

Division of Chemical Information (CINF) Poster Session
National ACS Meeting, San Francisco, April 6, 1992

KEDIT:

Text Editor For Post-Processing Searches

Thomas E. Wolff,
Amoco Corporation
P. O. Box 3011, M/C F-1
Naperville, Illinois 60566
(708)420-4662
mhs!amoco!thomas_e_wolff@attmail.com

The online information search involves many steps, from determining search requirements, formulating the search strategy and carrying out the online search, to reviewing and presenting the search results. Each step may be time-consuming, but the most repetitive ones and, therefore, most amenable to streamlining, are the post-search processing steps. We have found that text editor KEDIT is invaluable for editing captured search sessions and formulating search strategies for uploading to the online service. KEDIT (pronounced *KAY-edit*) is a full-screen text editor which runs on DOS- or OS/2-based personal computers and which works with standard DOS ASCII text files. While KEDIT is generally used by programmers, the reasons for their enthusiasm are at least equally valid for searchers. KEDIT is especially well suited for editing lengthy text database output, both because of the power of KEDIT and the regularity of the search session. The characteristics, uses and advantages of KEDIT for post-search processing will be presented.

KEDIT Basics

- KEDIT's screen layout
 - File Area - "usual" text generation and cursor movement
 - Current line
 - Command line
 - Scale and Tab lines
 - Informational areas: ID and Status lines
- KEDIT Commands
 - KEDIT functions, or KEDIT or user macros
 - Invoked by Enter when cursor is on the Command line
(For cursor in File Area, Enter acts as "carriage return")
 - Function keys (e.g., F1 - F12) and combination keystrokes (e.g., Ctrl-A, Alt-1, Shift-F6)
 - Commands include cursor movement, text editing (change; find and locate; copy, move, add, delete or sort lines; block manipulation), editor configuration, page formatting, printing, and many DOS commands
 - Commands act on Current line or on Cursor line per command or macro definition
 - Keystroke commands may be informative or global, e.g.:
 - help (F1)
 - quit the current file (F3)
 - redisplay previous command (F6, 15 command recall buffer)
 - reissue previous command (F9)

Typical KEDIT Screen

ID line
└─▶ c:\capture\911031.cap Line=10 Col=6 Size=19523 Alt=3,0

Scale
└─▶ |...+....1....+....2....+....3....+....4....+....5....+....6....+....7....]+....8
 * * * Top of File * * *

L6 ANSWER 1 OF 1198
 COPYRIGHT (C) 1991 AMERICAN CHEMICAL SOCIETY

Cursor └─▶ **Highlighted Current Line**
AN CA115(10):105213t
TI Coumarin determination in gasoline
AU Kataoka, Masayuki; Seta, Yasuhiro
CS Nippon Soda Co., Ltd.
SO Jpn. Kokai Tokkyo Koho, 3 pp.
PI JP 03087660 A2 12 Apr 1991 Heisei
AI JP 89-223213 31 Aug 1989
SC 80-6 (Organic Analytical Chemistry)
DT P
AB Coumarin concn. in gasoline is detd. by adding an alk. org.
 solvent to the gasoline, stirring, and irradiating the gasoline
 with UV radiation to measure fluorescence intensity. Coumarin
 concn. .gtoreq.0.01 ppm in gasoline can be detd. visually.

====>
└─▶ KEDIT 4.00D3 Files=6 Memory=4028K 4:31pm 'J'=4A/074

Status line

Command line

KEDIT Strengths - 1

- Large document size limits- up to 8 MEG in DOS (ca. 150K lines)
- KEDIT is Fast
 - Files in system RAM - "instantaneous" simple movement
 - Search a 100K line file for a five-letter word in ca. 17 sec (Compaq 386/25 MHz, KEDIT in DOS under Windows 3.0)
- Good documentation and technical support (telephone, BBS's)
 - Useful tutorial in the User's Guide
 - Comprehensive Reference Manual for advanced features
- Powerful Locate and Change Functionality
 - Number targets - absolute or relative, lines or columns (e.g, line no. 843, back 27 lines, forward 25 columns)
 - Character string targets, e.g. the line containing "US 3678124" or the column in which begins the word "Journal"
 - Selective access and editing of text areas
 - Set Zone - act on specified columns only, e.g., first 8 col.
 - Set Display/Scope - act on selected lines only, e.g., those containing index terms
 - All command - displays only lines with specified target, useful to create lists, e.g., patent or registry numbers, class codes, to sort for display or prepare search strategies for uploading
 - Boolean logic - "and", "or", "not", or combinations

