Local Search Algorithms for Random Satisfiability

Pekka Orponen (joint work with Sakari Seitz and Mikko Alava)

Laboratory for Theoretical Computer Science
Helsinki University of Technology

Outline

- Background
- WalkSAT and related algorithms
- Results and conjectures on WalkSAT
- Experiments on WalkSAT
- Record-to-Record Travel and variants
- Experiments on RRT
- Focused Metropolis Search
- Experiments on FMS
- Dynamics of FMS
- Analysis?

Background

Denote N = number of variables, M = number of clauses, $\alpha = M/N$.

Satisfiability transition at $\alpha_c \approx 4.267$ (Mitchell et al. 1992, $\ldots$, Braunstein et al. 2002).

Good experiences with local search methods in the satisfiable region $\alpha < \alpha_c$: e.g. GSAT (Selman et al. 1992), WalkSAT (Selman et al. 1996). Experiments 1996: $N \leq 2000$ at $\alpha \approx \alpha_c$.

GSAT

Selman et al. 1992 $\ldots$ 1996.

Denote by $E = E_F(s)$ the number of unsatisfied clauses in formula F under truth assignment s.

```
GSAT(F):
  s = initial truth assignment;
  while flips < max_flips do
    if s satisfies F then output s & halt, else:
    - find a variable x whose flipping causes
      largest decrease in E (if no decrease is
      possible, then smallest increase);
    - flip x.
```

NoisyGSAT

GSAT augmented by a fraction p of random walk moves.

```
NoisyGSAT(F,p):
  s = initial truth assignment;
  while flips < max_flips do
    if s satisfies F then output s & halt, else:
    - with probability p, pick a variable x
      uniformly at random and flip it;
    - with probability (1-p), do basic GSAT move:
      - find a variable x whose flipping causes
        largest decrease in E (if no decrease is
        possible, then smallest increase);
      - flip x.
```

WalkSAT

NoisyGSAT *focused* on the unsatisfied clauses.

```
WalkSAT(F,p):
  s = initial truth assignment;
  while flips < max_flips do
    if s satisfies F then output s & halt, else:
    - pick a random unsatisfied clause C in F;
    - if some variables in C can be flipped without
      breaking any presently satisfied clauses,
      then pick one such variable x at random; else:
    - with probability p, pick a variable x
      in C at random;
    - with probability (1-p), pick an x in C
      that breaks a minimal number of presently
      satisfied clauses;
    - flip x.
```

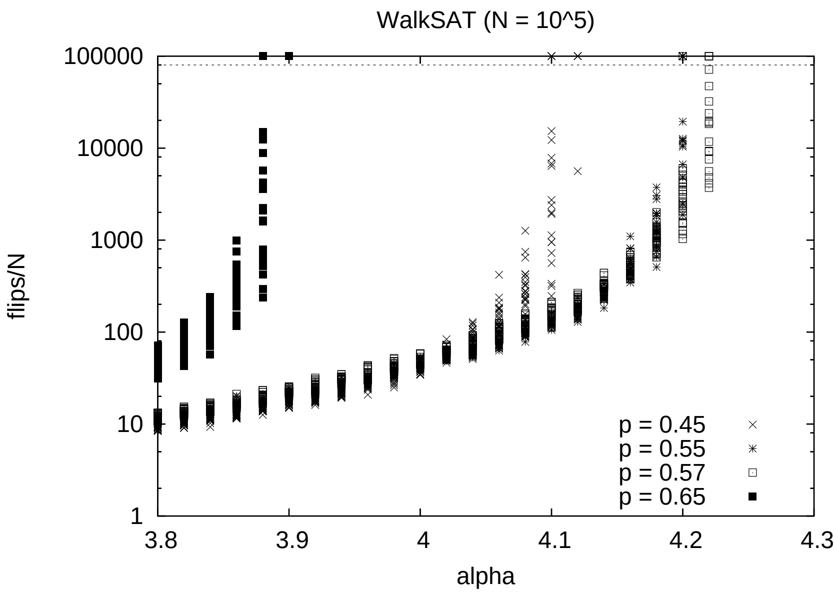
WalkSAT vs. NoisyGSAT

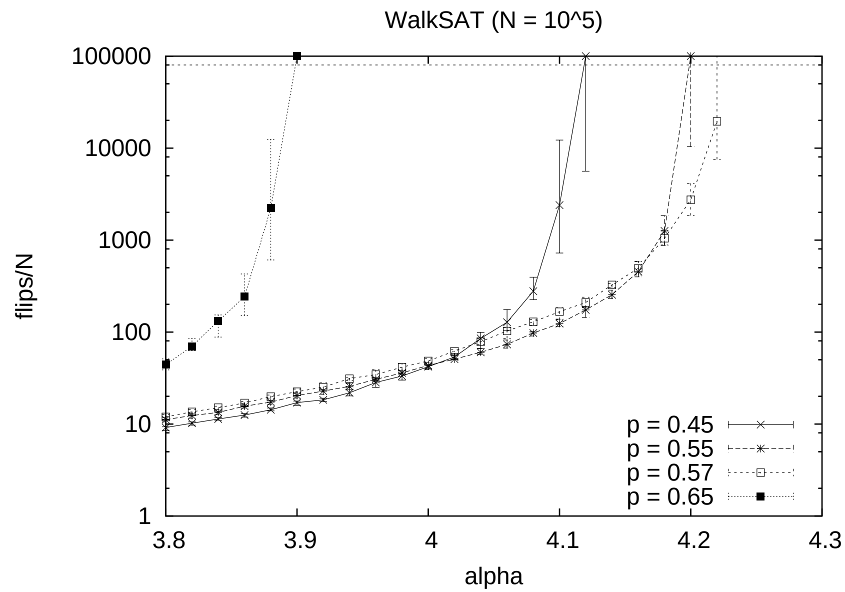
The focusing seems to be important: in the (unsystematic) experiments in Selman et al. (1996), WalkSAT outperforms NoisyGSAT by several orders of magnitude.

Recent results and conjectures

- Barthel et al. (2003): numerical experiments with WalkSAT (mainly with $p = 1$, some also with $p < 1$) at $N = 50,000$, $\alpha = 2.0 \ldots 4.0$. Observed transition in the dynamics at $\alpha_{\text{dyn}} \approx 2.7 - 2.8$. When $\alpha < \alpha_{\text{dyn}}$, satisfying assignments are found in linear time per variable (i.e. in a total of cN "flips"), when $\alpha > \alpha_{\text{dyn}}$ exponential time is required.
- Similar results obtained by Semerjian & Monasson (2003), though with smaller experiments ($N = 500$).
- Explanation: for $\alpha > \alpha_{\text{dyn}}$ the search equilibrates at a nonzero energy level, and can only escape to a ground state through a large enough random fluctuation.
- Conjecture: no local search algorithm works in linear time beyond the clustering transition at $\alpha_s \approx 3.92 - 3.93$ (Mézard, Monasson, Weigt et al.)

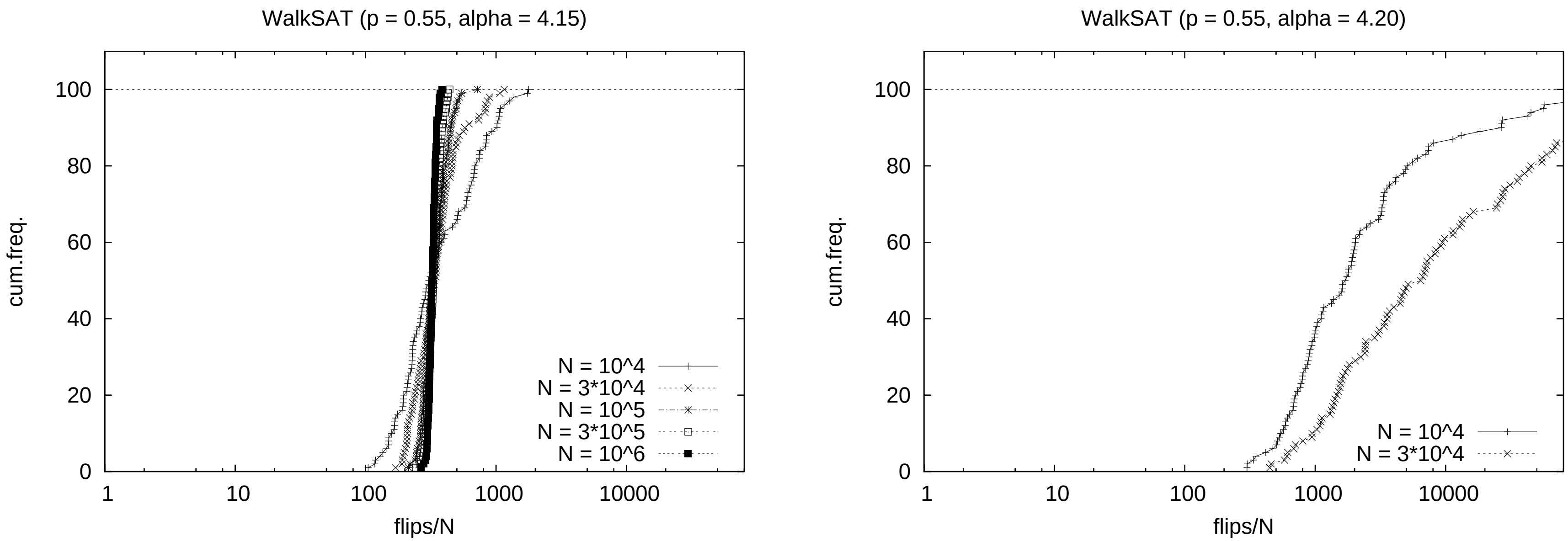
WalkSAT experiments (3-SAT)





