Problem Set 7
Due: 11:59pm, Sunday, April 21st

See homework submission instructions at
http://sites.fas.harvard.edu/~cs124/cs124/problem_sets.html

Problem 4 is worth 40% of this problem set, and problems 1-3 constitute the
remaining 60%.

For each problem where you are asked to give an algorithm, more points are given for
asymptotically faster algorithms. In judging the number of points to award a correct solution,
we only consider the running time in asymptotic notation, i.e. a writeup of an algorithm
taking \(1000n^2\) steps versus one taking \(.01n^2\) steps would receive the same score — both would
be simply treated as \(\Theta(n^2)\)-time solutions. For some problems we may also ask for memory
consumption, and using asymptotically less memory similarly is awarded more points.

Problem 1

In class we mentioned that for a universe \([U] := \{1, \ldots, U\}\) and positive integers \(m, p\) with
\(p \geq \max\{m, U\}\) a prime, the following is an almost universal family of functions mapping
\([U]\) to \(\{0, \ldots, m - 1\}\):

\[
H = \{h_a : h_a(x) = ((a \cdot x) \mod p) \mod m, \ 1 \leq a \leq p - 1\}. \tag{1}
\]

One issue here is that this universal family assumes inputs come from a finite universe,
which is not true if one wants a hash function that applies to some universe \(\Sigma^*\) (the set
of all finite arbitrary-length strings with characters in the alphabet \(\Sigma\)). As many of you
have probably seen though in your favorite programming languages, dictionaries do tend to
support keys of type \texttt{string}. In this problem you’ll show how, by proving that a certain
family mapping strings to \(\{0, \ldots, m - 1\}\) is, though not quite universal, still useful.

Let \(p\) be a prime with \(p > |\Sigma|\). For a fixed \(0 \leq b \leq p - 1\) define the hash function
\(h_b : \Sigma^* \to \{0, \ldots, p - 1\}\) which operates the following way on a string \(s = s_0s_1 \cdots s_k\):

\[
h_b(s) = \left(\sum_{i=0}^{k} s_i b^i\right) \mod p.
\]

In other words, we treat the characters in \(s\) as coefficients of a polynomial, then \(s\) hashes to
the evaluation of that polynomial on \(b\) taken mod \(p\). Now define

\[
H_{\text{polyhash}} = \{h_b : 0 \leq b \leq p - 1\}.
\]
(a) (8 points) Show that if \( s \neq s' \) are two strings in \( \Sigma^\ast \) each of length at most \( \ell \), then

\[
Pr_{h \in \mathcal{H}_{\text{polyhash}}} (h(s) = h(s')) < \frac{\ell}{p}.
\]

You may use the following fact without proof: a degree-\( d \) polynomial \( q(x) \) with coefficients in \( \{0, 1, \ldots, p-1\} \) has at most \( d \) roots modulo \( p \). That is, there are at most \( d \) values \( z \in \{0, 1, \ldots, p-1\} \) satisfying \( q(z) = 0 \mod p \). (If you want to learn why, take Math 122 and/or 123.)

(b) (4 points) How many bits of memory are needed to specify an \( h \in \mathcal{H}_{\text{polyhash}} \)? Justify.

(c) (8 points) Suppose you want to map strings into \( \{0, 1, \ldots, m\} \) where \( m \) is an arbitrary positive integer that may not be prime. Describe how you would use the results of (a) to devise such a scheme, which is \( 2 \)-almost universal when restricted to the domain of all strings of length at most \( \ell \) over the alphabet \( \Sigma \). You should try to optimize the logarithm of the size of your hash family (which corresponds to the bits needed to store an \( h \) drawn from that family). Note that one could also treat any such string as an element of \( [U] \) for \( U = |(\Sigma \cup \{\bot\})^\ell| = (|\Sigma| + 1)^\ell \) (where \( \bot \) is used to pad strings of length less than \( \ell \)) and use the almost-universal family from class (Eq. (1)). Explain why this would lead to a solution that is worse, in terms of bits needed to represent \( h \).

