Search for satisfaction

Toby Walsh

Cork Constraint Computation Center

tw@4c.ucc.ie

Cork Constraint Computation Center (4C)

- Gene Freuder (director)
  - **□6M from SFI**
- Toby Walsh (PI)
  - **□1.5M from SFI**
- ~20 staff
  - **Gene's 3C group from NH (Wallace, …)**
  - **Existing Cork people (Bowen,O'Sullivan,…)**
  - **Lots of new staff (Beck & Little from ILOG, …, your name here)**

- Research themes
  - **Modelling**
    - eliminating the consultant
  - **Uncertainty**
  - **Robustness**
- Cork
  - **Ireland's 2nd city**
  - **Cultural capital**

Health warning

- To cover more ground, credit & references may not always be given
- Many active researchers in this area:
  Achlioptas, Boros, Chaynes, Dunne, Eiter, Franco, Gent, Gomes, Hogg, …, Walsh, …,Zhang

Search for satisfaction

- Multi-media survey
  - **"hot" research area**
- My next stop
  - **5th International Symposium on SAT (SAT-2002)**
  - **1 panel, 3 invited speakers, 9 competing systems, 50 talks**

Satisfaction

- Propositional satisfiability (SAT)
  - does a truth assignment exist that satisfies a propositional formula?
  - NP-complete

(x1 v x2) & (-x2 v x3 v -x4)

x1/ True, x2/ False, ...

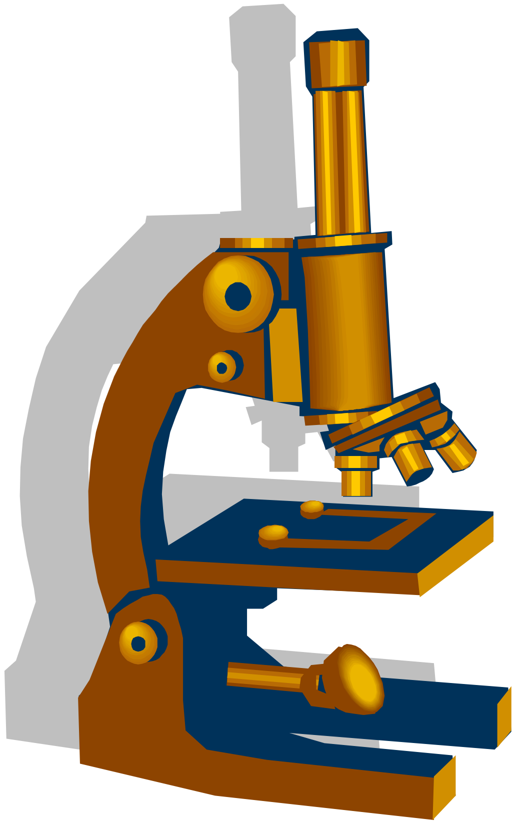
Satisfaction

- Propositional satisfiability (SAT)
  - **does a truth assignment exist that satisfies a propositional formula?**
  - **NP-complete**
- 3-SAT
  - **formulae in clausal form with 3 literals per clause**
  - **remains NP-complete**

(x1 v x2) & (-x2 v x3 v -x4)

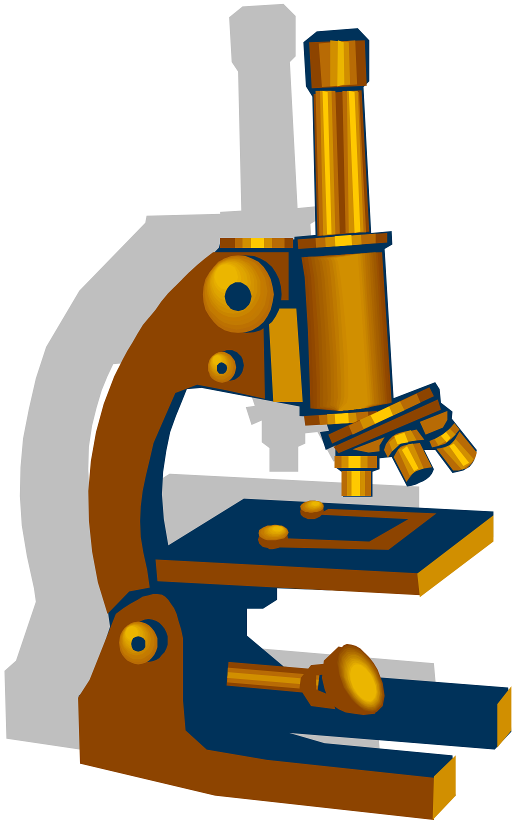
x1/ True, x2/ False, ...

Why search for satisfaction?



- Effective method to solve many problems
  - Model checking
  - Diagnosis
  - Planning
  - ...

Why search for satisfaction?

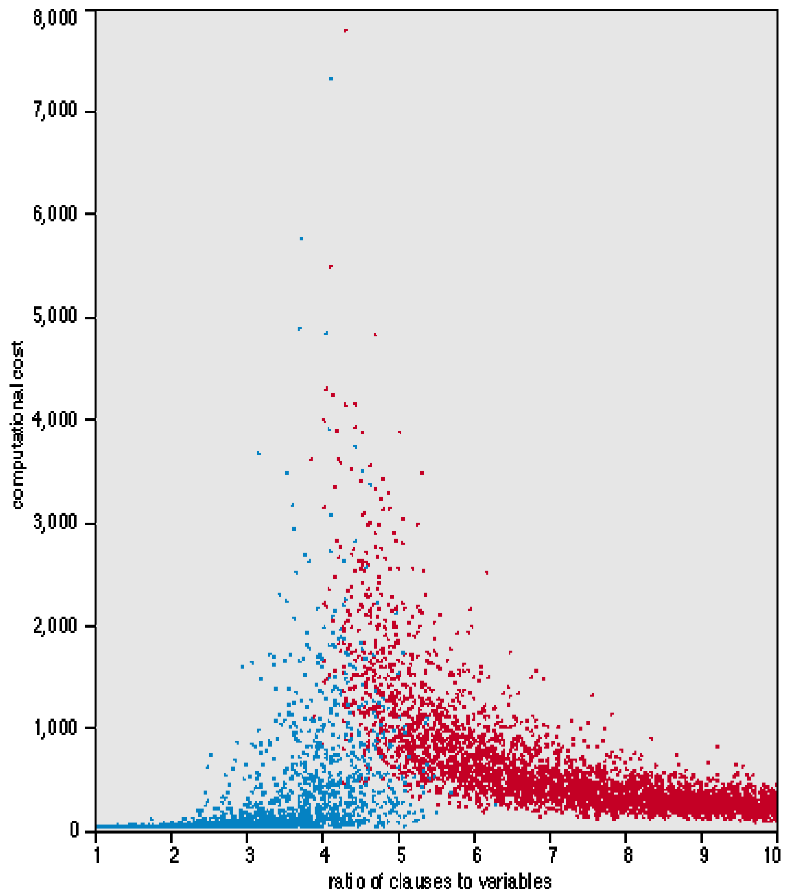


- Effective method to solve many problems
  - **Model checking**
  - **Diagnosis**
  - **Planning**
  - **...**
- Simple domain in which to understand
  - **Problem hardness**
  - **NP-hard search**
  - **...**

Outline

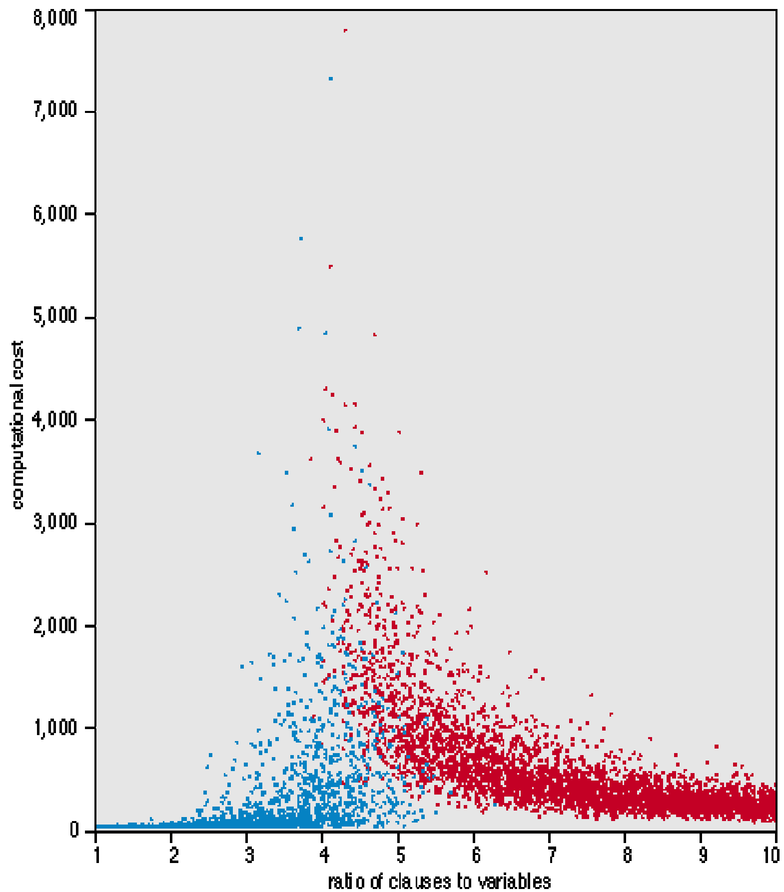
- SAT phase transition
  - **Why it might be important for you?**
- Problem structure
  - **Backbones**
- Real v random problems
  - **Small world graphs**
- Open problems
- Conclusions

Random 3-SAT



- Random 3-SAT
  - **sample uniformly from space of all possible 3-clauses**
  - ***n*** **variables,** ***l*** **clauses**

Random 3-SAT



- Random 3-SAT
  - **sample uniformly from space of all possible 3-clauses**
  - ***n* variables, *l* clauses**
- Which are the hard instances?
  - **around *l/n* = 4.3**

What happens with larger problems?

