The "Naive randomized algorithm for $MAX3SAT$", Markov's inequality, and the method of conditional expectations.

The naive randomized method for $MAX3SAT$ or for any $MAXSAT$ problem is simply (as explained in class and in section 13.4 of the text) to choose a random truth assignment $\tau : \{x_1, \ldots, x_n\} \rightarrow \{TRUE, FALSE\}$ and evaluate $Z = Z(\tau) = |\{C_i | \tau(C_i) = TRUE\}|$. In the weighted case one evaluates $w(Z) = \sum_i Z_i(\tau)$ where $Z_i(\tau) = 1$ if $\tau(C_i) = TRUE$ and 0 otherwise. Z is a random variable depending on the random choice of τ. As shown in class and in the text, $E[Z] = \sum_i E[Z_i] \geq \frac{7}{8}k$ where k = number of clauses in the $MAX3SAT$ formula F. For the weighted case $E[w(Z)] = \sum_i w_i \cdot E[Z_i] \geq \frac{7}{8} \sum_i W \geq \frac{7}{8} OPT$ where $W = \sum_i w_i$. We will stay with the unwieghted case but everything that follows also applies to the weighted case.

Having shown that a random truth assignment is expected (i.e. averaging over all assignments) to be "good", what do we do next?

- We can think of each random assignment τ as one trial and say that a trial is successful if $Z(\tau) \geq \frac{7}{8}k$. We can then ask how many trials are needed before we obtain a successful trial. The text (section 13.4) shows that $p = Prob[Z(\tau) \geq \frac{7}{8}k] \geq \frac{1}{8k}$. Since the probability of *not* having a successful trial is $q = 1 - p \leq 1 - \frac{1}{8k}$, if we perfrom t tirals then the probability of not obtaining a successful trial is $\leq q^t$. For the random variable T = number of trials before obtaining a successful trial, the analsyis in section 13.3 shows that $E[T] = 1/p$. Hence the expected number of trials is $\leq 8k$. As each trial takes $O(k)$, the expected time to obtain a good τ is $O(k^2)$.

- What can we achieve with expected time $O(k)$? Lets lower our standards and say (for example) that we just want to find a τ such that $Z(\tau) > \frac{3}{4}k$? Let $k - Z(\tau)$ be denoted by the random variable Y. We want to estimate $p = Prob[Z > \frac{3}{4}k]$. Equivalently we will estimate $q = 1 - p = Prob[Y \geq \frac{1}{4}k]$. Note that $E[Y] = \frac{1}{8}k$. Since $Y \geq 0$ we can use Markov's inequality which states that $Prob[Y \geq t] \leq \frac{E[Y]}{t}$. Setting $t = \frac{1}{4}k$ we obtain $q \leq \frac{(1/8)k}{(1/4)k}$ and hence $p \geq \frac{1}{2}$. So using the analysis of section 13.3 again, we obtain $E[T] = 2$; i.e. the expected number of trials to obtain a τ with $Z(\tau) \geq \frac{3}{4}k$ is 2. Hence the expected time to obtain a "good" solution is $O(k)$.

- Suppose we want an absolute guarantee on time and performance. Then we have to eliminate the randomization. It turns out that this randomized algorithm can be "de-randomized" by the method of conditional expedtations. Let F_1 (respectively, F_0) be the formula F (over propositional variable $x_2, \ldots, x_n$) when the variable x_1 is set to TRUE (respectively FALSE). Let Z_1 (respectively, Z_0) be the random variable corresponding to the number of satisfied clauses in F_1 (respectively, F_0) by a random truth assignment (over the variables $x_2, \ldots, x_n$).

  Then $E[Z] = E[Z_0] \cdot \frac{1}{2} + E[Z_1] \cdot \frac{1}{2}$. So either $E[Z_0] \geq \frac{7}{8}k$ or $E[Z_1] \geq \frac{7}{8}k$ (or both). The important observation is that all of these expectations can be easily calculated (in time $O(k)$) so we choose F_0 or F_1 whichever has the best expectation. Say that

Z_0 has the best expectation. Now having set $x_1 = FALSE$, we continue with F_0 and decide how to set x_2, etc. For our notation, we let (say) $F_{0,1}$ be formula F having set $x_1 = FALSE$ and $x_2 = TRUE$.

Here we will apply the method of conditional expectations to the example given in the lecture.

Let $F = (x_1 \vee x_2 \vee \overline{x_3}) \wedge (x_1 \vee \overline{x_2} \vee x_3) \wedge (\overline{x_1} \vee x_2 \vee x_3) \wedge (\overline{x_1} \vee \overline{x_2} \vee x_3) \wedge (\overline{x_1} \vee x_2 \vee \overline{x_3}) \wedge (x_1 \vee \overline{x_2} \vee \overline{x_3})$

We know that $E[F] \geq \frac{7}{8} \cdot 6$ and hence there is a truth assignment τ satisfying $\lceil \frac{7}{8} \cdot 6 \rceil = \lceil \frac{42}{8} \rceil = 6$ clauses; that is F is staisfiable.

We (arbirarily) consider $\tau(x_1) = TRUE$ and compute $E[F_1]$. We have $F_1 = TRUE \wedge TRUE \wedge (x_2 \vee x_3) \wedge (\overline{x_2} \vee x_3) \wedge (x_2 \vee \overline{x_3}) \wedge TRUE$. Hence $E[Z_1] = 1 + 1 + \frac{3}{4} + \frac{3}{4} + \frac{3}{4} + 1 = 3 + 9/4 = 42/8 = E[Z]$ so that we can take $x_1 = TRUE$. (We could also set $x_1 = FALSE$ since Z_1 and Z_0 have the same expectation.) Suppose we (again arbitrarily) try $\tau(x_2) = FALSE$. Then after substituting $x_2 = FALSE$ in F_1, we get the formula $F_{1,0} = TRUE \wedge TRUE \wedge (x_3) \wedge TRUE \wedge (\overline{x_3} \vee x_3) \wedge TRUE$ and $E[Z_{1,0}] = 1 + 1 + \frac{1}{2} + 1 + \frac{1}{2} + 1 = 5 < 42/8$. Hence we should take $\tau(x_2) = TRUE$. So we construct $F_{1,1} = TRUE \wedge TRUE \wedge TRUE \wedge (x_3) \wedge TRUE \wedge TRUE$ and $E[Z_{1,1}] = 5 + \frac{1}{2} = 44/8 > 42/8$ so that $\tau(x_1) = \tau(x_2) = \tau(x_3) = TRUE$ is a satisfying assignment. Note: Had we started with $\tau(x_1) = FALSE$ with $E[Z_0] = 42/8$ then we would have been led to the satisfying assignment $\tau(x_1) = \tau(x_2) = \tau(x_3) = FALSE$.