Lecturer: Tongyang, scribed by Shuo Zhou, Rui Yang
Date: October 27, 2023

Lecture 9

Unstructured Search

- Grover's algorithm
- Amplitude amplification

1 Recap & Preview

Deutsch-Jozsa $f : \{0,1\}^n \longmapsto \{0,1\}$. constant or balanced.
$\left.\begin{array}{c} 0\cdots 0 \\ \vdots \\ 1-1 \end{array}\right\} 2^n = N$ $s : 2^n$ binary string.

Simon's problem $f : \{0,1\}^n \longmapsto X$. $f(x) = f(y)$ iff $x = y$ or $x = y \otimes s$.
Less requirement than Deutsch-Jozsa but still structured.

Phase estimation $U\|\psi\rangle = e^{i\theta}\|\psi\rangle$. Find θ.
A genuinely quantum problem. $O(1/\epsilon)$ queries, w.p. $\geq \frac{8}{\pi^2}$. Going beyond Hadamard. QFT from $Z_2 \otimes \cdots \otimes Z_2$ to Z_{2^n}.
Application: Order finding $\Rightarrow$ Shor's algorithm.

For all of these, we have structural assumptions:

- Deutsch-Jozsa and Simon's problem: Special f.
- Order finding and Shor's algorithm: Cyclic group $\mathbb{Z}_N$.

How about we characterize "very general functions"?
Start with boolean functions.

Total function: A function $f : \{0,1\}^N \to \{0,1\}$ which has definition on all 2^N inputs.
Deutsch-Jozsa: $N = 2^n$, But the definition domain is only $\{s = s_1 \cdots s_{2^n} \mid \#$ of 1 in s_i is $0, 2^{n-1}$, or $2^n\}$.
If a function is defined on a proper subset of $\{0,1\}^n$. it's called a partial function.

A very typical problem is to compute the OR function (AND is symmetric).

$$f : \{0,1\}^n \to \{0,1\}, N = 2^n. \quad OR(s_1, \ldots, s_N) = s_1 \cup s_2 \ldots \vee s_N = \begin{cases} 0 & \text{if } s_1, \ldots s_n = 0, \\ 1 & \text{otherwise.} \end{cases}$$

Classically: Need $\Theta(n)$ queries to compute the value of OR.
Quantum: $|i, z\rangle \xrightarrow{U_f} |i, z \oplus s_i\rangle$, $\forall i \in [n], z \in \{0,1\}$ $(*)$.

Equivalent to having an oracle:

$$f(x) = s_x, \forall x \in [N] \ (f : [N] \to \{0,1\}) \qquad |x, z\rangle \xmapsto{U_f} |x, f(x) \otimes z\rangle, \quad \forall x \in [N], z \in \{0,1\}.$$

Main focus today: $O(\sqrt{N})$ quantum queries suffice.

2 Grover's algorithm

Phase kick-back: $|x\rangle|-\rangle \xrightarrow{U_f} (-1)^{f(x)}|x\rangle|-\rangle$.

"phase query": $|x\rangle \longmapsto (-1)^{f(x)}|x\rangle \qquad |\psi\rangle = \frac{1}{\sqrt{N}}\sum_{x=1}^{N}|x\rangle \xmapsto{\text{U}} \frac{1}{\sqrt{N}}\left(\sum_{x,f(x)=0}|x\rangle - \sum_{x\cdot f(x)=1}|x\rangle\right)$

How can we find x such that $f(x) = 1$?

This looks like a reflection.

For simplicity, consider a unique marked item w s.t. $f(w) = 1, \quad f(x) = 0 \ \ \forall x \neq w$.

$$\left.\begin{array}{l} U|w\rangle = -|w\rangle \\ U|x\rangle = |x\rangle \quad \forall x \neq w. \end{array}\right\} U = I - 2|w\rangle\langle w|.$$

$$\begin{aligned}
&(I - 2|w\rangle\langle w|)|w\rangle = |w\rangle - 2|w\rangle\langle w|w\rangle = -|w\rangle. \\
&(I - 2(w\rangle\langle w|)|x\rangle = |x\rangle - 2|w\rangle\langle w|x\rangle = |x\rangle \\
&(I - 2|w\rangle\langle w|)(I - 2|w\rangle\langle w|) = I - 2|w\rangle\langle w| - 2|w\rangle\langle w| + 4|w\rangle\langle w|w\rangle\langle w| \\
&= I.
\end{aligned}$$

We also consider $|\psi\rangle = \frac{1}{\sqrt{N}}\sum_{x=1}^{N}|x\rangle$.

