset equals to the endcharoffset, can be used to pause during or at the beginning of dynamic highlighting. The records shall be ordered and not overlap, as in the highlight record. The time in each record is the end time of this highlight event; the first highlight event starts at the indicated start-time offset from the start time of the sample. The time values are in the units expressed by the timescale of the track. The time values shall not exceed the duration of the sample.

The continuouskaraoke flag controls whether to highlight only those characters (continuouskaraoke = 0) selected by a karaoke entry, or the entire string from the beginning up to the characters highlighted (continuouskaraoke = 1) at any given time. In other words, the flag specifies whether karaoke should ignore the starting offset and highlight all text from the beginning of the sample to the ending offset.

Karaoke highlighting is usually achieved by using the highlight colour as the text colour, without changing the background.

At most one dynamic highlight ('krok') ~~atom~~box may occur in a sample.

```
class TextKaraokeBox() extends TextSampleModifierBox ('krok') {
    unsigned int(32)    highlight-start-time;
    unsigned int(16)    entry-count;
    for (i=1; i<=entry-count; i++) {
        unsigned int(32)    highlight-end-time;
        unsigned int(16)    startcharoffset;
        unsigned int(16)    endcharoffset;
    }
}
```

D.8a.17.1.4 Scroll Delay

'dlay' - Specifies a delay after a Scroll In and/or before Scroll Out. A 32-bit integer specifying the delay, in the units of the timescale of the track. The default delay, in the absence of this box, is 0.

```
class TextScrollDelayBox() extends TextSampleModifierBox ('dlay') {
    unsigned int(32)    scroll-delay;
}
```

D.8a.17.1.5 HyperText

'href' – HyperText link. The existence of the hypertext link is visually indicated in a suitable style (e.g. underlined blue text).

This box contains these values:

startCharOffset: – the start offset of the text to be linked

endCharOffset: – the end offset of the text (start offset + number of characters)

URLLength:– the number of bytes in the following URL

URL: UTF-8 characters – the linked-to URL

altLength:– the number of bytes in the following "alt" string

altstring: UTF-8 characters – an "alt" string for user display

The URL should be an absolute URL, as the context for a relative URL may not always be clear.

The “alt” string may be used as a tool-tip or other visual clue, as a substitute for the URL, if desired by the terminal, to display to the user as a hint on where the link refers.

Hypertext-linked text should not be scrolled; not all terminals can display this or manage the user interaction to determine whether user has interacted with moving text. It is also hard for the user to interact with scrolling text.

```
class TextHyperTextBox() extends TextSampleModifierBox (‘href’) {
    unsigned int(16)    startcharoffset;
    unsigned int(16)    endcharoffset;
    unsigned int(8) URLLength;
    unsigned int(8) URL[URLLength];
    unsigned int(8) altLength;
    unsigned int(8) altstring[altLength];
}
```

D.8a.17.1.6 Textbox

‘tbox’ – text box over-ride. This over-rides the default text box set in the sample description.

```
class TextboxBox() extends TextSampleModifierBox ('tbox') {
    BoxRecord   text-box;
}
```

D.8a.17.1.7 Blink

‘blnk’ – Blinking text. This requests blinking text for the indicated character range. Terminals are not required to support blinking text, and the precise way in which blinking is achieved, and its rate, is terminal-dependent.

```
class BlinkBox() extends TextSampleModifierBox ('blnk') {
    unsigned int(16)        startcharoffset;
    unsigned int(16)        endcharoffset;
}
```

D.8a.18 Combinations of features

Two modifier boxes of the same type shall not be applied to the same character (e.g. it is not permitted to have two href links from the same text). As the ‘hclr’, ‘dlay’ and ‘tbox’ are globally applied to the whole text in a sample, two modifier boxes of the same type shall not be present within a sample.

Table D.8 details the effects of multiple options:

Table D.8: Combinations of features

			First sample modifier ~~atom~~box				
		Sample description style record	**styl**	**hlit**	**krok**	**href**	**blnk**
Second sample modifier ~~atom~~box	**styl**	1	3				
	hlit			3			
	krok			4	3		
	href	2	2		5	3	
	blnk		6	6	6	6	6

1. The sample description provides the default style; the style records over-ride this for the selected characters.
2. The terminal over-rides the chosen style for HREF links.
3. Two records of the same type cannot be applied to the same character.

4. Dynamic and static highlighting must not be applied to the same text.
5. Dynamic highlighting and linking must not be applied to the same text.
6. Blinking text is optional, particularly when requested in combination with other features.

D.9 File Identification

3GPP multimedia files can be identified using several mechanisms. When stored in traditional computer file systems, these files should be given the file extension ".3gp" (readers should allow mixed case for the alphabetic characters). The MIME types "video/3gpp" (for visual or audio/visual content, where visual includes both video and timed text) and "audio/3gpp" (for purely audio content) are expected to be registered and used.

A file-type ~~atom~~box, as defined in the ~~JPEG 2000~~ISO base media file format specification [50~~36~~] shall be present in conforming files. The file type box 'ftyp' shall occur before any variable-length box (e.g. movie, free space, media data). Only a fixed-size box such as a file signature, if required, may precede it.

The brand identifier for this specification is '3gp5'. This brand identifier must occur in the compatible brands list, and may also be the primary brand. If the file is also conformant to release 4 of this specification, it is recommended that the Release 4 brand '3gp4' also occur in the compatible brands list; if 3gp4 is not in the compatible brand list the file will not be processed by a Release 4 reader. Readers should check the compatible brands list for the identifiers they recognize, and not rely on the file having a particular primary brand, for maximum compatibility. Files may be compatible with more than one brand, and have a 'best use' other than this specification, yet still be compatible with this specification.

Table D.9: The File-Type ~~atom~~box

Field	Type	Details	Value
~~Atom~~Box**Header**.Size	Unsigned int(32)		
~~Atom~~Box**Header**.Type	Unsigned int(32)		'ftyp'
Brand	Unsigned int(32)	The major or 'best use' of this file	
MinorVersion	Unsigned int(32)		
CompatibleBrands	Unsigned int(32)	A list of brands, to end of the ~~atom~~box	

Brand: Identifies the 'best use' of this file. The brand should match the file extension. For files with extension '.3gp' and conforming to this specification, the brand shall be '3gp~~4~~5'.

MinorVersion: This identifies the minor version of the brand. For files with brand '3gpZ', where Z is a digit, and conforming to release Z.x.y, this field takes the value x*256 + y.

CompatibleBrands: a list of brand identifiers (to the end of the ~~atom~~box). '3gp5' shall be a member of this list.

3GPP TSG-SA4 Meeting #23 **Tdoc ⌘S4-020626**
Montreal, Canada, 30 September–4 October 2002

CR-Form-v7

CHANGE REQUEST

⌘ **26.234** **CR** **042** ⌘ **rev** **2** ⌘ Current version: **4.4.0** ⌘

*For **HELP** on using this form, see bottom of this page or look at the pop-up text over the ⌘ symbols.*

Proposed change affects: UICC apps⌘ ☐ ME ☒ Radio Access Network ☐ Core Network ☐

Title:	⌘	Correction regarding undefined IPv6 support in SDP
Source:	⌘	TSG SA WG4
Work item code:	⌘	PSTREAM
Date:	⌘	10/12/2002
Category:	⌘	**F**
Release:	⌘	Rel-4

Use one of the following categories:
F *(correction)*
A *(corresponds to a correction in an earlier release)*
B *(addition of feature),*
C *(functional modification of feature)*
D *(editorial modification)*
Detailed explanations of the above categories can be found in 3GPP TR 21.900.

Use one of the following releases:

2	*(GSM Phase 2)*
R96	*(Release 1996)*
R97	*(Release 1997)*
R98	*(Release 1998)*
R99	*(Release 1999)*
Rel-4	*(Release 4)*
Rel-5	*(Release 5)*
Rel-6	*(Release 6)*

Reason for change:	⌘	In June 2002, RFC 3266, "Support for IPv6 in Session Description Protocol" has been published by IETF. This RFC defines how IPv6 addresses shall be handled in an SDP description. To make sure that PSS fully follows IETF standards when used over IPv6 operated network, IPv6 support in SDP must follow the same RFC.
Summary of change:	⌘	- A text added to the SDP section, related to the IPv6 support - A reference to RFC 3266 is included in the References section.
Consequences if not approved:	⌘	IPv6 support in PSS SDP descriptions remains undefined. 3GPP PSS implementations will not by default follow the standard way defined by IETF – potentially causing interoperability problems.

