CSB 439 Lecture on Chs 9 & 10 ①

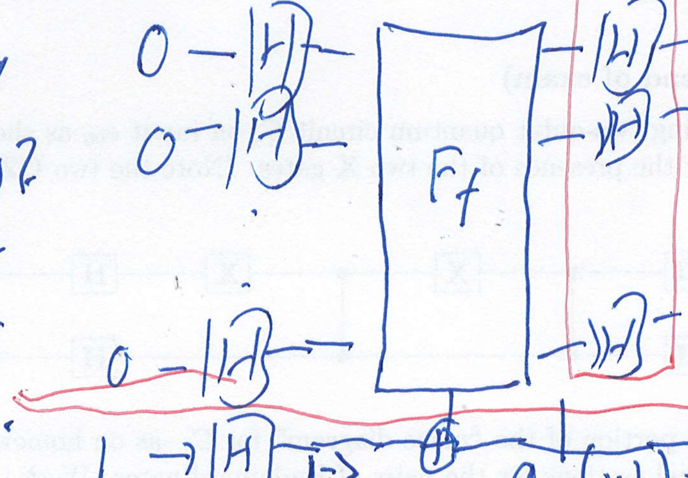
Deutsch Jozsa

$f = \{0,1\}^n \to \{0,1\}$

promise: either

f is constant or

f is balanced.



$$F_f(xb) = \begin{bmatrix} X \\ b \oplus f(x) \end{bmatrix} \qquad a_0 = |0^n 1\rangle \qquad a_1 = |+\rangle^n |-\rangle$$

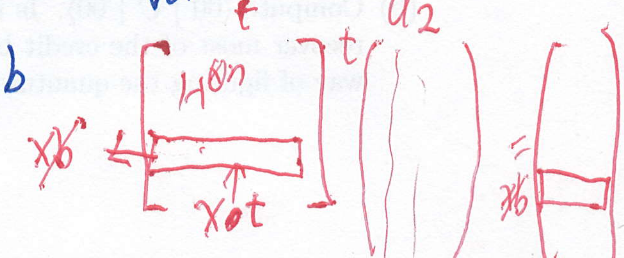
$$a_1(xb) = \frac{1}{\sqrt{2N}} (-1)^b$$

$$a_2(xb) = \frac{1}{\sqrt{2N}} (-1)^{b \oplus f(x)}$$

$$a_3(xb) = \boxed{\frac{1}{\sqrt{N}} \sum_t (-1)^{x \cdot t} a_2(tb)}$$

$$= \frac{1}{N\sqrt{2}} \sum_t (-1)^{x \cdot t} (-1)^{f(t) \oplus b}$$

continued see below



$$= \frac{1}{N\sqrt{2}} \sum_t (-1)^{x \cdot t} (-1)^b (-1)^{f(t)}$$

If f is constant $= c$, then this equals

$$a_3(xb) = \frac{1}{N\sqrt{2}} \sum_t (-1)^{x \cdot t} (-1)^b (-1)^c = \frac{1}{\sqrt{2}} \frac{(-1)^{b+c}}{N} \sum (-1)^{x \cdot t}$$

Analysis is about measuring on on the first n qubits - i.e. getting $0^n 0$ or $0^n 1$

$$a_3(0^n b) = \frac{1}{\sqrt{2}} \frac{(-1)^{b+c}}{N} \sum_{t \in \{0,1\}^n} (+1) = \frac{1}{\sqrt{2}} (-1)^{b+c}$$

$\therefore$ Half the probability is on $0^n 0$, the other half on $0^n 1$

$\therefore$ With certainty, we will get 0^n on the first n qubits

To make the distinction we need that if f is balanced, then we never get 0^n on the first n qubits

Generally, $a_3 \cancel{(0^n b) = \frac{1}{\sqrt{2}} (-1)^b (-1)^{f(0^n)}}$

$$a_3(xb) = \frac{1}{N\sqrt{2}} (-1)^b \sum_t (-1)^{x \cdot t} (-1)^{f(t)}$$

$$\therefore a_3(0^n b) = \frac{1}{N\sqrt{2}} (-1)^b \sum_t (+1) \cdot (-1)^{f(t)}.$$

This finishes the proof.

If f is balanced, the sum cancels, leaving zero amplitude on $a_3(0^n b)$ ($b = 0$ or $b = 1$)

No classical algorithm able to query only $f(x)$ is able to get certainty. But you can get "w high prob."