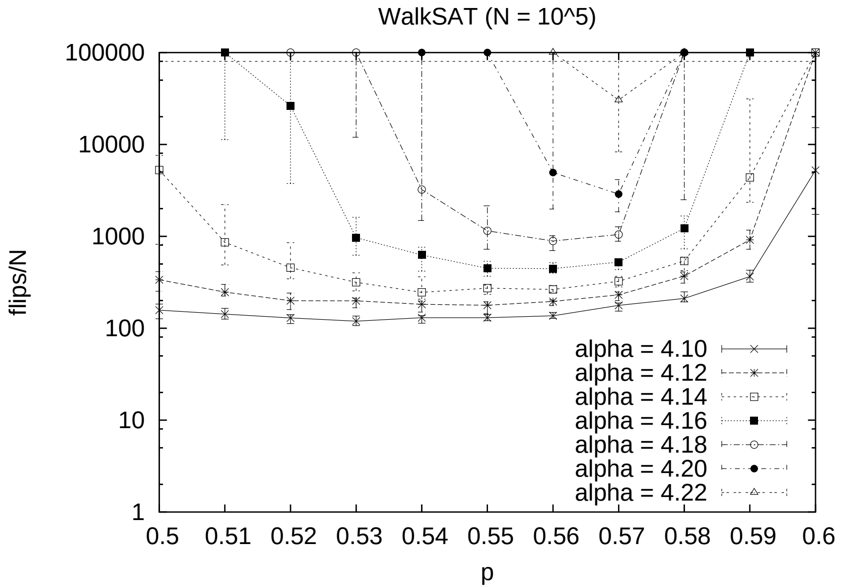
Normalised solution times for WalkSAT, $\alpha = 3.8 \ldots 4.3$.
Left: complete data; right: medians and quartiles.

WalkSAT linear scaling



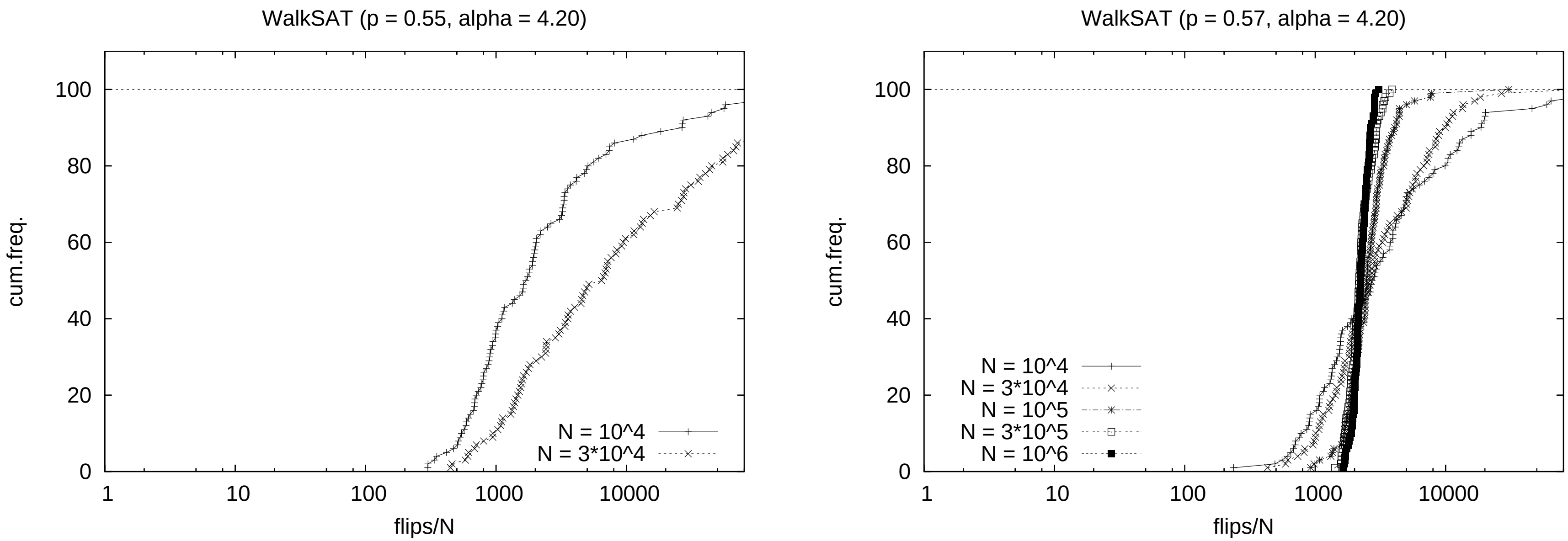
Cumulative solution time distributions for WalkSAT with $p = 0.55$.

WalkSAT optimal noise level?



Normalised solution times for WalkSAT with $p = 0.50 \ldots 0.60$, $\alpha = 4.10 \ldots 4.22$.

WalkSAT sensitivity to noise



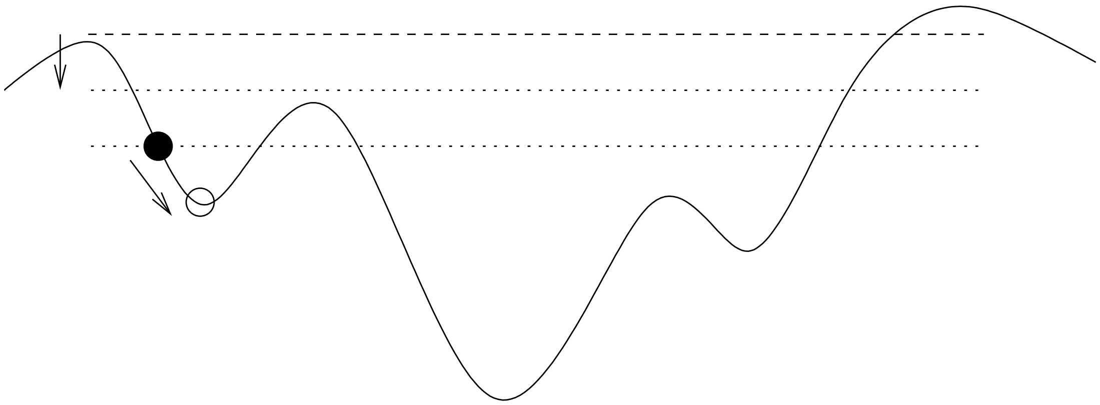
Cumulative solution time distributions for WalkSAT at $\alpha = 4.20$ with $p = 0.55$ and $p = 0.57$.

Record-to-Record Travel (RRT)

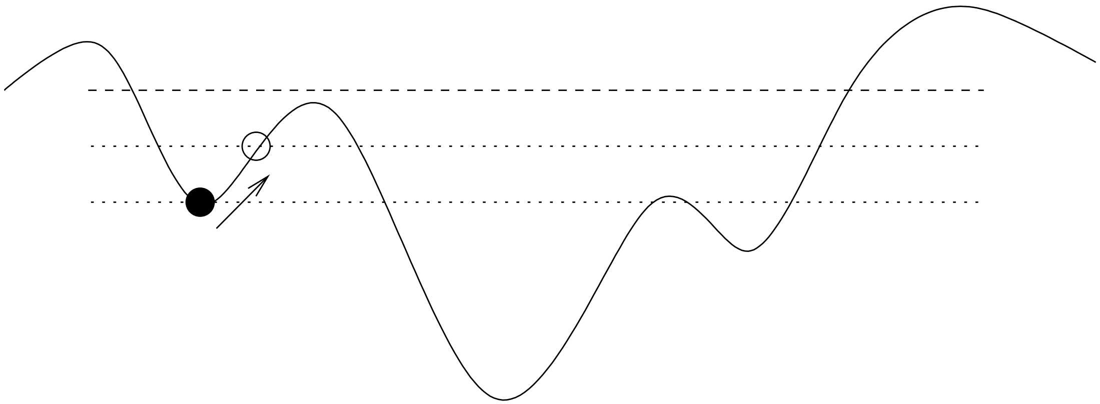
Very simple stochastic local optimisation algorithm introduced by Dueck (1993). Dueck claimed good results on solving 442-city and 532-city TSP's; after that little used.

```
RRT(E,d):
  s = initial feasible solution;
  s* = s; E* = E(s);
  while moves < max_moves do
    if s is a global min. of E then output s & halt,
    else:
      pick a random neighbour s' of s;
      if E(s') <= E* + d then let s = s';
      if E(s') < E* then:
        s* = s'; E* = E(s').
```