Clauses affected:	⌘	2, 5.3.3			
Other specs affected:	⌘	**Y**	**N**		
			N	Other core specifications	⌘
			N	Test specifications	
			N	O&M Specifications	
Other comments:	⌘				

How to create CRs using this form:
Comprehensive information and tips about how to create CRs can be found at http://www.3gpp.org/specs/CR.htm.
Below is a brief summary:

1) Fill out the above form. The symbols above marked ⌘ contain pop-up help information about the field that they are closest to.

2) Obtain the latest version for the release of the specification to which the change is proposed. Use the MS Word "revision marks" feature (also known as "track changes") when making the changes. All 3GPP specifications can be downloaded from the 3GPP server under ftp://ftp.3gpp.org/specs/ For the latest version, look for the directory name with the latest date e.g. 2001-03 contains the specifications resulting from the March 2001 TSG meetings.

3) With "track changes" disabled, paste the entire CR form (use CTRL-A to select it) into the specification just in front of the clause containing the first piece of changed text. Delete those parts of the specification which are not relevant to the change request.

2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

[1] (void)

[2] 3GPP TS 26.233: "End-to-end transparent streaming service; General description".

[3] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications".

[4] IETF RFC 1738: "Uniform Resource Locators (URL)", Berners-Lee, Masinter & McCahill, December 1994.

[5] IETF RFC 2326: "Real Time Streaming Protocol (RTSP)", Schulzrinne H., Rao A. and Lanphier R., April 1998.

[6] IETF RFC 2327: "SDP: Session Description Protocol", Handley M. and Jacobson V., April 1998.

[7] IETF STD 0006: "User Datagram Protocol", Postel J., August 1980.

[8] IETF STD 0007: "Transmission Control Protocol", Postel J., September 1981.

[9] IETF RFC 1889: "RTP: A Transport Protocol for Real-Time Applications", Schulzrinne H. et al., January 1996.

[10] IETF RFC 1890: "RTP Profile for Audio and Video Conferences with Minimal Control", Schulzrinne H. et al., January 1996.

[11] IETF RFC 3267: " RTP payload format and file storage format for the Adaptive Multi-Rate (AMR) Adaptive Multi-Rate Wideband (AMR-WB) audio codecs ", March 2002.

[12] void

[13] IETF RFC 3016: "RTP Payload Format for MPEG-4 Audio/Visual Streams", Kikuchi Y. et al., November 2000.

[14] IETF RFC 2429: "RTP Payload Format for the 1998 Version of ITU-T Rec. H.263 Video (H.263+)", Bormann C. et al., October 1998.

[15] IETF RFC 2046: "Multipurpose Internet Mail Extensions (MIME) Part Two: Media Types", N. Freed, N. Borenstein, November 1996.

[16] IETF RFC 3236: " The 'application/xhtml+xml' Media Type ", Baker M. and Stark P., January 2002.

[17] IETF RFC 2616: "Hypertext Transfer Protocol - HTTP/1.1", Fielding R. et al., June 1999.

[18] 3GPP TS 26.071: "Mandatory Speech Codec speech processing functions; AMR Speech Codec; General description".

[19] 3GPP TS 26.101: "Mandatory Speech Codec speech processing functions; AMR Speech Codec; Frame Structure".

[20] ITU-T Recommendation G.722.2 (2002) Wideband coding of speech at around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR-WB).

[21] ISO/IEC 14496-3:2001, "Information technology -- Coding of audio-visual objects -- Part 3: Audio".

[22] ITU-T Recommendation H.263: "Video coding for low bit rate communication".

[23] ITU-T Recommendation H.263 (annex X): "Annex X, Profiles and levels definition".

[24] ISO/IEC 14496-2:2001, "Information technology -- Coding of audio-visual objects -- Part 2: Visual".

[25] ISO/IEC 14496-2:2001/Amd 2:2002, "Streaming video profile".

[26] ITU-T Recommendation T.81 (1991) | ISO/IEC 10918-1 (1992): "Information technology - Digital compression and coding of continuous-tone still images - Requirements and guidelines.

[27] "JPEG File Interchange Format", Version 1.02, September 1, 1992.

[28] W3C Recommendation: "XHTML Basic", http://www.w3.org/TR/2000/REC-xhtml-basic-20001219, December 2000

[29] ISO/IEC 10646-1 (2000): "Information technology - Universal Multiple-Octet Coded Character Set (UCS) - Part 1: Architecture and Basic Multilingual Plane".

[30] The Unicode Consortium: "The Unicode Standard", Version 3.0 Reading, MA, Addison-Wesley Developers Press, 2000, ISBN 0-201-61633-5.

[31] W3C Recommendation: "Synchronized Multimedia Integration Language (SMIL 2.0)", http://www.w3.org/TR/2001/REC-smil20-20010807/, August 2001.

[32] CompuServe Incorporated: "GIF Graphics Interchange Format: A Standard defining a mechanism for the storage and transmission of raster-based graphics information", Columbus, OH, USA, 1987.

[33] CompuServe Incorporated: "Graphics Interchange Format: Version 89a", Columbus, OH, USA, 1990.

[34] ISO/IEC 14496-1 (2001): "Information technology - Coding of audio-visual objects - Part 1: Systems".

[35] 3GPP TS 23.140: "Multimedia Messaging Service (MMS), Functional description stage 2/3".

[36] ISO/IEC 15444-1 (2000): "Information technology - JPEG 2000 image coding system: Core coding system; Annex I: The JPEG 2000 file format".

[37] ~~(void)~~ IETF RFC 3266: "Support For IPv6 in Session Description Protocol (SDP)", Olson S., Camarillo G. and Roach A. B., June 2002

[38] (void)

5.3.3 SDP

RTSP requires a presentation description. SDP shall be used as the format of the presentation description for both PSS clients and servers. PSS servers shall provide and clients interpret the SDP syntax according to the SDP specification [6] and appendix C of [5]. The SDP delivered to the PSS client shall declare the media types to be used in the session using a codec specific MIME media type for each media. MIME media types to be used in the SDP file are described in clause 5.4 of the present document.

The SDP [6] specification requires certain fields to always be included in an SDP file. Apart from this a PSS server shall always include the following fields in the SDP:

- "a=control:" according to clauses C.1.1, C.2 and C.3 in [5];
- "a=range:" according to clause C.1.5 in [5];
- "a=rtpmap:" according to clause 6 in [6];
- "a=fmtp:" according to clause 6 in [6].

The bandwidth field in SDP can be used to indicate to the PSS client the amount of bandwidth that is required for the session and the individual media in the presentation. Therefore, a PSS server should include the "b=AS:" field in the SDP (both on the session and media level) and a PSS client shall be able to interpret this field. For RTP based applications, AS gives the RTP "session bandwidth" (including UDP/IP overhead) as defined in section 6.2 of [9].

IPv6 addresses in SDP descriptions shall be supported according to RFC 3266[37].

NOTE: The SDP parsers and/or interpreters shall be able to accept NULL values in the 'c=' field (e.g. 0.0.0.0 in IPv4 case). This may happen when the media content does not have a fixed destination address. For more details, see Section C.1.7 of [5] and Section 6 of [6].

3GPP TSG-SA Meeting S4#24
Redmond, USA, Nov 11-15, 2002

Tdoc S4-020723

CR-Form-v7

CHANGE REQUEST

⌘ **26.234** **CR** **043** ⌘ **rev** **2** ⌘ Current version: **4.4.0** ⌘

*For **HELP** on using this form, see bottom of this page or look at the pop-up text over the ⌘ symbols.*

Proposed change affects: UICC apps⌘ ☐ ME ☒ Radio Access Network ☐ Core Network ☐

Title:	⌘	SMIL Authoring Instructions (Rel.4)		
Source:	⌘	TSG SA WG4		
Work item code:	⌘	PSTREAM	***Date:*** ⌘	10/12/2002
Category:	⌘	**F**	***Release:*** ⌘	Rel-4

Use one of the following categories:
F *(correction)*
A *(corresponds to a correction in an earlier release)*
B *(addition of feature),*
C *(functional modification of feature)*
D *(editorial modification)*
Detailed explanations of the above categories can be found in 3GPP TR 21.900.

