Stochastic Control by Entropy Compression

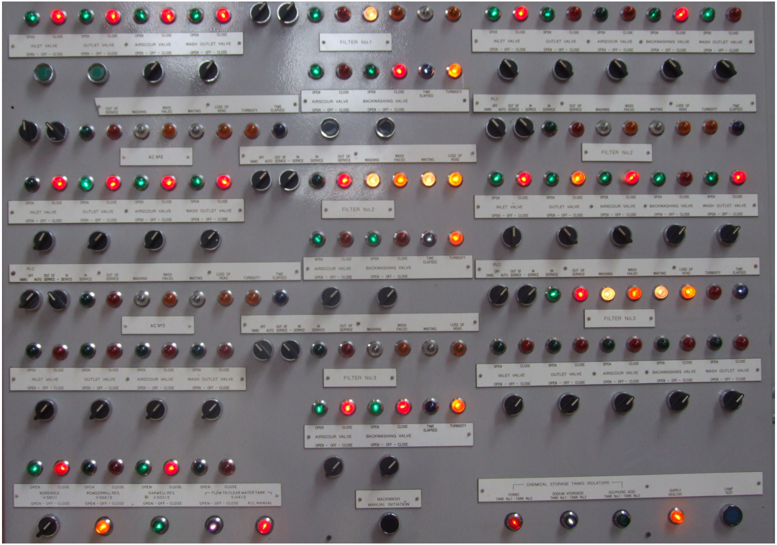
Dimitris Achlioptas [1] Fotis Iliopoulos [2] Nikos Vlassis [3]

[1]UC Santa Cruz

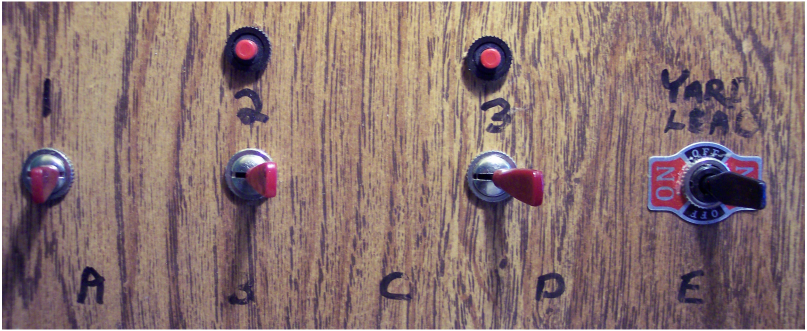
[2]UC Berkeley

[3]Adobe Research

Stochastic Control



Stochastic Control



The Probabilistic Method

- Probability space + Collection $B = \{\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_m\}$ of "bad" events.
- If $\{\mathcal{E}_i\}$ are independent, $\Pr[\text{Nothing bad happens}] = \prod_{i=1}^{m}(1 - p_i)$.
- What if avoiding some bad events boosts some other bad events ?

Example: $\Omega = \{0,1\}^3$ with uniform measure, $F = (x_1 \vee x_2) \wedge (\overline{x_2} \vee x_3)$.

The Probabilistic Method

- Probability space + Collection $B = \{\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_m\}$ of “bad” events.
- If $\{\mathcal{E}_i\}$ are independent, $\Pr[\text{Nothing bad happens}] = \prod_{i=1}^{m}(1-p_i)$.
- What if avoiding some bad events boosts some other bad events ?

Example: $\Omega = \{0,1\}^3$ with uniform measure, $F = (x_1 \vee x_2) \wedge (\overline{x_2} \vee x_3)$.

Symmetric LLL (Erdős, Lovász ’75)

Assume that every bad event has probability at most p and is independent of all but at most Δ bad events. If

$$p\Delta \leq 1/\mathrm{e} \ ,$$

then $\Pr[\text{Nothing bad happens}] > 0$.

A Tight Example and a Breakthrough

Example

Every k-CNF formula in which each clause shares variables with at most $\Delta \leq 2^k/\mathrm{e}$ other clauses is satisfiable. **Proof:** $2^{-k}\Delta \leq 1/\mathrm{e}$.