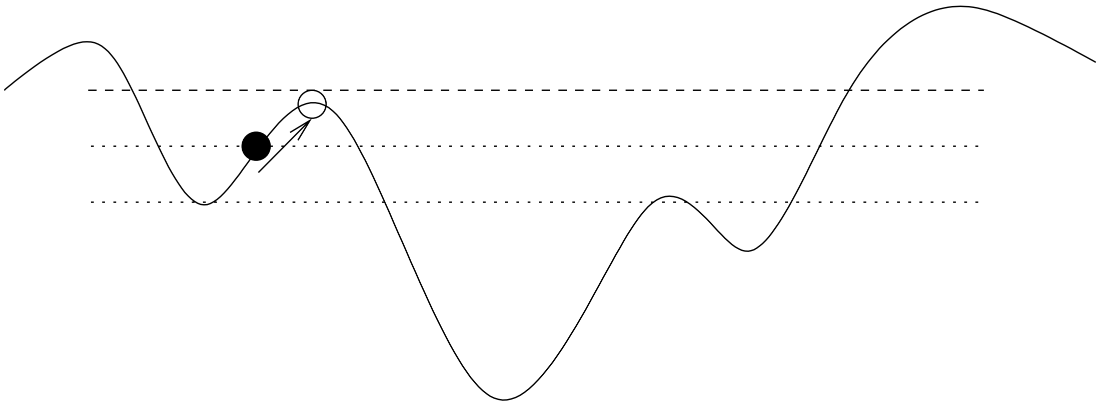
RRT in action (d = 2)



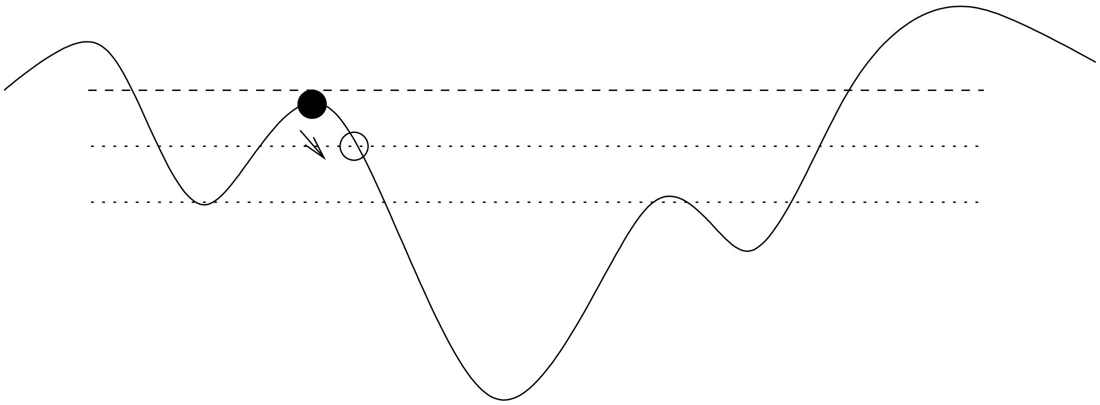
RRT in action (d = 2)



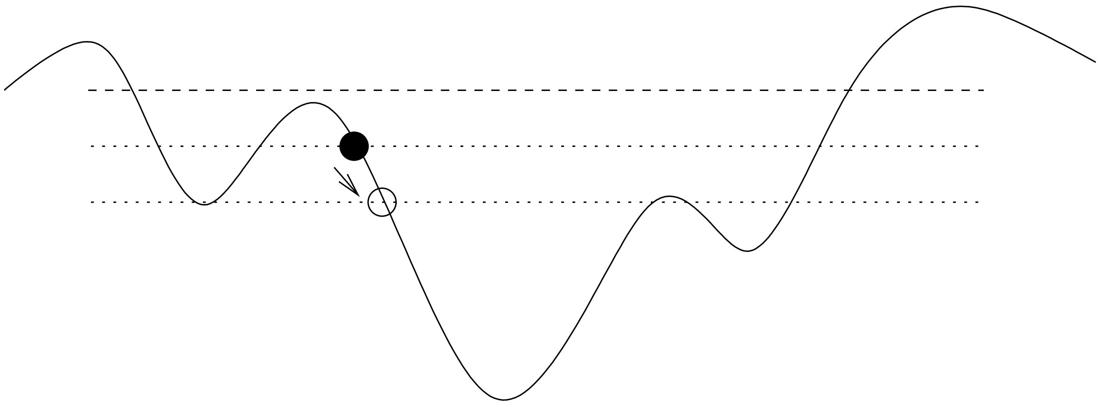
RRT in action (d = 2)



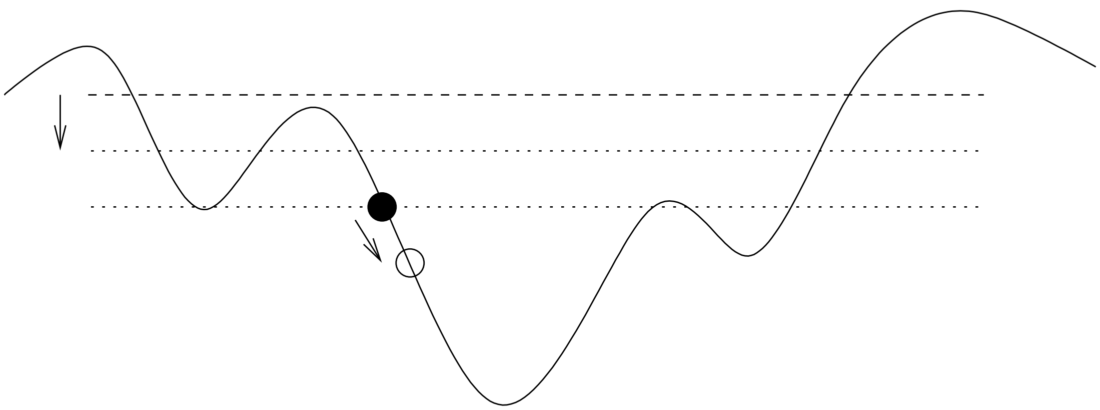
RRT in action (d = 2)



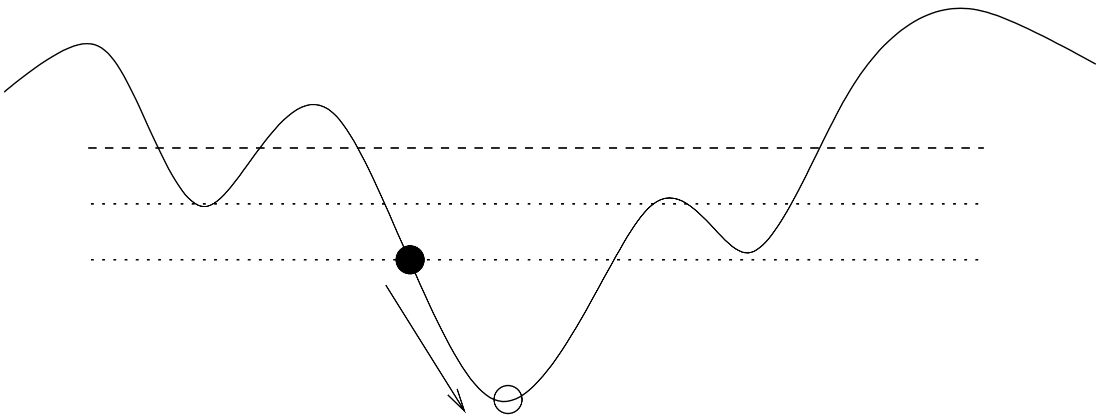
RRT in action (d = 2)



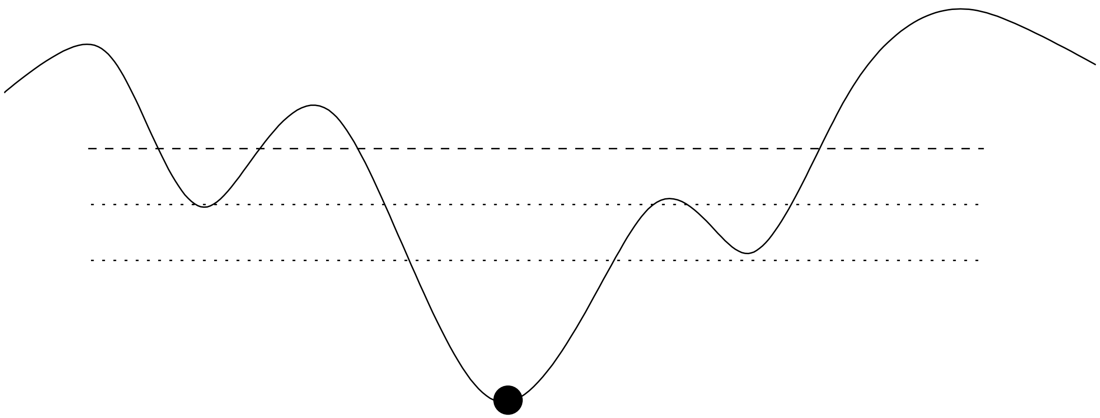
RRT in action (d = 2)



