
31-MAR-75 09:48:14,6278

Date: 31 MAR 1975 0948-PST

From: LEDERBERG

Subject: FIRST BLOCK OF COMMENTS ON CONGEN ORGANIZATION AND DOCUMENTATION

- - - -

PRUNE should not eliminate, but rather segregate a set of graphs into the "accepted" / "rejected" categories. The default selection for further processing would be the last accepted set; but the others would still be available unless cleared by the user.

+

1; <NORMAL>ECL-DUMP.SET;1 Fri 22-Sep-78 8:09PM

PAGE 411:78

> The User should also have the option of requesting a reverse- or de-IMBED function (when the constraint is a BL or GL item, the program has already done almost all the necessary work. A new command -- clear up redundancies -- may then be needed, to be sure. In complex structures, it may not be so easy for the user to recognize a superatom.

Each new aggregate should have a name to be used as a handle perhaps == an external file name; As well as names for single structures in an aggregate

CONSTRAINTS might well include some graph-oriented criteria. Besides: NON-Planar already mentioned, perhaps also

SINGLY-connected or with specific reference to a particular SUBstructure
**re (if 1-connected)

or even a singly-connected composition.

The user might even want to specify a 1-connected low-res mol. wt. as something he can infer from his data.

GRAPHS I have been running some tests on CONGEN by way of building n12 and ? n14 vertex graphs. What is the generator algorithm now implemented? Do you have some #SAVED files of such graphs?

DETACH: I did try (GCGAG NIL) and detach...continue apparently successfully on one of these bigish problems.

INTERRUPT Is it possible to break during generation And resume?

IS THERE ANY EASY WAY TO GET THE EQUIVALENT OF SYSTAT FOR JUST ONES OWN JOBS --E.G. TO MONITOR A DETACHED ONE? WHO WASTES TIME PRINTING THE WHOLE USER SET.

(Not for implementation)

A complete structure-manipulating language, on analogy of TECO would allow for the user construction of macros from the primitive commands. Not too virtuous here.

But I had in mind the analogy of a "line" in a text editor, to a structure U/e superatom in CONGEN. For example, in ES, a command should allow the recopying of a given structure n times, or other simple sequences of operations.

It wouldn't hurt if the general format of the ES interaction more generally resembled that of a polished text-edit language like TECO or TV-EDIT.

Is that a pipe-dream?

CONGEN.DOC COMMENTS on the documentation

+

1; <NORMAL>ECL-DUMP.SET;1 Fri 22-Sep-78 8:09PM

PAGE 411:79

p.6:2 ALL is not that clearly explained. One soon discovers that it means the other commands will be prompted.

p.7:3 para.5 "bond multiplicities are taken literally" - should be more carefully explained. I assume it means that a = is regarded as an additive conjunction of two -.

p.7:5 ATNAME This is the offending statement that Polynames... should not be used for superatoms. By implication, for superatoms to be used in composition lists. ES does Also refer to superatoms.

p.7:1 last line/ "no structures can possess" -- would be clearer as "no structures are permitted to contain". U "can" may be ambiguous. Indeed "may possess" perhaps already suffices.e

p.7:10 HINTS --par. 3 I haven't tried recursion What happens?

p.7:12 2) Explain significance of the query for intra-superatom bonds

How does CONGEN compare with PROPHET (what does latter do that may be similar)?

 CONGEN will not handle H2 correctly. (Nor have most versions of DENDRAL. But
 do you understand just why?)

Would it be possible, during ES, to get disconnected graphs ADRAWn or NDRAWn
 across the page instead of vertically? A small point, but it is rather frustrating
 trying to keep track of these complex structures while ESing them from a
 CRT terminal., and they are rolling by.

BREAKPOINT These computations are so long that we really do have to
 take account of that in engineering. There should be a breakpoint
 facility that can do autosaves of the current state of the generation
 perhaps every n garbage-collects, and which can be interrogated by
 the user to see what is happening. I realize all of these features
 take additional time and space; but much more will be wasted in tracking
 mysterious faults. All of this obviously will require further
 attention to compact representations.

Now that I have a better feel for the operation, I would very much like
 to dig into more of the details with you, viz-a-viz possible
 improvements in time-efficacy. Perhaps we need some more heuristic
 approaches to problems like solving the group; as well as deeper study
 of dictionary/compute-time tradeoffs.

STEREoisomers I persevere that there are really exciting possibilities
 here. Perhaps cis-trans geometric isomerism around $>C=C<$ is the
 simpler case to analyze and display.

+

1; <NORMAL>ECL-DUMP.SET;1 Fri 22-Sep-78 8:09PM

PAGE 411:80

DISPLAY. I really do favor more open formats. Look at what happens
 to RING 8+ in the ES ADRAW.

 Comments on memo 26-Mar

Much of above was composed before remarks about TCONGEN; which I
 have yet to try, but soon will. I was glad to note things like
 RAISE and the d effect. Yes, doubtless there should be an explicit RESET
 command.

If we could write macros like

do n

drawsome (n)

if tty: = then jump

or == in editor-like language, we would not have to make rigid decisions. But I think the SCAN alternative suggested is in the right ball park. If SCAN pauses for input at each DRAWing, you don't need SCANSOME -- that is one of the options available at the pause, i.e., what to go to next.

EXTENSIONS I foresee that one application of CONGEN will be to filter a file of known structures Ue.g. in interactive lookupe -- and this might be on quite a large scale. That's why I raised questions about merging files, etc. Ucall previous marks about analogy with a text-edit languagee. Perhaps we should again review what CA, ISI ARE doing in this field. UAnd ISI --surely not CA -- and perhaps also NLM might be interested in supporting further work to make CONGEN more efficient if it could serve in that kind of intelligent retrieval mide as well.]]e