A Tight Example and a Breakthrough

Example

Every k-CNF formula in which each clause shares variables with at most $\Delta \leq 2^k/\mathrm{e}$ other clauses is satisfiable. **Proof:** $2^{-k}\Delta \leq 1/\mathrm{e}$.

Theorem (Gebauer, Szabo, Tardos '11)

There exist unsatisfiable formulas with $\Delta = (1 + \delta_k)2^k/\mathrm{e}$, where $\delta_k \to 0$.

A Tight Example and a Breakthrough

Example

Every k-CNF formula in which each clause shares variables with at most $\Delta \leq 2^k/\mathrm{e}$ other clauses is satisfiable. **Proof:** $2^{-k}\Delta \leq 1/\mathrm{e}$.

Theorem (Gebauer, Szabo, Tardos '11)

There exist unsatisfiable formulas with $\Delta = (1+\delta_k)2^k/\mathrm{e}$, where $\delta_k \to 0$.

Algorithmic LLL: a s.t.a can be found efficiently if $\Delta \leq 2^{k/4}$

[Beck 91], [Alon 91], [Molloy, Reed 98], [Czumaj, Scheideler 00], [Srinivasan 08]

A Tight Example and a Breakthrough

Example

Every k-CNF formula in which each clause shares variables with at most $\Delta \leq 2^k/\mathrm{e}$ other clauses is satisfiable. **Proof:** $2^{-k}\Delta \leq 1/\mathrm{e}$.

Theorem (Gebauer, Szabo, Tardos '11)

There exist unsatisfiable formulas with $\Delta = (1+\delta_k)2^k/\mathrm{e}$, where $\delta_k \to 0$.

Algorithmic LLL: a s.t.a can be found efficiently if $\Delta \leq 2^{k/4}$

[Beck 91], [Alon 91], [Molloy, Reed 98], [Czumaj, Scheideler 00], [Srinivasan 08]

Theorem (Moser '09)

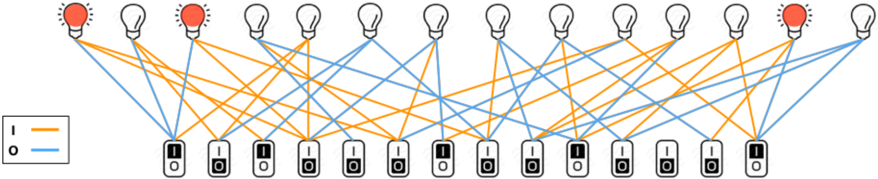
If $\Delta(F) < 2^{k-5}$ a sat assignment can be found in $O(|V| + |F|\log|F|)$.

Moser's ideas, with more care, yield $2^k/\mathrm{e}$. [Messner, Thierauf 11]

Moser's Algorithm

Resample

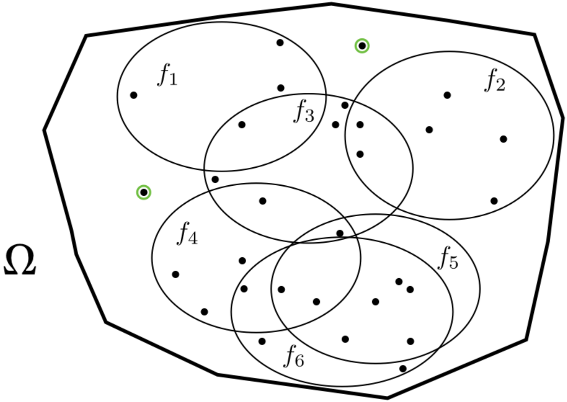
Start at an arbitrary truth assignment
while violated clauses exist **do**
  Select a random violated clause c
  for each variable v of c independently **do**
    Set v to $0/1$ with equal probability



The Flaws/Actions Framework [A., Iliopoulos FOCS'14/JACM'16]

Let Ω be an arbitrary finite set.