RRT in action (d = 2)



RRT in action (d = 2)



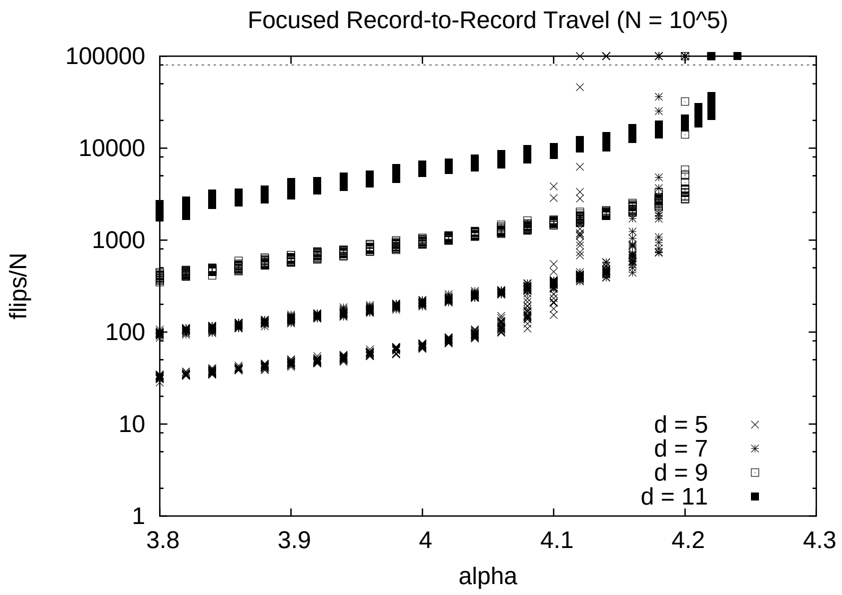
Focused RRT

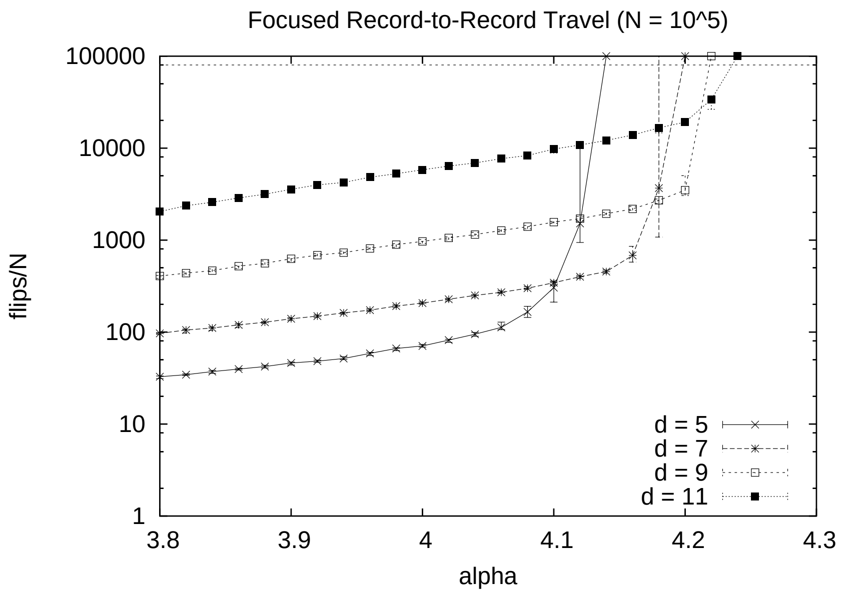
In applying RRT to SAT, $E(s) =$ number of clauses unsatisfied by truth assignment s. Single-variable flip neighbourhoods.

Focusing: flipped variables chosen from unsatisfied clauses. (Precisely: one unsatisfied clause is chosen at random, and from there a variable at random.)

$\Rightarrow$ FRRT = focused RRT.

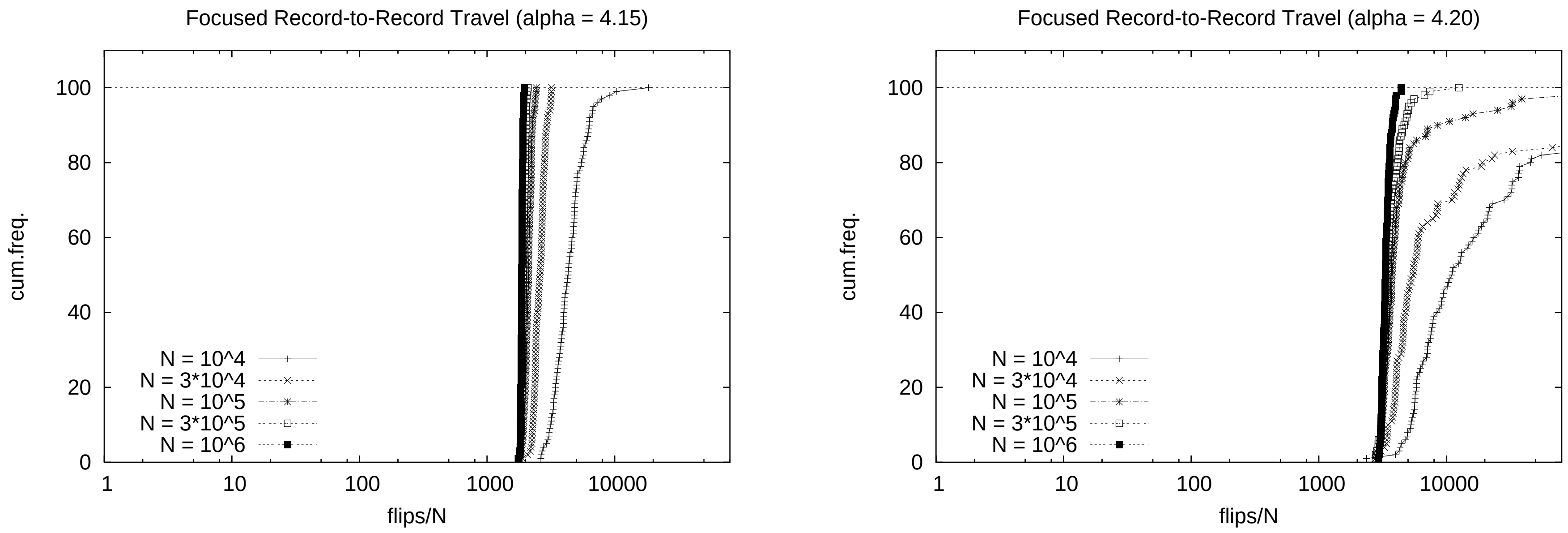
FRRT experiments (3-SAT)





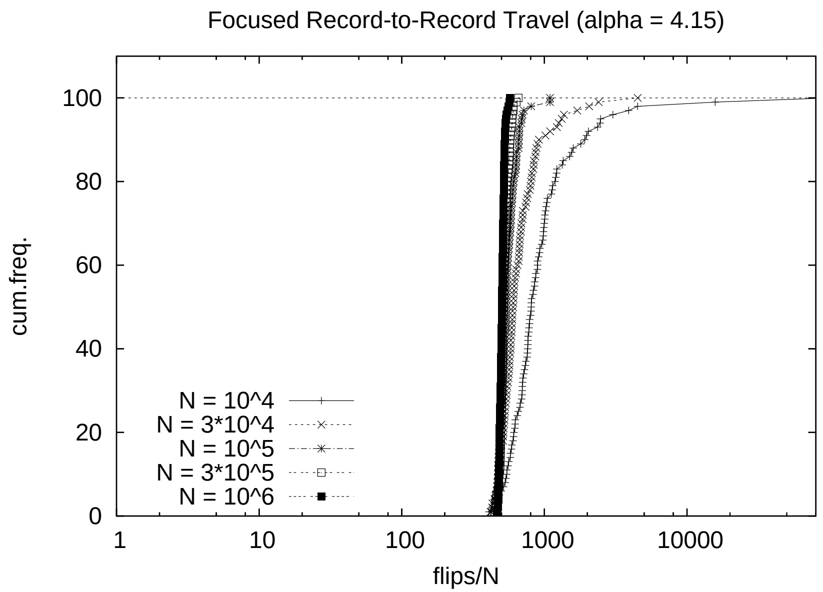
Normalised solution times for FRRT, $\alpha = 3.8 \ldots 4.3$.
Left: complete data; right: medians and quartiles.