Problem 2

In class we analyzed a toy model in which vertices \( 0, 1, \ldots, n+1 \) are connected in a path, we start at some vertex \( i \), and in every time step we move to a uniformly chosen random neighbor of our current location. We gave an exact expression for the expected number of steps to reach 0, starting at \( i \), for an \( 0 \leq i \leq n \). We then showed via a coupling argument that the random walk 2SAT algorithm finds a satisfying assignment of a satisfiable formula in at most \( O(n^2) \) time steps.

Use coupling to show the 3SAT random walk algorithm also finds a satisfying assignment of a satisfiable formula in a number of steps at most that of the corresponding toy model.

Problem 3

Complete the reading at the end of this problem set for the Embedded EthiCS module, then answer the questions at the end.

Programming Problem


Hint: Consider using \( \mathcal{H}_{\text{polyhash}} \) from Problem 1. Also: for a long string \( c \), how long does it take to compute \( h_b(c_2c_3\cdots c_{r+1}) \) if we already know \( h_b(c_1c_2\cdots c_r) \)?
PLEASE READ: In your upcoming Embedded EthiCS module for CS124 you will be thinking about the connections between fairness and flow. Don’t worry if that doesn’t make sense yet! For now, let me preface the module by saying: there is a way of using flow problems as resource allocation models. Whenever we think about resource allocation, questions of fairness can be raised. We will be considering some of those questions.

These reading assignments are intended to help provide some background information on concepts pertaining to fairness and justice more generally. Their purpose is to get you thinking about some ethical concepts that will be touched on in the module. Excerpt (1) paraphrases a crucial part of John Rawls concept of the original position and the associated veil of ignorance, whilst excerpt (2) provides some primary text from Rawls A Theory of Justice. The idea is that (1) will provide some context for (2). Excerpt (3) moves from these very abstract concepts/scenarios to some concrete definitions of two kinds of justice (or fairness): procedural vs. substantive.

Once you have read the excerpts carefully, please prepare answers to the questions assigned below. Aim to spend 30 minutes carefully reading the excerpts and 30 minutes on your answers.

EXCERPTS:


“...The original position is a central feature of John Rawls’s social contract account of justice, “justice as fairness,” set forth in A Theory of Justice (TJ). It is designed to be a fair and impartial point of view that is to be adopted in our reasoning about fundamental principles of justice. In taking up this point of view, we are to imagine ourselves in the position of free and equal persons who jointly agree upon and commit themselves to principles of social and political justice. The main distinguishing feature of the original position is “the veil of ignorance”: to insure impartiality of judgment, the parties are deprived of all knowledge of their personal characteristics and social and historical circumstances. They do know of certain fundamental interests they all have, plus general facts about psychology, economics, biology, and other social and natural sciences. The parties in the original position are presented with a list of the main conceptions of justice drawn from the tradition of social and political philosophy, and are assigned the task of choosing from among these alternatives the conception of justice that best advances their interests in establishing conditions that enable them to effectively pursue their final ends and fundamental interests. Rawls contends that the most rational choice for the parties in the original position are two principles of justice: The first guarantees the equal basic rights and liberties needed to secure the fundamental
interests of free and equal citizens and to pursue a wide range of conceptions of the good. The second principle provides fair equality of educational and employment opportunities enabling all to fairly compete for powers and positions of office; and it secures for all a guaranteed minimum of all-purpose means (including income and wealth) individuals need to pursue their interests and to maintain their self-respect as free and equal persons...

