The Impact of Structure

Toby Walsh
Cork Constraint Computation Centre
http://4c.ucc.ie/~tw

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 graphs tend to have short path lengths
  - **as do random graphs**
- Real graphs tend to be clustered
  - **unlike sparse random graphs**

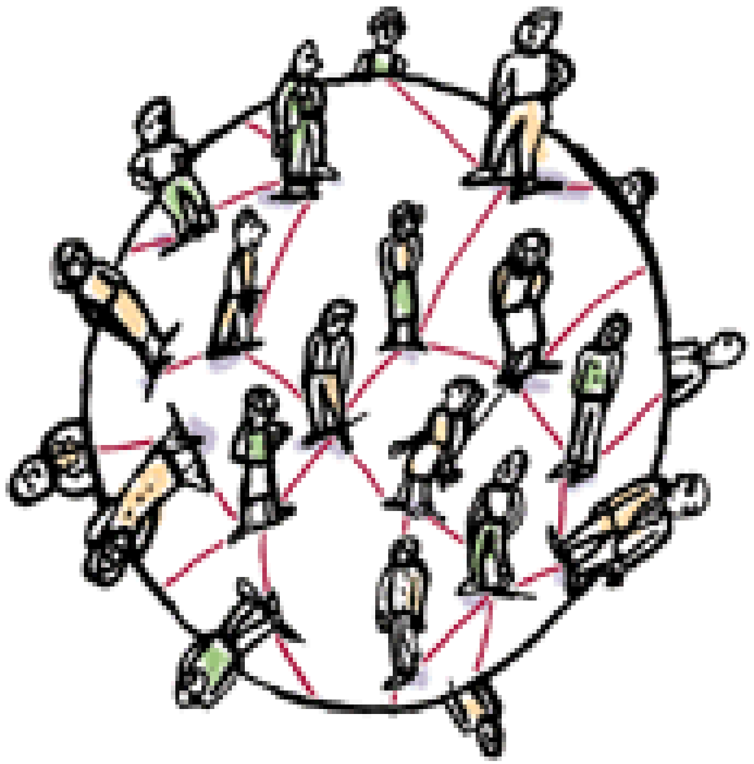
L, average path length

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

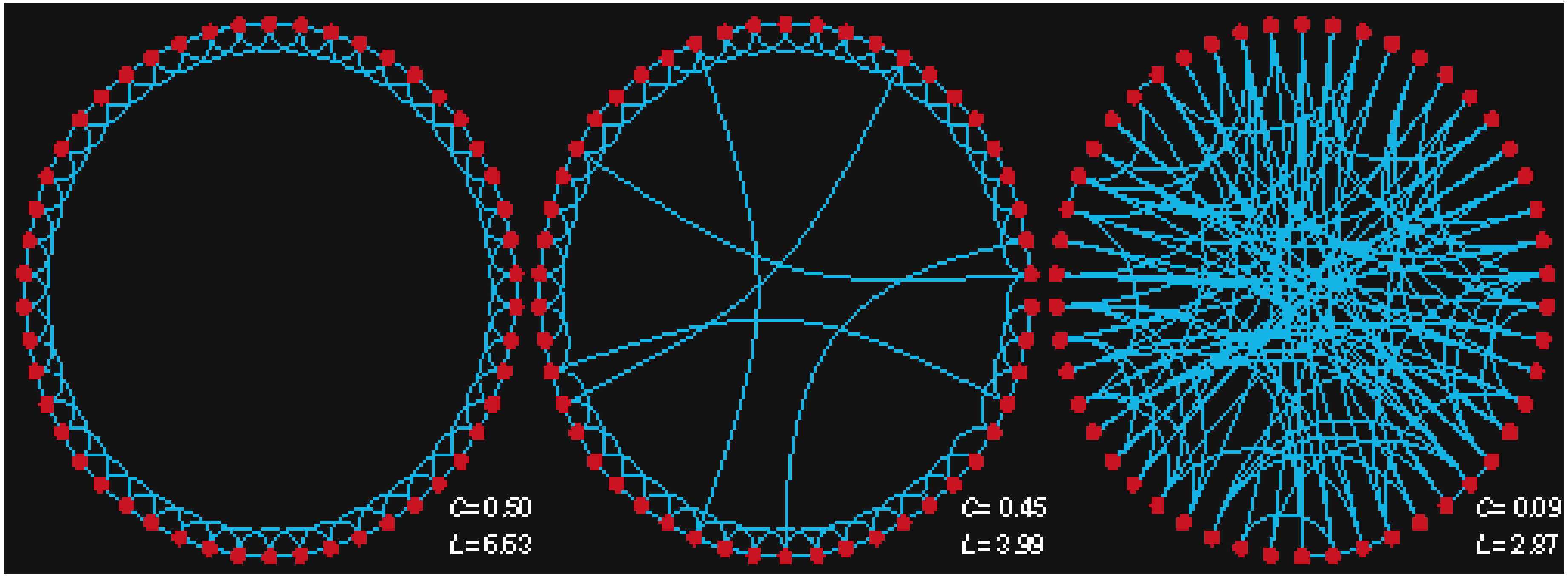
- 1994 exam timetable at Edinburgh University
  - **59 nodes, 594 edges so relatively sparse**
  - **but contains 10-clique**
- less than 10^-10 chance in a random graph
  - **assuming same size and density**
- clique totally dominated cost to solve problem