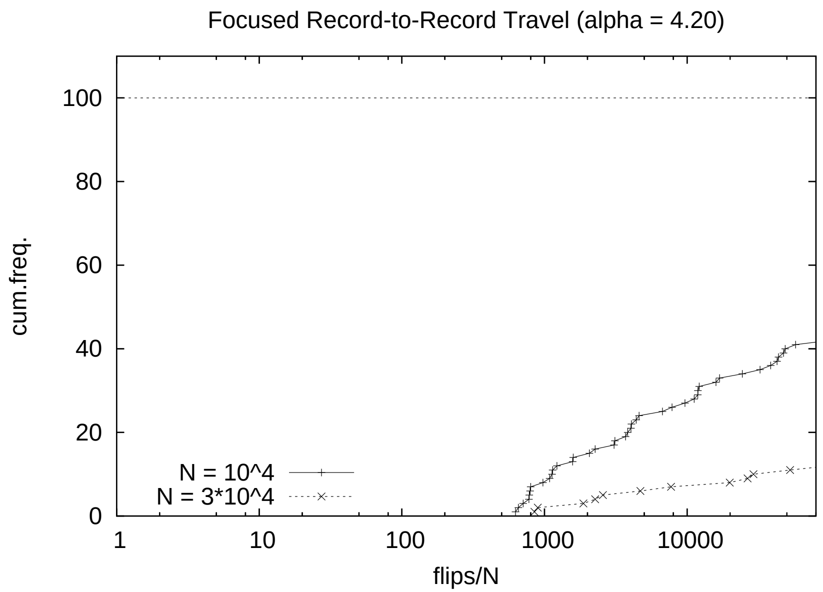
FRRT linear scaling



Cumulative solution time distributions for FRRT with $d = 9$.

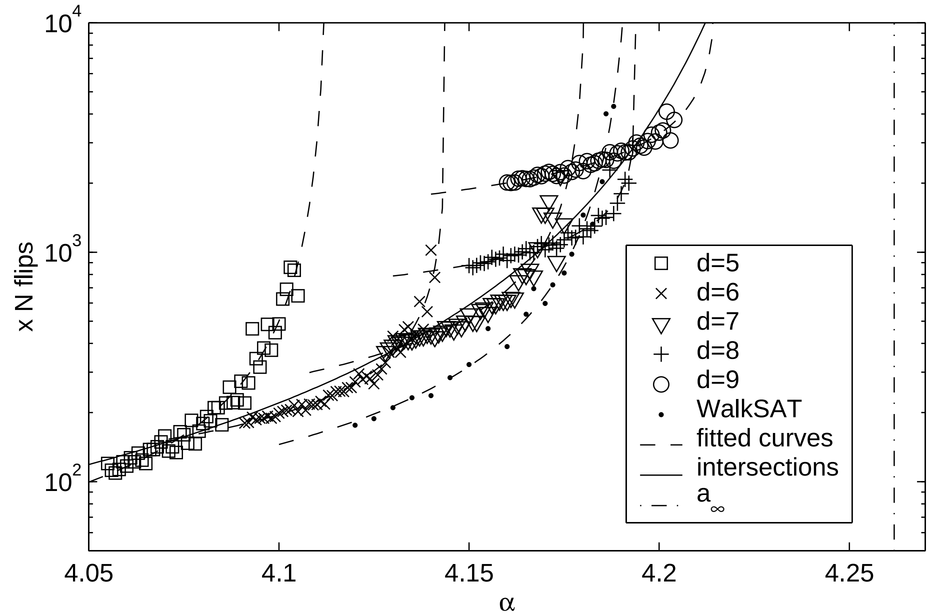
FRRT linear scaling (cont'd)





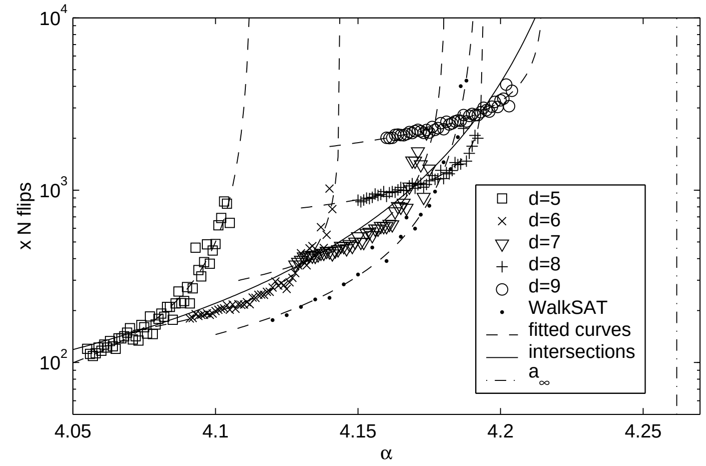
Cumulative solution time distributions for FRRT with $d = 7$.

FRRT qualitative observations



For each fixed d there seems to be a transition value α_d s.th. for $\alpha < \alpha_d$ the algorithm runs in linear time per variable, and for $\alpha > \alpha_d$ requires exponential time per variable. (Empirical estimates: $\alpha_5 \approx 4.11$, $\alpha_6 \approx 4.14$, $\alpha_7 \approx 4.18$, $\alpha_8 \approx 4.19$, $\alpha_9 \approx 4.21$.)

FRRT qualitative observations cont'd

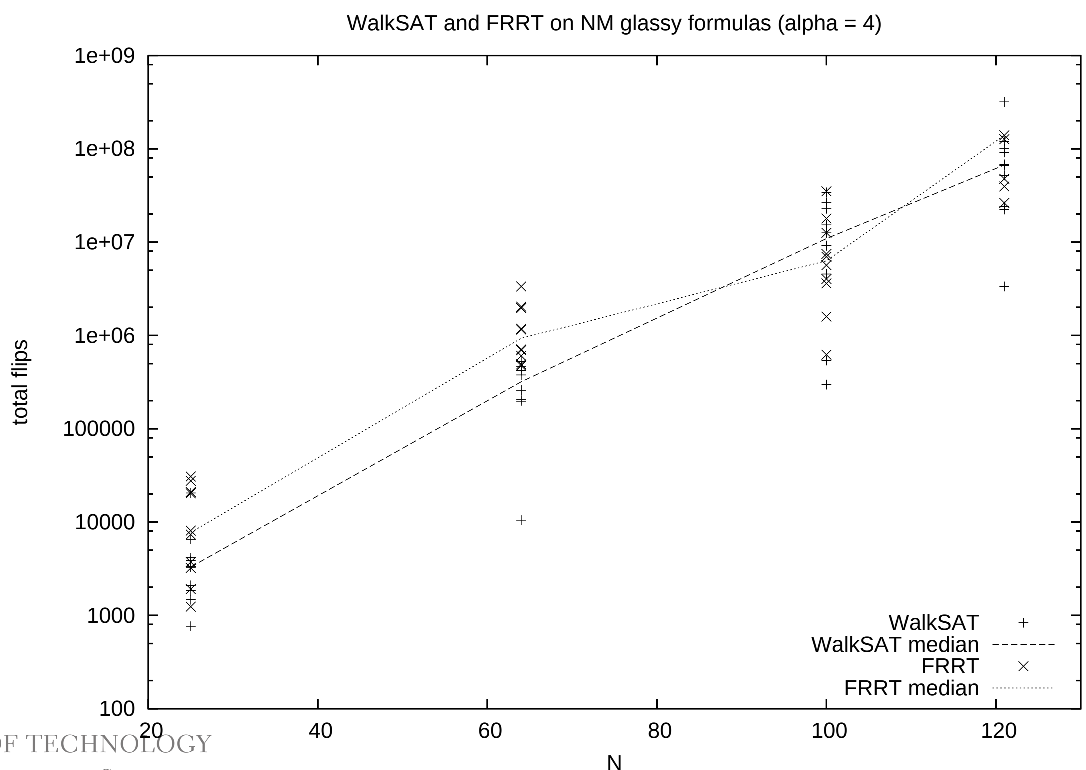


An empirical fit of the transition points α_d suggests that for large d they converge towards $\alpha_\infty \approx 4.26$.

Comparative experiments using WalkSAT with near optimal parameter settings ($p = 0.55$) yield estimate $\alpha_{\text{dyn}} \approx 4.19$ for WalkSAT's transition point.

WalkSAT & FRRT on structured problems

Jia, Moore & Selman (2004) tested WalkSAT and FRRT on highly structured "glassy" 3-SAT formulas. Here the number of variables is always of the form $N = L \times L$; values of $L = 5, 8, 10, 11, 16$ were tried out. At $L = 16$ WalkSAT no longer converged; FRRT did, but only for $d = 5$.