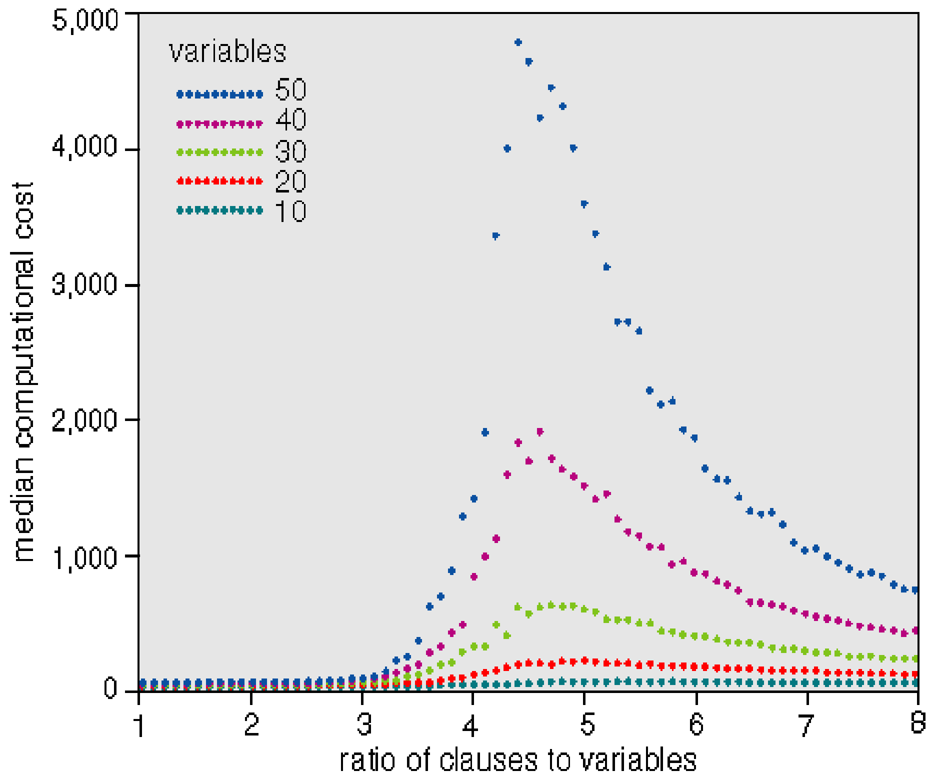
Why are some dots red and others blue?

Random 3-SAT

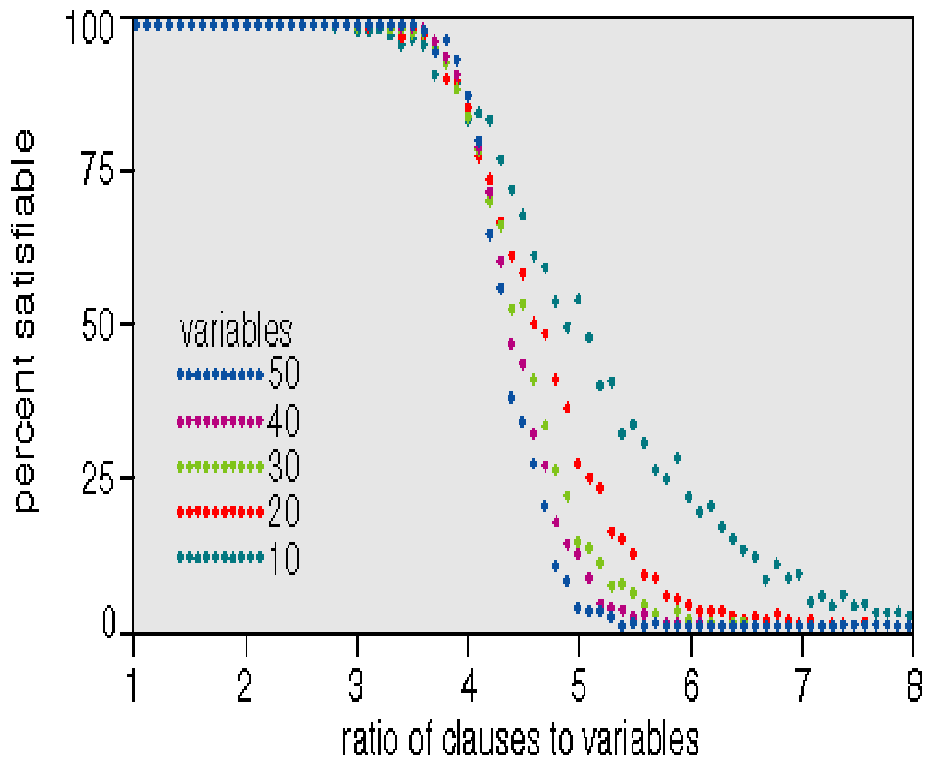
- Varying problem size, *n*

- Complexity peak appears to be largely invariant of algorithm
  - **backtracking algorithms like Davis-Putnam**
  - **local search procedures like GSAT**

What’s so special about 4.3?

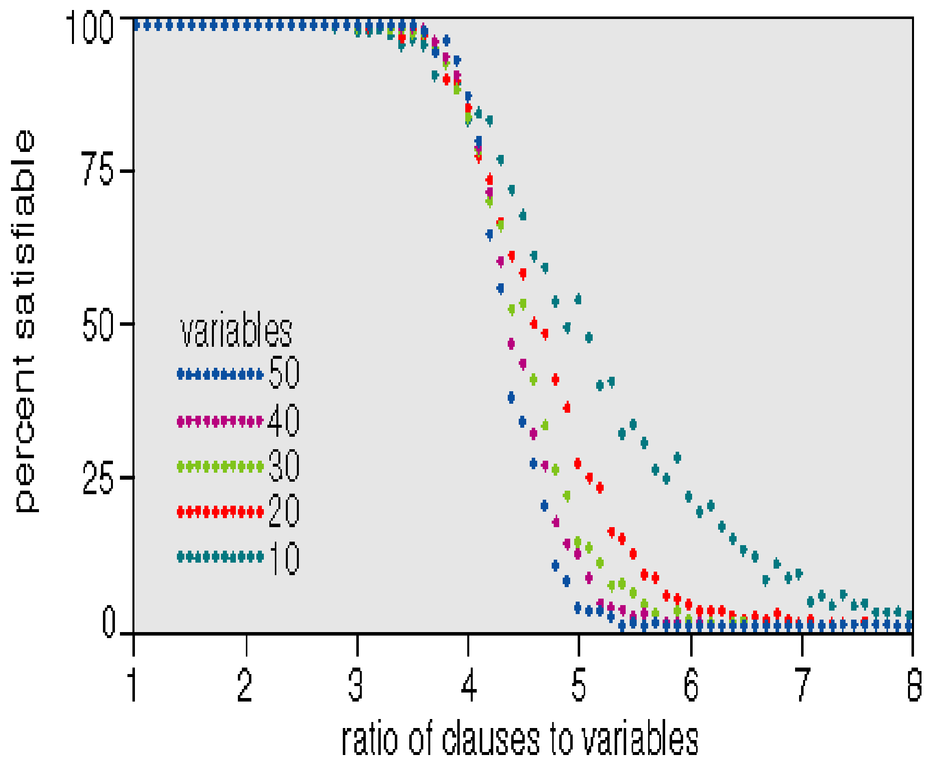


Random 3-SAT



- Complexity peak coincides with solubility transition
  - **l/n < 4.3 problems under-constrained and SAT**
  - **l/n > 4.3 problems over-constrained and UNSAT**

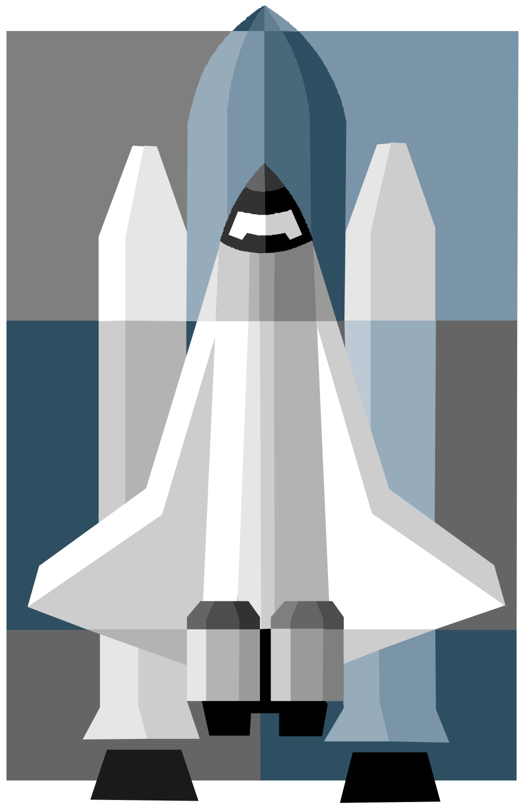
Random 3-SAT



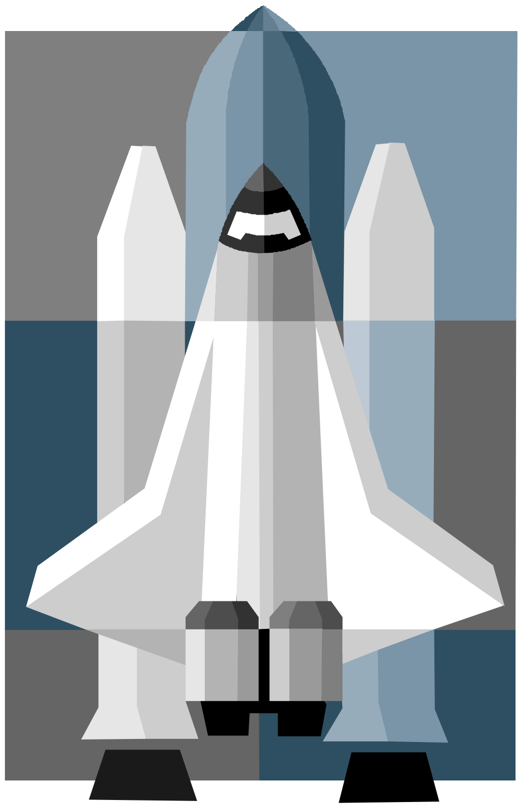
- Complexity peak coincides with solubility transition
  - **l/n < 4.3 problems under-constrained and SAT**
  - **l/n > 4.3 problems over-constrained and UNSAT**
  - **l/n=4.3, problems on “knife-edge” between SAT and UNSAT**

So, what's the relevance?

