MAKEDATABASE

A Facility for Automating the Generation and Maintenance
of Masterscope and Whereis Databases

author: Jon L White
files: [Eris]<Lisp...Library> -- MAKEDATABASE and PAGEHOLD
documentation: MAKEDATABASE.tedit/press
modified: 10-Mar-84 and 9-Oct-84 by JonL

This package contains two functions of interest -- **STARTMAKINGDATABASE** and **CONTINUEMAKINGDATABASE** -- each of which takes three arguments:

FILESLST	- a list of files to be masterscope analyzed
OUTPUTFILENAME	- the name of the resultant output file
CONSCOUNTLIMIT	- the "boxcount" at which the in-core database will be dumped out.

The effect is to produce a **.DATABASE** file of all the source files in **FILESLST**; masterscope primitives will be called to analyze each file, but after every so many "cons"'s, it will incrementally dump out the results. This is particularly useful when the time required to analyze all the files in **FILESLST** is on the order of hours, or when the amount of information in such analyses exceeds the capacity of Interlisp-D's limits (e.g., address space exhaustion, or excessive swapping due to smaller real memories).

STARTMAKINGDATABASE opens a new file for output; **CONTINUEMAKINGDATABASE** assumes that **OUTPUTFILENAME** is an existing file with some previous number of outputtings already on it. Thus if there is a crash after analyzing, say, the first 50 files of **FILESLST**, one can start up calling **CONTINUEMAKINGDATABASE** with the remainder of the file names, and not have to repeat the analysis of the first 50 files.

Currently, the output file is a text file of data, as produced by **DUMPDATABASE** (q.v.); as soon as a reliable masterscope hash file facility is operable in Interlisp-D, it will connect into that format. The format of the text output file is such that it can always be **LOAD**'d in to Lisp; it always has a **STOP** at the end, and the **DUMPDATABASE** phase is somewhat protected so that it doesn't leave a partial dump at the end of the file. Furthermore, each group of files whose analyses are dumped out is preceeded by a line in the output file like
(* files analyzed: <Name1> <Name2> . . . <Namen>)
where the various **<Name¡>** are the root filenames of the files analyzed and dumped. The dumped information appears just after this comment-like form. The reason for this line of commentary is so that if a crash occurs, one can search the partial output file (using the function **BFILEPOS** found in the library package BSEARCH) for the last occurence in the file of a pattern like
"
(* files analyzed:"
and then using **COPYBYTES** simply view that line to see the names of the last few files successfully analyzed.