CodeIndex.c

KEDS Project (Kansas Event Data System)

DESCRIPTION

This program produces an index of the codes from a TABARI .actors or .verbs dictionary. The output is effectively the reverse of the TABARI dictionaries: these associate phrases with codes, whereas this associates codes with phrases.

The output is formatted hierarchically using CAMEO actor and event coding conventions. The output can be either a text file with indentation done using tabs, or an HTML <table>.

RUNNING THE PROGRAM

The program runs as a *terminal* application on the Macintosh (or as a command-line application in any other Unix system).

USER INPUT

The program asks for

1. Name of the file to be indexed

2. Whether this is an actor (A) or verb file (V)

3. Whether the complete input string should be stored (Y), or just the pattern (N). The first option includes comments in the output

4. Whether the output file should be an HTML table (Y), or tabbed text (N)

Error checking is *not* done in steps (2) - (4)—the program defaults to "A" and "Y" if unexpected input is received.

INPUT FILES

The input files are standard TABARI .actors or .verbs dictionaries.

SCREEN FEEDBACK

Program shows the entry that is currently being evaluated -- this is useful in the event that the program crashes due to some unexpected input. On a Mac G5, it processed a 6,000 line verbs file in about two seconds, so if it is taking noticeable time, something is wrong. There is reasonable error checking for situations where there could be memory problems.

OUTPUT FILES

The output has the name of the input file followed by the suffix ".idx.html" or ".idx.html". Actor codes are indented when they match on 3-character blocks (following the CAMEO actor coding conventions), and verb (event) codes on the

third and fourth character (following WEIS and CAMEO conventions). The verb in a pattern is prefixed with *, so for example

```
ACCEPT [081] ;jw    11/14/91
– NOT * PROPOSAL              [111] ; pas 98.3.2
– DOES NOT *                  [111]
```

will show

```
– NOT *ACCEPT PROPOSAL
– DOES NOT *ACCEPT
```

under the code 111.

SYSTEM REQUIREMENTS:

This program was originally developed on the Macintosh operating system (OS 10.x) but contains no Macintosh-specific features and will presumably work on any system that has a gcc compiler. Current version was compiled under Macintosh OS-10.4. The source code is reasonably well commented.

Compiler: gcc
Language: ANSI C
To compile source code: gcc CodeIndex.c
To run: ./a.out

REVISION HISTORY

July-05: Initial version (0.8B1), based very loosely on OneADayFilter.c

PROVENANCE

Programmer:
  Philip A. Schrodt
  Center for International Political Analysis
  University of Kansas
  Lawrence, Kansas, 66044 U.S.A.
  http://www.ku.edu/~keds
  email: schrodt@ku.edu

Report bugs to: schrodt@ku.edu