Let $F = \{f_1, f_2, \ldots, f_m\}$ be arbitrary subsets of Ω called flaws.



The Flaws/Actions Framework [A., Iliopoulos FOCS'14/JACM'16]

Let Ω be an arbitrary finite set.

Let $F = \{f_1, f_2, \ldots, f_m\}$ be arbitrary subsets of Ω called flaws.

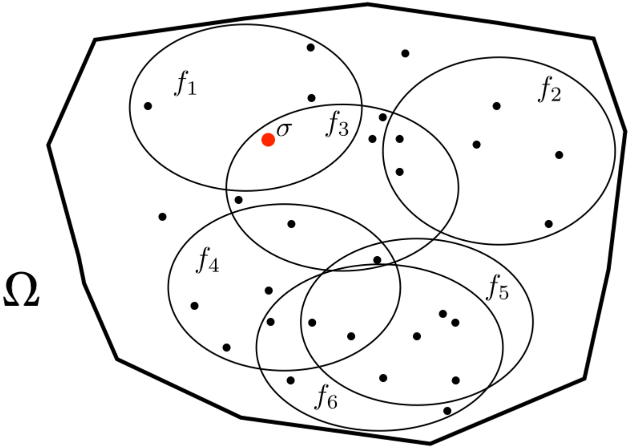
Goal

When flawless objects exist, find one in much less time than $|\Omega|$.

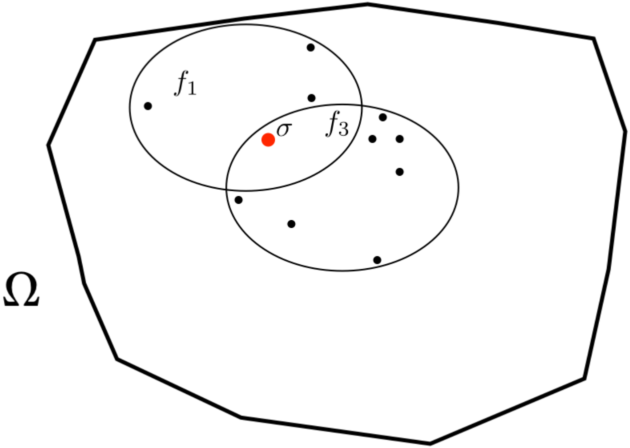
How?

- Specify a directed graph D on Ω such that:
  - Every flawed object has outdegree at least 1.
  - Every flawless object has outdegree 0.
- Start at an arbitrary $\sigma_1 \in \Omega$
- Take a random walk on D until you reach a sink.

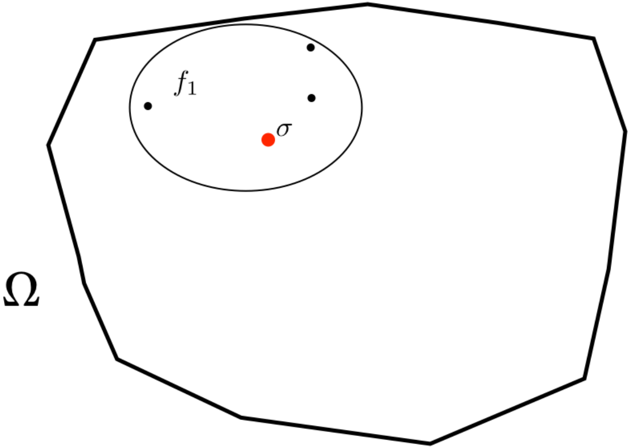
Address a Random Flaw of the Current State