Use one of the following releases:

2	*(GSM Phase 2)*
R96	*(Release 1996)*
R97	*(Release 1997)*
R98	*(Release 1998)*
R99	*(Release 1999)*
Rel-4	*(Release 4)*
Rel-5	*(Release 5)*
Rel-6	*(Release 6)*

Reason for change: ⌘

Layout is a major interoperability issue when composing SMIL content for a wide range of devices. One of the main reasons are the different resolution of the display surface used for rendering SMIL content. These issues can not be addressed in the SMIL language profile specification because it is not a problem of the language but an issue how the content is authored.

Interoperability issues can be largely avoided if authors create their SMIL content in a correct way. The purpose of authoring instructions is to recommend authors how to create SMIL content that will work well with a wide range of devices (and thereby expose good interoperability). The SMIL authoring instructions in Annex B are informative.

Experience gained since writing the original SMIL specification (i.e. chapter 8 and annex B) shows that the SMIL authoring instructions regarding SMIL layout in TS 26.234-410 section B.3 are insufficient and partly misleading. For this reason we request this section to be updated as detailed below.

Summary of change: ⌘

1. SMIL ContentControl allow definition of multiple layouts in a SMIL presentation. The SMIL player will choose the most suitable layout based in the rules encoded in the SMIL presentation. E.g. the SMIL presentation can define that the layout be choosen based on the size of the rendering surface of the device. Recommendation to use ContentControl is essential for defining the layout because this is the most powerful tool in SMIL to make presentation layout adapt to a device. Such recommendation needs to be added to section B.3.

2. The recommendations how to use the fit attribute are partly counter-productive for authoring SMIL presentations that work well accorss devices

		with different form factors. Concretely, a SMIL presentation that sets the fit attribute to 'hidden' will scale badly over different display sizes. Requesting the player to scale the content (i.e. setting the fit attribute to "meet") is a much better recommendation. According to SMIL2 specifications a player than can not scale the content must use the default option, i.e. it will clip the content anyhow. Therefore, the recommendation to authors should generally be to use scaling for media objects. Text is a special case as explained below. The other exception that could be made for this recommendation is video. 3. As currently recommended scrolling content (fit="scroll") should generally not be used. However, from the experience with MMS text scrolling is needed. The reasons are as follows: (1) a SMIL player should not scale text (i.e. the font size should not be changed from the default), and (2) an author can not predict how much space some text will take when rendered on a specific device because of the diverging default font and font metrix on different devices. Hence, the recommendation is to use fit="scroll" for text. This change is particularly relevant for SMIL use in MMS.
Consequences if not approved:	⌘	Sub-optimal and partly even misleading recommendations will be given to authors on how to create SMIL content. This may lead to sub-optimal rendering of SMIL presentation in the terminal compared to situation where SMIL content was authored correctly.

Clauses affected:	⌘	B.3			
Other specs affected:	⌘	**Y** / **N**		⌘	
			X Other core specifications		
			X Test specifications		
			X O&M Specifications		
Other comments:	⌘	Annex B.3 is identical in Release 4 and Release 5. Therefore, the same changes need to be made to Release 5 of this TS. A separate companion CR has been submitted for this purpose.			

How to create CRs using this form:

Comprehensive information and tips about how to create CRs can be found at http://www.3gpp.org/specs/CR.htm. Below is a brief summary:

1) Fill out the above form. The symbols above marked ⌘ contain pop-up help information about the field that they are closest to.

2) Obtain the latest version for the release of the specification to which the change is proposed. Use the MS Word "revision marks" feature (also known as "track changes") when making the changes. All 3GPP specifications can be downloaded from the 3GPP server under ftp://ftp.3gpp.org/specs/ For the latest version, look for the directory name with the latest date e.g. 2001-03 contains the specifications resulting from the March 2001 TSG meetings.

3) With "track changes" disabled, paste the entire CR form (use CTRL-A to select it) into the specification just in front of the clause containing the first piece of changed text. Delete those parts of the specification which are not relevant to the change request.

B.3 BasicLayout

When defining the layout of a SMIL presentation, a content author needs to be aware that the targeted devices might have diverse properties that effect how the content can be rendered. The different sizes of the display area that can be used to render content on the targeted devices should be considered for defining the layout of the SMIL presentation. The root-layout window might represent the entire display or only parts of it.

Content authors are encouraged to create SMIL presentations that will work well with different resolutions of the rendering area. As mentioned in the SMIL2 recommendation content authors should use SMIL ContentControl functionality for defining multiple layouts for their SMIL presentation that are tailored to the specific needs of the whole range of targeted devices. Furthermore, authors should include a default layout (i.e. a layout determined by the SMIL player) that will be used when none of the author-defined layouts can be used.

Using relative position and size attributes in the definition of a region is also helpful for making SMIL presentations more portable across different display sizes; these features should also be used.

A 3GPP SMIL player should use the layout definition of a SMIL presentation for presenting the content whenever possible. When the SMIL player fails to use the layout information defined by the author it is free to present the content using a layout it determines by itself.

The "fit" attribute defines how different media should be fitted into their respective display regions.

The rendering and layout of some objects on a small display might be difficult and all mobile devices may not support features such as scroll bars~~; in addition, the root-layout window may represent the full screen of the display~~. Therefore "fit=scroll" should not be used except for text content.

Due to hardware restrictions in mobile devices, operations such that scaling of a video sequence~~, or even images~~, may be very difficult to achieve. According to the SMIL 2.0 specification SMIL players may in these situations clip the content instead. To be sure of that the presentation is displayed as the author intended, video content should be encoded in a size suitable for the terminals intended and it is recommended to use "fit=hidden".

3GPP TSG-SA Meeting S4#24
Redmond, USA, Nov 11-15, 2002

Tdoc S4-020634

CR-Form-v7

CHANGE REQUEST

⌘ **26.234** **CR** **044** ⌘ **rev** **1** ⌘ Current version: **5.2.0** ⌘

*For **HELP** on using this form, see bottom of this page or look at the pop-up text over the ⌘ symbols.*

Proposed change affects: UICC apps⌘ ☐ ME ☒ Radio Access Network ☐ Core Network ☐

Title:	⌘	SMIL Authoring Instructions (Rel.5)			
Source:	⌘	TSG SA WG4			
Work item code:	⌘	PSS-E	***Date:*** ⌘	10/12/2002	
Category:	⌘	**A**	***Release:*** ⌘	Rel-5	
		Use one of the following categories:		*Use one of the following releases:*	
		F *(correction)*		*2*	*(GSM Phase 2)*
		A *(corresponds to a correction in an earlier release)*		*R96*	*(Release 1996)*
		B *(addition of feature),*		*R97*	*(Release 1997)*
		C *(functional modification of feature)*		*R98*	*(Release 1998)*
		D *(editorial modification)*		*R99*	*(Release 1999)*
		Detailed explanations of the above categories can be found in 3GPP TR 21.900.		*Rel-4*	*(Release 4)*
				Rel-5	*(Release 5)*
				Rel-6	*(Release 6)*