KEDIT Strengths - 2

- KEDIT has most standard word processors functions
 - Text blocks: line, box, or stream; functions such as copy, move, delete, case change, etc.
 - Paragraph targets, text between two blank lines, especially useful for search output
 - Can not control font changes, headers, footers, or pagination
 - Printing - use KEDIT, utility programs, or import into word processor (can add control characters via KEDIT)
- Customize almost any function or setting
 - Personal profile for each session
 - Substitute almost any keystroke combination
 - Modify screen layout or color scheme
 - Over 80 Set commands available, but defaults are satisfactory
- KEDIT Macros in KEXX (or REXX) Language
 - Carry out simple changes in KEDIT defaults or complex programming
 - Initial investment to learn, but macros easily maintained
 - Comparable in power to BASIC
 - Strength in interpreting and manipulating character strings
 - "Keystroke" macros - store in KEDIT Macro Library (KML) files
 - "Command line" macros - input parameters after macro name

Keystroke Macros for Post-search Processing

- **Scrolling by paragraph** (Ctrl-D, Ctrl-S) - the first line of the paragraph brought to the current line - facilitates browsing
- **Half screen scrolling** (Ctrl-CursorUp, Ctrl-CursorDown) - complements scrolling by paragraph, especially for long citations
- **Marking paragraph blocks** (Ctrl-G) - Since database records often correspond to paragraphs, single or groups of records can be easily manipulated, e.g. deleted; copied or moved within the file or to other files; shifted right or left; changed to upper or lower case.
- **Classification** - Transfer paragraph blocks selectively into twelve files using the Alt-F1 through Alt-F12 keys, then reassemble.
- **Pagination** - manually or automatically add "form feed" characters before citations so that none continues to the next page unless unavoidable (e.g., citation is >56 lines long). Check for citations of 57-60 lines to see if editing would be appropriate.

Command Line Macros - Manipulation of Output Fields

- **Macro WEED - Selective removal of information fields**
 - Accepts 1-15 ARGuments including:
 - BLANK - Eliminates all "extra" blank lines
 - ANSWER(S) - Deletes all lines with "ANSWER" in col. 1-17
 - COPYRIGHT - Deletes all lines with "COPYRIGHT" in col. 1-17
 - Field codes with indicator of source online system (=, -, *)
 - "WXYZ" - STN ("AU __", "PRAI _", etc. in col. 1-5)
 - "WXYZ=" - ORBIT ("TI __ - _", "XRAM- _" etc. in col. 1-6)
 - "WZ-" - Dialog tagged ("PA- _", "AB- _", etc. in col. 1-4)
 - "WZ*" - pre-1990 STN ("AN __", "SC __", etc. in col. 1-4)
 - Arguments entered lower case will be converted to upper case
 - Fields with more than one line of text are completely deleted
- **Macro SHOW - Selective field display**
 - Accepts 1-15 field codes similar to WEED macro
 - Blank lines retained between database records, i.e. paragraphs
 - Manipulate whole paragraphs while viewing selected fields only
 - Macro KEEP to make Show permanent
 - Macro UNSHOW or All to undo SHOW
- **Copyright macros** - Add copyright information to each citation, e.g., AN field of CA/STN, or abstract field of PTS PROMT/Dialog)

Macro WEED - Excerpt of KEXX Code - 1

* Check for ARGument passed to WEED macro, otherwise give HELP

```
if arg() = 0 then do
  'msg HELP: Type "weed xyz" where xyz = "blank", "ti", etc.'
  exit
end
```

* Read in fields and separate into individual entries

```
fields = arg(1)
do i = 1 to 15
  field.i = word(fields,i)
  if length(field.i) > 1 then imax = i
end
```

* Loop through the imax number of fields and weed them out

```
do i = 1 to imax
```

* Weed out fixed length fields (2-4 chars) in 1st several columns

```
lengfld = length(field.i)
```

* Determine if field is STN, ORBIT, DIALOG tagged, pre-1990 STN

```
ftype = 0
if substr(field.i,lengfld,lengfld) = '=' then ftype = 1
if substr(field.i,lengfld,lengfld) = '-' then ftype = 2
if substr(field.i,lengfld,lengfld) = '*' then ftype = 3
```

