

TUSTEP Version 2000 für Windows und Unix

Seit Mitte Oktober 1999 ist die Version 2000 von TUSTEP für Windows 95 / Windows 98 / Windows NT und für Unix (einschließlich Linux) verfügbar. Da in der Version 1999 noch Y2K-Probleme entdeckt wurden, raten wir dazu, die neue Version noch vor der Jahreswende zu installieren.

Im übrigen wurde für die Version 2000 viel Aufwand in die Abrundung und Konsolidierung der Erweiterungen der Vorgängerversionen (Kommandomakros; CGI-Schnittstelle; XML-Unterstützung im Satz) gesteckt. Darüber hinaus wurde der Tatsache Rechnung getragen, dass TUSTEP in immer größerem Umfang als Datenbanksystem, auch für den Zugriff über das WWW, eingesetzt wird. Dies erfordert u. a. Möglichkeiten zur Koordination von konkurrierenden (lesenden und schreibenden) Zugriffen auf Datenbestände. Schließlich wurden die Schnittstellen für den Datenaustausch den internationalen Entwicklungen angepaßt.

Zu den wichtigsten diesbezüglichen Neuerungen zählen:

- Dateityp RAN zur Verringerung des Zeitbedarfs beim Zugriff auf einzelne Sätze oder Bereiche / Segmente über Satznummern. Bei großen Dateien (bis zu 7 Gigabyte pro Datei) ergibt sich eine Geschwindigkeitssteigerung um bis zu drei Größenordnungen gegenüber Dateien vom Typ SEQ.
- Mega-Segmentdateien mit bis zu 1.000.000 Segmenten (statt 10.000 für »normale« Segmentdateien).
- Anweisungen in den Kommandomakros für konfliktfreien schreibenden und lesenden Zugriff auf Dateien durch mehrere gleichzeitig aktive Benutzer.
- Erleichterung der Datensicherung durch das Standard-Makro `#*MBUPDATE`.
- Unterstützung von UNICODE (UTF-8 und 16-bit).

Die neuen Leistungen sind in der Online-Version des Handbuchs beschrieben. Diese ist auf drei Wegen zugänglich:

a) Kommando `#HILFE` bzw. im Editor Anweisung *hilfe* oder Steueranweisung `HELP (Strg+V)` für systematisches Nachschlagen im Handbuch

b) Kommando `#SUCHE` zum Nachschlagen im Handbuch über einen Stichwort-Index

c) Standardmakro `#*DRUBE` zum Ausdrucken des Handbuchs bzw. seiner Teile (mit dem Aufruf `#*DRUBE, AB, win-10, +` erhält man Information über die einzelnen Teile der Beschreibung auf dem Bildschirm; auf X-Ter-

minals ist die Angabe *win-10* durch *decms* zu ersetzen).

Für die wichtigsten Neuerungen gibt es außerdem Ergänzungshefte zum Handbuch von 1993, die ebenfalls über `#*DRUBE` ausgedruckt werden können oder bereits ausgedruckt für je 3,- DM im Sekretariat des ZDV erhältlich sind. Die Ergänzungshefte können auch im WWW über die TUSTEP-homepage konsultiert werden unter:

<http://www.uni-tuebingen.de/zdv/tustep>

Im Folgenden soll eine kurze Zusammenfassung der neuen Leistungen gegeben werden.

TUSTEP-Sitzungen

Eine TUSTEP-Sitzung wird nur gestartet, wenn sie nicht schon aktiv ist.

Das neue Kommando `#MODI, MODUS = RETTEN|HOLEN, PROTOKOLL = -STD-|-|+` ist zur Verwendung in Kommandofolgen gedacht. Es rettet in einen Stack bzw. holt aus diesem die von den Kommandos `#FEHLERHALT, #PARAMETER, #PROTOKOLL #WISCHEN` eingestellten Modi. Wird zu `PROTOKOLL -` oder `+` angegeben, so wird zusätzlich das Kommando `#PROTOKOLL,-` bzw. `#PROTOKOLL,+` ausgeführt. Falls am Ende der Kommandofolge der Stack noch nicht leer ist, wird er geleert; die zuerst geretteten Modi werden wieder eingestellt.

Mit dem Kommando `#LISTE, MODUS = TRAEGER` können die Namen der als Träger nutzbaren Systemvariablen und deren Inhalt ausgegeben werden, mit dem Kommando `#LISTE, MODUS = VARIABLEN` die Namen und der Inhalt aller für TUSTEP nutzbaren Systemvariablen (in den Kommandomakros leisten dies die Makrofunktionen `VOLUMES` bzw. `VARIABLES`).