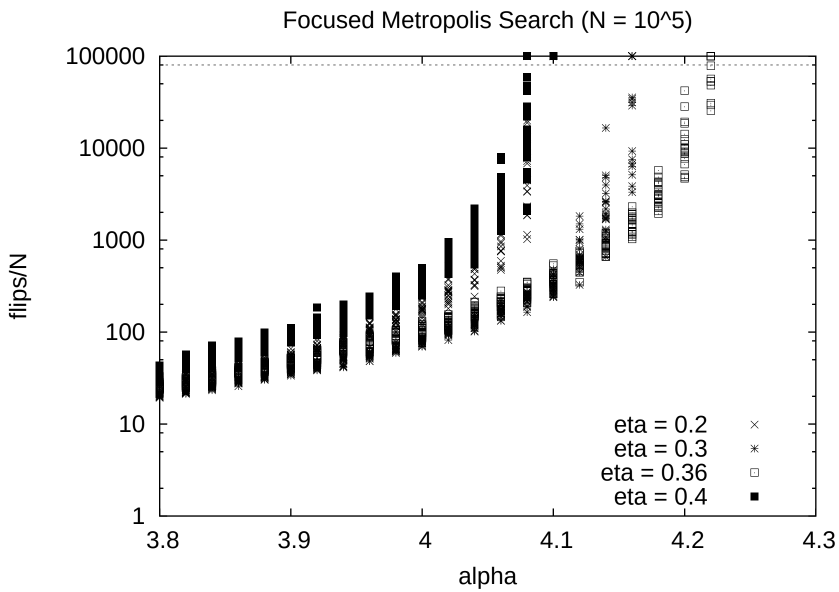


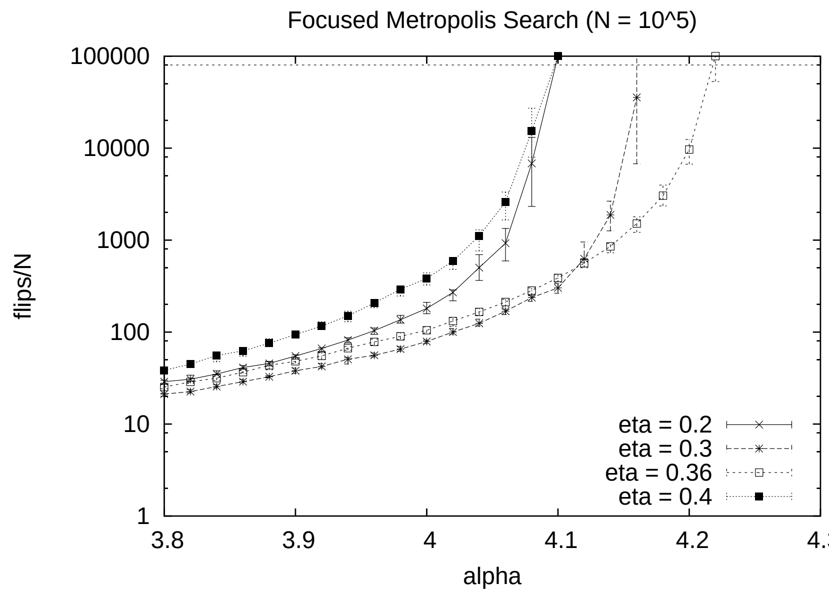
Focused Metropolis Search

Arguably the most natural focused local search algorithm. Variable flip acceptance probabilities determined by a parameter η, $0 \leq \eta \leq 1$.

```
FMS(F,eta):
  s = initial truth assignment;
  while flips < max_flips do
    if s satisfies F then output s & halt, else:
      pick a random unsatisfied clause C in F;
      pick a variable x in C at random;
      let x' = flip(x), s' = s[x'/x];
      if E(s') <= E(s) then flip x, else:
         flip x with prob. eta^(E(s')-E(s)).
```

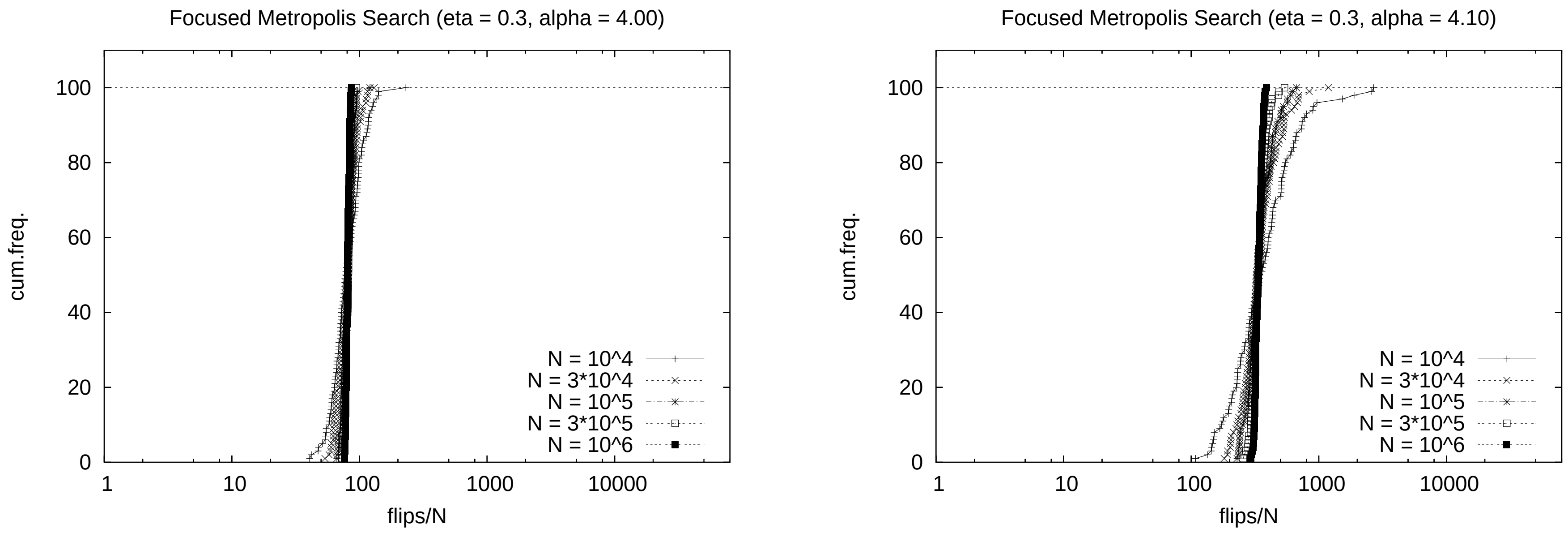
FMS experiments (3-SAT)





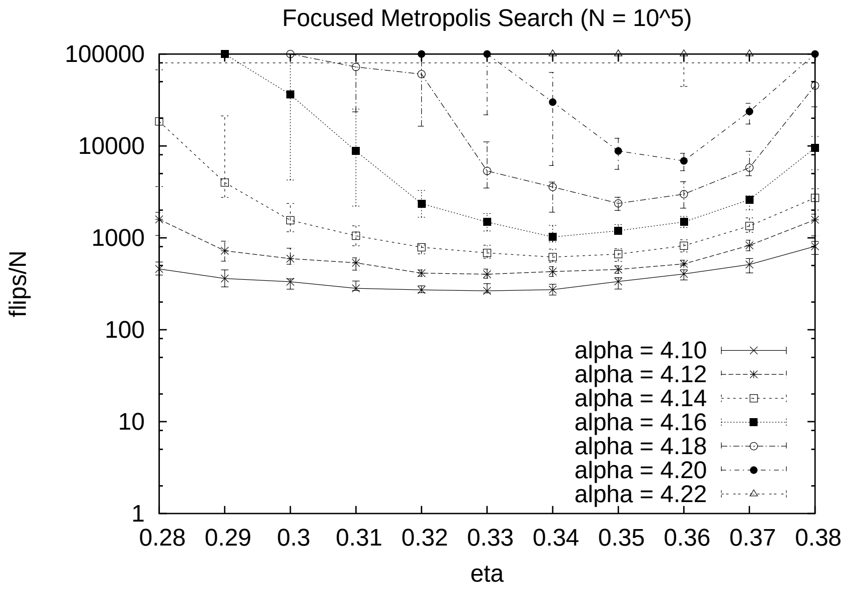
Normalised solution times for FMS, $\alpha = 3.8 \ldots 4.3$.
Left: complete data; right: medians and quartiles.

FMS linear scaling



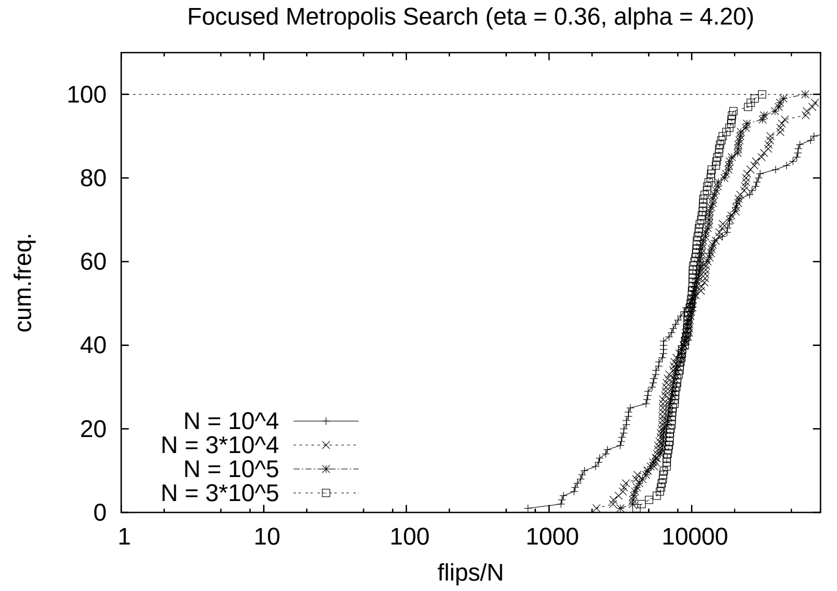
Cumulative solution time distributions for FMS with $\eta = 0.3$.

FMS optimal acceptance ratio?



Normalised solution times for FMS with $\eta = 0.28 \dots 0.38$, $\alpha = 4.10 \dots 4.22$.

