

Was ist "Unicode"?

Für wen ist diese Anleitung gedacht?

Diese Anleitung ist für diejenigen gedacht, welche OSDD/ME10 oder Modeling ab Version 15 betreiben und etwas über Unicode erfahren möchten. Beide Produkte unterstützen nun "Unicode" und eventuell haben Sie sich bereits gefragt, was sich hinter diesem Ausdruck verbirgt und welche Auswirkungen die Verwendung von Unicode hat.

Viel Spaß beim Lesen wünscht

Ihr TDWsoft-Team

1 Was ist Unicode?

Im Prinzip geht es bei Unicode darum, den Zeichen aller bekannten Schriftkulturen bzw. Zeichensysteme ein digitaler Code zuzuordnen. Das Ziel ist die weltweite Vereinheitlichung der digitalen Kennungen für (Schrift-)Zeichen. Mehr Informationen (z.B. über die unterschiedlichen Kodierungsarten) finden Sie unter:

<http://de.wikipedia.org/wiki/Unicode>

2 Mit welcher Kodierung arbeiten die CoCreate-Produkte?

Die CoCreate-Produkte arbeiten mit "UTF-8" (8-bit Unicode Transformation Format). Im Gegensatz zu ihren Schwestern (UTF-16 bzw. UTF-32) verwendet diese Kodierungsart für den Zeichenbereich 0 bis 127 nur ein Byte (dieser Bereich entspricht übrigens dem ASCII-Zeichensatz). Das macht die ASCII-Dateien etwas schlanker. Denn erst die Zeichen, welche ausserhalb des ASCII-Bereichs liegen, werden mit 2 Byte abgebildet.

Weiterführende Informationen:

<http://de.wikipedia.org/wiki/UTF-8>

<http://de.wikipedia.org/wiki/ASCII>

3 Welche Auswirkungen hat die Verwendung von Unicode?

Ganz abgesehen davon, dass Sie die Teilnamen bei Drafting bzw. bei Modeling nun auch in Japanisch angeben können, enthalten die von Modeling bzw. von Drafting erzeugten Dateien eine bestimmte Kennung am Dateianfang. Diese Kennung nennt sich "BOM" Das steht für "Byte Order Mark(er)" und besteht aus dieser Bytefolge:

```
EF BB BF
```

Mit dieser Bytefolge beginnen fast alle erzeugten Textdateien, so dass Editoren deren Kodierung gleich beim Öffnen erkennen. Öffnet man solche Dateien in einem Editor, erkennt man die UTF-8-Kennung (die ersten 3 Byte der Datei) an den folgenden Zeichen:

```
ï»¿
```

Abweichend davon erfolgt die Angabe der Kodierung je nach Dateityp unterschiedlich. In mit dem Modeling-Recorder erzeugten Dateien findet sich noch die folgende Kennung:

```
;; encoding: UTF-8
```

Bei ME10-Makros findet man den Befehl "DEFINE_ENCODING", den man auch in eigenen Makros anwenden kann:

```
DEFINE_ENCODING 'UTF-8'
```

Und in den ab V. 15 erzeugten Zeichnungs-Dateien ("MI-Dateien") findet sich am Dateianfang ebenfalls ein Hinweis auf die verwendete Kodierung:

```
#~1  
LANG:de  
ENCODING:UTF-8
```

Normalerweise werden Dateien, welche mit älteren Systemen erzeugt wurden, problemlos gelesen. Denn das System nimmt an, dass sie mit der früher üblichen Kodierung namens "ROMAN-8" erzeugt wurden.

4 Welche Bereiche sind von Unicode betroffen?

Betroffen sind prinzipiell alle Bereiche des Datenaustausches über Dateien. Das Lesen von Dateien der Vorversionen ist (inklusive der Lisp- bzw. Makrodateien) aus Gründen der Abwärtskompatibilität normalerweise kein Problem. Werden jedoch Dateien mit anderen Programmen ausgetauscht, kann es aufgrund unterschiedlicher Kodierungen zu Problemen kommen. Meist treten diese Probleme dann auf, wenn Sonderzeichen im Spiel sind. Hier bietet Modeling, als auch Drafting entsprechende Befehle bzw. Optionen zu bestehenden Befehlen an:

Bei Modeling: open, load, with-open-file (:external-format), oli.:sd-recode-textfile, usw.