* If field is an ORBIT format field with equals sign indicator,
* adjust field to 5 characters with space-filling as necessary

```
if ftype = 1 then do
  zoneleng = 6
  lengfld = lengfld - 1
  field.i = substr(field.i,1,lengfld)
  if lengfld = 2 then fcode = upper(field.i | ' - ')
  if lengfld = 3 then fcode = upper(field.i | ' - ')
  if lengfld = 4 then fcode = upper(field.i | '- ')
end
```

* IF field is in DIALOG tagged, pre-1990 or current STN format,
* adjust field with appropriate space-filling as for ORBIT
* field described above

Macro WEED - Excerpt of KEXX Code - 2

* Next, reject all field indicators not 2-4 letters long

```
if lengfld = 1 then do
  fieldmsg = upper('' || field.i || '')
  'msg ***** Fields must be 2-4 chars.; ' fieldmsg 'rejected'
  iterate
end
if lengfld > 4 then do
  fieldmsg = upper('' || field.i || '')
  'msg ***** Fields must be 2-4 chars.; ' fieldmsg 'rejected'
  iterate
end
```

* Delete all lines which are part of selected fields (2-4 chars.)

```
'top'
set zone 1 zoneleng
set case mixed respect respect
set wrap off
set hex on
do forever
  'nomsg locate/'fcode '/'
  if rc \= 0 then leave
  'delete'
  'sos current'
  'sos cup'
  linenol = cursor.3()
  do forever
    'nomsg locate /x''20202020''/'
    if rc \= 0 then leave
    'extract /cursor/'
    if blank() then leave
    if cursor.3 - linenol \= 1 then leave
    'delete'
    'sos cup'
  nop
  end
  'sos qc'
  'top'
end
fieldmsg = upper('' || fcode || '')
'sos beep'
'msg          All' fieldmsg 'lines deleted'
end
```

```
'set zone 1' trunc.1()
'set case mixed ignore respect'
'set wrap on'
'set hex off'
```

Command Line Macros - General Purpose

- **Removing extraneous characters** (macro Hexclean) - remove the ◀ (^Q, or ASCII 017) from Dialog output; exchange graphic highlighting characters in STN Express transcript files for text "***"
- **Reformatting citations from "regular" untagged output**
 - Use series of All commands to remove: ISSN, word count, text availability indicators, bylines, null dates or page numbers, etc.
 - Case change, e.g. citation title to uppercase characters
 - Move whole line, e.g. source information from end of citation to just below the title
 - Placing tab characters at beginning of each new paragraph
 - Close-pack right-justified information
- **Preparation of edited report for importing into word processors**
 - Insert of CR/LF (new paragraph) indicators
 - Add font control characters based on Rich Text Format used by Word for Windows - generally intuitive strings: e.g., "\tab", "\b{...}" (boldfaced) or "\i{...}" (italics)
- **Conversion of word processor ASCII output to clean flat file** - translate tab characters into blank spaces, weed out extra blank lines and format text within flat file margin settings, as required for e-mail or other applications.

PTS PROMT (Dialog File 16) Citation as Downloaded

03621094

Recycling insulates PS from landfills

It wasn't easy, incorporating recycled polystyrene into a new brand of insulation board for use in home and commercial construction. First, Amoco Foam Products Co. had to work with its sources of recycled PS to make sure their material would not 'gum up the works,' business manager Peter Sullivan said.

After several years of development, Amoco Foam has commercialized the board, called Amofoam-RCY. The insulation contains a minimum of 50 percent recycled material, including unspecified percentages of post-consumer and industrial scrap.

THIS IS AN EXCERPT: Copyright 1992 Crain Communications Inc.

FULL TEXT AVAILABLE IN FORMAT 9

WORD COUNT: 706

Plastics News February 10, 1992 p. 5

ISSN: 1042-802X

Modified by KEDIT Macros, Including Rich Text Format (RTF) Coding

<><

<>{\b RECYCLING INSULATES PS FROM LANDFILLS}

<>\tab {\i Plastics News, February 10, 1992, p. 5}

<>\tab It wasn't easy, incorporating recycled polystyrene into a new brand of insulation board for use in home and commercial construction. First, Amoco Foam Products Co. had to work with its sources of recycled PS to make sure their material would not 'gum up the works,' business manager Peter Sullivan said.

<>\tab After several years of development, Amoco Foam has commercialized the board, called Amofoam-RCY. The insulation contains a minimum of 50 percent recycled material, including unspecified percentages of post-consumer and industrial scrap.