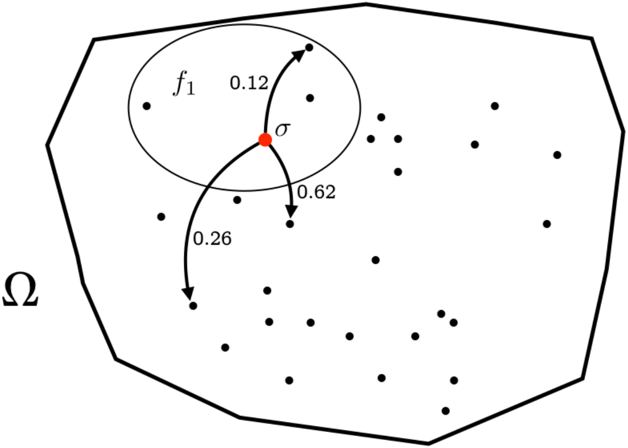
Address a Random Flaw of the Current State



Address a Random Flaw of the Current State



Address a Random Flaw of the Current State



Address a Random Flaw of the Current State

Moser's Algorithm on a k-CNF formula

If $F = c_1 \wedge \cdots \wedge c_m$ is a k-CNF formula with n variables:

- $\Omega = \{0,1\}^n$
- $f_i = \{\sigma \in \Omega : \sigma \text{ violates clause } c_i\}$
- $A(f_i, \sigma) = \{\text{The } 2^k \text{ mutations of } \sigma \text{ through } \mathrm{var}(c_i)\}$
- The 2^k actions in $A(f_i, \sigma)$ are equiprobable, for all (i, σ)

Measuring the Speed of State-Space Exploration

Local Entropy

Let $\rho_i(\sigma, \tau)$ denote the probability of $\sigma \to \tau$ when addressing f_i at σ. The local entropy of flaw f_i is

$$\min_{\sigma \in f_i} H[\rho_i(\sigma, \cdot)]$$

Example: The local entropy of every flaw in Moser's algorithm is k.

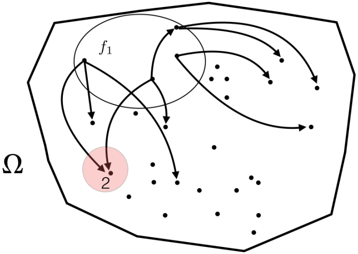
Measuring the Speed of State-Space Exploration

Local Entropy

Let $\rho_i(\sigma, \tau)$ denote the probability of $\sigma \to \tau$ when addressing f_i at σ. The local entropy of flaw f_i is

$$\min_{\sigma \in f_i} H[\rho_i(\sigma, \cdot)]$$

Example: The local entropy of every flaw in Moser's algorithm is k.



Measuring the Speed of State-Space Exploration

Local Entropy

Let $\rho_i(\sigma, \tau)$ denote the probability of $\sigma \to \tau$ when addressing f_i at σ. The local entropy of flaw f_i is

$$\min_{\sigma \in f_i} H[\rho_i(\sigma, \cdot)]$$

Example: The local entropy of every flaw in Moser's algorithm is k.

Congestion

Let $\mathrm{InDeg}_i(\tau) = |\sigma : \tau \in A(f_i, \sigma)|$. The congestion of flaw f_i is

$$\max_{\tau \in \Omega} \log_2[\mathrm{InDeg}_i(\tau)] \quad .$$

Example: The congestion of every flaw in Moser's algorithm is 0.

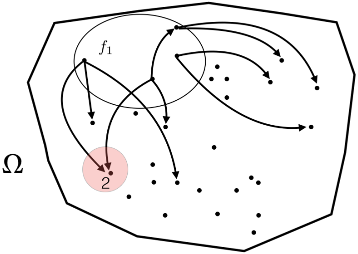
Measuring the Speed of State-Space Exploration

Local Entropy

Let $\rho_i(\sigma, \tau)$ denote the probability of $\sigma \to \tau$ when addressing f_i at σ. The local entropy of flaw f_i is

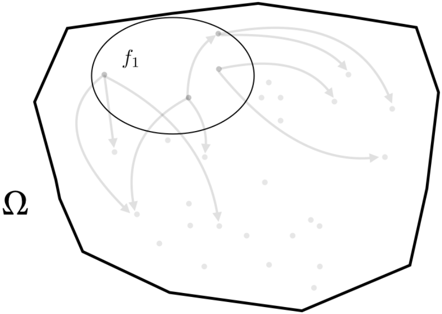
$$\min_{\sigma \in f_i} H[\rho_i(\sigma, \cdot)]$$

Example: The local entropy of every flaw in Moser's algorithm is k.



