Hi Wei Wei,

I was giving some talks last week. One issue that comes up quite a bit is about the "spectral properties" of the graphs underlying the formulas.

The distribution of eigen values of a graph (more specifically its Lapacian) do reveal important global properties.

For example, how well walksat / SA samples may be a function of what the formula spectral properties are. There is bound to be some correlation between the behavior of the algorithms and the spectral properties of the formulas.

This would be a fairly easy to obtain result that are quite interesting.

Can you write a C program that takes a cnf formula and creates a (sparse) matlab representation? It would be nicest to have a program "cnf2spectrum" that creates a .m that when executed plots the eigenvalue distribution.

I have not thought about the full details of the graph representation. Possibilities:

1) for each variable have a node and an undirected link between two variables if they occur in the same clause.

2) for each literal have a node and an undirected link between two variables if they occur in the same clause.

3) a bipartite graph. one node for each clause. one node for each literal. a link between a "clause node" and a "literal node" if the literal occurs in the clause.

There are probably still more options; in particular, one issue is how to represent the difference between x and ~x occurring. However, as a start 1) would be good. Then later we can add options to the program to get different representations. (Including the Laplacian.)

-- Bart

p.s. The standard representation of a graph would be its adjacency matrix: an nxn matrix for an n-node graph. in row i, the entry i,j would be 1 iff there is an edge from i to j. for an undirected graph the matrix is symmetric.

use sparse graph representation in matlab.

test on some small formulas. does "random" look very different from "chain"? how about "chain" with redundancies?