

INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC 1/SC 29/WG 3
CODING OF MOVING PICTURES AND AUDIO

ISO/IEC JTC 1/SC 29/WG 3 m yyyy

Milford, Ontario, Canada – March 2024

Title: Increasing file-size and bandwidth efficiency, and improving internationalization support, of OpenFont files using delta tables (DMAP)

Author: Peter Constable (Microsoft, Inc., pconstable@microsoft.com), Dave Crossland (Google Inc., dcrossland@google.com), Behdad Esfahbod (behdad@behdad.org), Laurence Penney (lorp@lorp.org), Liam Quin (Delightful Computing, liam@delightfulcomputing.com), Rod Sheeter (Google Inc., rsheetter@google.com)

Introduction

(This introduction is not part of the proposal)

The purpose of a DMAP table in a Collection is to avoid duplication of the cmap table, which for a CJK font can be large. It is a per-subfont table that is consulted *before* cmap, overriding cmap.

The ‘DMAP’ table was previously proposed by Ned Holbrook, Ken Lunde, Dwayne Robinson, and Peter Constable.

DMAP—Delta map table [5.6.15] [add at end of optional tables section, before 6 Advanced Open Font layout tables]

An optional ‘DMAP’ table, if present, shall take priority over ‘cmap’—that is, the character-to-glyph lookup shall first look in the ‘DMAP’ table (including any variation-selectors). If a match is not found, ‘cmap’ shall be consulted. The ‘DMAP’ table is identical in structure to the ‘cmap’ table.

This allows individual members of a TTC collection to override or supply parts of a shared character map, saving space in the font.

‘DMAP’ Header

Type	Name	Description
uint16	version	Set to zero
uint16	NumTables	Number of encoding tables that follow
EncodingRecord	encodingRecords[numTables]	Array of cmap-format subtables, measured from the beginning of the DMAP table.

The DMAP subtables are in the same formats as CMAP subtables.

The *language* field for a format 4, 12, or 13 DMAP subtable shall be set to zero.

Any subtables in a DMAP table, including Format 14 and Format 15, shall supersede any such subtables in the corresponding ‘cmap’ table in entirety: all character processing, including for Format 14 lookups, shall first consult DMAP.

Near the end of The Font Collection file structure [4.6.2] we add a paragraph:

The Font Collection file structure [4.6.2]

A font collection file consists of a single TTC header table, one or more table directories (each corresponding to a different font resource), and a number of OFF tables. The TTC header shall be located at the beginning of the TTC file.

The TTC file shall contain a complete table directory for each font resource. The same TableDirectory format is used for each font resource in a collection file as is used in a non-collection file. The table offsets in all table directories within a TTC file are measured from the beginning of the TTC file.

Each OFF table in a TTC file is referenced through the table directory of each font which uses that table. Some of the OFF tables must appear multiple times, once for each font included in the TTC; while other tables may be shared by multiple fonts in the TTC.

As an example, consider a TTC file which combines two Japanese fonts (Font1 and Font2). The fonts have different kana designs (Kana1 and Kana2) but use the same design for kanji. The TTC file contains a single 'glyf' table which includes both designs of kana together with the kanji; both fonts' table directories point to this 'glyf' table. But each font's table directory points to a different 'cmap' table, which identifies the glyph set to use. Font1's 'cmap' table points to the Kana1 region of the 'loca' and 'glyf' tables for kana glyphs, and to the kanji region for the kanji. Font2's 'cmap' table points to the Kana2 region of the 'loca' and 'glyf' tables for kana glyphs, and to the same kanji region for the kanji.

The tables that should have a unique copy per font are those that are used by the system in identifying the font and its character mapping, including 'cmap', 'name', and 'OS/2'. The tables that should be shared by fonts in the TTC are those that define glyph and instruction data or use glyph indices to access data: 'glyf', 'loca', 'hmtx', 'hdmx', 'LTSH', 'cvt', 'fpgm', 'prep', 'EBLC', 'EBDT', 'EBSC', 'maxp', and so on. In practice, any tables which have identical data for two or more fonts may be shared.

Each font in the TTC may contain its own DMAP table. In addition, a font using CMAP may support characters and GlyphIDs also defined by another font; in the case of overlapping coverage the order of precedence is first DMAP, then CMAP, and then finally cmap.

NOTE When building a collection file from separate font files, close attention needs to be paid to the issue of glyph renumbering in a font and the side effects that can result in the 'cmap' table and elsewhere. The fonts to be merged also need to have compatible TrueType instructions; that is, their preprograms, function definitions, and control values cannot conflict.

Collection files containing TrueType glyph outlines should use the filename suffix .TTC. Collection files containing CFF or CFF2 outlines should use the file extension .OTC.

In CMAP just before the cmap header (5.1.2.1)

An optional 'DMAP' table, if present, shall take priority over 'cmap'—that is, the character-to-glyph lookup shall first look in the 'DMAP' table (including any variation-selectors). If a match is not found, 'cmap' is to be consulted. The 'DMAP' table is identical in structure to the 'cmap' table.

[end]