EXERCISE FOR CSE202 – WEEK 8

Note: the exercise this week requires more computation than recent ones.

This exercise aims at getting a better understanding of the landscape where WalkSat is "walking", by looking more closely at the sum

$$\sum_{d=0}^{n} \binom{n}{d}(k-1)^{-d} \tag{1}$$

that gives the lower bound on the probability of success of the algorithm. For simplicity, we restrict the analysis to the case $k = 3$, but the general case would follow from the same approach.

The value $d = 0$ corresponds to a case when the random choice of assignment is exactly A and the success is immediate. There is only one such choice. As d increases, the probability of success for each random choice decreases fast, but as long as $d < n/2$, there are more and more assignments at distance d from A, given by the binomial coefficient in the sum. The question is to understand how these two conflicting phenomena combine.

Question 1. *Show that the assignments at distance $d \sim n/3$ are those that contribute most to the success of the algorithm. [Indication: consider the ratio between consecutive summands.]*

Solution. The ratio between consecutive summands is

$$r_d := \frac{\binom{n}{d+1}2^{-d-1}}{\binom{n}{d}2^{-d}} = \frac{1}{2}\left(\frac{n+1}{d+1} - 1\right).$$

As a function of d, this ratio decreases from $n/2$ to 0. Therefore the summands start from 1, first increase fast since the ratio is large initially, then reach a plateau when the ratio is close to 1, before they decrease down to 2^{-n} when $r_d < 1$. The maximum is reached when $r_d \sim 1$, ie,

$$d \sim \frac{n-2}{3} \sim \frac{n}{3}. \qquad \square$$

Question 2. *Show, as a preparation for the next question, that for any integer $n > 0$, positive real number $\beta < 1$, and $x \in (0, 1)$, one has*

$$\binom{n}{\lfloor \beta n \rfloor} < \left(\frac{1+x}{x^{\beta}}\right)^{n},$$

and that the choice $x = \beta/(1-\beta)$ minimizes that bound.

Solution. By the binomial theorem, for any $x > 0$, $(1 + x)^n$ is a sum of positive terms, one of which is $\binom{n}{\lfloor \beta n \rfloor}x^{\lfloor \beta n \rfloor}$. Therefore

$$\binom{n}{\lfloor \beta n \rfloor} < \frac{(1+x)^n}{x^{\lfloor \beta n \rfloor}} \le \frac{(1+x)^n}{x^{\beta n}},$$

where the last inequality follows from $x \in (0, 1]$ and $\lfloor \beta n \rfloor \le \beta n$.

This bound is optimized by choosing x that minimizes $f : x \mapsto (1+x)/x^\beta$. The logarithmic derivative is

$$\frac{1}{1+x} - \frac{\beta}{x},$$

showing that f' has only one 0, at $x = \beta/(1-\beta)$, where f is minimal. This value can be chosen for x as soon as $\beta \le 1/2$. Finally, we have obtained

$$\binom{n}{\lfloor \beta n \rfloor} < \left(\beta^\beta (1-\beta)^{1-\beta}\right)^{-n}, \qquad \beta \le 1/2. \qquad \square$$

Question 3. *Show that for any $\beta > 1/3$, the sum of the contributions to the sum* (1) *of all assignments at distance larger than βn is exponentially small compared to the sum* (1) *itself, as $n \to \infty$.*

Solution. From the computation in the answer to question 1, the summands in the range $k = \lfloor \beta n \rfloor, \dots, n$ decrease since $\lfloor \beta n \rfloor$ is larger than $n/3$ (this is true both asymptotically from Question 1, which is sufficient for this argument, or more generally by observing that $\beta > 1/3$ implies $\lfloor \beta n \rfloor \ge \beta n - 1 > (n-2)/3$, which is the maximum from Question 1). Also if $\beta \ge 1/2$, we can upper bound the sum by the sum with more terms obtained with a $\beta < 1/2$, so that it is sufficient to focus on that case. (The difficult part is when β is close to 1/3.) There are fewer than n summands, each upper bounded by the first one, so the whole sum is bounded by

$$n\binom{n}{\lfloor \beta n \rfloor} 2^{-\beta n} < n\left((2\beta)^\beta (1-\beta)^{1-\beta}\right)^{-n}.$$

The function $f : u \mapsto (2u)^u (1-u)^{1-u}$ is increasing for $u > 1/3$:

$$\frac{f'(u)}{f(u)} = \ln\left(\frac{2u}{1-u}\right)$$

is positive for $u > 1/3$ since $2u > 1-u$.

Thus from $f(1/3) = 2/3$, we obtain that for $\beta > 1/3$, this upper bound is exponentially smaller than the sum which is $(3/2)^n$. $\square$

In summary, the probability of success is concentrated on choices of initial assignments at distance less than $(1/3+\epsilon)n$ from a satisfying assignment for any $\epsilon > 0$, up to an exponentially small probability.