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**

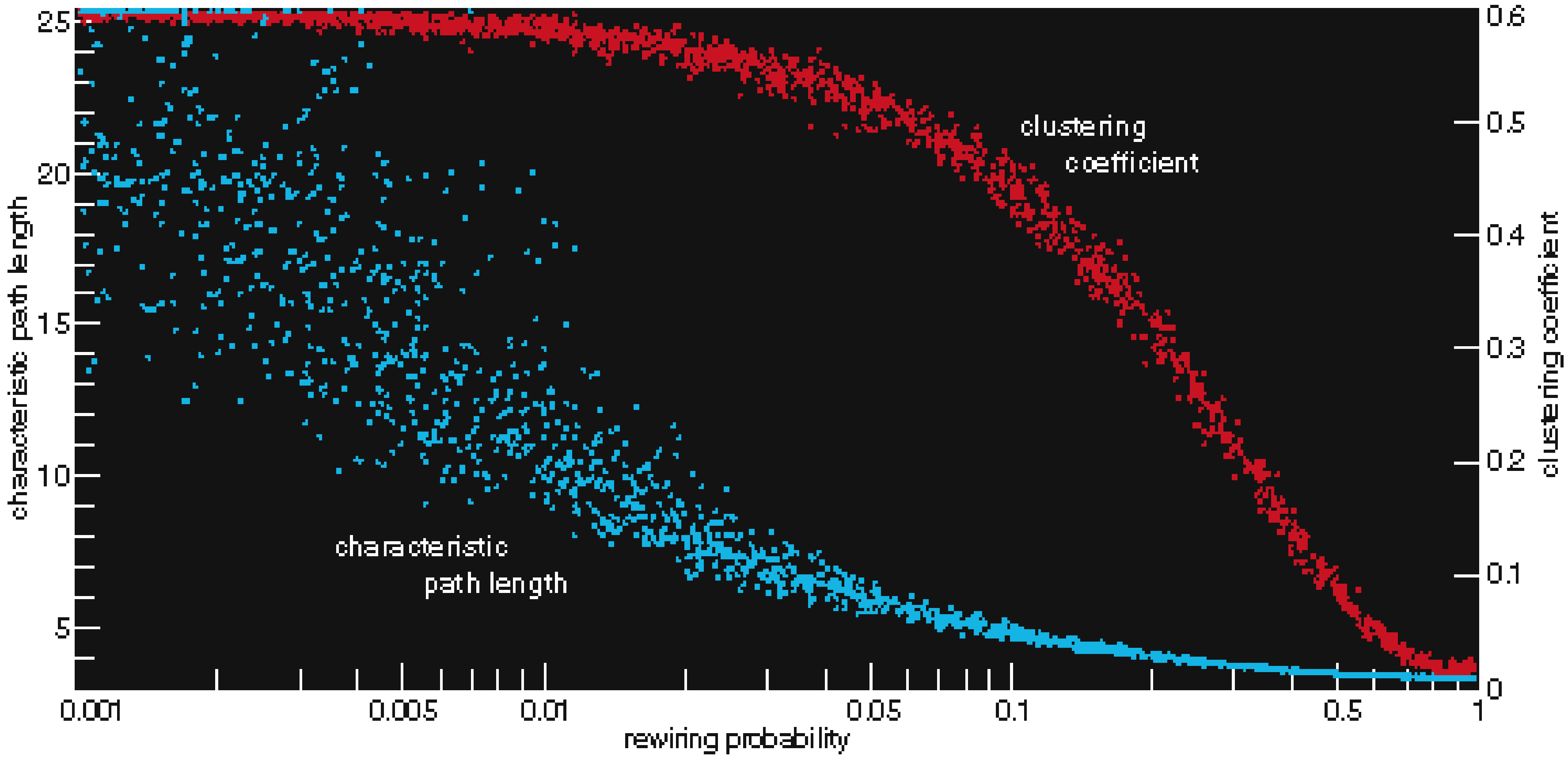
real problems often contain similar structure and stochastic components?