FMS optimal acceptance ratio cont'd



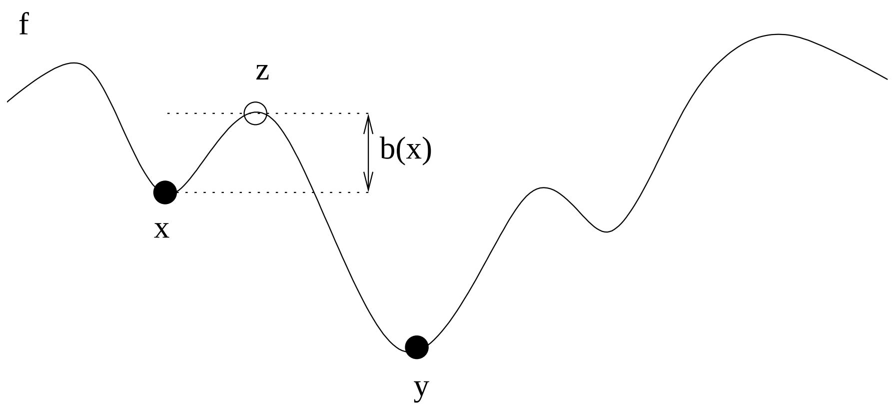
Cumulative solution time distributions for FMS with $\eta = 0.36$, $\alpha = 4.20$.

Analysis?

- For FRRT: landscape structure?
- For FMS: contact processes?

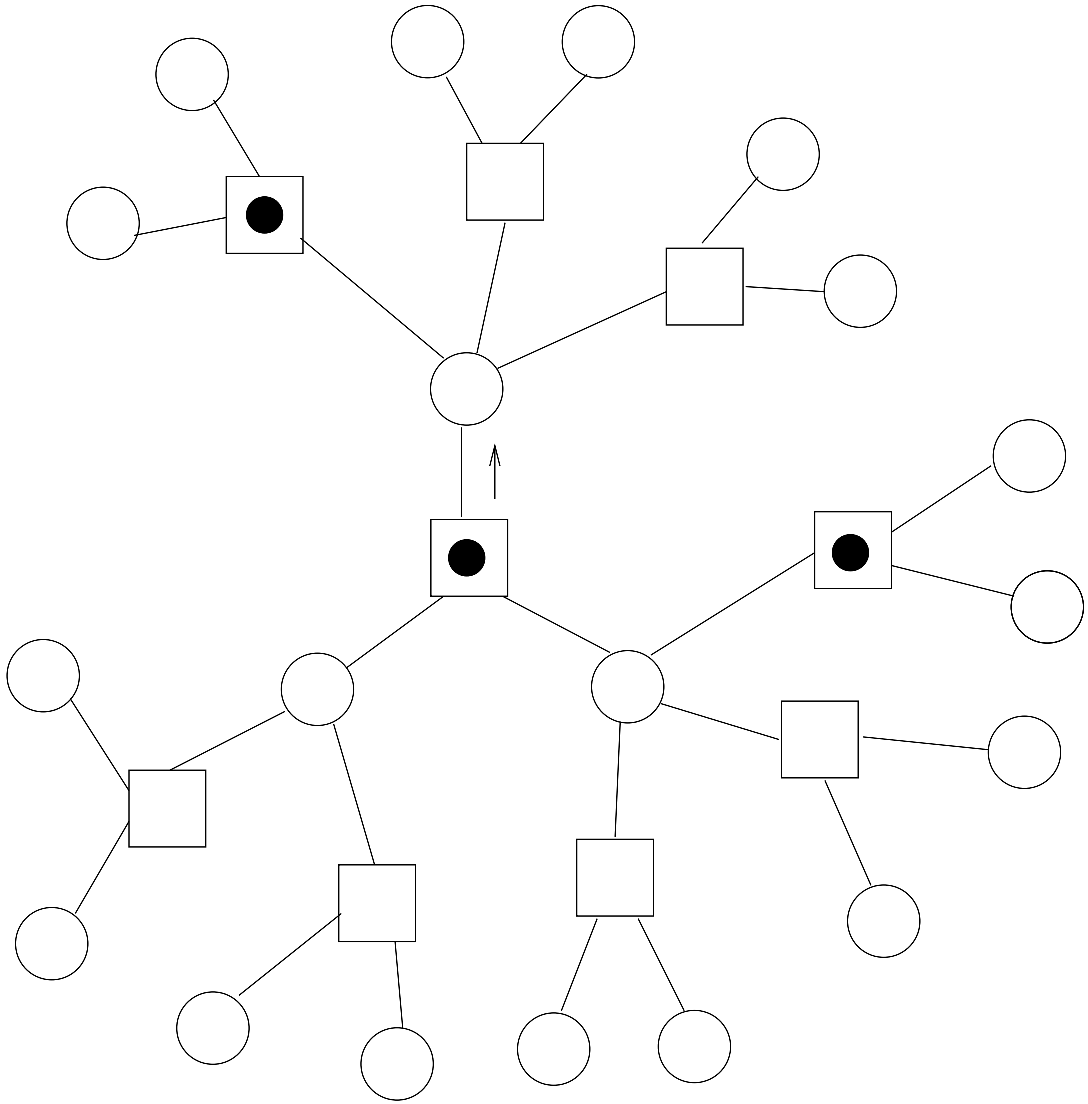
Combinatorial landscapes

Reidys & Stadler (SIAM Rev. 2002)

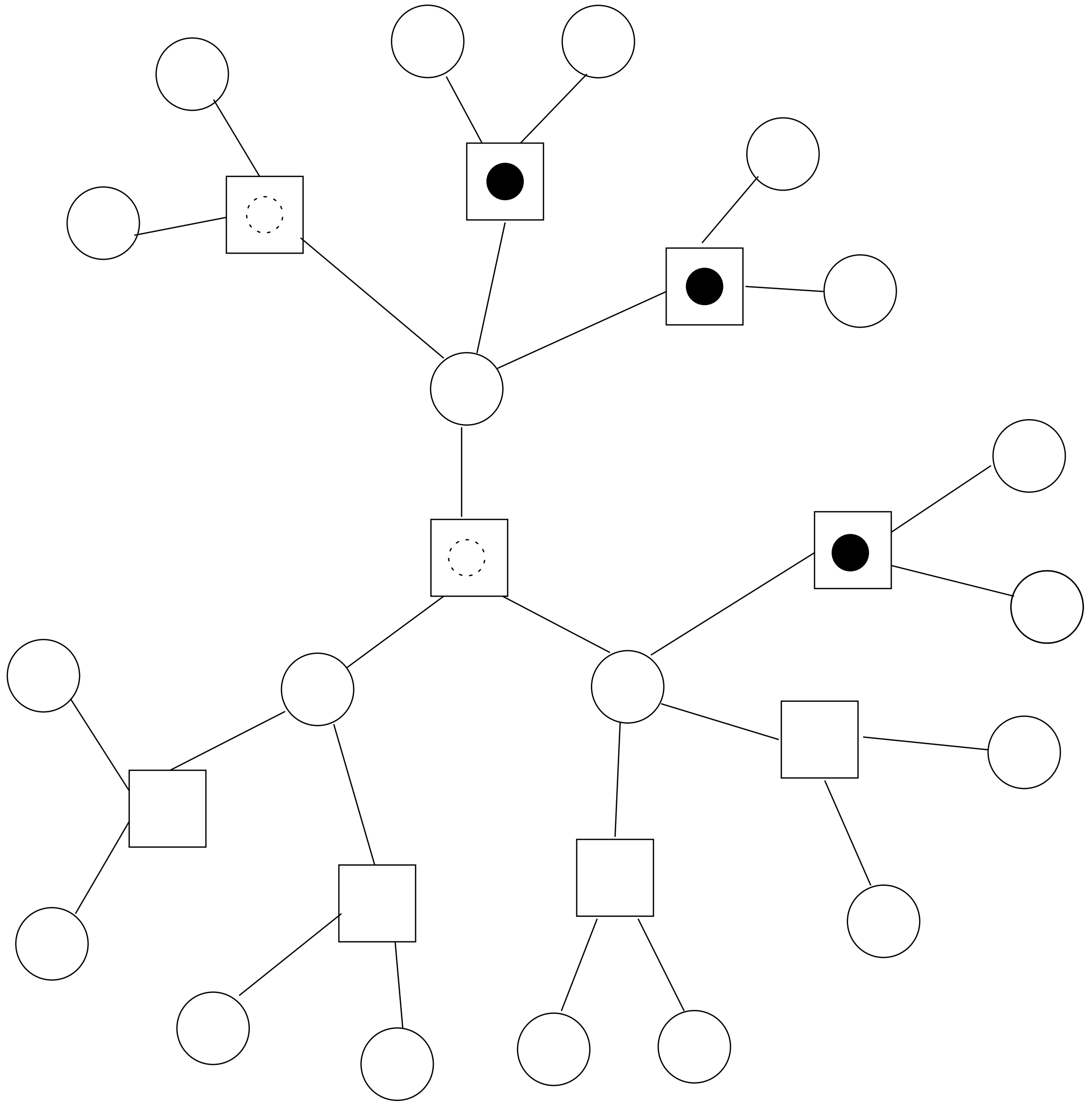


- *barrier height from x to y*: $b(x,y) = \min\{\max\{f(z) - f(x), 0 \mid z \in p\} \mid p$ an x-y path$\}$
- *barrier height of $x \notin$ Opt*: $b(x) = \min\{b(x,y) \mid f(y) < f(x)\}$
- *depth* of a landscape: $D_f = \max\{b(x) \mid x \notin \text{Opt}\}$

