Investigating 3SAT

(Guide presentation for 380CT Coursework 2)

Kamal Bentahar

December 6, 2016

Notation

Let $x_1, x_2, \ldots, x_n$ be Boolean **variables**, and let ϕ be a Boolean formula written in 3-cnf (Conjunctive Normal Form)

$$\phi = c_1 \wedge c_2 \wedge \cdots \wedge c_\ell,$$

where each **clause** $c_m = x_i \vee x_j \vee x_k$, for some $i, j, k = 1, 2, \ldots, n$ and $m = 1, \ldots, \ell$.

A **literal** can be x_i or $\neg x_i$ for some $i = 1, 2, \ldots, n$.

The ratio ℓ / n is important for experiments, and will be denoted by ρ.

Definition of the problem

Decisional 3SAT

Decide if ϕ is satisfiable.

Computational/Search 3SAT

If ϕ is satisfiable then find a satisfying assignment.

Optimization 3SAT (Max 3SAT)

Find an assignment that minimizes the number of non-satisfying clauses.

Sampling strategy

General 3SAT instances will be generated by selecting literals from

$$\{x_1, \neg x_1, x_2, \neg x_2, \ldots, x_n \neg x_n\}$$

uniformly at random.
For 'yes' instances, a random variable assignment is fixed first, then clauses are randomly constructed making sure each is satisfiable.

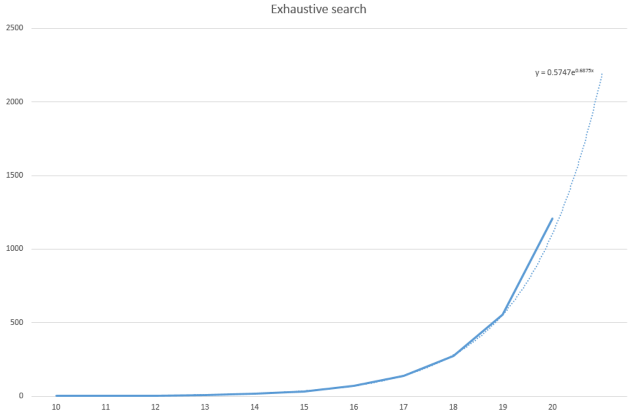
Exhaustive search – theory

for all possible variable assignments of $x_1, x_2, \ldots, x_n$ **do**
    if $\phi(x_1, x_2, \ldots, x_n)$ evaluates to True **then**
        return True
    end if
end for
return False

There are 2^n possible assignments, and each evaluation of ϕ costs $O(\ell)$. So this algorithm costs

$$O(\ell\, 2^n).$$

Exhaustive search – empirical results



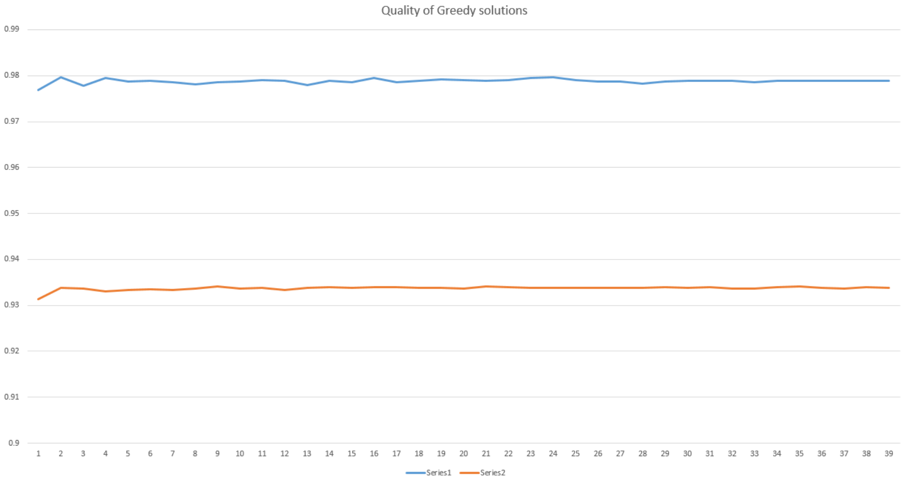
Average time in $100\times$ seconds [TODO: REDO EXPERIMENT] for randomly generated instances with $n = \ell$ for $n = 10, \ldots, 20$.
Dotted line: fitted exponential curve.

Greedy method

Find the variable that appears most often and assign it accordingly to maximize ...

$L \leftarrow \emptyset$
for $w \in \{x_1, \neg x_1, \ldots, x_n, \neg x_n\}$ **do**
    Count occurrences of w in ϕ
    Append pair $(w, \text{count of occurrences of } w \text{ in } \phi)$ to L
end for
Sort L with respect to the second component
for $(w, c) \in L$ **do**
    Set w to True $\triangleright$ If $w = \neg x_i$ then set x_i to False
end for
return count of satisfied clauses

Cost: $O(n \log n)$ assuming the use of an $O(n \log n)$ sorting algorithm.



Average ratio of clauses satisfied by Greedy. [TODO: WRONG X-AXIS VALUES] Blue when $\rho = 1$ giving about 98%, and orange when $\rho = 10$ dropping to about 93%.

References

- Hoos, H. and Stutzler, T. (2005) **Stochastic Local Search: Foundations and Applications.** Morgan Kaufmann
- Garey, S. and Johnson, D. (1979) **Computers and Intractability: A Guide to the Theory of NP-Completeness.** Freeman