Dateizugriffe

Während eine Datei in einer Sitzung schreibend bearbeitet wird, kann diese Datei in keiner anderen Sitzung angemeldet oder – wenn sie schon angemeldet war – gelesen oder verändert werden. Während eine Datei in einer Sitzung lesend bearbeitet wird, kann diese Datei in keiner anderen Sitzung verändert werden.

In Kommandomakros können Dateizugriffe u. a. über die neuen Anweisungen `LOCK` und `UNLOCK` (siehe unten) koordiniert werden.

Beim Kommando #LISTE (und analog dazu bei den Kommandos #ABMELDE, #ANMELDE, #LOESCHE) gelten jetzt für die Auswahl der Dateinamen über die Spezifikationen POSITIV und NEGATIV die von Windows und Unix gewohnten Konventionen.

Dateityp RAN

Beim Kommando #DATEI kann jetzt auch RAN als Dateityp angegeben werden. Dateien vom Typ RAN erlauben wahlfreien direkten Zugriff über die Satznummer. Hierfür erhalten solche Dateien einen Index mit 1024 Aufsatzpunkten für Seitennummern. Der Zugriff auf solche Dateien über Satznummern ist damit je nach Größe der Datei um bis zu drei Größenordnungen schneller als bei SEQ-Dateien.

Segmentdateien

Mit dem neuen Kommando

```
#HOLE, QUELLE, ZIEL, MODUS = name,  
LOESCHEN
```

mit dem auch ganze Dateien blockweise oder satzweise kopiert werden können, kann zu MODUS ein Name angegeben werden. In diesem Fall werden aus der Quelldatei die Daten des Segments mit dem angegebenen Namen in die ZIEL-Datei kopiert.

Wird beim Kommando #RETTE zu MODUS ein Name angegeben, so werden die Sätze der Quelldatei als Segment in die ZIEL-Datei (die entweder noch leer oder eine Segmentdatei sein muß) kopiert. Die Angabe zu LOESCHEN bezieht sich dann nicht auf die ganze Segmentdatei, sondern nur auf das Segment mit dem angegebenen Namen.

Außer »normalen« Segmentdateien, die 9.999 Segmente enthalten können, können jetzt Mega-Segmentdatei mit bis zu 999.999 Segmenten benutzt werden. Eine Segmentdatei wird zu einer Mega-Segmentdatei, indem sie mit dem Standardmakro #*SESO in eine zweite Datei kopiert wird und dabei MODUS = GENERIEREN / MEGA bzw. MODUS = SORTIEREN / MEGA angegeben wird. Die gleiche Wirkung hat die Angabe MODUS = + ; - bzw. MODUS = + ; + beim Kommando #RETTE. Soll eine Mega-Segmentdatei in eine normale Segmentdatei verwandelt werden, so kann bei #RETTE die Angabe MODUS = - ; - bzw. MODUS = - ; + gemacht bzw. beim Makro #*SESO der Zusatz / MEGA zu den Spezifikationen GENERIEREN bzw. SORTIEREN weggelassen werden.

Code-Pages

Bei den Kommandos #DEFINIERE und #UMWANDLE können zur Spezifikation CODE folgende Code-Pages zusätzlich angegeben werden:

```
CP852 (DOSLatin2, Eastern Europe)  
CP1250 (WinLatin2, Eastern Europe)  
CP1252 (WinLatin1, Obermenge von ISO  
8859-1)  
ANSI (wie CP1252)
```

Beim Kommando #UMWANDLE können zur Spezifikation CODE zusätzlich noch folgende Code-Pages angegeben werden:

```
CP10000 (Macintosh Roman)  
CP10029 (Macintosh Central Europe)
```

Datenimport / Datenexport

Beim Import von Fremddaten kommt es immer wieder vor, dass diese keine Zeileneinteilung enthalten. Mit

```
#UMWANDLE, . . . , SATZLAENGE = nn,  
KENNUNG = zf
```

können jetzt beim Umwandeln von SDF-Dateien in TUSTEP-Dateien lange Sätze beim nächsten Blank vor bzw. nach nn Zeichen unterteilt werden. Dabei wird an der Stelle der Unterteilung die Zeichenfolge zf eingefügt. Beim Umwandeln von TUSTEP- in SDF-Dateien werden Sätze, die am Zeilenende die zu KENNUNG angegebene Zeichenfolge enthalten, mit dem folgenden Satz zusammengefaßt, wobei diese Zeichenfolge entfernt wird.