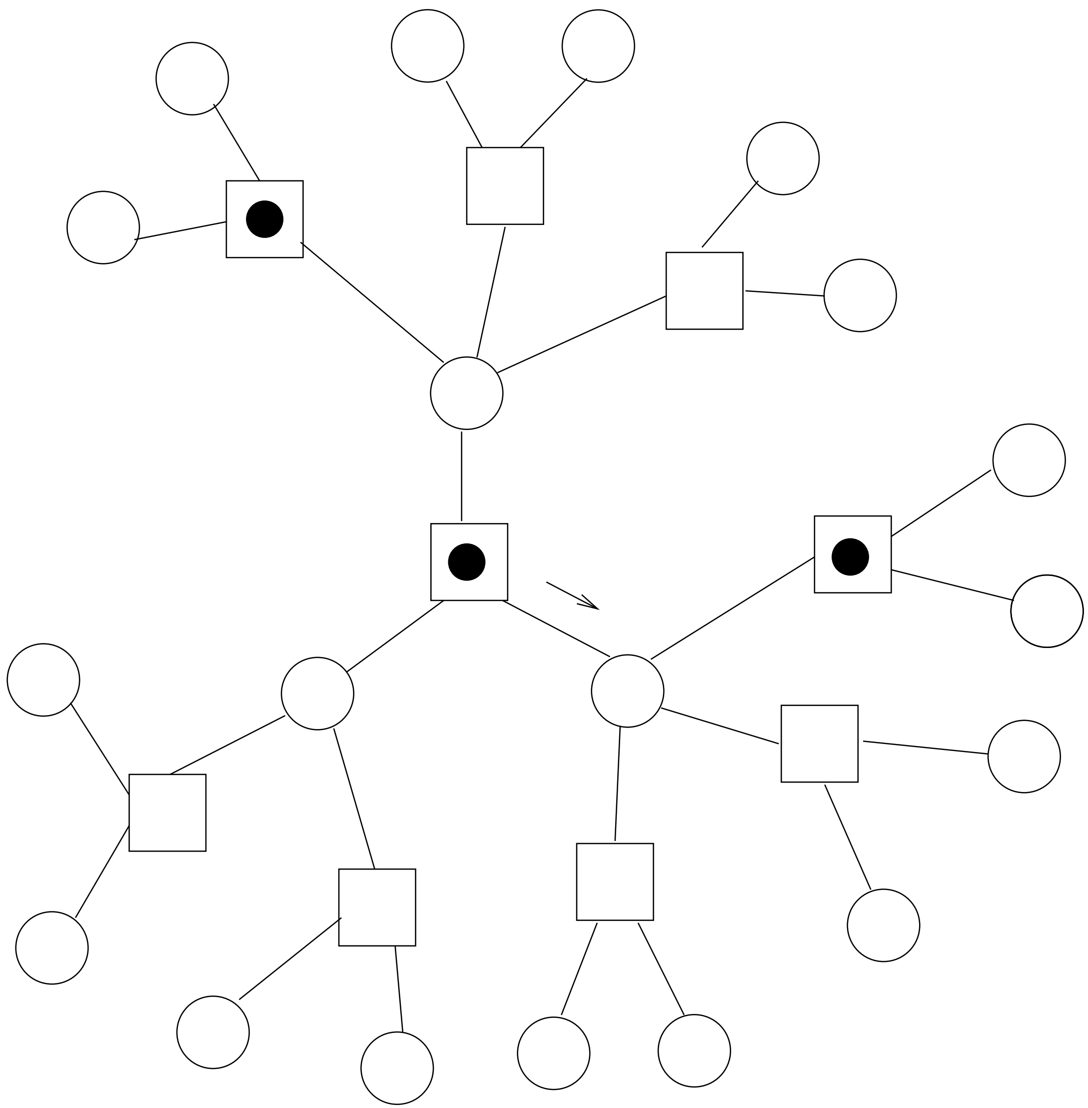
Focused search as a contact process



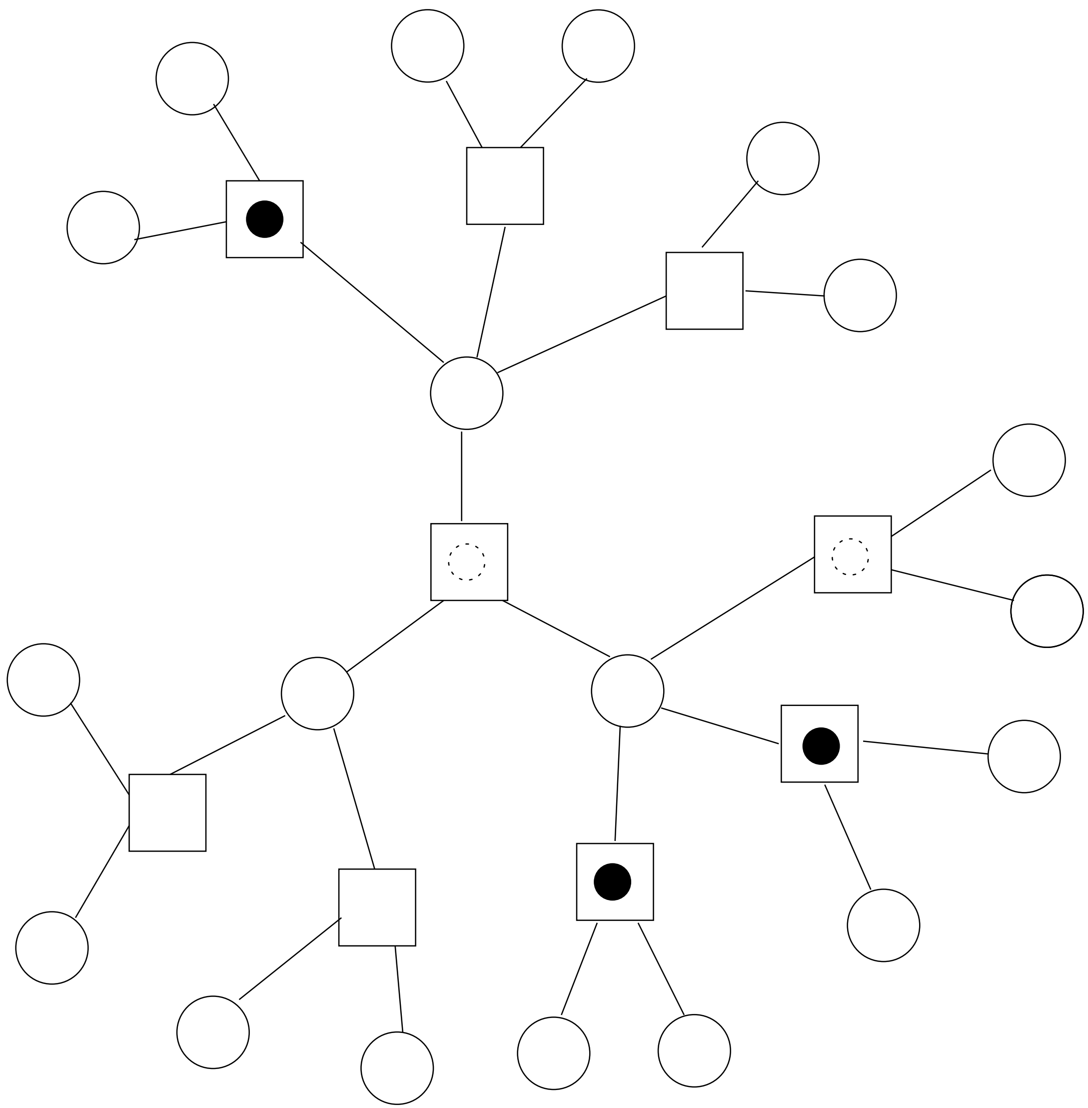
Focused search as a contact process



Focused search as a contact process



Focused search as a contact process



References

[1] S. Seitz, P. Orponen: An efficient local search method for random 3-satisfiability. *Proc. LICS'03 Workshop on Typical Case Complexity and Phase Transitions (Ottawa, Canada, June 2003).* Elsevier Electronic Notes in Discrete Mathematics Vol. 16.

[2] S. Seitz, M. Alava, P. Orponen: Threshold behaviour of WalkSAT and focused Metropolis search on random 3-satisfiability. *Proc. 8th Intl. Conf. on Theory and Applications of Satisfiability Testing (St. Andrews, Scotland, June 2005).* Springer-Verlag, Berlin, to appear.

[3] S. Seitz, M. Alava, P. Orponen: Focused local search algorithms for random 3-satisfiability. To appear.