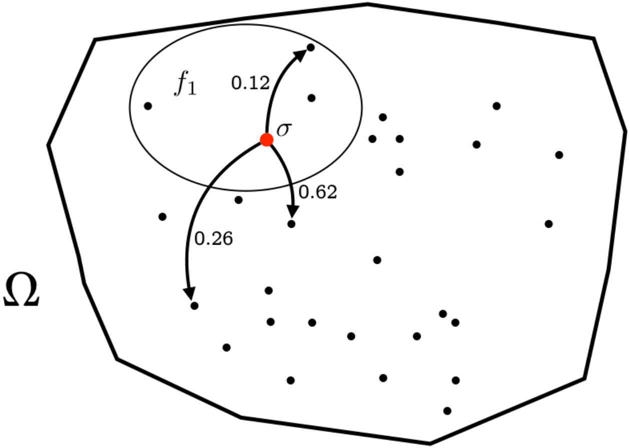
Amenability = Local Entropy minus Congestion

$$\text{Amenability}(f_i) = \text{Local Entropy}(f_i) - \text{Congestion}(f_i)$$
$$= \min_{\sigma \in f_i} H[\rho(\sigma, \cdot)] - \max_{\tau \in \Omega} \log_2[\text{Indeg}_i(\tau)]$$

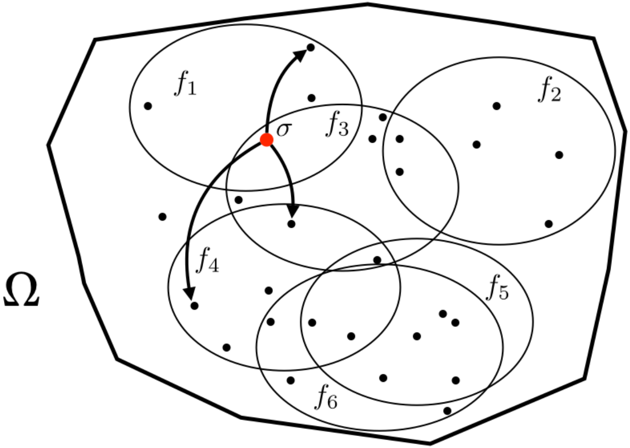


Example: The amenability of every flaw in Moser's algorithm is $k - 0 = k$

Potential Causality



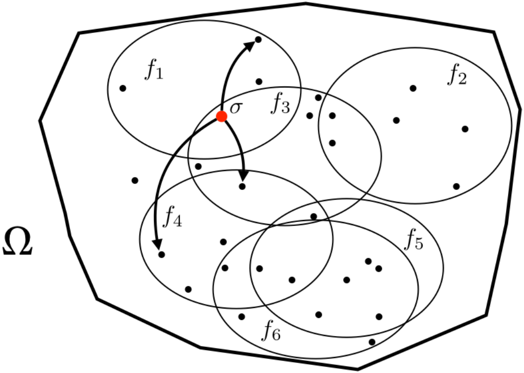
Potential Causality



Potential Causality

For each $\tau \in A(f_i, \sigma)$ we define the set of flaws $\Gamma_i(\sigma, \tau)$ to contain:

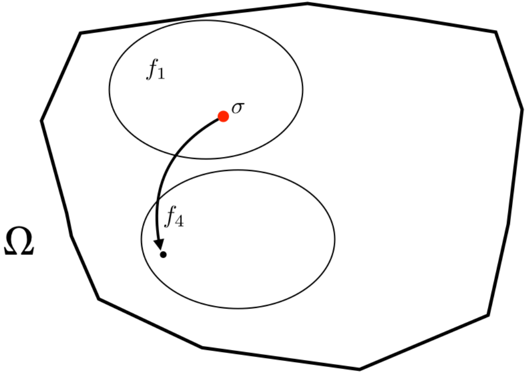
- Every flaw present in τ that was not present in σ, and,
- Flaw f_i itself, if $\tau \in f_i$.