- Livingstone model-based diagnosis system
  - **Deep Space One**
- Tough operating constraints
  - **Autonomous**
  - **Real time**
  - **Limited computational resources**
- Compiled down to propositional theory …

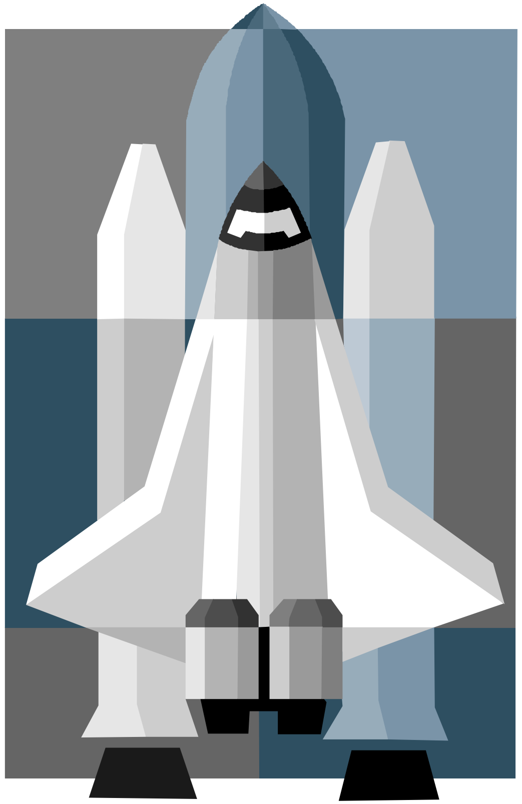


Deep Space One



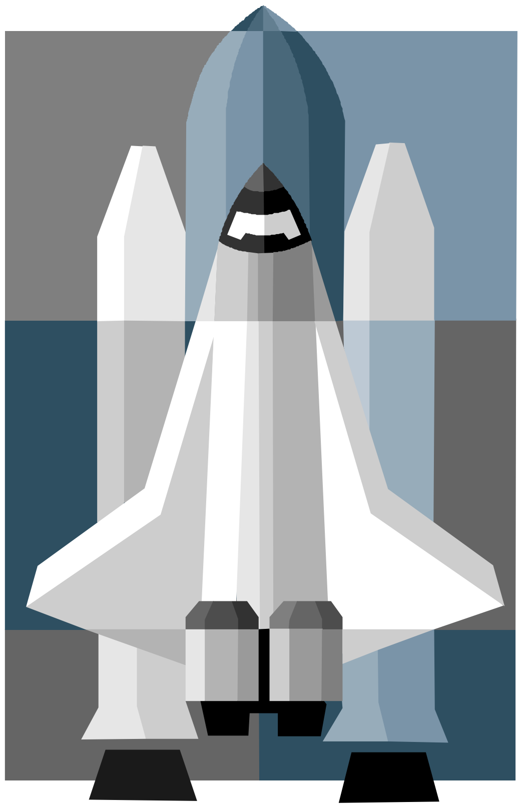
- Limited computational resources
- Deep Space One model has 2^160 states

Deep Space One



- Limited computational resources
- Deep Space One model has 2^{160} states
- Fortunately, far from phase boundary

Deep Space One



- Limited computational resources
- Deep Space One model has 2^160 states
- Fortunately, far from phase boundary
- Not so surprising
  - **Very over-engineered**

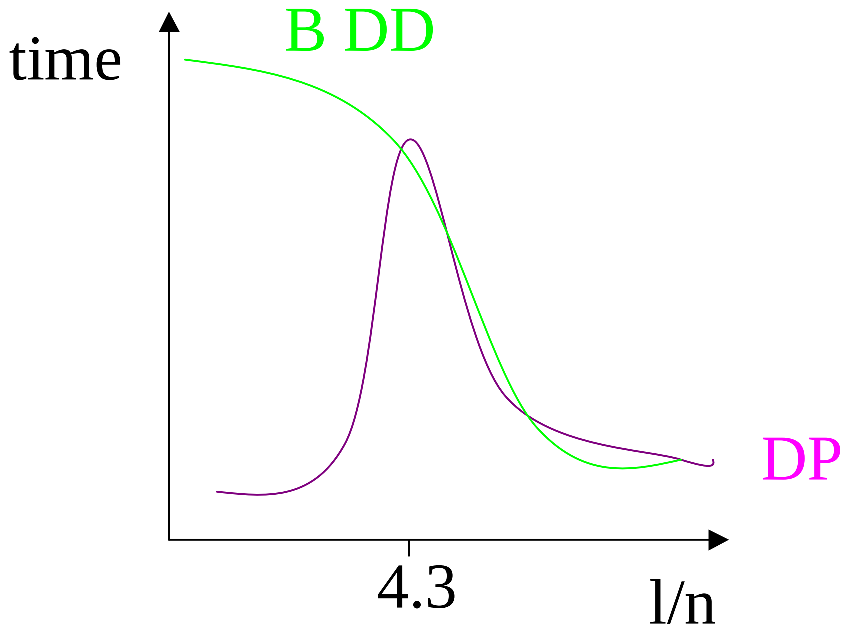
So what’s the relevance?

- Model checking
  - **Does an implementation satisfy a specification?**
  - **PSpace in general**
- So how can SAT help?
  - **It’s only NP-complete!**

So what's the relevance?

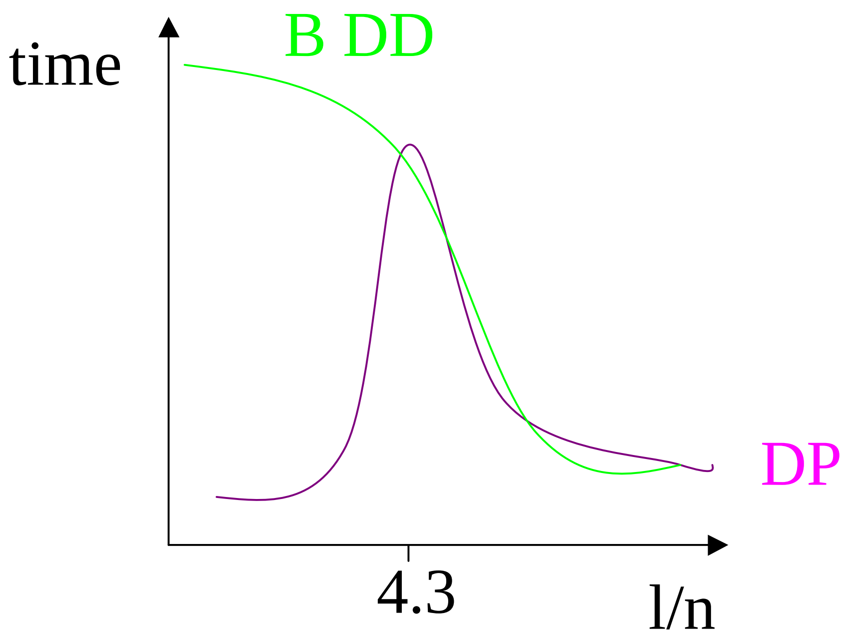
- Model checking
  - **Does an implementation satisfy a specification?**
  - **PSpace in general**
- So how can SAT help?
  - **Bounded model checking**
  - **Bound = path length in state transition diagram**

Model checking



- SAT solvers (e.g. Davis Putnam) very effective at finding bugs
  - **BDDs good at proving correctness**

Model checking



- SAT solvers (e.g. Davis Putnam) very effective at finding bugs
  - **BDDs good at proving correctness**
- Surprised it took so long to see benefits of SAT solvers
  - **DP is O(n) space, O(2^n) time**
  - **BDDs are O(2^n) space and time**
  - **Memory isn't that cheap**

“But phase transitions don’t occur in X?”

- X = some NP-complete problem
- X = real problems
- X = some other complexity class

Little evidence yet to support any of these claims!

“But it doesn’t occur in X?”

- X = some NP-complete problem
- Phase transition behaviour seen in:
  - **TSP problem (decision not optimization)**
  - **Hamiltonian circuits (but NOT a complexity peak)**
  - **number partitioning**
  - **graph colouring**
  - **independent set**
  - **...**

“But it doesn’t occur in X?”

- X = real problems

  No, you just need a suitable ensemble of problems to sample from?