Reason for change:	⌘	Layout is a major interoperability issue when composing SMIL content for a wide range of devices. One of the main reasons are the different resolution of the display surface used for rendering SMIL content. These issues can not be addressed in the SMIL language profile specification because it is not a problem of the language but an issue how the content is authored. Interoperability issues can be largely avoided if authors create their SMIL content in a correct way. The purpose of authoring instructions is to recommend authors how to create SMIL content that will work well with a wide range of devices (and thereby expose good interoperability). The SMIL authoring instructions in Annex B are informative. Experience gained since writing the original SMIL specification (i.e. chapter 8 and annex B) shows that the SMIL authoring instructions regarding SMIL layout in TS 26.234 section B.3 are insufficient and partly misleading. For this reason we request this section to be updated as detailed below.
Summary of change:	⌘	1. SMIL ContentControl allow definition of multiple layouts in a SMIL presentation. The SMIL player will choose the most suitable layout based in the rules encoded in the SMIL presentation. E.g. the SMIL presentation can define that the layout be choosen based on the size of the rendering surface of the device. Recommendation to use ContentControl is essential for defining the layout because this is the most powerful tool in SMIL to make presentation layout adapt to a device. Such recommendation needs to be added to section B.3. 2. The recommendations how to use the fit attribute are partly counter-productive for authoring SMIL presentations that work well accorss devices

		with different form factors. Concretely, a SMIL presentation that sets the fit attribute to 'hidden' will scale badly over different display sizes. Requesting the player to scale the content (i.e. setting the fit attribute to "meet") is a much better recommendation. According to SMIL2 specifications a player than can not scale the content must use the default option, i.e. it will clip the content anyhow. Therefore, the recommendation to authors should generally be to use scaling for media objects. Text is a special case as explained below. The other exception that could be made for this recommendation is video. 3. As currently recommended scrolling content (fit="scroll") should generally not be used. However, from the experience with MMS text scrolling is needed. The reasons are as follows: (1) a SMIL player should not scale text (i.e. the font size should not be changed from the default), and (2) an author can not predict how much space some text will take when rendered on a specific device because of the diverging default font and font metrix on different devices. Hence, the recommendation is to use fit="scroll" for text. This change is particularly relevant for SMIL use in MMS.
Consequences if not approved:	⌘	Sub-optimal and partly even misleading recommendations will be given to authors on how to create SMIL content. This may lead to sub-optimal rendering of SMIL presentation in the terminal compared to situation where SMIL content was authored correctly.

Clauses affected:	⌘	B.3			
Other specs affected:	⌘	**Y** / **N**	Other core specifications	⌘	
			Test specifications		
		X	O&M Specifications		
Other comments:	⌘	Annex B.3 is identical in Release 4 and Release 5. Therefore, the same changes need to be made to release 4 of this TS. A separate CR has been submitted for this purpose.			

How to create CRs using this form:

Comprehensive information and tips about how to create CRs can be found at http://www.3gpp.org/specs/CR.htm. Below is a brief summary:

1) Fill out the above form. The symbols above marked ⌘ contain pop-up help information about the field that they are closest to.

2) Obtain the latest version for the release of the specification to which the change is proposed. Use the MS Word "revision marks" feature (also known as "track changes") when making the changes. All 3GPP specifications can be downloaded from the 3GPP server under ftp://ftp.3gpp.org/specs/ For the latest version, look for the directory name with the latest date e.g. 2001-03 contains the specifications resulting from the March 2001 TSG meetings.

3) With "track changes" disabled, paste the entire CR form (use CTRL-A to select it) into the specification just in front of the clause containing the first piece of changed text. Delete those parts of the specification which are not relevant to the change request.

B.3 BasicLayout

When defining the layout of a SMIL presentation, a content author needs to be aware that the targeted devices might have diverse properties that effect how the content can be rendered. The different sizes of the display area that can be used to render content on the targeted devices should be considered for defining the layout of the SMIL presentation. The root-layout window might represent the entire display or only parts of it.

Content authors are encouraged to create SMIL presentations that will work well with different resolutions of the rendering area. As mentioned in the SMIL2 recommendation content authors should use SMIL ContentControl functionality for defining multiple layouts for their SMIL presentation that are tailored to the specific needs of the whole range of targeted devices. Furthermore, authors should include a default layout (i.e. a layout determined by the SMIL player) that will be used when none of the author-defined layouts can be used.

Using relative position and size attributes in the definition of a region is also helpful for making SMIL presentations more portable across different display sizes; these features should also be used.

A 3GPP SMIL player should use the layout definition of a SMIL presentation for presenting the content whenever possible. When the SMIL player fails to use the layout information defined by the author it is free to present the content using a layout it determines by itself.

The "fit" attribute defines how different media should be fitted into their respective display regions.

The rendering and layout of some objects on a small display might be difficult and all mobile devices may not support features such as scroll bars~~; in addition, the root-layout window may represent the full screen of the display~~. Therefore "fit=scroll" should not be used except for text content.

Due to hardware restrictions in mobile devices, operations such that scaling of a video sequence~~, or even images~~, may be very difficult to achieve. According to the SMIL 2.0 specification SMIL players may in these situations clip the content instead. To be sure of that the presentation is displayed as the author intended, video content should be encoded in a size suitable for the terminals intended and it is recommended to use "fit=hidden".

3GPP TSG-SA4 Meeting #23 **Tdoc ⌘S4-020588**
Montreal, Canada, 30 September–4 October 2002

CR-Form-v7

CHANGE REQUEST

⌘ **26.234** **CR** **045** ⌘ **rev** **1** ⌘ Current version: **5.2.0** ⌘

*For **HELP** on using this form, see bottom of this page or look at the pop-up text over the ⌘ symbols.*

Proposed change affects: UICC apps⌘ ☐ ME ☒ Radio Access Network ☐ Core Network ☐

Title:	⌘	Client usage of bandwidth parameter at the media level in SDP
Source:	⌘	TSG SA WG4
Work item code:	⌘	PSS-E
Date:	⌘	10/12/2002
Category:	⌘	**F**
Release:	⌘	Rel-5

Use one of the following categories:
F *(correction)*
A *(corresponds to a correction in an earlier release)*
B *(addition of feature),*
C *(functional modification of feature)*
D *(editorial modification)*
Detailed explanations of the above categories can be found in 3GPP TR 21.900.

Use one of the following releases:

2	*(GSM Phase 2)*
R96	*(Release 1996)*
R97	*(Release 1997)*
R98	*(Release 1998)*
R99	*(Release 1999)*
Rel-4	*(Release 4)*
Rel-5	*(Release 5)*
Rel-6	*(Release 6)*

Reason for change:	⌘	Release-5 PSS states that servers must provide bandwidth information ('b=AS:') at the media level in SDP. However, a Release-5 client cannot rely on receiving this parameter, since a Release-4 server is not required to provide it. This point is not currently clear in the PSS specification.
Summary of change:	⌘	It is clarified that the PSS client must handle the case where the bandwidth information is not provided by a (Release-4) server.
Consequences if not approved:	⌘	The lack of clarity on this point may result in Release-5 client implementations which do not handle the case where the bandwidth parameter is not provided, and such clients may not interwork with Release-4 servers.

Clauses affected:	⌘	5.3.3.1

	Y	N		
Other specs affected: ⌘		N	Other core specifications	⌘
		N	Test specifications	
		N	O&M Specifications	

Other comments: ⌘

How to create CRs using this form:
Comprehensive information and tips about how to create CRs can be found at http://www.3gpp.org/specs/CR.htm. Below is a brief summary:

1) Fill out the above form. The symbols above marked ⌘ contain pop-up help information about the field that they are closest to.

2) Obtain the latest version for the release of the specification to which the change is proposed. Use the MS Word "revision marks" feature (also known as "track changes") when making the changes. All 3GPP specifications can be

downloaded from the 3GPP server under ftp://ftp.3gpp.org/specs/ For the latest version, look for the directory name with the latest date e.g. 2001-03 contains the specifications resulting from the March 2001 TSG meetings.

3) With "track changes" disabled, paste the entire CR form (use CTRL-A to select it) into the specification just in front of the clause containing the first piece of changed text. Delete those parts of the specification which are not relevant to the change request.

…

5.3.3 SDP

5.3.3.1 General

RTSP requires a presentation description. SDP shall be used as the format of the presentation description for both PSS clients and servers. PSS servers shall provide and clients interpret the SDP syntax according to the SDP specification [6] and appendix C of [5]. The SDP delivered to the PSS client shall declare the media types to be used in the session using a codec specific MIME media type for each media. MIME media types to be used in the SDP file are described in clause 5.4 of the present document.