3. The Veil of Ignorance

Rawls calls his conception “justice as fairness.” His aim in designing the original position is to describe an agreement situation that is fair among all the parties to the hypothetical social contract. He assumes that if the parties to the social contract are fairly situated and take all relevant information into account, then the principles they would agree to are also fair. The fairness of the original agreement situation transfers to the principles everyone agrees to, and further that whatever laws or institutions are required by the principles of justice are also fair. The principles of justice chosen in the original position are in this way the result of a choice procedure designed to “incorporate pure procedural justice at the highest level” (CP, 310, cf. TJ 120/104).”

(2) Excerpts from John Rawls — A Theory of Justice

“24. THE VEIL OF IGNORANCE

The idea of the original position is to set up a fair procedure so that any principles agreed to will be just. The aim is to use the notion of pure procedural justice as a basis of theory. Somehow we must nullify the effects of specific contingencies which put men at odds and tempt them to exploit social and natural circumstances to their own advantage. Now in order to do this I assume that the parties are situated behind a veil of ignorance. They do not know how the various alternatives will affect their own particular case and they are obliged to evaluate principles solely on the basis of general considerations.

It is assumed, then, that the parties do not know certain kinds of particular facts. First of all, no one knows his place in society, his class position or social status; nor does he know his fortune in the distribution of natural assets and abilities, his intelligence and strength, and the like. Nor, again, does anyone know his conception of the good, the particulars of his rational plan of life, or even the special features of his psychology such as his aversion to risk or liability to optimism or pessimism. More than this, I assume that the parties do not know the particular circumstances of their own society. That is, they do not know its economic or political situation, or the level of civilization and culture it has been able to achieve. The persons in the original position have no information as to which generation they belong.” (p118)

(3) Excerpts from the Stanford Encyclopedia of Philosophy entry on ‘Justice’, in particular on ‘substantive’ vs. ‘procedural’ justice. NOTE — this is sometimes referred to as ‘outcome’ vs. ‘procedural’ fairness.

“2.3 Procedural versus Substantive Justice

A third distinction that must be drawn is between the justice of the procedures that might
be used to determine how benefits and burdens of various kinds are allocated to people, and
the justice of the final allocation itself. It might initially seem as though the justice of a
procedure can be reduced to the justice of the results produced by applying it, but this is not
so. For one thing, there are cases in which the idea of an independently just outcome makes
no sense. A coin toss is a fair way of deciding who starts a game, but neither the Blues nor
the Reds have a claim of justice to bat first or kick off. But even where a procedure has been
shaped by a concern that it should produce substantively just outcomes, it may still have
special properties that make it intrinsically just. In that case, using a different procedure
to produce the same result might be objectionable. In an influential discussion, John Rawls
contrasted perfect procedural justice, where a procedure is such that if it is followed a just
outcome is guaranteed (requiring the person who cuts a cake to take the last slice himself
is the illustration Rawls provides), imperfect procedural justice, where the procedure is such
that following it is likely, but not certain, to produce the just result, and pure procedural
justice, such as the coin-tossing example, where there is no independent way to assess the
outcome if we call it just, it is only on the grounds that it has come about by following the
relevant procedure (Rawls 1971, 1999, 14).”

QUESTIONS:

1. (3 points) What is the relationship between The Original Position and The Veil of
   Ignorance? (1 sentence)

2. (3 points) Why does Rawls think The Veil of Ignorance is necessary? (hint: think
   about what the goal of the VOI is) (1-2 sentences)

3. (2 points) Would someone behind The Veil of Ignorance know their race? (yes/no)

4. For the following scenarios, decide whether they would be examples of perfect proce-
dural justice, imperfect procedural justice or pure procedural justice and explain your
answer (2-3 sentences per scenario)

   (a) (3 points) Students are placed into schools on the basis of a random lottery

   (b) (3 points) Loan applicants are given loans on the basis of their credit score

   (c) (3 points) Jail sentences are delivered using a very detailed and genuinely binding-
in-all-cases system such that equal crimes serve equal time regardless of who
   perpetrates the crime

   (d) (3 points) A couple uses a ‘you split, I’ll pick’ policy for distributing naan bread
   between them when they get takeout