Small world graphs

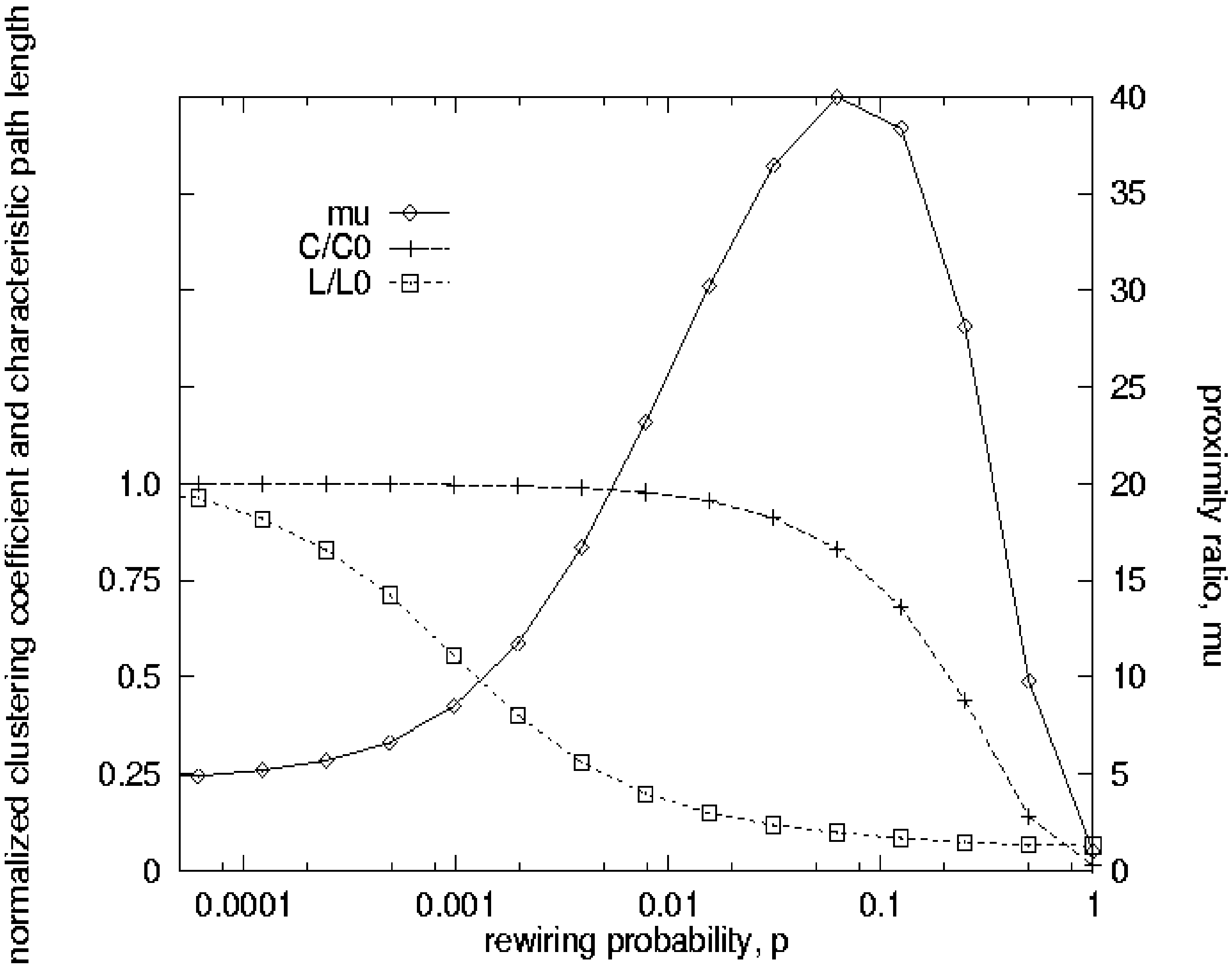


- ring lattice is clustered but has long paths
- random edges provide shortcuts without destroying clustering