Chapter 10: Simon's Algorithm $f(x) = y$

Now we are given $f: \{0,1\}^n \to \{0,1\}^n$ with the promise that there is a "hidden vector" $s \in \{0,1\}^n$ s.t. $\forall y, z \in \{0,1\}^n$, $\boxed{f(y) = f(z) \iff y = s \oplus z}$.

When $s = 0^n$, this says f is 1-1, else f is 2-1 in this special way.

Goal: Given ability to query $f(y)$ classically, $F(y,w) = [y, w \oplus f(y)]$ in quantum, compute s.

In particular, this distinguishes the case $s = 0^n$ from the case f is 2-1 with s.

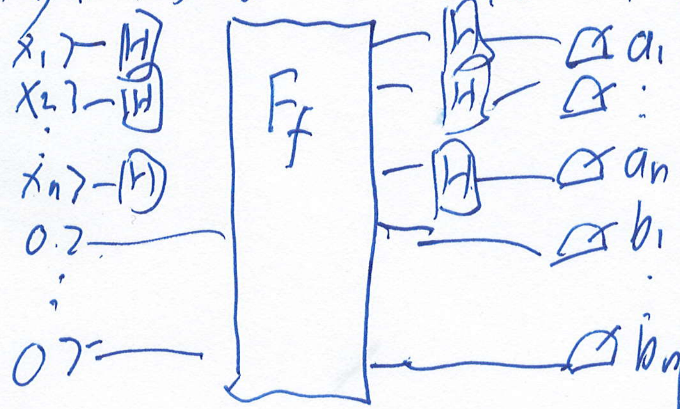
Simon's Theorems:

1. We can build a classical algorithm with a quantum sampling subroutine that computes s w.h.p. in $n^{O(1)}$ time.

Proof really long; skipped

2. No classical algorithm able to query $f(y) \to$ is able to distinguish $s = 0^n$ from $s \neq 0^n$ in $2^{o(n)}$ time w.h.p.

Begin Proof of (1): The quantum circuit is C W



Classical part

$E = \emptyset$;
while (rank(E) $< n$) {
  Sample $C \to \begin{bmatrix} \vec{a} \\ \vec{b} \end{bmatrix}$
  if $a \notin$ span(E) {
    $E = E \cup \{a\}$
  }
}

Inner Lemma: a always gives $a \bullet s = 0$

Thus when rank(E) $= n$ we can solve for s

Outer Lemma: Given that the a we measure is uniformly random s.t. $a \bullet s = 0$, with prob. at least 1/2 on each sample, $b \notin \langle$span $E\rangle$. So we make progress.

Thursday's lecture went to here.

Proof of Inner Lemma: The state after F_f is $\frac{1}{\sqrt{N}} \sum_x |x\rangle|f(x)\rangle$. In "function-vector" form:

$u(xy) = \begin{cases} 1/\sqrt{N} & \text{if } y = f(x) \\ 0 & \text{otherwise} \end{cases}$ The state v after the second Hadamard transform on the "x space" is:

$v(xy) = \sum_{t \in \{0,1\}^n} (-1)^{x \bullet t} u(ty) = \frac{1}{N} \sum_t \begin{cases} (-1)^{x \bullet t} & \text{if } y = f(t) \\ 0 & \text{otherwise} \end{cases}$

(with constant $\frac{1}{\sqrt{N}}$ again, $N = 2^n$)

Immediately we can deduce that the measurement outcome xy has nonzero amplitude only if $y \in R = \text{Ran}(f)$.

Continuing the proof of Simon's Algorithm: We measure the state

$$\mathbf{V}(xy) = \frac{1}{N} \sum_{t \in \{0,1\}^n} \begin{cases} (-1)^{x \bullet t} & \text{if } y = f(t) \\ 0 & \text{otherwise} \end{cases}$$

where $N = 2^n$ and we multiplied $\frac{1}{\sqrt{N}} \cdot \frac{1}{\sqrt{N}} = \frac{1}{N}$.

Now suppose we measure and get the result $\begin{bmatrix} a \\ b \end{bmatrix}$, $a \in \{0,1\}^n$, $b \in \{0,1\}^n$.

We know $b \in Ran(f)$. If f is 1-to-1, then $s = 0^n$ and automatically $a \bullet s = 0$. Moreover, the amplitude and hence probability for getting b will be the same: there is exactly one t such that $f(t) = b$, so the sum over t has only the single nonzero term $(-1)^{a \bullet t}$. Thus the b given as a sampling of the quantum circuit C will be a uniformly random string in the space of $b \in \{0,1\}^n$ such that $b \bullet s = 0$ which is all of $\{0,1\}^n$ in this case. Since f is 1-to-1, the a we get will be uniformly random as well.