Potential Causality

For each $\tau \in A(f_i, \sigma)$ we define the set of flaws $\Gamma_i(\sigma, \tau)$ to contain:

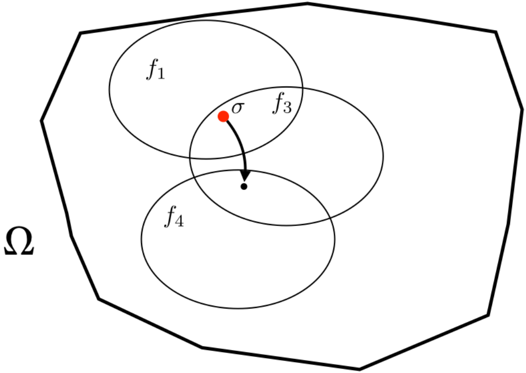
- Every flaw present in τ that was not present in σ, and,
- Flaw f_i itself, if $\tau \in f_i$.



Potential Causality

For each $\tau \in A(f_i, \sigma)$ we define the set of flaws $\Gamma_i(\sigma, \tau)$ to contain:

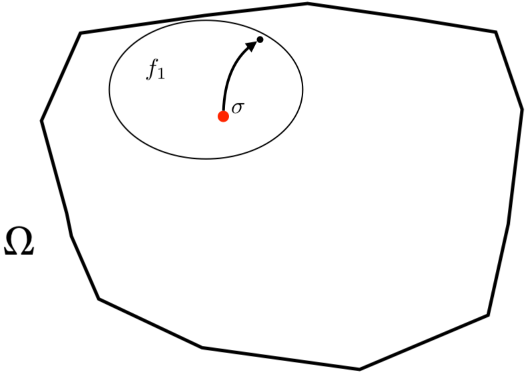
- Every flaw present in τ that was not present in σ, and,
- Flaw f_i itself, if $\tau \in f_i$.



Potential Causality

For each $\tau \in A(f_i, \sigma)$ we define the set of flaws $\Gamma_i(\sigma, \tau)$ to contain:

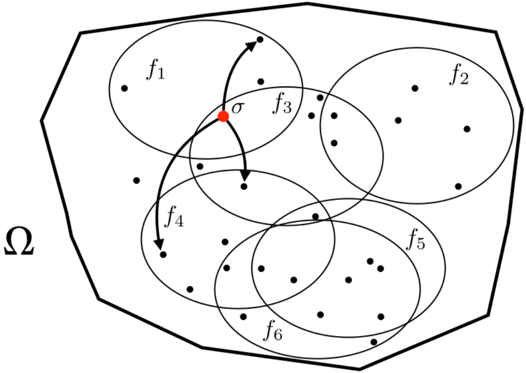
- Every flaw present in τ that was not present in σ, and,
- Flaw f_i itself, if $\tau \in f_i$.



Potential Causality

For each $\tau \in A(f_i, \sigma)$ we define the set of flaws $\Gamma_i(\sigma, \tau)$ to contain:

- Every flaw present in τ that was not present in σ, and,
- Flaw f_i itself, if $\tau \in f_i$.



Potential Causality

For each $\tau \in A(f_i, \sigma)$ we define the set of flaws $\Gamma_i(\sigma, \tau)$ to contain:

- Every flaw present in τ that was not present in σ, and,
- Flaw f_i itself, if $\tau \in f_i$.

Potential Causality

Let

$$\Gamma_i = \bigcup_{\substack{\sigma \in f_i \\ \tau \in A(f_i, \sigma)}} \Gamma_i(\sigma, \tau)$$

Example

In Moser's algorithm each clause potentially causes:

- Each clause with which it shares a variable with opposite sign
- Itself

Results: Noiseless Case

- Let $\sigma_1 \in \Omega$ be arbitrary.
- For $t = 1, 2, \ldots$
  - Let f_i be a random flaw present in σ_t.
  - Move to $\tau \in A(f_i, \sigma)$ with probability $\rho_i(\sigma, \tau)$.