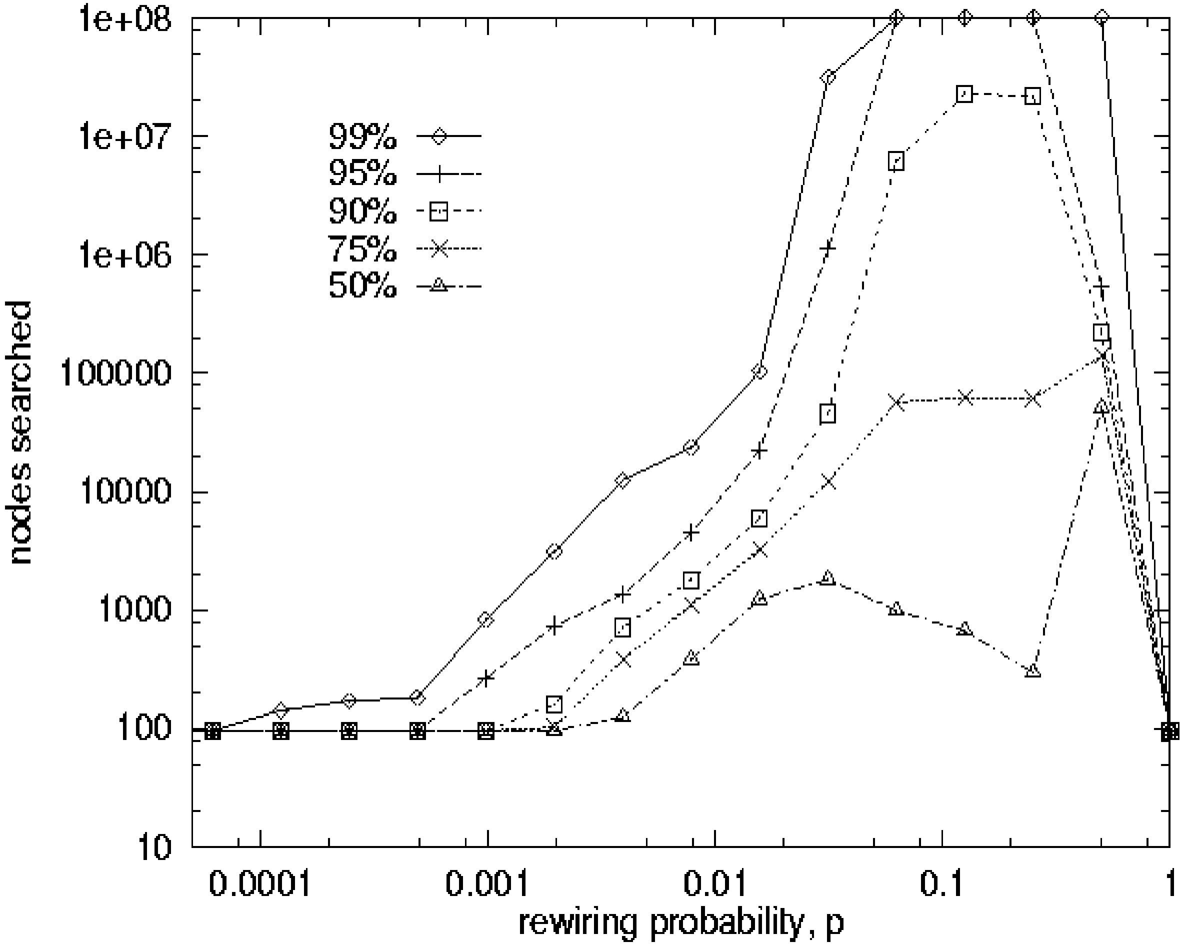
Small world graphs



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

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

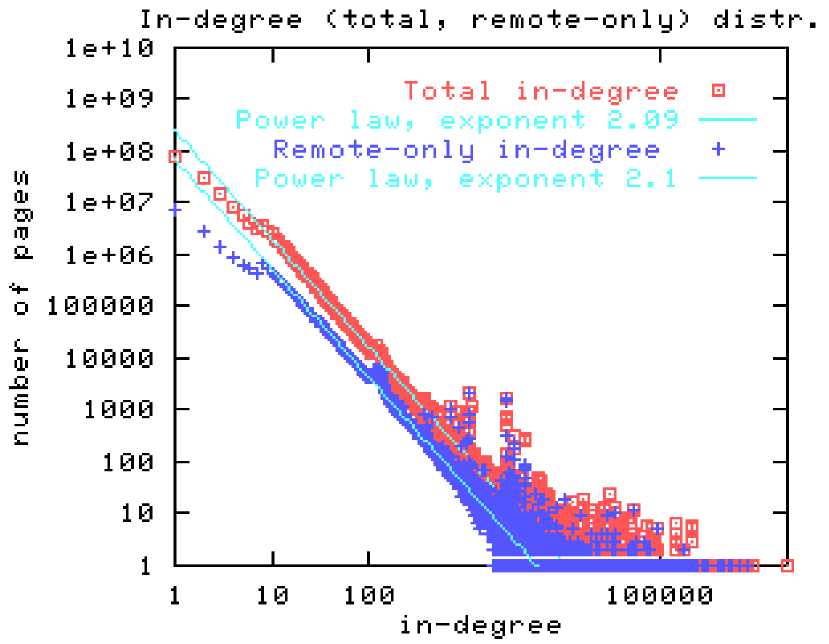
    prob(leading digit=i) = log(1+1/i)
  - **such clustering, makes number partitioning much easier**

High degree graphs

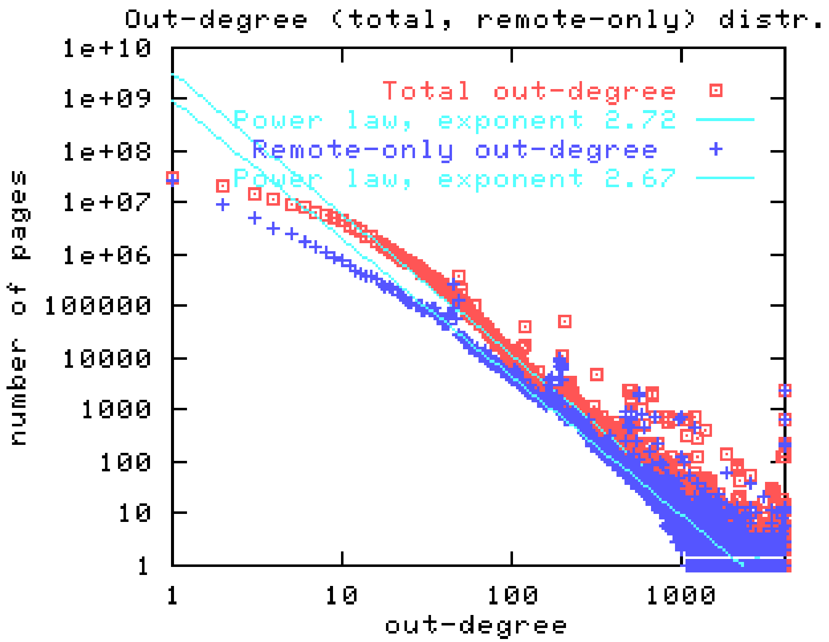
- Degree = number of edges connected to node
- Directed graph
  - **Edges have a direction**
  - **E.g. web pages = nodes, links = directed edges**
- In-degree, out-degree
  - **In-degree = links pointing to page**
  - **Out-degree = links pointing out of page**

In-degree of World Wide Web

- Power law distribution
  - **Pr(in-degree = k) = ak^-2.1**
- Some nodes of very high in-degree
  - **E.g. google.com, ...**

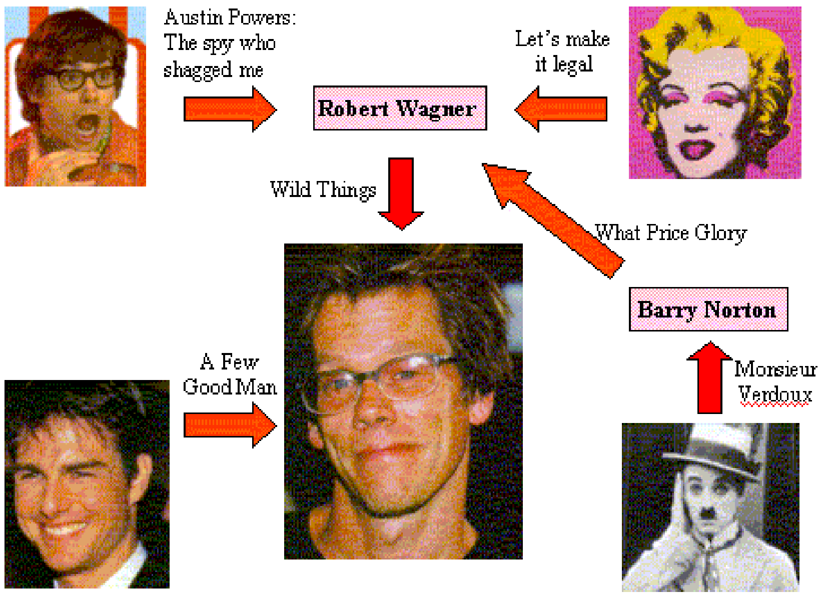


Out-degree of World Wide Web



- Power law distribution
  - **Pr(in-degree = k) = ak^-2.7**
- Some nodes of very high out-degree
  - **E.g. people in SAT**