The SDP [6] specification requires certain fields to always be included in an SDP file. Apart from this a PSS server shall always include the following fields in the SDP:

- "a=control:" according to clauses C.1.1, C.2 and C.3 in [5];
- "a=range:" according to clause C.1.5 in [5];
- "a=rtpmap:" according to clause 6 in [6];
- "a=fmtp:" according to clause 6 in [6].

The bandwidth field in SDP is needed by the client in order to properly set up QoS parameters. Therefore, a PSS server shall include the “b=AS:” field at the media level for each media stream in SDP, and a PSS client shall interpret this field. When a PSS client receives SDP, it should ignore the session level “b=AS:” parameter (if present), and instead calculate session bandwidth from the media level bandwidth values of the relevant streams. A PSS client shall also handle the case where the bandwidth parameter is not present, since this may occur when connecting to a Release-4 server.

Note that for RTP based applications , ‘b=AS:’ gives the RTP "session bandwidth" (including UDP/IP overhead) as defined in section 6.2 of [9].

NOTE: The SDP parsers and/or interpreters shall be able to accept NULL values in the 'c=' field (e.g. 0.0.0.0 in IPv4 case). This may happen when the media content does not have a fixed destination address. For more details, see Section C.1.7 of [5] and Section 6 of [6].

5.3.3.2 Additional SDP fields

…

3GPP TSG-SA Meeting S4#24 ***Tdoc* S4-020699**
November 11 – 15, 2002, Redmond, WA, USA

CR-Form-v7

CHANGE REQUEST

⌘ **26.234** **CR** **046** ⌘ **rev** **1** ⌘ Current version: **4.4.0** ⌘

*For **HELP** on using this form, see bottom of this page or look at the pop-up text over the ⌘ symbols.*

Proposed change affects: UICC apps⌘ ☐ ME ☒ Radio Access Network ☐ Core Network ☐

Title:	⌘	CR on SMIL Language Profile		
Source:	⌘	TSG SA WG4		
Work item code:	⌘	PSTREAM	***Date:*** ⌘	10/12/2002
Category:	⌘	**F**	***Release:*** ⌘	Rel-4

Use one of the following categories:
F *(correction)*
A *(corresponds to a correction in an earlier release)*
B *(addition of feature),*
C *(functional modification of feature)*
D *(editorial modification)*
Detailed explanations of the above categories can be found in 3GPP TR 21.900.

Use one of the following releases:

2	*(GSM Phase 2)*
R96	*(Release 1996)*
R97	*(Release 1997)*
R98	*(Release 1998)*
R99	*(Release 1999)*
Rel-4	*(Release 4)*
Rel-5	*(Release 5)*
Rel-6	*(Release 6)*

Reason for change: ⌘

The purpose for this CR is to clarify on the use of the LinkingAttribute module in the 3GPP SMIL language profile specification (clause 8.1 of TS 26.234)

Even though the LinkingAttributes module is clearly used in the 3GPP SMIL language profile the module is missing from the list of of modules in section 8.2.1. The specification would be clearer if this inconsistency was removed by adding the LinkingAttributes module to that list. The module should also be mentioned in section 8.2.4.3 .

We believe this inconsistemcy is due to an earlier oversight that some SMIL attributes are not included in the specification of the BasicLinking module but defined in their own module, i.e. LinkingAttributes module. The attributes of the LinkingAttributes module are 'sourcePlaystate', 'DestinationPlaystate', 'show', 'sourceLevel', 'destinationLevel', 'external', and 'actuate'.

The reasons why it is clear that the LinkingAttributes module is included in 3GPP SMIL are as follows:

- BasicLinking module is included in 3GPP SMIL.2 The SMIL 2.0 REC [31] defines that inclusion of BasicLinking requires also including LinkingAttributes module (see section 6.5 of [31]). BasicLinking relies on attributes defined by the LinkingAttributes module.
- Section 8.2.4.3 of TS 26.234 "Linking Module" includes specification on the 'sourcePlayState' and 'show' attributes.
- The content model in section 8.2.5 allows 'LINKING-ATTS' set of attributes for the 'a' and 'area' elements. 'LINKING-ATTS' includes the

		attributes defined by LinkingAttributes module (section A.1.4 of [31]).
Summary of change:	⌘	Apply the following minimal additions to the specification: • Add LinkingAttributes module name in the list of modules in section 8.2.1. • Mention LinkingAttributes module name in section 8.2.4.3.
Consequences if not approved:	⌘	Without this clarification the aforementioned inconsistencies in the specification would remain.

Clauses affected:	⌘	Clause 8.2.1 an clause 8.2.4.3			
Other specs affected:	⌘	**Y** / **N** **N**	Other core specifications	⌘	
		N	Test specifications		
		N	O&M Specifications		
Other comments:	⌘	An identical CR needs to be applied to release 5 of TS 26.234. The initial revision of this CR was accidently based on 26.234 rel.5. We are sorry for any inconvenience this might have caused to the group.			

How to create CRs using this form:

Comprehensive information and tips about how to create CRs can be found at http://www.3gpp.org/specs/CR.htm. Below is a brief summary:

1) Fill out the above form. The symbols above marked ⌘ contain pop-up help information about the field that they are closest to.

2) Obtain the latest version for the release of the specification to which the change is proposed. Use the MS Word "revision marks" feature (also known as "track changes") when making the changes. All 3GPP specifications can be downloaded from the 3GPP server under ftp://ftp.3gpp.org/specs/ For the latest version, look for the directory name with the latest date e.g. 2001-03 contains the specifications resulting from the March 2001 TSG meetings.

3) With "track changes" disabled, paste the entire CR form (use CTRL-A to select it) into the specification just in front of the clause containing the first piece of changed text. Delete those parts of the specification which are not relevant to the change request.

8.2 3GPP PSS4 SMIL Language Profile

8.2.1 Introduction

3GPP PSS4 SMIL is a markup language based on SMIL Basic [31] and SMIL Scalability Framework.

3GPP PSS4 SMIL shall consist of the modules required by SMIL Basic Profile (and SMIL 2.0 Host Language Conformance) and additional MediaAccessibility, MediaDescription, MediaClipping, MetaInformation, PrefetchControl and EventTiming modules. All in all the following modules are included:

- SMIL 2.0 Content Control Modules – BasicContentControl, SkipContentControl and PrefetchControl
- SMIL 2.0 Layout Module -- BasicLayout
- SMIL 2.0 Linking Module – BasicLinking, LinkingAttributes
- SMIL 2.0 Media Object Modules – BasicMedia, MediaClipping, MediaAccessibility and MediaDescription
- SMIL 2.0 Metainformation Module -- Metainformation
- SMIL 2.0 Structure Module -- Structure
- SMIL 2.0 Timing and Synchronization Modules -- BasicInlineTiming, MinMaxTiming, BasicTimeContainers, RepeatTiming and EventTiming

8.2.2 Document Conformance

A conforming 3GPP PSS4 SMIL document shall be a conforming SMIL 2.0 document.

All 3GPP PSS4 SMIL documents use SMIL 2.0 namespace.

<smil xmlns="`http://www.w3.org/2001/SMIL20/Language`">

3GPP PSS4 SMIL documents may declare requirements using systemRequired attribute:

EXAMPLE 1:
```
<smil xmlns="http://www.w3.org/2001/SMIL20/Language"
      xmlns:EventTiming="http://www.w3.org/2000/SMIL20/CR/EventTiming "
      systemRequired="EventTiming">
```

Namespace URI http://www.3gpp.org/SMIL20/PSS4/ identifies the 3GPP PSS4 SMIL. Authors can use this URI to indicate requirement for exact 3GPP PSS4 SMIL semantics for a document or a subpart of a document:

EXAMPLE 2:
```
<smil xmlns="http://www.w3.org/2001/SMIL20/Language"
      xmlns:pss4="http://www.3gpp.org/SMIL20/PSS4/"
      systemReqzuired="pss4">
```

The content authors generally should choose not to include the PSS requirement in the document unless the SMIL document relies on PSS specific semantics that are not part of the W3C SMIL. The reason for this is that SMIL players that are not conforming 3GPP PSS user agents may not recognize the PSS4 URI and thus refuse to play the document.

8.2.3 User Agent Conformance

A conforming 3GPP PSS4 SMIL user agent shall be a conforming SMIL Basic User Agent.