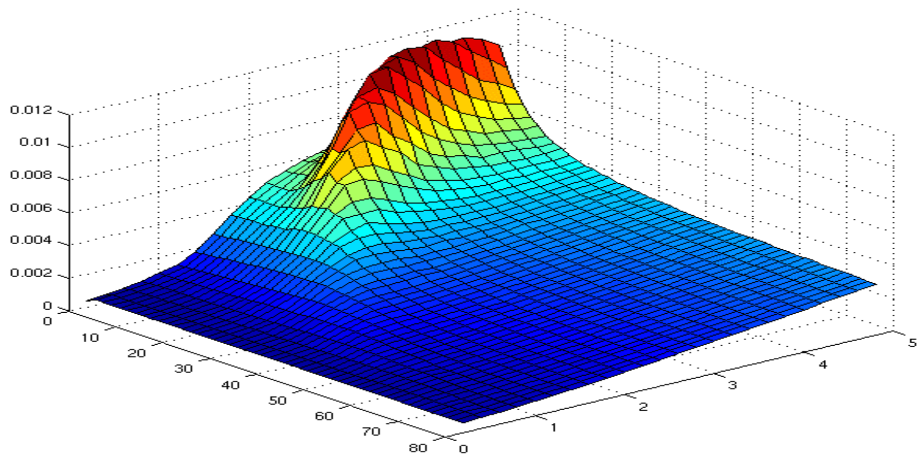
- Phase transition behaviour seen in:
  - **job shop scheduling problems**
  - **TSP instances from TSPLib**
  - **exam timetables @ Edinburgh**
  - **Boolean circuit synthesis**
  - **Latin squares (alias sports scheduling)**
  - **...**

“But it doesn’t occur in X?”

- X = some other complexity class

  Ignoring trivial cases (like O(1) algorithms)

- Phase transition behaviour seen in:
  - **polynomial problems like arc-consistency**
  - **PSPACE problems like QSAT and modal K**
  - **...**

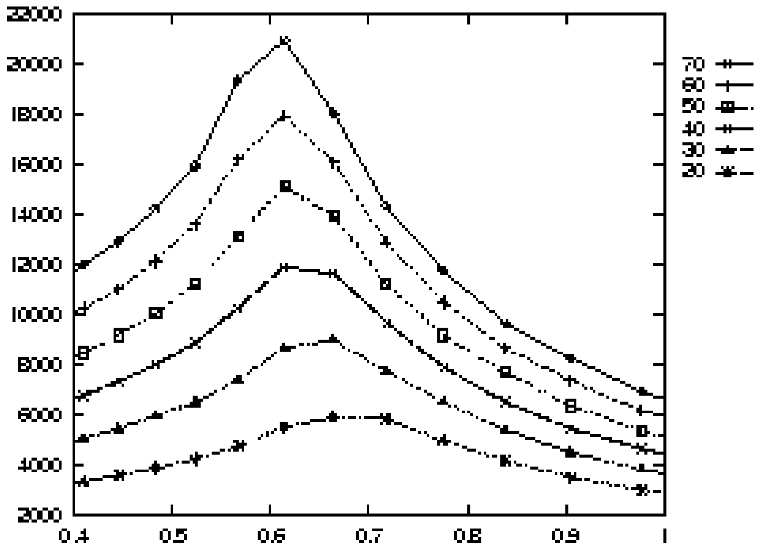


Random 2-SAT

- 2-SAT is P
  - **linear time algorithm**
- Random 2-SAT displays "classic" phase transition
  - **c/n < 1, almost surely SAT**
  - **c/n > 1, almost surely UNSAT**
  - **complexity peaks around c/n=1**

x1 v x2, -x2 v x3, -x1 v x3, …

Phase transitions in P

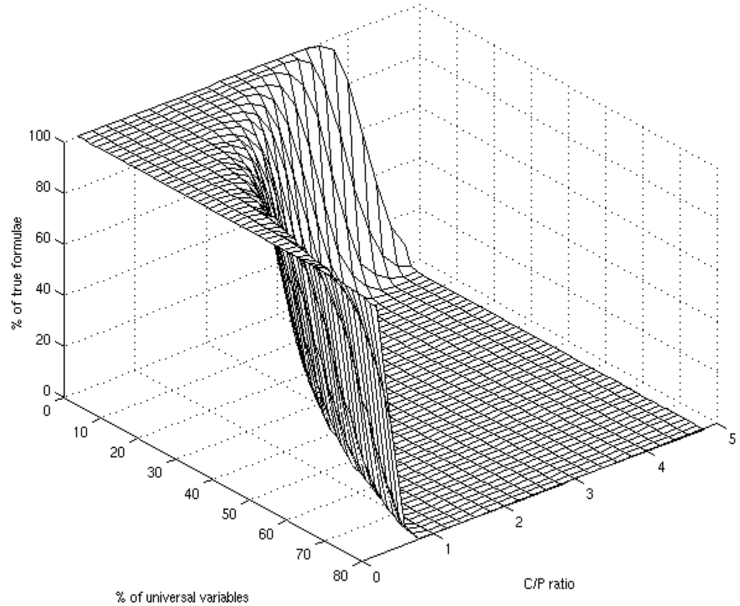


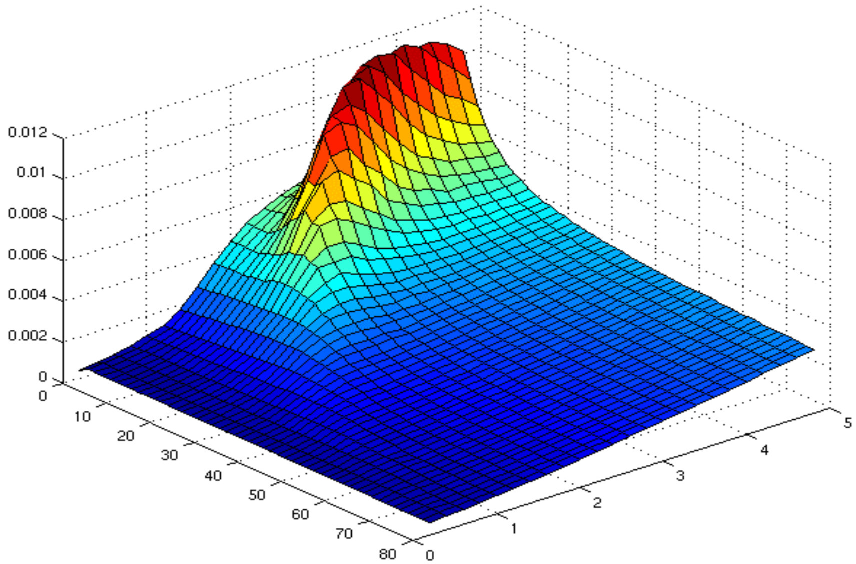
- 2-SAT
  - **c/n=1**
- Horn SAT
  - **transition not “sharp”**
- Arc-consistency
  - **rapid transition in whether problem can be made AC**
  - **peak in (median) checks**

Phase transitions above NP

- PSpace
  - QSAT (SAT of QBF)

  ∀x1 ∃x2 ∀x3 . x1 v x2 & -x1 v x3





Phase transitions above NP

- PSpace-complete
  - QSAT (SAT of QBF)
  - stochastic SAT
  - modal SAT
- PP-complete
  - polynomial-time probabilistic Turing machines
  - counting problems
  - #SAT(>= 2^n/2)

  [Bailey, Dalmau, Kolaitis IJCAI-2001]

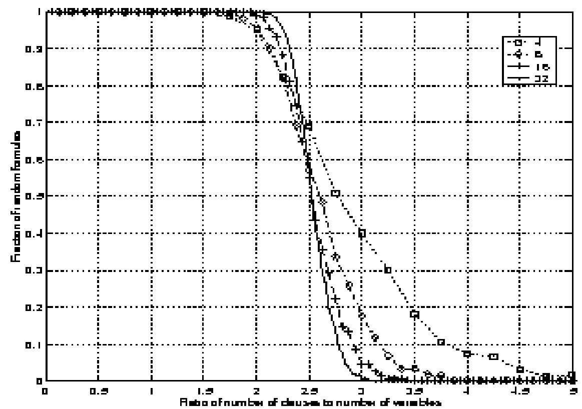


Figure 1: Phase Transition Graphs

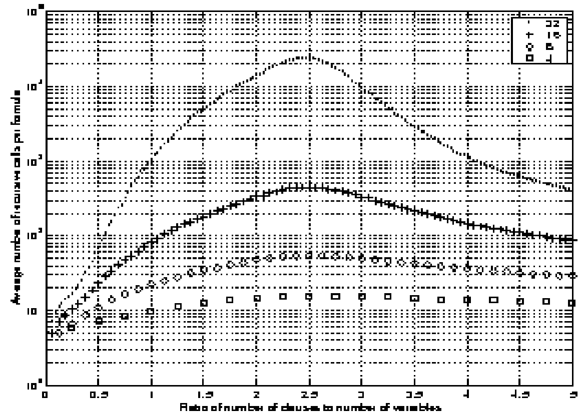


Figure 2: Performance Graphs

Exact phase boundaries in NP

- Random 3-SAT is only known within bounds
  - **3.26 < c/n < 4.506**

Are there any NP phase boundaries known exactly?

- Recent result gives an exact NP phase boundary
  - **1-in-k SAT at c/n = 2/k(k-1)**
  - **2nd order transition (like 2-SAT and unlike 3-SAT)**

1st order transitions not a characteristic of NP as has been conjectured

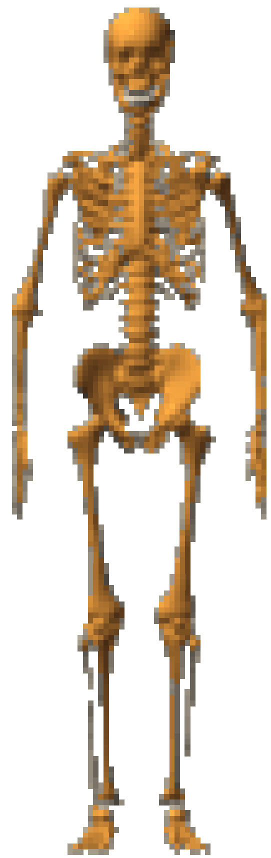
Structure

What structures makes problems hard?