UNICODE

Mit der Version 2000 unterstützt TUSTEP auch UNICODE. Der Kommandoaufruf

```
#UMWANDLE, . . . , CODE = UNICODE | UTF8,  
NL = LS | PS
```

wandelt UNICODE-Dateien in TUSTEP-Dateien um und umgekehrt. Die Umwandlungsrichtung ergibt sich aus den Dateitypen der zu QUELLE bzw. ZIEL angegebenen Dateien.

Bei der Angabe CODE = UNICODE werden 16-bit-Codes interpretiert bzw. erzeugt, bei CODE = UTF8 wird das UCS Transformation Format interpretiert bzw. erzeugt.

Bei CODE = UNICODE kann zur Spezifikation NL auch LS bzw. PS angegeben werden. Bei NL = LS (»Line Separator«) wird die Codefolge 2028 (hex) für den Zeilenwechsel verwendet, bei NL = PS (»Paragraph Separator«) die Codefolge 2029 (hex).

Die Unterstützung von UNICODE ist ausführlicher in dem Beitrag »TUSTEP und UNICODE« in BI 99 / 7–10, S. 3 f. behandelt.

Prüfen von Dateien

Das neue Kommando #TESTE prüft die Identität zweier Dateien (MODUS=-STD-) bzw. vergleicht (bei MODUS=name) das Segment name einer Segmentdatei mit dem Inhalt der TEST-Datei; bei Ungleichheit wird das Fehlerflag gesetzt und ggf. eine Fehlermeldung ins Ablaufprotokoll ausgegeben.

Das neue Kommando #MBTEST vergleicht (Festplatten-)Dateien mit entsprechenden auf einem (logischen) Band gesicherten Dateien. Bei Ungleichheit wird das Fehlerflag gesetzt und ggf. eine Fehlermeldung ins Ablaufprotokoll ausgegeben.

Analoge Funktionen bietet der neue MODUS=TESTEN für das Überprüfen der automatischen Sicherung beim Standardmakro #*MBUPDATE.

Editor

Bei den Editoranweisungen gibt es folgende Änderungen bzw. Erweiterungen:

- u! wirkt bei MODUS = p wie u, ,1 und anschließend »j«, bei MODUS = t wie u, (.) ,1 und anschließend »j«. Die Stern-Position wird logisch beibehalten.
- h Die Anweisung h ist aus Sicherheitsgründen nicht mehr möglich, wenn die Editor-Datei eine Segmentdatei ist.
- a Bei Austauschweisungen (und anderen erweiterten Anweisungen wie e, k, l, u, bei denen vor der Ausführung eine Rückfrage erfolgt) wird der Kontext der auszutauschenden Zeile mit angezeigt. Die Zeilennummer wird hervorgehoben.
- m,v+ schaltet den Viewer (mit TUSTEP-Farben) ein
- m,v- schaltet den Viewer aus
- m,w+ schaltet den Viewer (mit Windows-Farben) ein
- m,w- schaltet den Viewer aus.

Zum Zweck der farbigen Hervorhebung von Zeichenfolgen (»Colorierung«) beim Anzeigen im Editor können bis zu 9 Farbgruppen c1 bis c9 mit jeweils bis zu 9 Farben definiert werden. Dies geschieht mit der Anweisung

cn,m=xx:zfstab

(z. B.: c1,1=AF:/zf1/zf2//zf3/), die für die Farbgruppe cn als m-te Farbe die Farbe xx definiert, in der die Zeichenfolgen angezeigt werden sollen, die der Zeichenfolgen-Such-Tabelle zfstab entsprechen. Die möglichen Codes für xx sind der Tabelle zu entnehmen, die durch die Tastenkombination CTRL+F angezeigt wird. Bereits definierte Farbgruppen

können mit den Anweisungen c1 bis c9 ausgewählt werden. Mit der Anweisung c0 wird die Colorierung wieder ausgeschaltet.

#KOPIERE

Die maximale Abschnittslänge wurde auf 256.000 Zeichen heraufgesetzt (gilt auch für #RVORBEREITE). Für Parameter (z. B. ZF+, XX, TTT) steht ebenfalls etwa 5 Mal mehr Platz zur Verfügung als bisher. Es können also längere Such- und Austausch Tabellen angegeben werden.