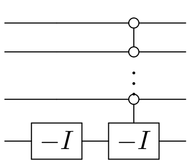
We consider the unitary $V = 2|\psi\rangle\langle\psi| - I$. V is independent of the queries, and V can be efficiently implemented with cost $O(\log N)$.

Say $N = 2^n$. Otherwise find the smallest power of 2 larger than N ($2^n < N < 2^{n+1}$), and set $f(N+1)\ldots, f\left(2^{n+1}\right) = 0$.

$$\frac{1}{\sqrt{2^n}}\sum_{x=1}^{2^n}|x\rangle = H^{\otimes n}|0\rangle \quad V = H^{\otimes n}R_0H^{\otimes n}, \text{ where } R_0 = 2\left|0^n\right\rangle\left\langle 0^n\right| - I.$$

$$H^{\otimes n}\left(2\left|0^n\right\rangle\left\langle 0^n\right| - I\right)H^{\otimes n}$$

$R_0 : \left|0^n\right\rangle \to \left|0^n\right\rangle, |x\rangle \to -|x\rangle \quad \forall x \in \{0,1\}^n / \left\{0^n\right\}$.



$$\begin{aligned}
&\left|0^n\right\rangle|0\rangle \longmapsto -\left|0^n\right\rangle|0\rangle \longmapsto \left|0^m\right\rangle|0\rangle \\
&|x\rangle|0\rangle \longmapsto -|x\rangle|0\rangle \longmapsto -|x\rangle|0\rangle (x \neq 0^n)
\end{aligned}$$

As a conclusion, V can be implemented with cost $O(\log N)$. Same as preparing $|\psi\rangle = H^{\otimes n}|0\rangle$.

Grover algorithm:

- \- Prepare $|\psi\rangle$
- \- Repeat $t = \lceil \frac{\pi}{4}\sqrt{n} \rceil$ times

  Apply U;

  Apply V;
- \- Measure in the computational basis

$$U|\psi\rangle = (I - 2|\omega\rangle\langle\omega|)|\psi\rangle = |\psi\rangle - 2\langle\omega|\psi\rangle|\omega\rangle = |\psi\rangle - \frac{2}{\sqrt{N}}|\omega\rangle.$$

$$U\,|\omega\rangle = -\,|\omega\rangle$$

$$V\,|\psi\rangle = (2|\psi\rangle\langle\psi| - I)\,|\psi\rangle = |\psi\rangle$$

$$V|\omega\rangle = (2|\psi\rangle\langle\psi| - I)|\omega\rangle = 2(\langle\psi|\omega\rangle|\psi\rangle - |\omega\rangle) = \frac{2}{\sqrt{N}}|\psi\rangle - |\omega\rangle$$

Therefore, the subspace span $\{|\psi\rangle, |\omega\rangle\}$ is invariant under U and V.
However, $\langle\psi|\omega\rangle \neq 0$. It will read better to consider an orthonormal basis span $\{|\omega\rangle, |\omega^\perp\rangle\}$:

$$|\omega^\perp\rangle = \frac{|\psi\rangle - \langle\omega|\psi\rangle\,|\omega\rangle}{\text{normalization}}, \quad \langle\omega|\omega^\perp\rangle = \langle\omega|\psi\rangle - \langle\omega|\psi\rangle\,\langle\omega|\omega\rangle = 0.$$

$$U = I - 2|w\rangle\langle w|,$$
$$U|w\rangle = (I - 2|w\rangle\langle w|)|w\rangle = -1|w\rangle + 0\left|w^\perp\right\rangle$$
$$U\left|w^\perp\right\rangle = (I - 2|w\rangle\langle w|)\left|w^\perp\right\rangle = 0|w\rangle + 1|w^\perp\rangle$$