High degree graphs

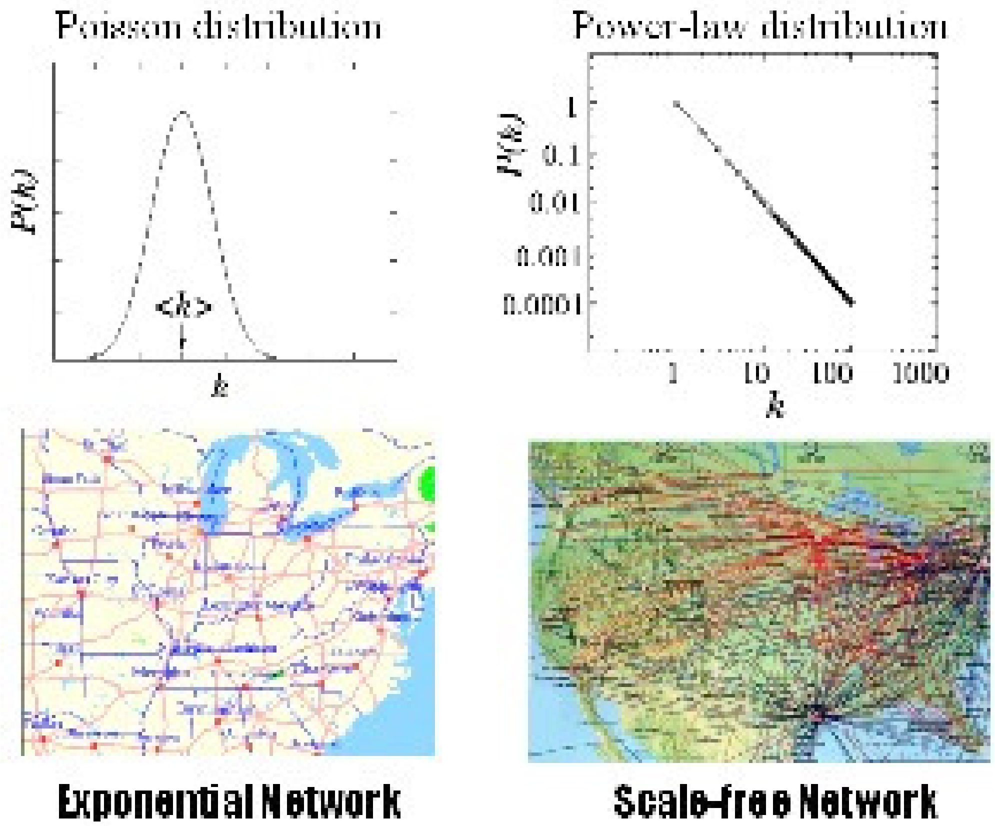


- World Wide Web
- Electricity grid
- Citation graph
  - **633,391 out of 783,339 papers have < 10 citations**
  - **64 have > 1000 citations**
  - **1 has 8907 citations**
- Actors graph
  - **Robert Wagner, Donald Sutherland, …**

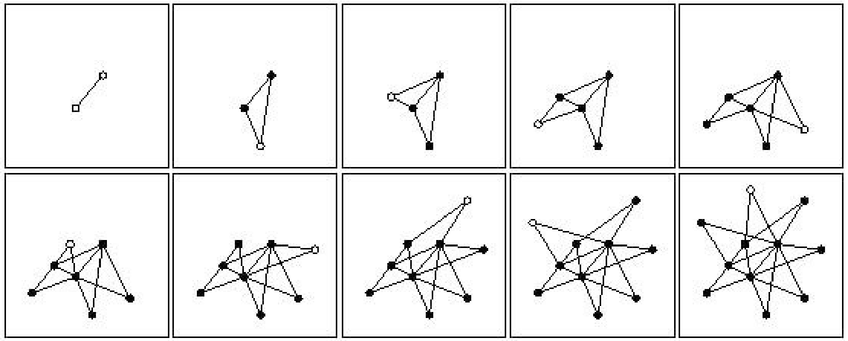
High degree graphs

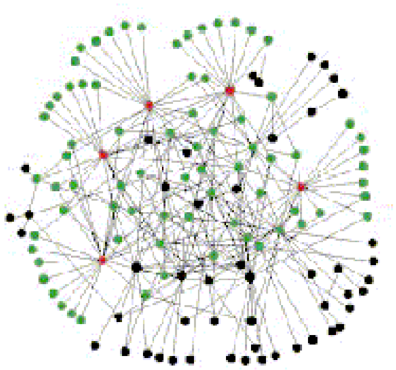
- Power law in degree distribution
  - **Pr(degee = k) = ak^-b where b typically around 3**
- Compare this to random graphs
  - **Gnm model**
    - n nodes, m edges chosen uniformly at random
  - **Gnp model**
    - n nodes, each edge included with probability p
  - **In both, Pr(degree = k) is a Poisson distribution**
    - tightly clustered around mean

Random v high degree graphs



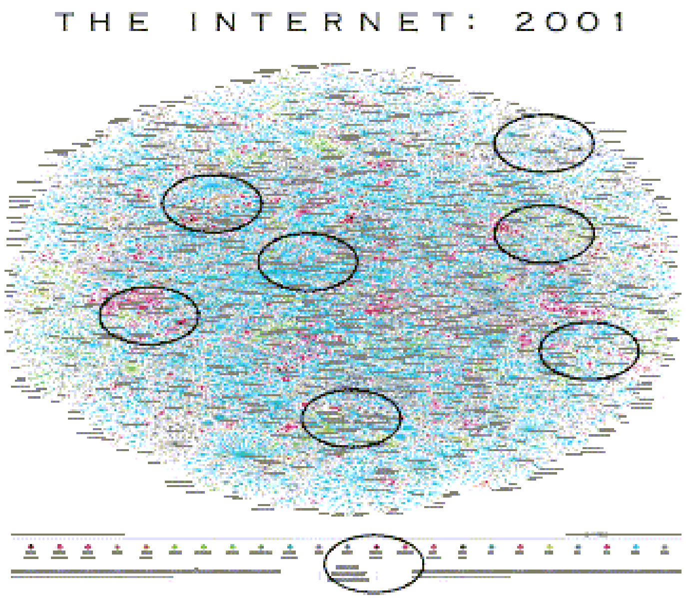
Generating high-degree graphs





- Grow graph
- Preferentially attach new nodes to old nodes according to their degree
  - **Prob(attach to node j) proportional to degree of node j**
  - **Gives Prob(degree = k) = ak^-3**

High-degree = small world?