Results: Noiseless Case

- Let $\sigma_1 \in \Omega$ be arbitrary.
- For $t = 1, 2, \ldots$
  - Let f_i be a random flaw present in σ_t.
  - Move to $\tau \in A(f_i, \sigma)$ with probability $\rho_i(\sigma, \tau)$.

Theorem

If for every flaw f_i,

$$\sum_{f_j \in \Gamma_i} 2^{-\mathrm{Amen}(f_j)} < \frac{1}{4} \ ,$$

then the probability we don't reach a flawless state within $O(T_0 + s)$ steps is less than 2^{-s}, where $T_0 = \log_2 |\Omega| + |F|$.

Results: Noiseless Case

- Let $\sigma_1 \in \Omega$ be arbitrary.
- For $t = 1, 2, \ldots$
  - Let f_i be a random flaw present in σ_t.
  - Move to $\tau \in A(f_i, \sigma)$ with probability $\rho_i(\sigma, \tau)$.

Theorem

If for every flaw f_i,

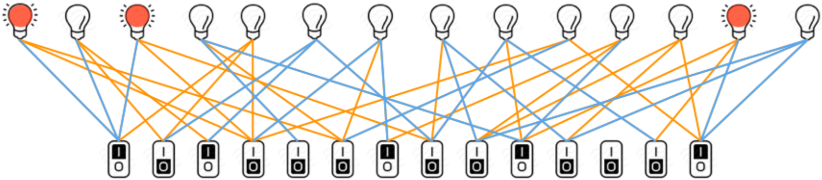
$$\sum_{f_j \in \Gamma_i} 2^{-\mathrm{Amen}(f_j)} < \frac{1}{4} ,$$

then the probability we don't reach a flawless state within $O(T_0 + s)$ steps is less than 2^{-s}, where $T_0 = \log_2 |\Omega| + |F|$.

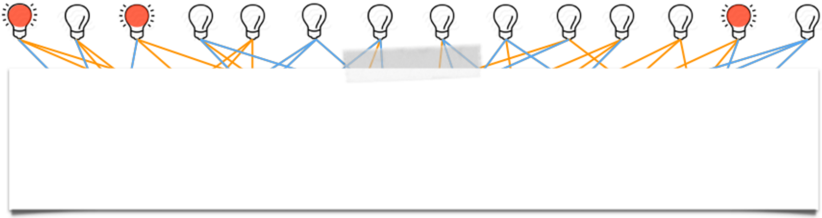
Example

For Moser's algorithm on k-CNF formulas we get $\Delta(F) < 2^{k-2}$.

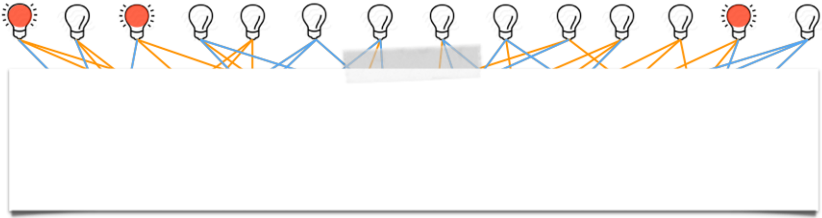
Let's Add Some Noise



Let's Add Some Noise



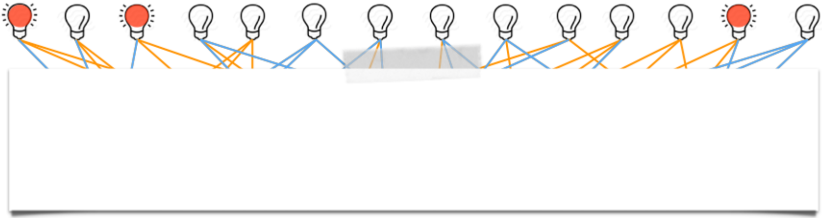
Let's Add Some Noise



Question: Can we still find a satisfying assignment if:

- The lightbulbs are faulty, having both false positives and negatives.
- When we reset a variable it doesn't always happen.
- Variables change values on their own, silently.