Bei Drafting: SET_FONT_ENCODING, UCHR, UNUM, OPEN_OUTFILE_ENCODING, usw.

5 Gibt es nun auch neue Schriftarten (Fonts)?

Ja, die gibt es. Sowohl für Drafting, als auch für Annotation gibt es nun die folgenden neuen Schriftarten:

- » osd_default
- » osd_default2
- » osd_default3

Diese neuen Schriftarten sind jedoch nichts anderes, als Zusammenfassungen von 1- bzw. 2-Byte-Schriftarten. Daher können mit ihnen auch japanische bzw. chinesische Schriftzeichen gemeinsam mit den lateinischen Buchstaben innerhalb einer Zeichnung unter Verwendung einer Schriftart erstellt werden.

Welche bisherigen Schriftarten nun in den neuen Schriftarten zusammengefasst wurden zeigt die folgende Tabelle:

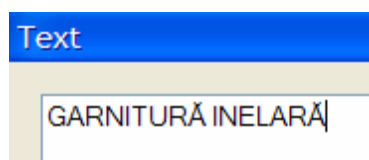
osd_default	hp_i3098_v hp_kjis_v hp_kan2_c
osd_default2	hp_i3098_c hp_kan2_c
osd_default3	hp_d17_v hp_kan2_c

Die Schriftart "osd_default" ist - sofern man keine eigenen Anpassungen mit anderen Einstellungen lädt - voreingestellt. Daher müsste man sich gleich fragen, was passiert, wenn nun Zeichnungen mit dieser Schriftart für frühere Versionen heruntergespeichert werden. Schließlich kennen die früheren OSD- bzw. ME10-Versionen die Schriftart "osd_default" noch nicht. Beim Herunterspeichern wird allen Texten, welche z.B. mit der Schriftart "osd_default" erstellt wurden, die Schriftart "hp_i3098_v" zugewiesen. Somit ist die Abwärtskompatibilität gewährleistet.

6 Kann ich mit den eingebauten Schriftarten von ME10/Drafting nun alle Sonderzeichen wie unter Windows darstellen?

Die neuen Schriftarten enthalten zwar mehr Zeichen, doch nicht unbedingt jene, die für die Darstellung bestimmter Sprachen (wie z.B. Polnisch, Russisch oder Rumänisch) benötigt werden. Schließlich kamen ja nur die asiatischen Zeichen zum bisherigen Zeichen-Umfang von "hp_i3098" hinzu.

Angenommen, Sie möchten also den rumänischen Ausdruck für "O-Ring" in Ihre Zeichnung einbringen und Sie geben den folgenden Text in den Editor ein:



Dann werden Sie nach dem Absetzen des Textes feststellen, dass ein Zeichen (das A mit dem Dach – "großes A mit Breve") nicht dargestellt werden kann:



Mit einem kleinen Trick schafft man es trotzdem, den Text fehlerfrei in die Zeichnung einzubringen, wie die folgende Abbildung zeigt:




Stellt man nämlich für den frisch erstellten Text die Schriftart auf eine Windows-Schriftart um (z.B. auf "Arial" oder auf "Courier New"), dann können die Zeichen korrekt dargestellt werden. Das liegt daran, dass die Windows-Schriftarten natürlich wesentlich mehr Zeichen enthalten, wie die hauseigenen ME10-Schriftarten. Allerdings enthalten selbst die Windows-Schriftarten nicht immer alle benötigten Zeichen. Die Windows-Schriftart "Arial Unicode MS" wird übrigens ihrem Namen dadurch gerecht, als dass sie alle derzeit definierten Zeichen enthält.

Stellt man für einen Text unter ME10 eine Windows-Schriftart ein, so läuft vom Anwender meist unbemerkt eine Hintergrund-Konvertierung ab. ME10 wandelt die TrueType-Schriftart von Windows in eine ME10-Schriftart um. Die erstellten (binären) Schriftartdateien werden im Windows-Profil des Anwenders abgelgt:

```
%APPDATA%\CoCreate\Shared\15.00\fonts
```

Somit sind wir auch schon bei einem Nachteil der eben vorgestellten Technik: Diese Schriftarten liegen auf der lokalen Festplatte des jeweiligen Arbeitsplatzrechners. Sofern der Anwender eine Schriftart gewählt hat, welche bei allen Windows-Versionen vorhanden ist, dürfte das kein Problem sein. Beim Austausch mit anderen Systemen über die typischen Austauschformate wie z.B. DXF bzw. DWG könnte es da schon mehr Probleme geben. Immerhin kann man mittlerweile die in einer Zeichnung verwendeten Nicht-ME10-Schriftarten beim Speichern einbetten (was größere MI-Dateien zur Folge hat).