<>\tab THIS IS AN EXCERPT: Copyright 1992 Crain Communications Inc.

Output by Word Processor after no Additional Formatting

RECYCLING INSULATES PS FROM LANDFILLS

Plastics News, February 10, 1992, p. 5

It wasn't easy, incorporating recycled polystyrene into a new brand of insulation board for use in home and commercial construction. First, Amoco Foam Products Co. had to work with its sources of recycled PS to make sure their material would not 'gum up the works,' business manager Peter Sullivan said.

After several years of development, Amoco Foam has commercialized the board, called Amofoam-RCY. The insulation contains a minimum of 50 percent recycled material, including unspecified percentages of post-consumer and industrial scrap.

THIS IS AN EXCERPT: Copyright 1992 Crain Communications Inc.

T. E. Wolff

#11

Amoco Corporation

April 1992

Command Line Macros - Editing of Downloaded Files for Importing into Personal Bibliographic Databases

- Translate patent and application numbers in CA (STN) to Derwent (Dialog) format
(e.g., CA: CC YY-NNNNNNN == > Derwent: CC YYNNNNNNN)
- Translate ISO Dates (YYMMDD) to STN format (DD MMM YY)
- Edit field descriptors and create unique fields names (Dialog Files)
(e.g., "PI- <AD,AN>" == > "AN- " in IFI file)
- Parse original patent field(s) into patent number and date fields
Derwent and IFI: parse the PN field;
CA: rearrange PI and SO fields
- Prepare date-sorted lists of equivalents, applications
(one entry per line)
- Determine the earliest application priority and make separate earliest application number and date fields --
Derwent and IFI: PR field; CA: AI and/or PRAI fields
- Add commas after author first names for Library Master
- Delete, add, change or move separator characters
(e.g., virgule, double slashes in author, index term fields))

**Derwent World Patent Index (Dialog) Citation as
Downloaded in Tagged Format 4**

8/4/8

AX- 87-110218/16|

AX- <XRAM> C88-088483|

TI- Rhodium-carbon catalyst prepn. for use in terephthalic acid purification|

PA- (STAD)_AMOCO CORP|

AU- <INVENTORS> SCHROEDER H; WITTMAN R L|

PN- <BASIC> EP 219288 A 870422 8716|

PN- <EQUIVALENTS> JP 62121647 A 870602 8727; US 4728630 A 880301 8812; CN 8606590 A 870422 8828; EP 219288 B 890906 8936; DE 3665416 G 891012 8942; ES 2011014 B 891216 9007|

AN- <PRIORITIES> US 905758 (860909); US 785055 (851007)|

AN- <APPLICATIONS> JP 86238986 (861007); EP 86307678 (861003); EP 85785055 (851007)|

AB- <BASIC> EP0219288

Prepn. of a Rh-carbon catalyst (III) is effected by contacting a porous carbonaceous material (I), having a surface area of at least about 600 m²/g and pH 9-11 in aq. suspension, with a pH 1-4 aq. Rh³⁺ salt (V) soln. (II) such that (I) contains about 0.01-2 (pref. 0.5) wt.% Rh pref. to a depth of at least 5 microns (more pref. 10-20 micron), and pH (I) plus (II) is about 12-13.5.

The pref. pH of the impregnated wet catalyst compsn. is about 6-8. The wt. ratio of solids to liquids is at least about 0.5 (pref. 0.5-20, more pref. about 0.5-2). The contacting of (I) with (II) is by either soaking or spraying.

USE/ADVANTAGE - (III) is used in the purification of terephthalic acid (IV) which is the starting material ... for prepn. of polyester fibres, films and resins. (III) has enhanced selectivity for decarboxylation of 4-carboxybenzaldehyde (the principal impurity in the prepn. of (IV)) and removes colour and fluorescence. @(12pp Dwg.No.0/0)|

AB- <US> 8812 US4728630

Rhodium-on-carbon compsn. is made using an alkaline granular porous carbonaceous catalyst support material of surface area 600 sq. m. per g. or more and pH 9-11.

Process comprises contacting support material with an aq. Rh (3+) salt soln. of pH 1-4 for a time to form a wetted support material contg. 0.01-2 wt.% Rh (as metal w.r.t. dry catalyst). Support material and salt soln. are preselected so that sum of pH in each is 12-13.5. Pref. Rh-salt is rhodium trinitrate or -trichloride.

USE - For purification of terephthalic acid under reducing conditions. @(7pp)|