A conforming user agent shall implement the semantics of the language as described in this document.

A conforming user agent shall recognize the URIs of all included SMIL 2.0 modules. It shall also recognize URI http://www.3gpp.org/SMIL20/PSS4/ as referring to all modules and semantics of 3GPP SMIL language.

8.2.4 3GPP SMIL Language Profile

3GPP PSS4 SMIL is based on SMIL 2.0 Basic language profile [31]. This chapter defines the content model and integration semantics of the included modules where they differ from those defined by SMIL Basic.

8.2.4.1 Content Control Modules

3GPP PSS4 SMIL shall include the content control functionality of the BasicContentControl, SkipContentControl and PrefetchControl modules of SMIL 2.0. PrefetchControl is not part of SMIL Basic and is an additional module in this profile.

All BasicContentControl attributes listed in the module specification shall be supported.

NOTE: The SMIL specification [31] defines that all functionality of PrefetchControl module is optional. This mean that even that PrefetchControl is mandatory user agents may implement semantics of PrefetchControl module only partially or not to implement them at all. PrefetchControl module adds the **prefetch** element to the content model of SMIL Basic **body, switch, par** and **seq** elements.

The **prefetch** element has the attributes defined by the PrefetchControl module (**mediaSize, mediaTime and bandwidth**), the **src** attribute, the BasicContentControl attributes and the **skip-content** attribute.

8.2.4.2 Layout Module

3GPP PSS4 SMIL shall use the BasicLayout module of SMIL 2.0 for spatial layout. The module is part of SMIL Basic.

Default values of the width and height attributes for root-layout shall be the dimensions of the device display area.

8.2.4.3 Linking Module

3GPP PSS4 SMIL shall use the SMIL 2.0 BasicLinking and LinkingAttributes modules for providing hyperlinks between documents and document fragments. ~~This~~ The BasicLinking module is from SMIL Basic.

When linking to destinations outside the current document, implementations may ignore values "play" and "pause" of the 'sourcePlaystate' attribute and values "new" and "pause" of the 'show' attribute, instead using the semantics of values "stop" and "replace" respectively. When the values of 'sourcePlaystate' and 'show' are ignored the player may also ignore the 'sourceLevel' attribute since it is of no use then

3GPP TSG-SA Meeting S4#24 ***Tdoc* S4-020700**
November 11 – 15, 2002, Redmond, WA, USA

CR-Form-v7

CHANGE REQUEST

⌘ **26.234** **CR** **047** ⌘ **rev** **1** ⌘ Current version: **5.2.0** ⌘

*For **HELP** on using this form, see bottom of this page or look at the pop-up text over the ⌘ symbols.*

Proposed change affects: UICC apps⌘ ☐ ME ☒ Radio Access Network ☐ Core Network ☐

Title:	⌘	CR on SMIL Language Profile		
Source:	⌘	TSG SA WG4		
Work item code:	⌘	PSS-E	***Date:*** ⌘	10/12/2002
Category:	⌘	**A**	***Release:*** ⌘	Rel-5

Use one of the following categories:
F *(correction)*
A *(corresponds to a correction in an earlier release)*
B *(addition of feature),*
C *(functional modification of feature)*
D *(editorial modification)*
Detailed explanations of the above categories can be found in 3GPP TR 21.900.

Use one of the following releases:

2	*(GSM Phase 2)*
R96	*(Release 1996)*
R97	*(Release 1997)*
R98	*(Release 1998)*
R99	*(Release 1999)*
Rel-4	*(Release 4)*
Rel-5	*(Release 5)*
Rel-6	*(Release 6)*

Reason for change: ⌘

The purpose for this CR is to clarify on the use of the LinkingAttribute module in the 3GPP SMIL language profile specification (clause 8.1 of TS 26.234)

Even though the LinkingAttributes module is clearly used in the 3GPP SMIL language profile the module is missing from the list of of modules in section 8.2.1. The specification would be clearer if this inconsistency was removed by adding the LinkingAttributes module to that list. The module should also be mentioned in section 8.2.4.3 .

We believe this inconsistemcy is due to an earlier oversight that some SMIL attributes are not included in the specification of the BasicLinking module but defined in their own module, i.e. LinkingAttributes module. The attributes of the LinkingAttributes module are 'sourcePlaystate', 'DestinationPlaystate', 'show', 'sourceLevel', 'destinationLevel', 'external', and 'actuate'.

The reasons why it is clear that the LinkingAttributes module is included in 3GPP SMIL are as follows:

- BasicLinking module is included in 3GPP SMIL.2 The SMIL 2.0 REC [31] defines that inclusion of BasicLinking requires also including LinkingAttributes module (see section 6.5 of [31]). BasicLinking relies on attributes defined by the LinkingAttributes module.
- Section 8.2.4.3 of TS 26.234 "Linking Module" includes specification on the 'sourcePlayState' and 'show' attributes.
- The content model in section 8.2.5 allows 'LINKING-ATTS' set of attributes for the 'a' and 'area' elements. 'LINKING-ATTS' includes the

		attributes defined by LinkingAttributes module (section A.1.4 of [31]).
Summary of change:	⌘	Apply the following minimal additions to the specification: • Add LinkingAttributes module name in the list of modules in section 8.2.1. • Mention LinkingAttributes module name in section 8.2.4.3.
Consequences if not approved:	⌘	Without this clarification the aforementioned inconsistencies in the specification would remain.

Clauses affected:	⌘	Clause 8.2.1 an clause 8.2.4.3			
Other specs affected:	⌘	**Y** / **N**		⌘	
			N Other core specifications		
			N Test specifications		
			N O&M Specifications		
Other comments:	⌘	An identical CR will be be applied to release 4 of TS 26.234. Revision 1 of this CR corrects a minor oversight in section 8.2.4.3.			

How to create CRs using this form:

Comprehensive information and tips about how to create CRs can be found at http://www.3gpp.org/specs/CR.htm. Below is a brief summary:

1) Fill out the above form. The symbols above marked ⌘ contain pop-up help information about the field that they are closest to.

2) Obtain the latest version for the release of the specification to which the change is proposed. Use the MS Word "revision marks" feature (also known as "track changes") when making the changes. All 3GPP specifications can be downloaded from the 3GPP server under ftp://ftp.3gpp.org/specs/ For the latest version, look for the directory name with the latest date e.g. 2001-03 contains the specifications resulting from the March 2001 TSG meetings.

3) With "track changes" disabled, paste the entire CR form (use CTRL-A to select it) into the specification just in front of the clause containing the first piece of changed text. Delete those parts of the specification which are not relevant to the change request.

8.2 3GPP PSS SMIL Language Profile

8.2.1 Introduction

3GPP PSS SMIL is a markup language based on SMIL Basic [31] and SMIL Scalability Framework.

3GPP PSS SMIL consists of the modules required by SMIL Basic Profile (and SMIL 2.0 Host Language Conformance) and additional MediaAccessibility, MediaDescription, MediaClipping, MetaInformation, PrefetchControl, EventTiming and BasicTransitions modules. All of the following modules are included:

- SMIL 2.0 Content Control Modules -- BasicContentControl, SkipContentControl and PrefetchControl
- SMIL 2.0 Layout Module -- BasicLayout
- SMIL 2.0 Linking Module – BasicLinking, LinkingAttributes
- SMIL 2.0 Media Object Modules – BasicMedia, MediaClipping, MediaAccessibility and MediaDescription
- SMIL 2.0 Metainformation Module -- Metainformation
- SMIL 2.0 Structure Module -- Structure
- SMIL 2.0 Timing and Synchronization Modules -- BasicInlineTiming, MinMaxTiming, BasicTimeContainers, RepeatTiming and EventTiming
- SMIL 2.0 Transition Effects Module -- BasicTransitions

8.2.2 Document Conformance

A conforming 3GPP PSS SMIL document shall be a conforming SMIL 2.0 document.

All 3GPP PSS SMIL documents use SMIL 2.0 namespace.