Das Konvertieren der Windows-Schriftarten hat bringt jedoch noch (je nach Sichtweise) einen Nachteil kosmetischer Art mit sich: Sie werden (im Gegensatz zu fast allen eingebauten ME10-Schriftarten) als so genannte "Outline-Fonts" erzeugt. Das sind Schriftarten, deren Zeichen über Konturen repräsentiert werden. Daher können sie auch wahlweise gefüllt dargestellt werden:



Das Programm "PE-Lingua" ist übrigens in der Lage alle osteuropäischen Sprachen (Polnisch, Russisch, Slowakisch, Rumänisch usw.) und sogar Japanisch so abzubilden, dass keine Outline-Schriftarten von Windows verwendet werden müssen:

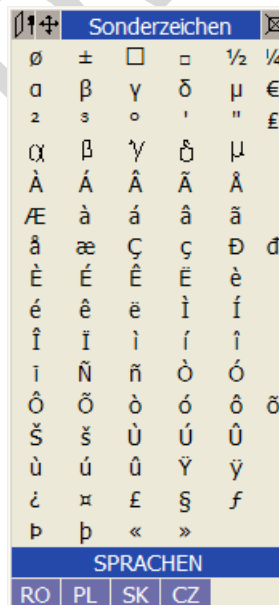


Neben dem manuellen Erzeugen von Texten können komplette Zeichnungen (auch in Annotation) automatisiert übersetzt werden. Die Pflege des firmenspezifischen Wörterbuches erfolgt ganz komfortabel über eine Excel-Tabelle:

B226		SCHNITT A-A	
A	B	V	X
2	Deutsch	Slowakisch	Rumänisch
3		hp_13098_v_ee	hp_13098_v_ee
217	214 NUT UND GEWINDE SIND ZUEINANDER LAGEABHAENGIG	DRÁŽKA A ZÁVIT SÚ ZÁVISLÉ OD VZÁJOMNEJ POLOHY	214 POZIȚIA CANELURII DEINDE DIRECT DE
218	215 MIT DAUERELASTISCHEM MATERIAL AUSGEFUGT	VYŠPÁROVANÉ TRVALO ELASTICKÝM MATERIÁLOM	215 ROSTUIT CU MATERIAL CU ELASTICITAI
219	216 NACH DEM SCHLEIFEN GETRENNT	ODDELENÉ PO BRÚSENÍ	216 SEPARAT DUPĂ ȘLEFUIRE
220	217 O-RING	O-KRÚŽOK	217 GARNITURĂ INELARĂ
221	218 PASSLAENGE	DĹŽKA PRIECHODU	218 LUNGIME DE PĂSUIRE
222	219 PASSUNGLAENGE	DĹŽKA LICOVANIA	219 LUNGIME DE PĂSUIRE
223	220 ROHR	RÚRKA	220 TEAVĂ
224	221 ROT LACKIERT	NALAKOVANÝ NA ČERVENO	221 VOPSIT CU ROȘU

Dort können Sie die unter Windows üblichen Sonderzeichen der jeweiligen Sprache verwenden. Um die Zuordnung dieser Zeichen zu den ME10- bzw. Annotation-Schriftarten kümmert sich dann PE-Lingua.

FACE 18/2007 erleichtert das Auffinden und Einfügen von Sonderzeichen bestimmter Sprachen erheblich. Übersichtliche Auswahlménüs (geordnet nach Sprache) ermöglichen die gezielte Auswahl. Nützliche Einfügeoptionen (vor bestehenden Text, hinter bestehenden Text usw.) erleichtern das Einbringen der Zeichen in die Zeichnung:



7 Wo finde ich weitere Informationen?

Bei Modeling findet man einige weiterführende Informationen, indem man den Suchbegriff "Unicode" im Suchfenster der Onlinehilfe eingibt. Im Programmverzeichnis von ME10 findet man die Dateien "Readme.htm", "Whatsnew.htm" und "Appendix.htm", welche weiterführende Informationen und Hinweise enthalten.

