Discrete Structures
CS 280

Example application of probability:

MAX 3-SAT

MAX 3-SAT

Consider a propositional logical formula on N Boolean variables in conjunctive normal form (CNF), i.e., a conjunction (logical AND) of disjunctions (logical OR).

Example: $$(x_1 \vee \neg x_2 \vee x_3) \wedge (x_2 \vee x_3) \wedge (\neg x_3)$$

The truth assignment with x_1 and x_2 assigned to **True** and x_3 assigned to **False** satisfies this formula. Each disjunction is also referred to as a "clause".

If each clause contains exactly k variables, the formula is a *k-CNF formula*.

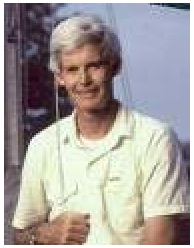
MAX 3-SAT cont.

Problem: **MAX-3-SAT**

Given a 3-CNF formula F, find a truth assignment that satisfies as many clauses as possible.

The MAX 3-SAT problem is a so-called NP-hard problem; it is generally believed that no efficient (i.e., polynomial time) algorithm exists for solving such problems.

[The $1M Clay Millennium prize, click on **P=/=NP**]



Stephen Cook

Leonid Levin

Note that we have a search space of 2^N truth assignments.

So, finding a maximally satisfying assignment is (most likely) computationally very hard.

However, it's surprisingly easy to find a reasonable good assignment, satisfying 7/8th (87.5%) of the clauses in expectation. **How??**

Thm. Given a 3-CNF formula with k clauses, the expected number of clauses satisfied by a random assignment is $\frac{7}{8}k$.

Proof. (by linearity of expectation)

A random assignment is obtained by setting each variable $x_1,\ldots, x_n$ independently to True or False with probability ½ each.

Let Z denote the r.v. equal to the number of satisfied clauses. Z can be written as a sum of random indicator variables Z_i , one for each clause.

I.e.

$$Z = Z_1 + Z_2 + \ldots + Z_k$$

with $Z_i = 1$ if the i^{th} clause is satisfied, and 0 otherwise.

Now, we have by linearity of expectation

$$E[Z] = E[Z_1] + E[Z_2] + \ldots + E[Z_k]$$

(Remember this holds *no matter how* the random variables Z_i are correlated!)

What is $E[Z_i]$?

The probability that a clause is *not* satisfied is $(1/2)^3 = 1/8$. So, the probability that a clause is satisfied by the random assignment is 1 – 1/8 = 7/8. So, $E[Z_i] = 7/8$

And, therefore: $$E[Z] = \frac{7}{8}k$$

QED

So, we can actually find a pretty good assignment, in expectation, very easily, even though it's believed intractable to find the maximally satisfying assignment. We can obtain yet another surprise from our analysis.

Note that a random variable has to assume a value at least as large as its expectation at some point in the sample space. This observation immediately leads us to the following result.

Thm. Given a 3-CNF formula, there must exist a truth assignment that satisfies at least a $7/8^{th}$ fraction of the clauses.

So, from the analysis of a random event (a randomly sampled truth assignment), we have now obtained a statement that does not involve any randomness or probability!

The technique we used to prove this result is more generally referred to as the "probabilistic method", which can be used to show the existence of certain combinatorial objects (in this case, a truth assignment satisfying $7/8^{th}$ of the clauses) by showing that a random construction produces the desired object with non-zero probability.

The probabilistic method (link) is a non-constructive proof technique. The method was pioneered by the famous mathematician Paul Erdos (link).

End note: We showed that a randomly generated assignment satisfies $7/8^{th}$ of the clauses, *in expectation*. Hmm… How often do we have to guess to be sure to have an assignment satisfying $7/8^{th}$ of the clauses? It can be shown that the expected number of guesses grows only polynomially in *N*, the number of Boolean variables.