<smil xmlns="http://www.w3.org/2001/SMIL20/Language">

3GPP PSS SMIL documents may declare requirements using systemRequired attribute:

EXAMPLE 1:
```
<smil xmlns="http://www.w3.org/2001/SMIL20/Language"
      xmlns:EventTiming="http://www.w3.org/2000/SMIL20/CR/EventTiming "
      systemRequired="EventTiming">
```

Namespace URI http://www.3gpp.org/SMIL20/PSS5/ identifies the version of the 3GPP PSS SMIL profile described in the present document. Authors may use this URI to indicate requirement for exact 3GPP PSS SMIL semantics for a document or a subpart of a document:

EXAMPLE 2:
```
<smil xmlns="http://www.w3.org/2001/SMIL20/Language"
      xmlns:pss5="http://www.3gpp.org/SMIL20/PSS5/"
      systemRequired="pss5">
```

The content authors should generally not include the PSS requirement in the document unless the SMIL document relies on PSS specific semantics that are not part of the W3C SMIL. The reason for this is that SMIL players that are not conforming 3GPP PSS user agents may not recognize the PSS URI and thus refuse to play the document.

8.2.3 User Agent Conformance

A conforming 3GPP PSS SMIL user agent shall be a conforming SMIL Basic User Agent.

A conforming user agent shall implement the semantics 3GPP PSS SMIL as described in clauses 8.2.4 and 8.2.5 (including subclauses).

A conforming user agent shall recognise

- the URIs of all included SMIL 2.0 modules;
- the URI http://www.3gpp.org/SMIL20/PSS5/ as referring to all modules and semantics of the version of the 3GPP PSS SMIL profile described in the present document;
- the URI http://www.3gpp.org/SMIL20/PSS4/ as referring to all modules and semantics of the 3GPP PSS SMIL profile defined in Release 4 of the present document.

NOTE: The difference between PSS4 and PSS5 is that the BasicTransitions module has been added in PSS5.

8.2.4 3GPP PSS SMIL Language Profile definition

3GPP PSS SMIL is based on SMIL 2.0 Basic language profile [31]. This chapter defines the content model and integration semantics of the included modules where they differ from those defined by SMIL Basic.

8.2.4.1 Content Control Modules

3GPP PSS SMIL includes the content control functionality of the BasicContentControl, SkipContentControl and PrefetchControl modules of SMIL 2.0. PrefetchControl is not part of SMIL Basic and is an additional module in this profile.

All BasicContentControl attributes listed in the module specification shall be supported.

NOTE: The SMIL specification [31] defines that all functionality of PrefetchControl module is optional. This mean that even although PrefetchControl is mandatory user agents may implement semantics of PrefetchControl module only partially or not to implement them at all.

PrefetchControl module adds the **prefetch** element to the content model of SMIL Basic **body, switch, par** and **seq** elements. The **prefetch** element has the attributes defined by the PrefetchControl module (**mediaSize, mediaTime and bandwidth**), the **src** attribute, the BasicContentControl attributes and the **skip-content** attribute.

8.2.4.2 Layout Module

3GPP PSS SMIL includes the BasicLayout module of SMIL 2.0 for spatial layout. The module is part of SMIL Basic.

Default values of the width and height attributes for root-layout shall be the dimensions of the device display area.

8.2.4.3 Linking Module

3GPP PSS SMIL includes the SMIL 2.0 BasicLinking and LinkingAttributes modules for providing hyperlinks between documents and document fragments. ~~This~~ The BasicLinking module is from SMIL Basic.

When linking to destinations outside the current document, implementations may ignore values "play" and "pause" of the 'sourcePlaystate' attribute and values "new" and "pause" of the 'show' attribute, instead using the semantics of values "stop" and "replace" respectively. When the values of 'sourcePlaystate' and 'show' are ignored the player may also ignore the 'sourceLevel' attribute since it is of no use then

3GPP TSG-SA4 Meeting #24
Redmond, USA, 11 –15 November 2002

Tdoc ⌘S4-020724

CR-Form-v7

CHANGE REQUEST

⌘ **26.234** **CR** **050** ⌘ **rev** **1** ⌘ Current version: **5.2.0** ⌘

*For **HELP** on using this form, see bottom of this page or look at the pop-up text over the ⌘ symbols.*

Proposed change affects: UICC apps⌘ ☐ ME ☒ Radio Access Network ☐ Core Network ☐

Title:	⌘	CR on Usage of Multiple Media Sample Entries in Media Tracks of 3GP files (Rel-5)
Source:	⌘	TSG SA WG4
Work item code:	⌘	PSS-E
Date:	⌘	10/12/2002
Category:	⌘	**F**
Release:	⌘	Rel-5

Use one of the following categories:
F *(correction)*
A *(corresponds to a correction in an earlier release)*
B *(addition of feature),*
C *(functional modification of feature)*
D *(editorial modification)*
Detailed explanations of the above categories can be found in 3GPP TR 21.900.

Use one of the following releases:

2	*(GSM Phase 2)*
R96	*(Release 1996)*
R97	*(Release 1997)*
R98	*(Release 1998)*
R99	*(Release 1999)*
Rel-4	*(Release 4)*
Rel-5	*(Release 5)*
Rel-6	*(Release 6)*

Reason for change:	⌘	3GPP file format is based on ISO media file format, which enables the usage of multiple media sample entry boxes in an audio or visual media track. This kind of usage brings up potential playback and rendering difficulties with it on the client side. A restriction on such a usage is already introduced and accepted for AMR and AMR-WB. This CR enlarges the scope of the restriction of using a single media sample entry for other media types as well. It also makes sure that the client is free to not playback or skip media tracks with multiple media sample entries in a 3GP file. The recommendation does not cover the timed text track, where multiple sample entries are definitely needed for efficiency.
Summary of change:	⌘	A text added to Clause 9.2.3 and a sentence removed from Annex D.5.
Consequences if not approved:	⌘	Clients may receive 3GP files with complex combinations of multiple media samples and try to playback the file, although they might not be able to cope with the complexity. It also stands as a guideline for content ctreators.

Clauses affected:	⌘	Clause 9.2.3 and Annex D.5

Other specs affected:	⌘	Y	N		⌘
			N	Other core specifications	
			N	Test specifications	
			N	O&M Specifications	

Other comments:	⌘	The CR is based on CR 041 rev 2 approved at SA4#23.

How to create CRs using this form:

Comprehensive information and tips about how to create CRs can be found at http://www.3gpp.org/specs/CR.htm. Below is a brief summary:

1) Fill out the above form. The symbols above marked ⌘ contain pop-up help information about the field that they are closest to.

2) Obtain the latest version for the release of the specification to which the change is proposed. Use the MS Word "revision marks" feature (also known as "track changes") when making the changes. All 3GPP specifications can be downloaded from the 3GPP server under ftp://ftp.3gpp.org/specs/ For the latest version, look for the directory name with the latest date e.g. 2001-03 contains the specifications resulting from the March 2001 TSG meetings.

3) With "track changes" disabled, paste the entire CR form (use CTRL-A to select it) into the specification just in front of the clause containing the first piece of changed text. Delete those parts of the specification which are not relevant to the change request.

9 3GPP file format (interchange format for MMS)

9.1 General

The 3GPP file format (3GP) is based on the ISO base media file format [50] and is defined in this specification. It is mandated in [35] to be used for continuous media along the entire delivery chain envisaged by the MMS, independent on whether the final delivery is done by streaming or download, thus enhancing interoperability.

In particular, the following stages are considered:

- upload from the originating terminal to the MMS proxy;
- file exchange between MMS servers;
- transfer of the media content to the receiving terminal, either by file download or by streaming. In the first case the self-contained file is transferred, whereas in the second case the content is extracted from the file and streamed according to open payload formats. In this case, no trace of the file format remains in the content that goes on the wire/in the air.

Additionally, the 3GPP file format should be used for the storage in the servers and the "hint track" mechanism may be used for the preparation for streaming.

Clause 9.2 of the present document gives the necessary requirements to follow for the 3GPP file format used in MMS. These requirements will guarantee PSS to interwork with MMS as well as the 3GPP file format to be used internally within the MMS system. For PSS servers not interworking with MMS there is no requirement to follow these guidelines.