How does such structure affect phase transition behaviour?

Backbone

- Variables which take fixed values in all solutions
  - alias unit prime implicates



Backbone

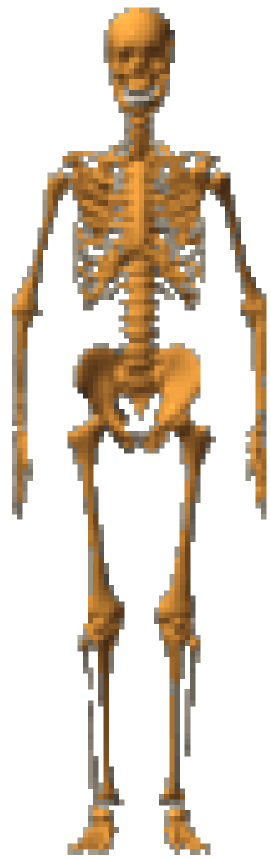
- Variables which take fixed values in all solutions
  - **alias unit prime implicates**

- Let f_k be fraction of variables in backbone
  - **in random 3-SAT**

    $c/n < 4.3$, f_k vanishing (otherwise adding clause could make problem unsat)

    $c/n > 4.3$, $f_k > 0$

    discontinuity at phase boundary (1st order)!



Backbone

- Search cost correlated with backbone size
  - **if f_k non-zero, then can easily assign variable “wrong” value**
  - **such mistakes costly if at top of search tree**
- One source of “thrashing” behaviour
  - **can tackle with randomization and rapid restarts**

Can we adapt algorithms to offer more robust performance guarantees?

Backbone

- Backbones observed in structured problems
  - **quasigroup completion problems (QCP)**

    colouring partial Latin squares

- Backbones also observed in optimization and approximation problems
  - **coloring, TSP, blocks world planning …**

    see [Slaney, Walsh IJCAI-2001]

Can we adapt algorithms to identify and exploit the backbone structure of a problem?

2+p-SAT

- Morph between 2-SAT and 3-SAT
  - **fraction p of 3-clauses**
  - **fraction (1-p) of 2-clauses**

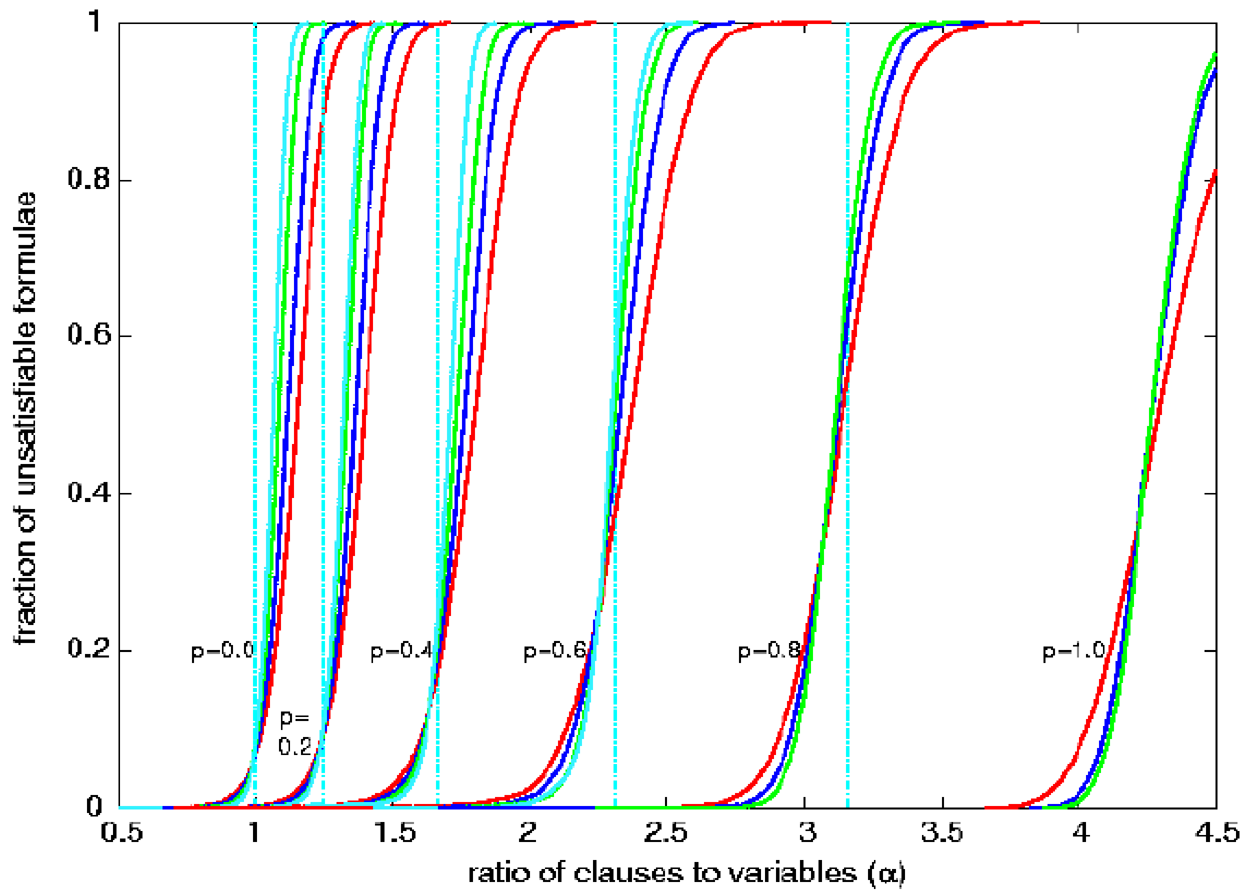
2+p-SAT

- Morph between 2-SAT and 3-SAT
  - **fraction p of 3-clauses**
  - **fraction (1-p) of 2-clauses**
- 2-SAT is polynomial (linear)
  - **phase boundary at $c/n = 1$**
  - **but no backbone discontinuity here!**

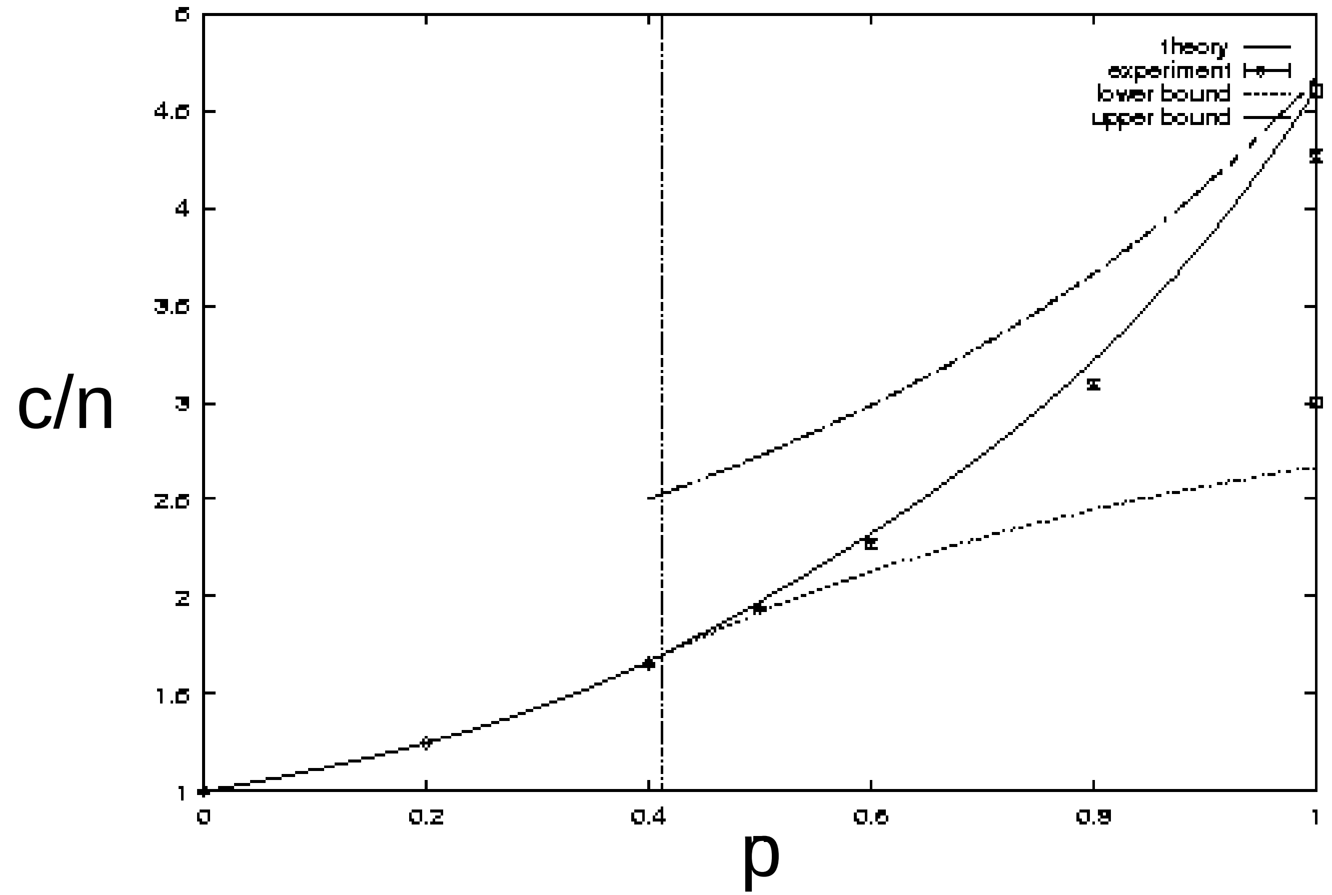
2+p-SAT

- Morph between 2-SAT and 3-SAT
  - **fraction p of 3-clauses**
  - **fraction (1-p) of 2-clauses**
- 2-SAT is polynomial (linear)
  - **phase boundary at c/n =1**
  - **but no backbone discontinuity here!**
- 2+p-SAT maps from P to NP
  - **p>0, 2+p-SAT is NP-complete**