If f is 2-to-1, then there are two distinct t_1 and t_2 such that $f(t_1) = b$ and $f(t_2) = b$, and $t_2 = t_1 \oplus s$. These will be the only two terms of the sum for $\mathbf{V}(ab)$ that can give a nonzero result. So we get

$$V(ab) = \frac{1}{N}\left((-1)^{a \bullet t_1} + (-1)^{a \bullet t_2}\right) = \frac{1}{N}\left((-1)^{a \bullet t_1} + (-1)^{a \bullet (t_1 \oplus s)}\right)$$

The $\bullet$ denotes the "Boolean dot product of binary strings modulo 2," but we can still use the distributive law over $\oplus$, which is addition modulo 2. So we get:

$$= \frac{1}{N}\left((-1)^{a \bullet t_1} + (-1)^{a \bullet t_1}(-1)^{a \bullet s}\right)$$

Now if $a \bullet s = 1$, then the second term is the negative of the first term. So they cancel, so $\mathbf{V}(ab) = 0$. This means ab has zero amplitude — so this means: we could not have gotten $\begin{bmatrix} a \\ b \end{bmatrix}$ as a result of the measurement.

$$\boxed{\text{The only } \begin{bmatrix} a \\ b \end{bmatrix} \text{ we can get from the measurement are cases where } a \bullet s = 0.}$$

Moreover the ~~amplitude~~ magnitude has the same for any a: it is $\frac{1}{N}\left(2 \cdot (-1)^{a \bullet t_1}\right)$ which is $\frac{\pm 2}{N}$ depending on whether $a \bullet t_1 = 0$ or 1. So the probability is $\frac{4}{N^2}$ for any a. Thus a is uniformly at random from the subspace of a such that $a \bullet s = 0$. ☒

Technote: We do not get $f(a) = b$, only that $b \in Ran(f)$. There are $\frac{1}{2}N$ such b and $\frac{1}{2}N$ a's such that $a \bullet s = 0$. So we get $\frac{1}{4}N^2$ possible outputs, each equally likely. So the probabilities do sum to 1.

Proof of the Outer Lemma: First suppose f is 1-to-1, i.e., $s = 0^n$. Let $a_1, \ldots, a_m$ be the n-vectors sampled thus far. They are members of the vector space $\mathbb{Z}_2^n$ with addition modulo 2. Let $A = \langle a_1, \ldots, a_m \rangle$ be the subspace spanned by the sampled vectors, and let $r = \dim(A)$. We can calculate r as the *rank* of the $m \times n$ matrix with $a_1, \ldots, a_m$ as its rows. Then $r \leq n$.

- If $r = n$, then we *know* that f is 1-to-1, and can say so.
- If $r < n$, then A is not the entire space. Since it is a linear subspace, its cardinality is at most half of the space's $|\mathbb{Z}_2^n| = 2^n$. Since C gives a *random member of the whole space* in this case, with probability at least $1/2$ ~~the~~ the next sampled vector a_{m+1} is not in A. Then the new r goes up by 1, so we *make progress*.

It follows that the *expected* number of iterations to get $r = n$ is $\leq$ *$2n$*. (Initially the chance of making progress is much better than $1/2$. It only equals $\frac{1}{2}$ on the last step when $r = n-1$.) *Furthermore*, if you do *$4n$* iterations and still don't have $r = n$, then you can "give up" and conclude *whp* that f is not 1-to-1.

If f is 2-to-1 (with *period* s), then $S = \{x : x \bullet s = 0\}$ is a subspace of dimension $n-1$. Since $a_1, \ldots, a_m, \ldots$ always belong to S, it follows that we will *never* get $r = n$. Thus when we "give up" after $4n$ iterations, we will be correct in this case. Moreover, by reasoning similar to the first part, (if $r < n-1$) we always have at least a $1/2$ chance of incrementing r. This is because when $r = \dim(A)$ is $< n-1$, then A is at most half the subspace S. Since the inner lemma gives a *uniformly random* $a_{m+1} \in S$ on the next iteration, it has at least a $1/2$ chance of giving $a_{m+1} \in S \setminus A$, which means r goes up by *1*. Hence with high probability we get $r = n-1$, and the $n-1$ equations *allow us to calculate* s. ☒