9.2 3GPP file format conformance

The 3GPP file format, used in this specification for timed multimedia (such as video, associated audio and timed text), is structurally based on the ISO base media file format defined in [50]. However, the conformance statement for 3GP files is defined in the present document by addressing the registration of codecs, file identification, file extension and MIME type definition.

NOTE: Codecs or functionalities not conforming to a 3GP file may be ignored.

9.2.1 Registration of codecs

MPEG-4 video and AAC audio code streams, as well as the non-ISO code streams AMR narrow-band speech, AMR wideband speech, H.263 encoded video and timed text can be included in 3GP files as described in annex D of the present document.

9.2.2 Hint tracks

Hint tracks are a mechanism that a server implementation may choose to use in preparation for the streaming of media content contained in 3GP files. However, it should be observed that the usage of hint tracks is an internal implementation matter for the server, and it falls outside the scope of the present document.

9.2.3 Limitations to the ISO base media file format

The following limitations to the ISO base media file format [50] shall apply to a 3GP file of this specification:

- there shall be no references to external media outside the file, i.e. a 3GP file shall be self-contained;
- the maximum number of tracks shall be one for video, one for audio and one for text;

- the maximum number of sample entries shall be one per track for video and audio (but unrestricted for text);
- compact sample sizes ('stz2') shall not be used;
- movie fragments shall not be used.

NOTE: If a file contains video or audio tracks with more than one sample entry per track, a reader may skip those tracks or the entire file.

9.2.4 MPEG-4 systems specific elements

For the storage of MPEG-4 media specific information in 3GP files, this specification refers to MP4 [51], which is also based on the ISO base media file format. However, tracks relative to MPEG-4 system architectural elements (e.g. BIFS scene description tracks or OD Object descriptors) are optional in 3GP files and shall be ignored. The inclusion of MPEG-4 media does not imply the usage of MPEG-4 systems architecture. The receiving terminal is not required to implement any of the specific MPEG-4 system architectural elements.

9.2.5 Interpretation of 3GPP file format

All index numbers used in the 3GPP file format start with the value one rather than zero, in particular "first-chunk" in Sample to chunk box, "sample-number" in Sync sample box and "shadowed-sample-number", "sync-sample-number" in Shadow sync sample box.

--- (skipped text) ---

D.5 AMRSampleEntry box

For narrow-band AMR, the box type of the AMRSampleEntry Box shall be 'samr'. For AMR wideband (AMR-WB), the box type of the AMRSampleEntry Box shall be 'sawb'. ~~Each AMR or AMR-WB track shall be associated with a single AMRSampleEntry.~~

The AMRSampleEntry Box is defined as follows:

AMRSampleEntry ::= BoxHeader

Reserved_6

Data-reference-index

Reserved_8

Reserved_2

Reserved_2

Reserved_4

TimeScale

Reserved_2

AMRSpecificBox

Table D.4: AMRSampleEntry fields

Field	Type	Details	Value
BoxHeader.Size	Unsigned int(32)		
BoxHeader.Type	Unsigned int(32)		'samr' or ‘sawb’
Reserved_6	Unsigned int(8) [6]		0
Data-reference-index	Unsigned int(16)	Index to a data reference that to use to retrieve the sample data. Data references are stored in data reference boxes.	
Reserved_8	Const unsigned int(32) [2]		0
Reserved_2	Const unsigned int(16)		2
Reserved_2	Const unsigned int(16)		16
Reserved_4	Const unsigned int(32)		0
TimeScale	Unsigned int(16)	Copied from media header box of this media	
Reserved_2	Const unsigned int(16)		0
AMRSpecificBox		Information specific to the decoder.	

If one compares the MP4AudioSampleEntry Box - AMRSampleEntry Box the main difference is in the replacement of the ESDBox, which is specific to MPEG-4 systems, with a box suitable for AMR and AMR-WB. The **AMRSpecificBox** field structure is described in clause D.7.

3GPP TSG-SA4 Meeting #24
Redmond, USA, 11–15 November 2002

Tdoc ⌘S4-020712

CR-Form-v7

CHANGE REQUEST

⌘ **26.234** **CR** **51** ⌘ **rev** **1** ⌘ Current version: **5.2.0** ⌘

*For **HELP** on using this form, see bottom of this page or look at the pop-up text over the ⌘ symbols.*

Proposed change affects: UICC apps⌘ ☐ ME ☒ Radio Access Network ☐ Core Network ☐

Title:	⌘	Progressive download of 3GP files
Source:	⌘	TSG SA WG4
Work item code:	⌘	PSS-E
Date:	⌘	10/12/2002
Category:	⌘	**F**
Release:	⌘	Rel-5

Use one of the following categories:
F *(correction)*
A *(corresponds to a correction in an earlier release)*
B *(addition of feature),*
C *(functional modification of feature)*
D *(editorial modification)*
Detailed explanations of the above categories can be found in 3GPP TR 21.900.

Use one of the following releases:

2	*(GSM Phase 2)*
R96	*(Release 1996)*
R97	*(Release 1997)*
R98	*(Release 1998)*
R99	*(Release 1999)*
Rel-4	*(Release 4)*
Rel-5	*(Release 5)*
Rel-6	*(Release 6)*

Reason for change:	⌘	The 3GPP file format (3GP) is a versatile file format. However, when reading the TS 26.234, one may be led to the wrong conclusion that the file format can only be used for MMS messages and for download. We think it is important to inform potential implementers that it is very easy to use the 3GPP file format also for progressive download.
Summary of change:	⌘	Subclause 9.2.6 added.
Consequences if not approved:	⌘	An important feature of the 3GPP file format will not be implemented by clients.

Clauses affected:	⌘	Clause 9.2

Other specs affected:	⌘	Y	N		⌘
			N	Other core specifications	
			N	Test specifications	
			N	O&M Specifications	

Other comments:	⌘	The CR is based on CR 041 rev 2 approved at SA4#23.

How to create CRs using this form:
Comprehensive information and tips about how to create CRs can be found at http://www.3gpp.org/specs/CR.htm.
Below is a brief summary:

1) Fill out the above form. The symbols above marked ⌘ contain pop-up help information about the field that they are closest to.

2) Obtain the latest version for the release of the specification to which the change is proposed. Use the MS Word "revision marks" feature (also known as "track changes") when making the changes. All 3GPP specifications can be

downloaded from the 3GPP server under ftp://ftp.3gpp.org/specs/ For the latest version, look for the directory name with the latest date e.g. 2001-03 contains the specifications resulting from the March 2001 TSG meetings.

3) With "track changes" disabled, paste the entire CR form (use CTRL-A to select it) into the specification just in front of the clause containing the first piece of changed text. Delete those parts of the specification which are not relevant to the change request.

9.2.5 Interpretation of 3GPP file format

All index numbers used in the 3GPP file format start with the value one rather than zero, in particular “first-chunk” in Sample to chunk box, “sample-number” in Sync sample box and “shadowed-sample-number”, “sync-sample-number” in Shadow sync sample box.

9.2.6 3GPP file authoring guidelines for progressive download

The present document specifies the 3GPP file format to be used for distribution of continuous media clips using MMS. The same file format can also be used for file download from any server. However, to achieve a better response to a user request for media, it is often advantageous if the client can start playing the media before the full file is downloaded. This scenario is known as progressive download and is provided by many proprietary media solutions. It is the purpose of this clause to point out that this is also easily achievable by using the 3GPP file format. This possibility has been inherent in the file format from the first version in Release 4, and the only thing that is needed is that the content creator follows the guidelines provided here.

The principles behind progressive download are that the session information should be put at the beginning of the file and that the media tracks should be interleaved within the file. In practice, this leads to the following guidelines for the creation of 3GP files:

- the 'moov' box should be placed at the start of the file, right after the 'ftyp' box;
- the media tracks should be interleaved inside the file. The typical interleaving length is one second.

For the release-4 file format, the boxes are called atoms but, except for that, everything applies equally well.

It should be noted that no change is needed at the server side, and that a client that does not support progressive download can always play the file once it has been completely downloaded. A progressive download client can start playing a 3GP file that has been created along the progressive download guidelines once it has received a first chunk of all media in the session. If the file has not been prepared for progressive download, the client will always need to wait for the full download.