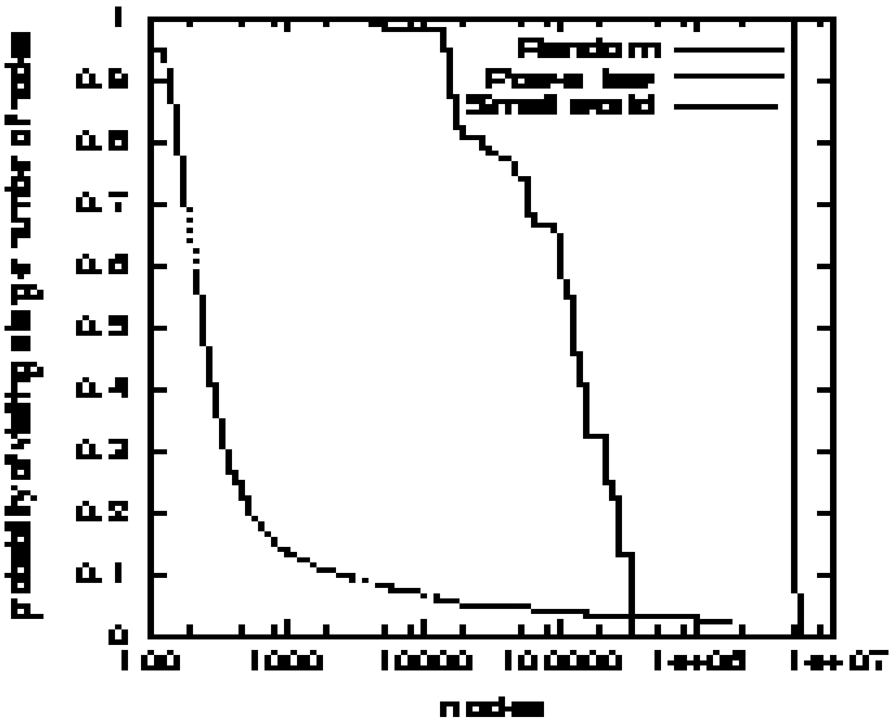


- Preferential attachment model
  - **n=16, mu=1**
  - **n=64, mu=1.35**
  - **n=256, mu=2.12**
  - **…**

Small world topology thus for large n!

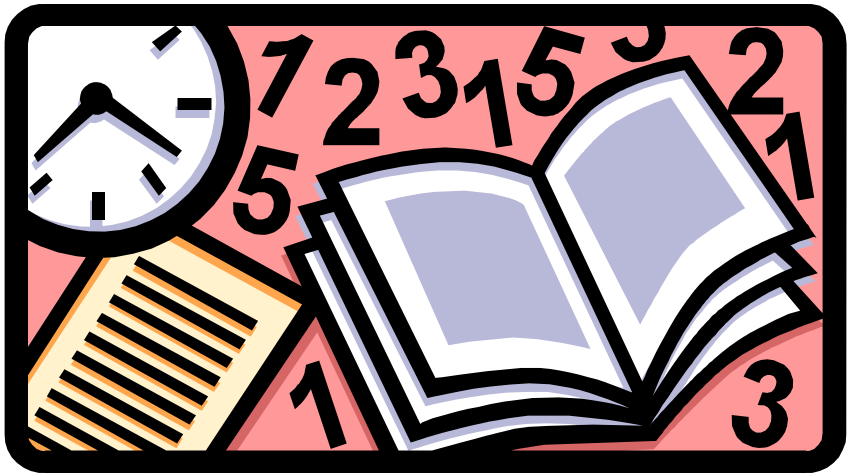
Search on high degree graphs

- Random
  - **Uniformly hard**
- Small world
  - **A few long runs**
- High degree
  - **More uniform**
  - **Easier than randor**



What about numbers?

- So far, we've looked at structural features of graphs
- Many problems contain numbers
  - **Do we see phase transitions here too?**



Number partitioning

- What's the problem?
  - dividing a bag of numbers into two so their sums are as balanced as possible
- What problem instances?
  - n numbers, each uniformly chosen from (0,l]
  - other distributions work (Poisson, …)

Number partitioning

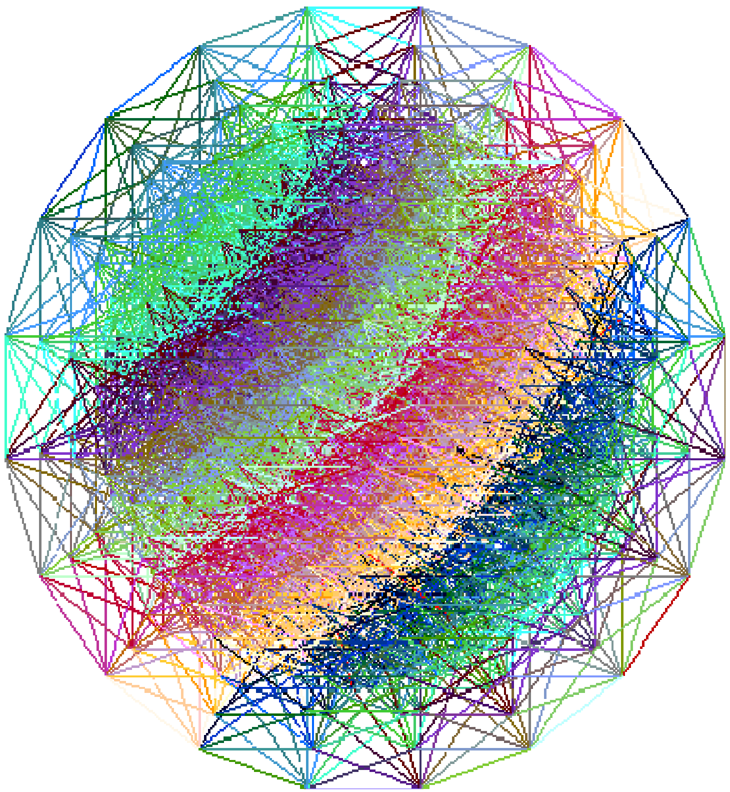
- Identify a measure of constrainedness
  - **more numbers => less constrained**
  - **larger numbers => more constrained**
  - **could try some measures out at random (*l/n, log(l)/n, log(l)/sqrt(n), …*)**