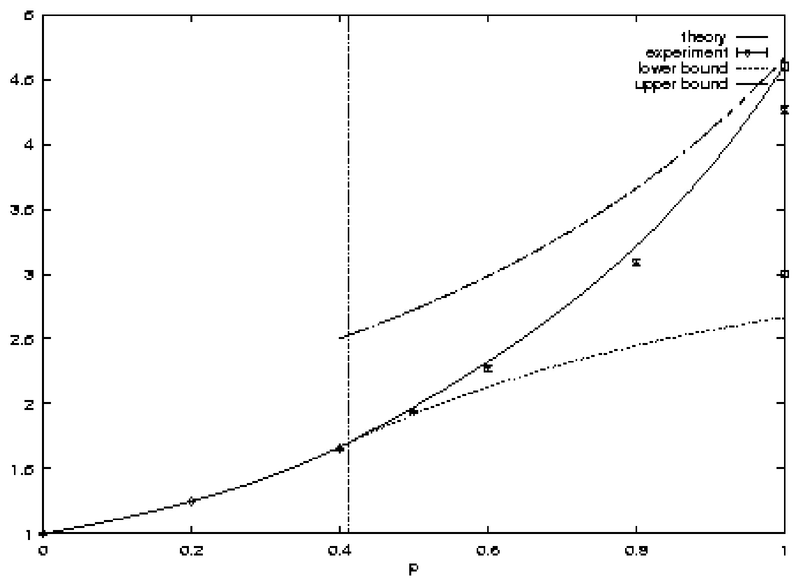
2+p-SAT phase transition



2+p-SAT phase transition

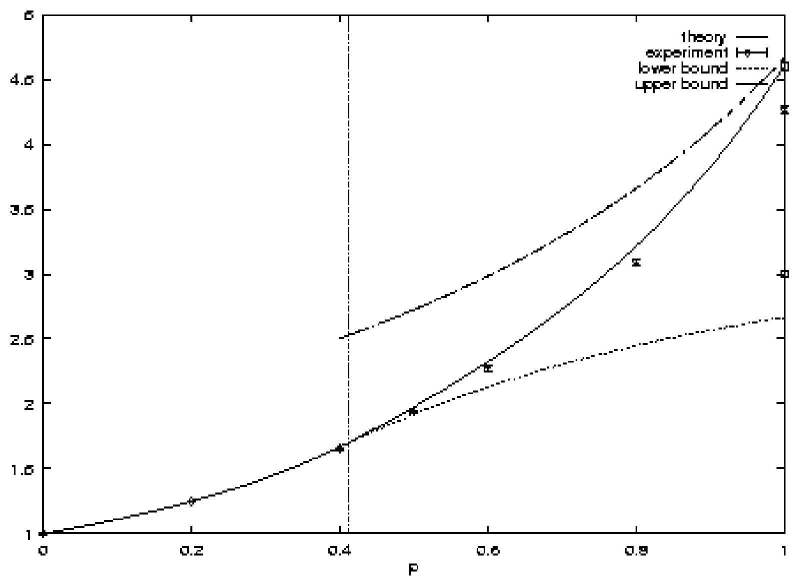


2+p-SAT phase transition



- Lower bound
  - **are the 2-clauses (on their own) UNSAT?**
  - **n.b. 2-clauses are much more constraining than 3-clauses**

2+p-SAT phase transition

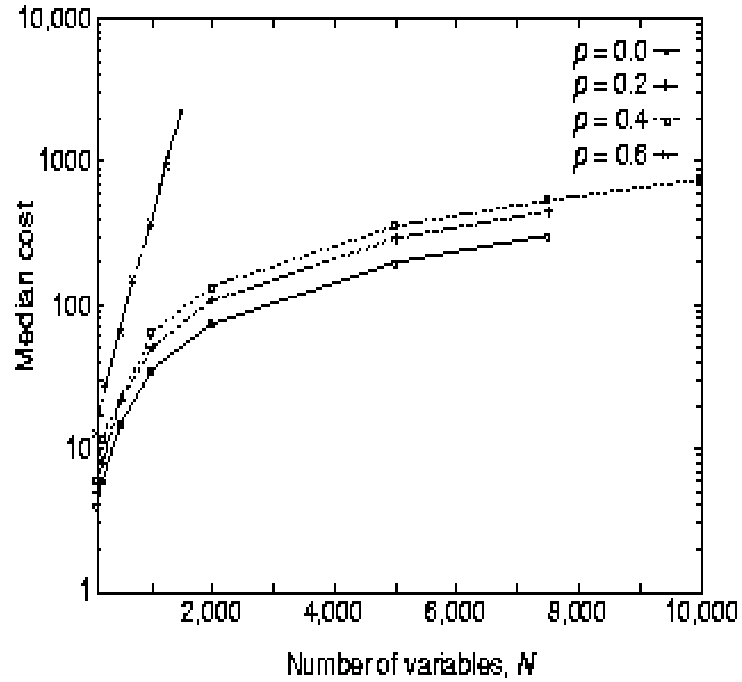


- Lower bound
  - **are the 2-clauses (on their own) UNSAT?**
  - **n.b. 2-clauses are much more constraining than 3-clauses**
- $p <= 0.4$
  - **transition occurs at lower bound**
  - **3-clauses are not contributing!**

2+p-SAT backbone

- f_k becomes discontinuous for p>0.4
  - **but NP-complete for p>0 !**
- search cost shifts from linear to exponential at p=0.4
- similar behavior seen with local search algorithms

Search cost against n



Structure

How do we model structural features found in real problems?

How does such structure affect phase transition behaviour?

The real world isn't random?

- Very true!

  Can we identify structural features common in real world problems?

- Consider graphs met in real world situations
  - **social networks**
  - **electricity grids**
  - **neural networks**
  - **...**

Real versus Random

- Real graphs tend to be sparse
  - **dense random graphs contains lots of (rare?) structure**

Real versus Random

- Real graphs tend to be sparse
  - **dense random graphs contains lots of (rare?) structure**
- Real graphs tend to have short path lengths
  - **as do random graphs**

Real versus Random

- Real graphs tend to be sparse
  - **dense random graphs contains lots of (rare?) structure**
- Real graphs tend to have short path lengths
  - **as do random graphs**
- Real graphs tend to be clustered
  - **unlike sparse random graphs**

Real versus Random

- Real graphs tend to be sparse
  - **dense random graphs contains lots of (rare?) structure**
- Real graphs tend to have short path lengths
  - **as do random graphs**
- Real graphs tend to be clustered
  - **unlike sparse random graphs**

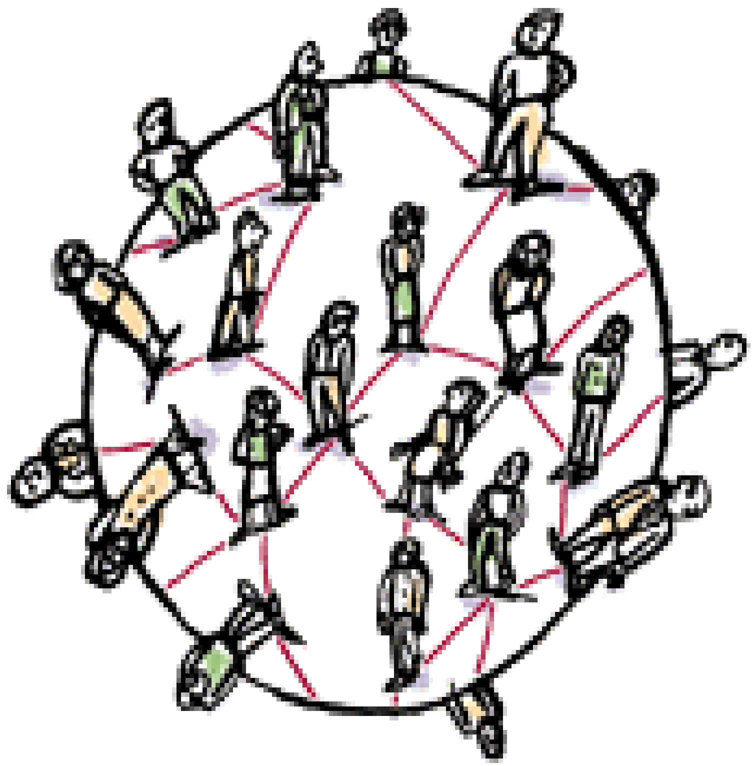
L, average path length

C, clustering coefficient

(fraction of neighbours connected to each other, cliqueness measure)

mu, proximity ratio is C/L normalized by that of random graph of same size and density

Small world graphs



- Sparse, clustered, short path lengths
- Six degrees of separation
  - **Stanley Milgram's famous 1967 postal experiment**
  - **recently revived by Watts & Strogatz**
  - **shown applies to:**
    - actors database
    - US electricity grid
    - neural net of a worm
    - ...