$$U = \begin{array}{cc} \begin{array}{cc} |w\rangle & |w^\perp\rangle \end{array} & \\ \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} & \begin{array}{c} |w\rangle \\ |w^\perp\rangle \end{array} \end{array}$$

$$|\psi\rangle = \frac{1}{\sqrt{N}}|\omega\rangle + \sqrt{1 - \frac{1}{N}}\,|\omega^\perp\rangle = \sin\theta|\omega\rangle + \cos\theta\,|\omega^\perp\rangle \quad \left(\theta = \arcsin\frac{1}{\sqrt{N}}\right).$$

$$V = 2\,|\psi\rangle\,\langle\psi| - I = 2\begin{pmatrix} \sin\theta \\ \cos\theta \end{pmatrix}(\sin\theta \quad \cos\theta) - I = \begin{pmatrix} 2\sin^2\theta - 1 & 2\sin\theta\cos\theta \\ 2\sin\theta\cos\theta & 2\cos^2\theta - 1 \end{pmatrix} = \begin{pmatrix} -\cos 2\theta & \sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{pmatrix}$$

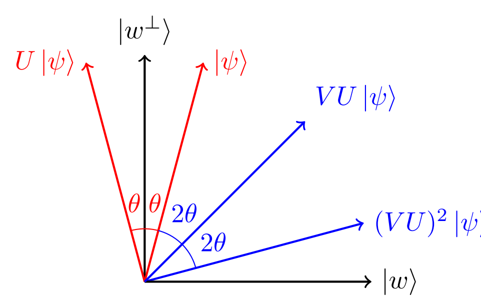
$$VU = \begin{pmatrix} -\cos 2\theta & \sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ -\sin 2\theta & \cos 2\theta \end{pmatrix}.$$

$$(VU)^t = \begin{pmatrix} \cos(2(t-1)\theta) & \sin(2(t-1)\theta) \\ -\sin(2(t-1)\theta) & \cos(2(t-1)\theta) \end{pmatrix}\begin{pmatrix} \cos 2\theta & \sin 2\theta \\ -\sin 2\theta & \cos 2\theta \end{pmatrix}.$$
$$= \begin{pmatrix} \cos(2(t-1)\theta)\cos 2\theta - \sin(2(t-1)\theta)\sin 2\theta & \cos(2(t-1)\theta)\sin 2\theta + \sin(2(t-1)\theta)\cos 2\theta \\ -\sin(2(t-1)\theta)\cos 2\theta - \cos(2(t-1)\theta)\sin 2\theta & -\sin(2(t-1)\theta)\sin 2\theta + \cos(2(t-1)\theta)\cos 2\theta \end{pmatrix}$$
$$= \begin{pmatrix} \cos 2t\theta & \sin 2t\theta \\ -\sin 2t\theta & \cos 2t\theta \end{pmatrix}.$$

$$(VU)^t|\psi\rangle = \begin{pmatrix} \cos 2t\theta & \sin 2t\theta \\ -\sin 2t\theta & \cos 2t\theta \end{pmatrix}\begin{pmatrix} \sin\theta \\ \cos\theta \end{pmatrix} = \begin{pmatrix} \cos 2t\theta\sin\theta + \sin 2t\theta\cos\theta \\ -\sin 2t\theta\sin\theta + \cos 2t\theta\cos\theta \end{pmatrix} = \begin{pmatrix} \sin(2t+1)\theta \\ \cos(2t+1)\theta \end{pmatrix} \begin{array}{c} |w\rangle \\ |w^\perp\rangle \end{array}$$

$$\Longrightarrow \Pr(\omega) = \sin^2((2t+1)\theta)$$

Pr(w) is close to 1 when $(2t+1)\theta \approx \frac{\pi}{2} \Rightarrow t \approx \frac{\pi}{4\theta} - \frac{1}{2} \approx \frac{\pi}{4}\sqrt{N}$.