- Better still, use kappa!
  - **(approximate) theory about constrainedness**
  - **based upon some simplifying assumptions**

    e.g. ignores structural features that cluster solutions together

Theory of constrainedness

- Consider state space searched
  - **see 10-d hypercube opposite of 2^10 possible partitions of 10 numbers into 2 bags**
- Compute expected number of solutions, <*Sol*>
  - **independence assumptions often useful and harmless!**



Theory of constrainedness

- Constrainedness given by:

  kappa= 1 - log2(<Sol>)/n

  where *n* is dimension of state space

- kappa lies in range [0,infty)
  - ***kappa=0,*** ***<Sol>=2^n*****,** **under-constrained**
  - ***kappa=infty,*** ***<Sol>=0*****,** **over-constrained**
  - ***kappa=1,*** ***<Sol>=1*****,** **critically constrained phase boundary**

Phase boundary

- Markov inequality
  - ***prob(Sol) < <Sol>***

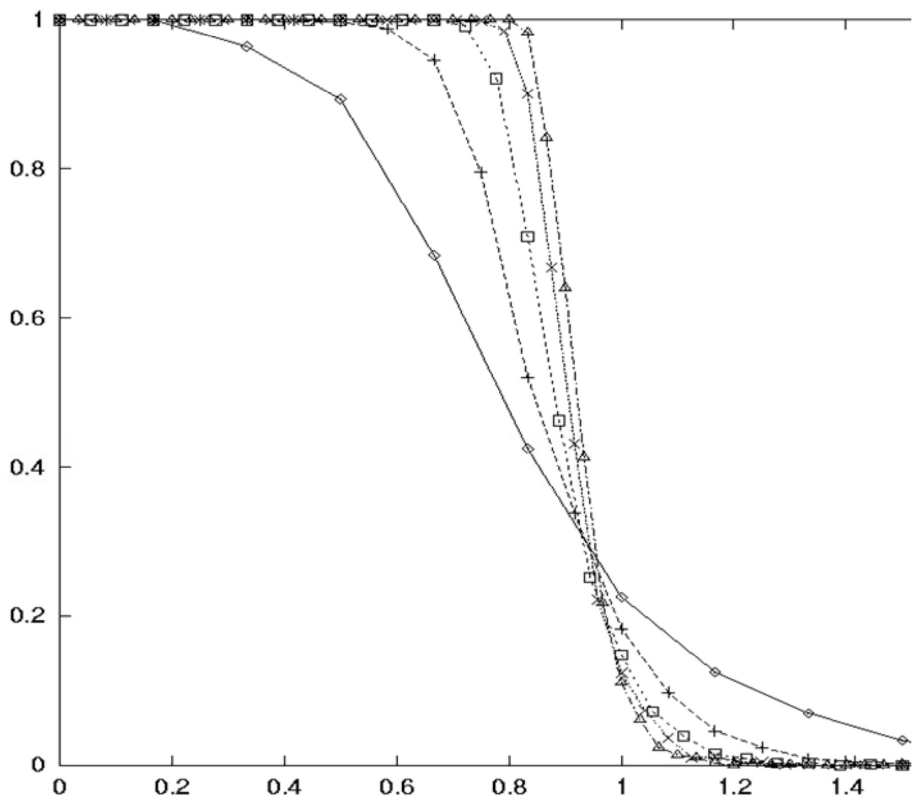
  Now, *kappa > 1* implies *<Sol> < 1*
  Hence, *kappa > 1* implies *prob(Sol) < 1*

- Phase boundary typically at values of *kappa* slightly smaller than *kappa=1*
  - **skew in distribution of solutions (e.g. 3-SAT)**
  - **non-independence**

Examples of kappa

- 3-SAT
  - **kappa = l/5.2n**
  - **phase boundary at kappa=0.82**
- 3-COL
  - **kappa = e/2.7n**
  - **phase boundary at kappa=0.84**
- number partitioning
  - **kappa = log2(l)/n**
  - **phase boundary at kappa=0.96**