An example

- 1994 exam timetable at Edinburgh University
  - **59 nodes, 594 edges so relatively sparse**
  - **but contains 10-clique**
- less than 10^-10 chance in a random graph
  - **assuming same size and density**

An example

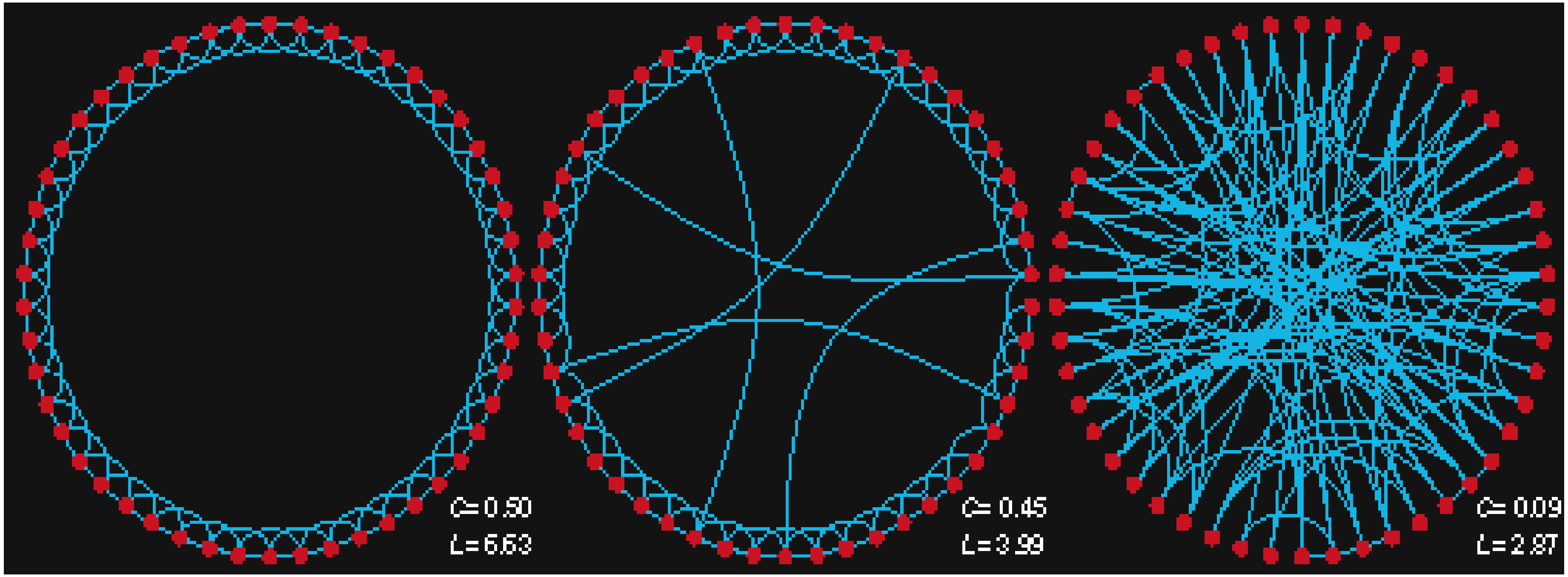
- 1994 exam timetable at Edinburgh University
  - **59 nodes, 594 edges so relatively sparse**
  - **but contains 10-clique**
- less than 10^{-10} chance in a random graph
  - **assuming same size and density**
- clique totally dominated cost to solve problem

Small world graphs

- To construct an ensemble of small world graphs
  - **morph between regular graph (like ring lattice) and random graph**
  - **prob p include edge from ring lattice, 1-p from random graph**

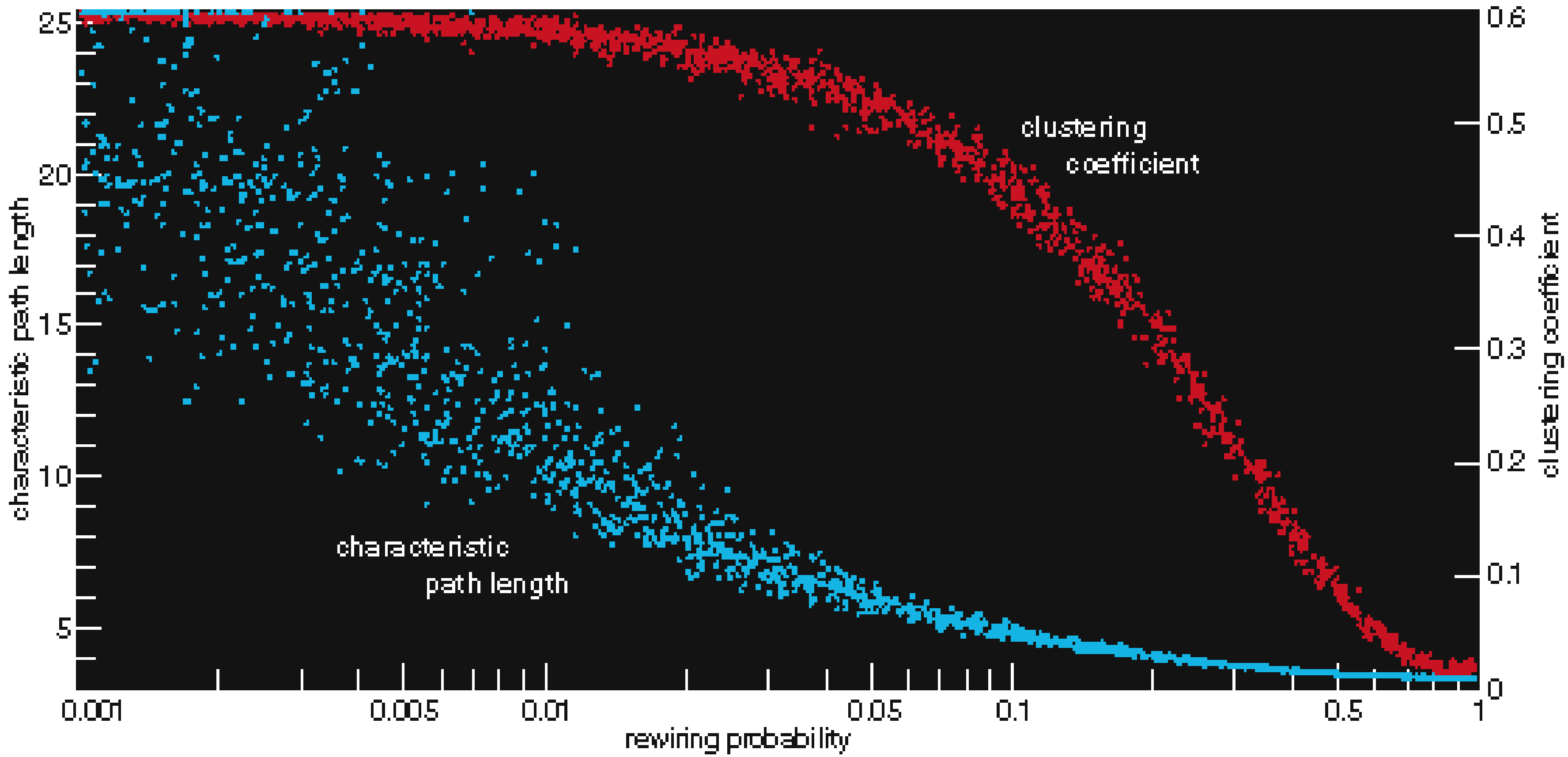
real problems often contain similar structure and stochastic components?