Let's Add Some Noise

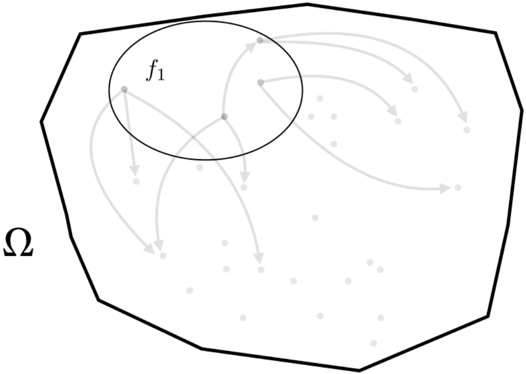


Question: Can we still find a satisfying assignment if:

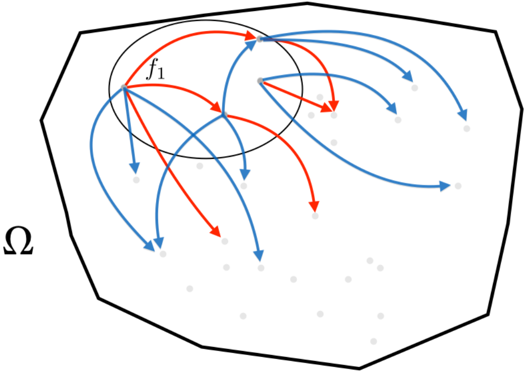
- The lightbulbs are faulty, having both false positives and negatives.
- When we reset a variable it doesn't always happen.
- Variables change values on their own, silently.

Answer: Yes! Lack of "internal conflict" implies "noise resistance".

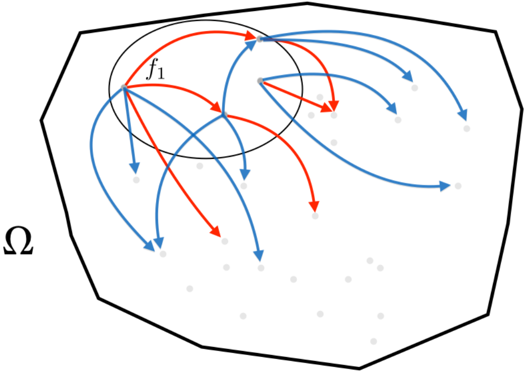
Modeling Noise



Modeling Noise



Modeling Noise



In each step:

- With probability $1 - p$ the system acts normally
- With probability p it acts according to the adversary's chain

Thinking of the adversary's chain as just some other algorithm, we define

$$\begin{aligned} q_i(p) &= p\left[|\Gamma_i^*|\left(b^* + \frac{5}{2} + h(p)\right) - 2 - h(p)\right] \\ &\sim p\,|\Gamma_i^*|\,b^* \ , \end{aligned}$$

where $b^* = \max_i \text{Congestion}(f_i)$.

Thinking of the adversary's chain as just some other algorithm, we define

$$q_i(p) = p\left[|\Gamma_i^*|\left(b^* + \frac{5}{2} + h(p)\right) - 2 - h(p)\right]$$
$$\sim p\,|\Gamma_i^*|\,b^* \quad ,$$

where $b^* = \max_i \text{Congestion}(f_i)$.

Theorem

If for every flaw f_i,

$$\sum_{f_j \in \Gamma_i} 2^{-\text{Amen}(f_j) + q_j(p)} < \frac{1}{4} 2^{-h(p)} \quad ,$$

then the probability we don't reach a flawless state within $O(T_0 + s)$ steps is less than 2^{-s}, where $T_0 = \log_2|\Omega| + m$.

Thanks!