Number partition phase transition



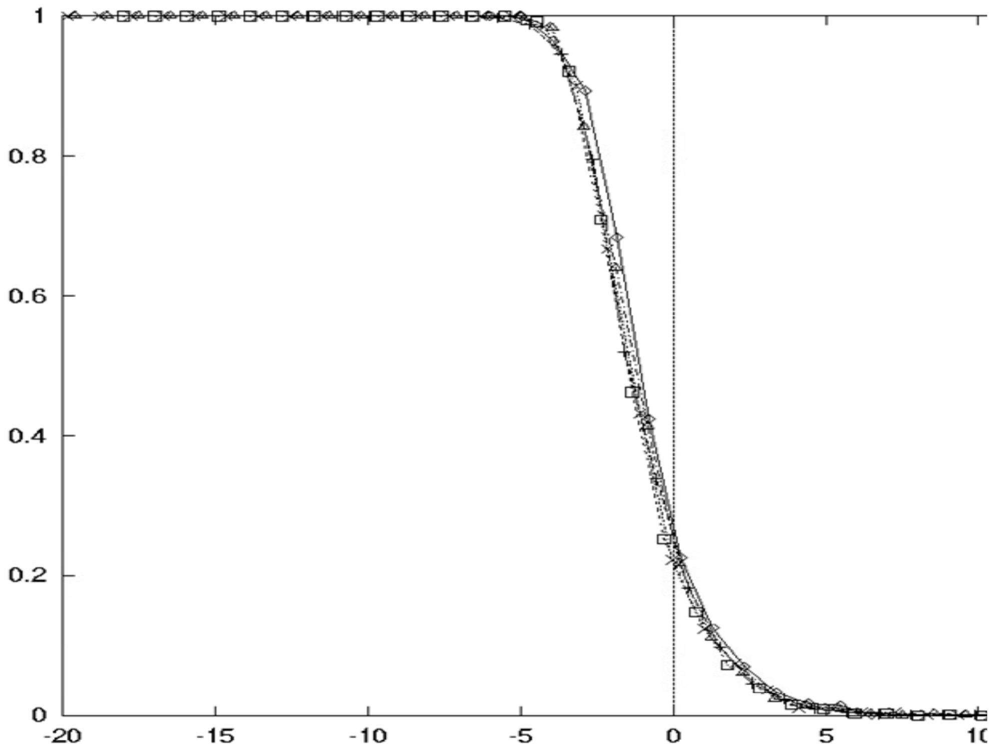
Prob(perfect partition) against kappa

Finite-size scaling

- Simple "trick" from statistical physics
  - **around critical point, problems indistinguishable except for change of scale given by simple power-law**

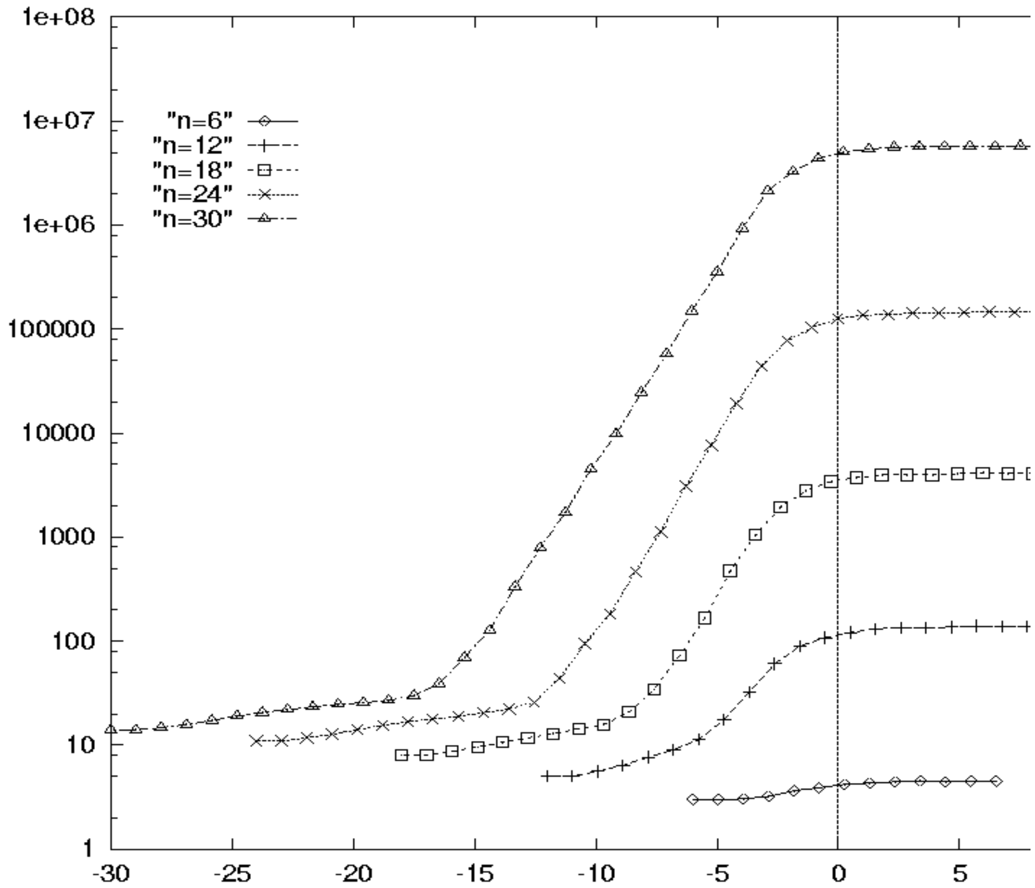
- Define rescaled parameter
  - ***gamma = kappa-kappac . n^1/v***
    kappac
  - **estimate *kappac* and *v* empirically**
    - e.g. for number partitioning, *kappac*=0.96, *v*=1

Rescaled phase transition



Prob(perfect partition) against gamma

Rescaled search cost



Optimization cost against gamma

Easy-Hard-Easy?

- Search cost only easy-hard here?
  - **Optimization not decision search cost!**
  - **Easy if (large number of) perfect partitions**
  - **Otherwise little pruning (search scales as 2^0.85n)**

- Phase transition behaviour less well understood for optimization than for decision
  - **sometimes optimization = sequence of decision problems (e.g branch & bound)**
  - **BUT lots of subtle issues lurking?**

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 $3.42 < l/n < 4.506$**

- 2+p-COL
  - **Prove problem changes around p=0.8**
  - **What happens to colouring backbone?**

Open questions

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

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

- AI becoming more of an experimental science?
  - **theory and experiment complement each other well**
  - **increasing use of approximate/heuristic theories to keep theory in touch with rapid experimentation**
- Phase transition behaviour is FUN
  - **lots of nice graphs as promised**
  - **and it is teaching us lots about complexity and algorithms!**

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 & 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

Walsh, *Search on High Degree Graphs*, Proc. of IJCAI-2001.

Walsh, *From P to NP: COL, XOR, NAE, 1-in-k, and Horn SAT*, Proc. of AAAI-2001.

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