Small world graphs

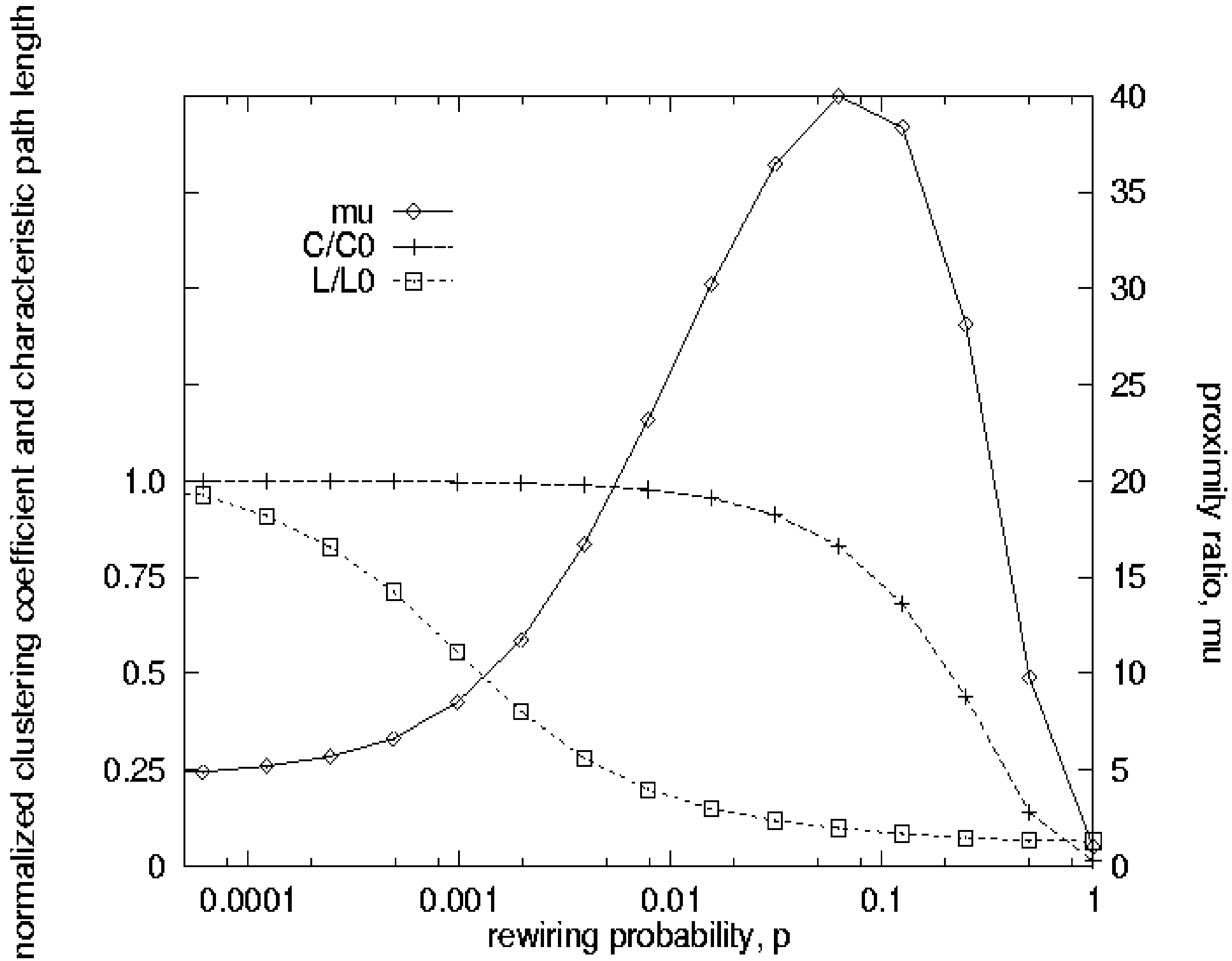


- ring lattice is clustered but has long paths
- random edges provide shortcuts without destroying clustering

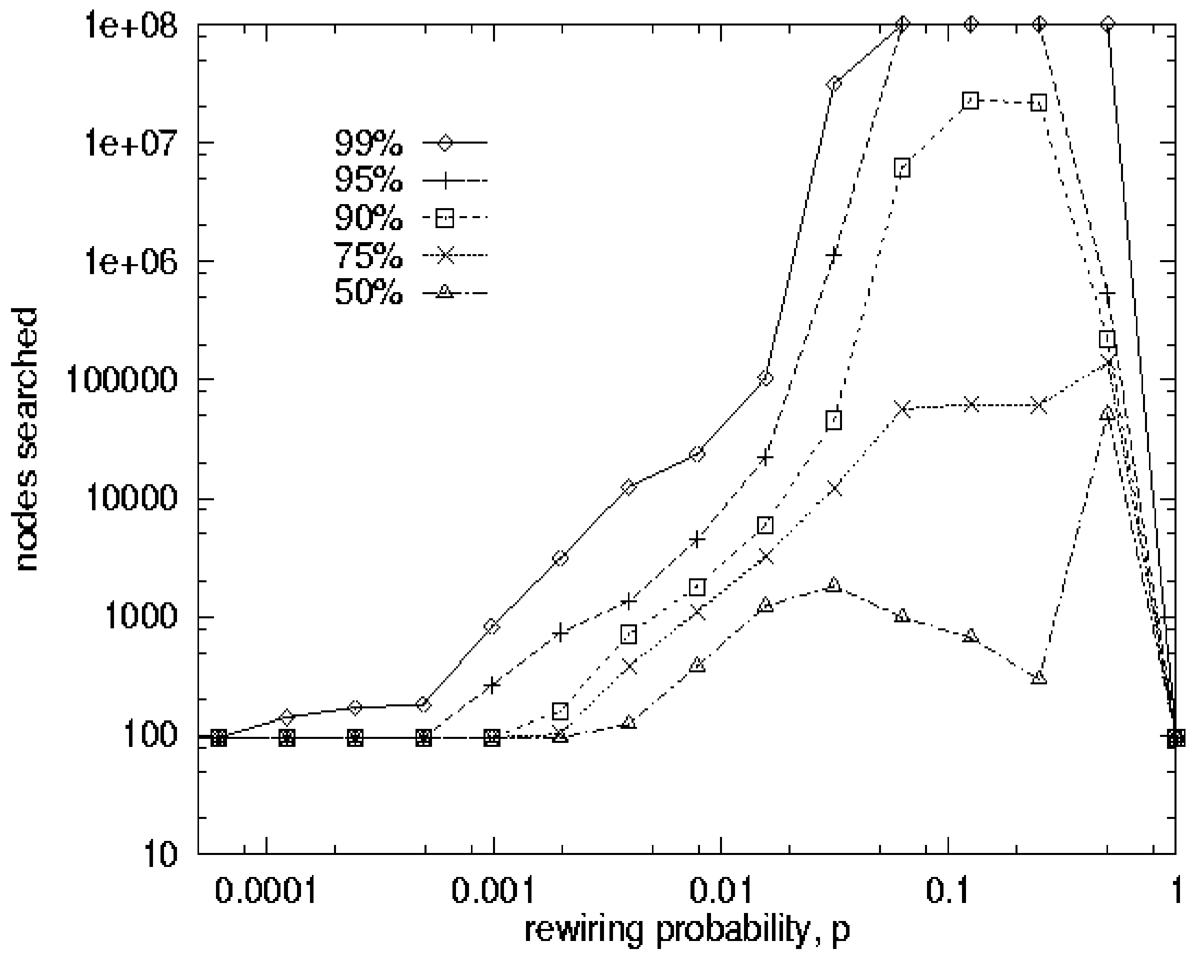
Small world graphs



Small world graphs



Colouring small world graphs



Small world graphs

- Other bad news
  - **disease spreads more rapidly in a small world**
- Good news
  - **cooperation breaks out quicker in iterated Prisoner's dilemma**

Other structural features

It's not just small world graphs that have been studied

- Large degree graphs
  - **Barbasi et al's power-law model** ***[Walsh, IJCAI 2001]***
- Ultrametric graphs
  - **Hogg's tree based model**
- Numbers following Benford's Law
  - **1 is much more common than 9 as a leading digit!**

    prob(leading digit=i) = log(1+1/i)
  - **such clustering, makes number partitioning much easier**

The future?

What open questions remain?

Where to next?

Open questions

- Prove random 3-SAT occurs at $l/n = 4.3$
  - **random 2-SAT proved to be at $l/n = 1$**
  - **random 3-SAT transition proved to be in range $3.26 < l/n < 4.506$**
  - **random 3-SAT phase transition proved to be "sharp"**

Open questions

- Impact of structure on phase transition behaviour
  - **some initial work on quasigroups (alias Latin squares/sports tournaments)**
  - **morphing useful tool (e.g. small worlds, 2-d to 3-d TSP, …)**
- Optimization *v* decision
  - **some initial work by Slaney & Thiebaux**
  - **economics often pushes optimization problems naturally towards feasible/infeasible phase boundary**

Open questions

- Does phase transition behaviour give help answer P=NP?
  - **it certainly identifies hard problems!**
  - **problems like 2+p-SAT and ideas like backbone also show promise**
- Problems away from phase boundary can be hard
  - **over-constrained 3-SAT region has exponential resolution proofs**
  - **under-constrained 3-SAT region can throw up occasional hard problems (early mistakes?)**

Research directions in SAT

- Algorithm development
  - **Fast but cheap solvers (chaff from Princeton)**
    - Basic operations are constant time (e.g. branching heuristic, finding unit clauses, ..)
  - **Nogood learning**
  - **Randomization and restarts**
    - Learning across restarts
- Domain enlargement
  - **New encodings into SAT**
  - **Beyond the propositional (QBF, modal SAT, …)**

Summary

That’s nearly all from me!

Conclusions

- Phase transition behaviour ubiquitous
  - **decision/optimization/...**
  - **NP/PSpace/P/...**
  - **random/real**
- Phase transition behaviour gives insight into problem hardness
  - **suggests new branching heuristics**
  - **ideas like the backbone help understand branching mistakes**

Conclusions

- Propositional satisfiability (SAT)
  - **Very active research area**
    - SAT2002
  - **Useful for understanding source of problem hardness**
  - **Useful also for solving problems**
    - E.g. Planning as SAT, model checking via SAT, …
  - **Developing new algorithms**
    - E.g. Randomization and restarts, learning, non-chronological backtracking, ..

Very partial bibliography

Cheeseman, Kanefsky, Taylor, *Where the really hard problem are*, Proc. of IJCAI-91

Gent et al, *The Constrainedness of Search*, Proc. of AAAI-96

Gent et al, *SAT 2000*, IOS Press, Fronteirs in Artificial Intelligence, 2000

Gent & Walsh, *The TSP Phase Transition*, Artificial Intelligence, 88:359-358, 1996

Gent & Walsh, *Analysis of Heuristics for Number Partitioning*, Computational Intelligence, 14 (3), 1998

Gent & Walsh, *Beyond NP: The QSAT Phase Transition*, Proc. of AAAI-99

Gent et al, *Morphing: combining structure and randomness*, Proc. of AAAI-99

Hogg & Williams (eds), special issue of *Artificial Intelligence*, 88 (1-2), 1996

Mitchell, Selman, Levesque, *Hard and Easy Distributions of SAT problems*, Proc. of AAAI-92

Monasson et al, *Determining computational complexity from characteristic 'phase transitions'*, Nature, 400, 1998

Walsh, *Search in a Small World*, Proc. of IJCAI-99

Watts & Strogatz, *Collective dynamics of small world networks*, Nature, 393, 1998

See http://www.cs.york.ac.uk/~tw/Links/ for more