Kommandomakros

Für die Beschreibung der Neuerungen in den Kommandomakros sei auf die Online-Beschreibung (oder auf das auch im WWW zugängliche Ergänzungsheft) verwiesen. Die wichtigsten Neuerungen betreffen die Koordination des Dateizugriffs über Makros: Mit Hilfe der Funktionen LOCK und UNLOCK können Dateien für Zugriffe anderer Benutzer gesperrt und wieder freigegeben werden. Mit diesen Funktionen und durch zusätzliche Angaben zu den Funktionen OPEN und CHECK und zu ACCESS können gleichzeitige Zugriffswünsche koordiniert werden. Dies ist Voraussetzung für einen konfliktfreien schreibenden Zugriff auf Dateien, die von mehreren Anwendern (z. B. über CGI) gleichzeitig benutzt werden.

Das Absuchen der Datei mit der FIND-Anweisung kann dadurch weiter eingegrenzt werden, dass hinter den Suchbedingungen, durch WHILE eingeleitet, weitere Suchbedingungen angegeben werden. Sobald diese hinter WHILE angegebenen Suchbedingungen erstmals nicht erfüllt sind, wird die Suche beendet.

Für das Absuchen von Dateien mit Hilfe von R_TABLES erlaubt die neue Option USER beim Aufbau dieser Tabellen, in der Suchbedingung selbst logische Operatoren zu verwenden.

Ansonsten lassen sich weitere Neuerungen bei folgenden Anweisungen bzw. Funktionen in der Beschreibung bzw. der Online-Hilfe finden:

- Anweisungen: EXECUTE/QUIET
- Makrofenster / Feldtypen: INPUT/SHIFT; INPUT/EDIT
- Dateizugriff: READ/CHECK
- Dateitransfer: UPLOAD; DOWNLOAD
- Bedingungen: »Zflg« .EQ. PRINTER; WAIT
- Makrofunktionen: FILE »dateiname« +n | -n; FILE »dateiname« segmentname +n | -n; VOLUMES; TIME_3; VARIABLES; EDIT

Satzprogramm

Im Satzprogramm gibt es viele Detailverbesserungen, die im Online-Handbuch oder im Ergänzungsheft nachgelesen werden können. Hier seien nur die wichtigsten aufgeführt.

Mit dem Parameter SCH können jetzt bis zu 16 Schriften angegeben werden, die gleichzeitig in einem Satzauftrag benutzt werden sollen. Auch für Sonderzeichen, die mit `&!##mmmm/nnn` angegeben werden, sind zusätzlich 16 statt bisher zusätzlich 8 Schriften möglich.

In der Schriftfamilie *AGaramond* werden jetzt echte Kapitälchen (31454) sowie in den Schriftschnitten normal (31451), kursiv (31452) und halbfett (31453) Mediaevalziffern und die Ligaturen ff, fi, fl, ffi, ffl unterstützt.

Zur Wahl der kulturell korrekten Form der mit " kodierten Anführungszeichen wurde der Parameter FRZ durch den Parameter AFZ ersetzt. Näheres dazu ist der Online-Beschreibung zu entnehmen.

Da die Satzausgabe immer häufiger direkt auf die Druckplatte statt auf Film belichtet wird und auf der Druckplatte nicht mehr retuschiert

werden kann, war es sinnvoll, die Anweisungen zum Unterdrücken von Seitennummern und Kolumnentiteln zu überarbeiten. Die Steueranweisungen `&!q-` `&!q.` `&!l-` `&!l.` `&!n-` `&!n.` wirken deshalb jetzt immer erst ab der nächsten Spalte/Seite, unabhängig von der Lage von Seitenzahl und Kolumnentitel und von deren Art.

Beim Makro `#*PSAUS` kann die Angabe zu RAHMEN jetzt auch in mm (mit bis zu zwei Dezimalstellen) gemacht werden. In diesem Fall wird die sogenannte `HiResBoundingBox` in die PostScript-Datei eingesetzt (auch dies ist ggf. für Belichtung direkt auf die Druckplatte erforderlich).

Beim Makro `#*GRAFIK` wird, wenn zwei Dateien zur Spezifikation ZIEL angegeben sind, die `BoundingBox` der Grafik in die erste Zielfeile als Kommentar in die vierte Zeile zur jeweiligen Grafik mit ausgegeben. Dies erleichtert die (automatische) Berechnung des Platzbedarfs für diese Grafiken.

Wilhelm Ott
ott@zdv.uni-tuebingen.de