What if there are M marked items?

$$|\omega\rangle \longmapsto \frac{1}{\sqrt{M}} \sum_{x:f(x)=1} |x\rangle \quad \sin\theta = \langle\psi|w\rangle = \langle\psi| \frac{1}{\sqrt{M}} \sum_{x:f(x)=1} |x\rangle = \frac{1}{\sqrt{N}} \cdot \frac{1}{\sqrt{M}} \cdot M = \sqrt{\frac{M}{N}}.$$

$\Rightarrow O\left(\sqrt{\frac{N}{M}}\right)$ steps suffice for Grover's algorithm.

Remark 1. What if M is unknown? Without knowing M, we may "overshoot".

Research paper: Yoder, Low, Chuang. Fixed-point quantum search with an optimal number of queries. PRL 2014. arxiv: 1409.3305.

Contribution: An algorithm without overshooting, nor knowing M. Cost: $O\left(\sqrt{\frac{N}{M}}\right)$.

Or naively: Guess M using a geometric series: $M = \frac{n}{2}$ $\quad \frac{\pi}{4}\sqrt{\frac{N}{M}}$. $\quad M \leftarrow M \cdot \frac{3}{4}$. $\quad$ Iteration: $\log N$.

3 Amplitude amplification

In general, we can consider a form of having a unitary U acting on l qubits s.t.

$$U\left|0^l\right\rangle = \sqrt{p}|1\rangle\left|\psi_1\right\rangle + \sqrt{1-p}|0\rangle\left|\psi_0\right\rangle$$

where $|\psi_1\rangle$ and $|\psi_0\rangle$ are normalized $(l-1)$-qubit quantum states.

We can think of $|1\rangle|\psi_1\rangle$ as the "good state" and $|0\rangle|\psi_0\rangle$ as the "bad state".

Similar to Grover:

1. A reflection with respect to the bad state $|0\rangle|\psi_0\rangle$.
2. A refection with respect to $U\left|0^l\right\rangle$.

1: This is basically putting a "-" in front of $|1\rangle|\psi_1\rangle$ and leaving $|0\rangle|\psi_0\rangle$ alone.
Solution: Put a Z on the first quit.

$$Z|0\rangle = |0\rangle$$
$$Z|1\rangle = -|1\rangle$$

2: Same to Grover, apply $UR_0U^\dagger$.

Applying 1 and 2 alternatively for t times, we get

$$\sin((2t+1)\theta)\left|1\right\rangle\left|\psi_1\right\rangle + \cos((2t+1)\theta)|0\rangle\left|\psi_0\right\rangle, \qquad \text{where } \theta = \arcsin\sqrt{p} \approx \sqrt{p}.$$

Taking $(2t+1)\theta \approx \frac{\pi}{2} \Rightarrow t = O\left(\frac{1}{\sqrt{p}}\right)$, we can (approximately) get $|1\rangle|\psi_1\rangle$.

This is known as amplitude amplification.

Remark 2. In fact, there's also a procedure called amplitude estimation, where we can quantitatively output a $\tilde{p}$ s.t. $|\tilde{p} - p| \leqslant \varepsilon p$, with high probability (say $2/3$). using $O(1/\varepsilon)$ queries to U.

On the other hand: Classically, tossing a coin, getting a head with prob. p, need $O(1/\varepsilon^2)$ tries to estimate such a $\tilde{p}$. Quantum: Amp Est has quadratic speedup.

Amp Amp: Classically, tossing a coin, getting a head with prob. p, need $O(1/p)$ tries to see the first head.

$$\begin{aligned} \Pr[\text{fail}] &= 1-(1-p)^m \quad (1-p)^{\frac{1}{p}} \approx \frac{1}{e} \\ &\geqslant \frac{2}{3} \quad m \sim \frac{1}{p} \end{aligned}$$

Quantum: Amp Amp has quadratic speedup.

Amp Est: Brassard, Hoyer, Mosca, and Tapp. Quantum Amplitude Amplification and Estimation. Contemporary Mathematics, 2002. arxiv: quant-ph/0005055.