8 Zwei Tipps zum Schluss

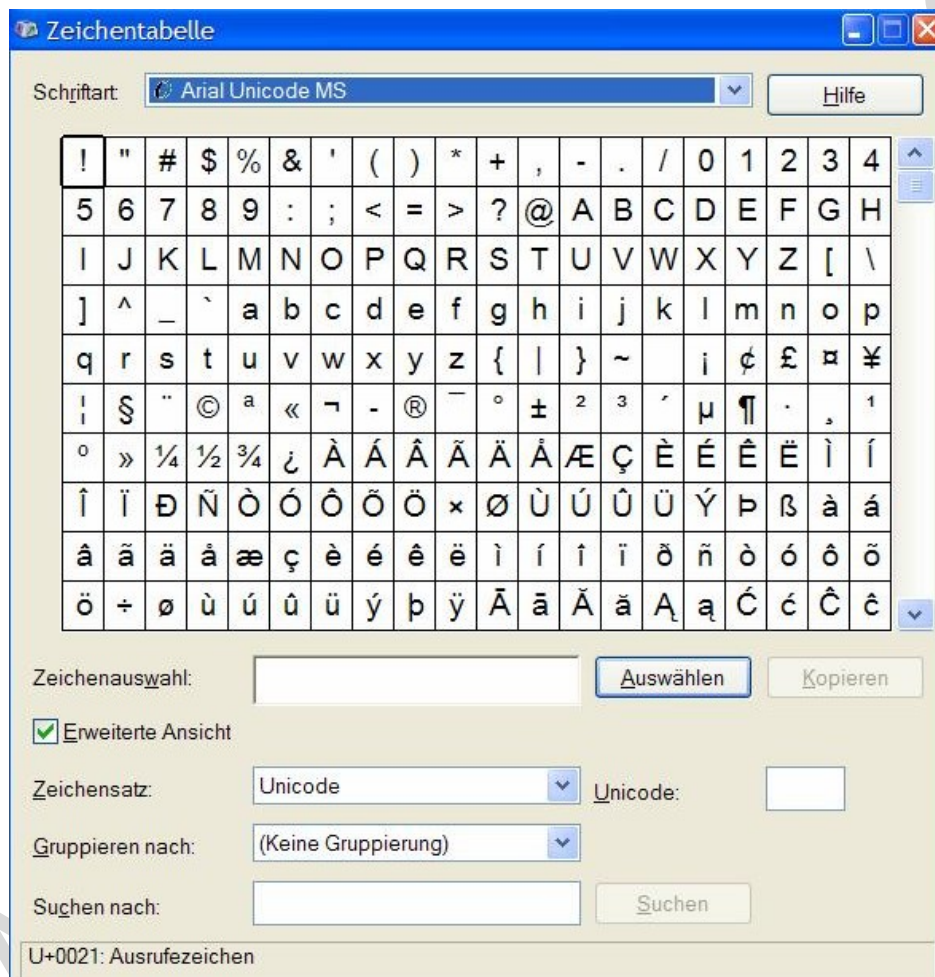
Nach soviel Theorie nun zwei Tipps zum Abschluss:

TIPP Nr. 1

Wenn Sie Texte in einer anderen Sprache in Zeichnungen verwenden müssen, könnten Sie durch die Großschreibung eventuell Sonderzeichen einsparen. Das funktioniert z.B. bei Französisch. Dort akzeptiert man es zumindest, wenn die Großbuchstaben keine Akzente enthalten.

TIPP Nr. 2

Wenn Sie sich eine Übersicht der in einer Windows-Schriftart enthaltenen Zeichen verschaffen möchten, können Sie das Windows-Programm "Charmap" verwenden (zu Deutsch: "Zeichentabelle"):



Dieses Programm wird folgendermaßen gestartet: "Start/Ausführen" und charmap.exe eingeben.

Viel Spaß beim Erforschen der Schriftarten wünscht

Ihr TDWsoft-Team

[<http://www.tdwsoft.com>] • [<http://tdwsoft.blogspot.com>]