CMSC 858L: Quantum Complexity

Instructor: Daniel Gottesman

Spring 2023

5 Lecture 5: Solovay-Kitaev Theorem, non-universal gate sets

For a proof of the Solovay-Kitaev theorem, see Ch. 8 of Kitaev, Shen, and Vyalyi, *Classical and Quantum Computation* or Appendix 3 of Nielsen and Chuang, *Quantum Computation and Quantum Information*. The inverse-free Solovay-Kitaev theorem is Bouland and Giurgica-Tiron, "Efficient Universal Quantum Compilation: An Inverse-free Solovay-Kitaev Algorithm," arXiv:2112.02040 [quant-ph].

5.1 Solovay-Kitaev

Theorem 1 (Solovay-Kitaev)**.** *Let $\mathcal{G}$ be a universal set of gates. Then for any unitary V in a fixed Hilbert-space dimension D, there exists a classical algorithm to find, for any $\epsilon > 0$, a quantum circuit of size $O(\text{poly}(\log(1/\epsilon)))$ which realizes U_ϵ such that $\|V - U_\epsilon\| < \epsilon$. The classical algorithm runs in time $O(\text{poly}(\log(1/\epsilon)))$ as well.*

The original version of the Solovay-Kitaev theorem requires that $\mathcal{G}$ is closed under inverses, but there is a recent improvement that removes that requirement. It is also worth noting that certain gate sets (such as $\{H, R_{\pi/8}\}$) can achieve this approximation more efficiently (i.e., with a lower exponent of the logarithm) than the general case.

Also, very importantly, we are fixing the Hilbert space dimension. While the theorem works for any dimension, the number of gates needed is certainly exponential in the number of qubits. The algorithm is only efficient in the accuracy needed.

Proof sketch. The first step of the proof is to get a very rough approximation, with a constant degree of accuracy ϵ_0. We can discover how to do this by simply trying out all the possible circuits up to a certain (constant) size. There are a constant number of such circuits. Because the gate set is universal, *eventually* we will get ϵ_0-close to every unitary. The size of circuit needed to do this will be dependent on the exact gate set, but it is independent of the eventual target accuracy ϵ, which hasn't shown up yet.

An ϵ-net S for a subset $B \subseteq SU(D)$ is a set such that for all $V \in B$, $\exists U \in S$ such that $\|U - V\| < \epsilon$. (Here $SU(D)$ is the special unitary group of dimension D, the set of all $D \times D$ unitary matrices of determinant 1; unitaries always have determinant with absolute value 1, so any unitary is related to an element of $SU(D)$ up to a global phase, which is physically irrelevant.) So this first stage of the proof is to create an ϵ_0-net for $SU(D)$.

Given the target V for the theorem, we can therefore pick a U_0 from the ϵ_0-net such that $\|V - U_0\| < \epsilon_0$. Let $V_1 = VU_0^{-1}$. Then

$$\|V_1 - I\| = \|(V - U_0)U_0^{-1}\| = \|V - U_0\| < \epsilon_0 \tag{1}$$

since the distance measures we are likely to use are invariant under unitary rotations.

Our next task is to find a good approximation U_1 to V_1. Suppose $\|V_1 - U_1\| < \epsilon_1 < \epsilon_0$. Then

$$\|V - U_1U_0\| = \|V_1U_0 - U_1U_0\| = \|V_1 - U_1\| < \epsilon_1, \tag{2}$$

again using invariance of the distance under unitaries, so we have found a better approximation to V. We keep doing this with smaller and smaller ϵ_i to find the desired circuit.

So how to do we find this better approximation? Suppose we have an ϵ_i-net S_i in a ball of radius δ_i around the identity. We want to use this to find a ϵ_{i+1}-net S_{i+1} in a ball of radius δ_{i+1} around the identity, with $\epsilon_{i+1} < \epsilon_i$ and $\delta_{i+1} < \delta_i$.

In the case of unitaries, we write $U = e^{i\eta H} \approx I + i\eta H$, where H is Hermitian (i.e., $H = H^\dagger$) and traceless ($\operatorname{Tr} H = 0$). Suppose we have $U_1 = e^{-i\eta H_1}$ and $U_2 = e^{-i\eta H_2}$. By the Baker-Campbell-Hausdorff lemma,

$$U_1 U_2 = e^{-i\eta(H_1+H_2)-\eta^2[H_1,H_2]/2+O(\eta^3)}, \tag{3}$$

with $[H_1, H_2] = H_1H_2 - H_2H_1$. In particular, when η is small, we can reasonably approximate the product $U_1U_2 = e^{-i\eta(H_1+H_2)+O(\eta^2)}$, whereas the group commutator

$$U_1U_2U_1^\dagger U_2^\dagger = e^{-\eta^2[H_1,H_2]+O(\eta^3)}. \tag{4}$$

Lemma 1. *The set of unitaries $U_1U_2U_1^\dagger U_2^\dagger$ where U_1 and U_2 are arbitrary elements of S_i form an $O(\epsilon_i\delta_i)$-net in a ball of radius $O(\delta_i^2)$ provided $\epsilon_i < \delta_i$ and $\epsilon_i = \Omega(\delta_i^2)$.*

Proof of lemma. Suppose we have an element W in a ball of radius $O(\delta_i^2)$. Then we can write $W = e^{-i\delta_i^2 H}$ and pick traceless Hermitian H_1, H_2 such that $\delta_i^2 H = -i[\delta_i H_1, \delta_i H_2] + O(\delta_i^3)$. (This is not obvious, but follows from properties of the Lie algebra for $SU(D)$.). Since $e^{-i\delta_i H_1}$ and $e^{-i\delta_i H_2}$ are in a ball of radius δ_i around the identity, we can approximate them with elements of S_i U_1 and U_2. We have $U_j = e^{-i\delta_i H_j'}$, with $H_j' = H_j + O(\epsilon_i)$ $(j = 1, 2)$. Then

$$\begin{aligned} U_1U_2U_1^\dagger U_2^\dagger &= e^{-\delta_i^2[H_1',H_2']+O(\delta_i^3)} && (5)\\ &= e^{-\delta_i^2[H_1,H_2]+O(\epsilon_i\delta_i)+O(\delta_i^3)} && (6)\\ &= e^{-i\delta_i^2 H+O(\epsilon_i\delta_i)+O(\delta_i^3)} && (7)\\ &= W + O(\epsilon_i\delta_i) + O(\delta_i^3). && (8) \end{aligned}$$

When $\epsilon_i = \Omega(\delta_i^2)$, we therefore have an approximation of W to an accuracy $O(\epsilon_i\delta_i)$. □

It's actually possible to get a somewhat tighter error bound for ϵ_i than in this lemma due to cancellations in the commutator. Unfortunately, the net does not cover enough area for our purposes, so we need to expand it to cover a larger ball using the same approximation technique we are planning to use for V (shift to near I and then approximate).

Putting all of these components together, we can get finer and finer nets with $\epsilon_i = O(\epsilon_{i-1}^c)$ for some constant c, so $\epsilon_i = O(\epsilon_0^{c^i})$. Each net uses more gates, however, by a constant factor d. To achieve the desired approximation ϵ, then, we need to have $c^i = \log\epsilon/\log\epsilon_0 = O(\log(1/\epsilon))$ (since $\epsilon_0 < 1$) and a total number of gates

$$d^i = O(\log(1/\epsilon)^{\log d/\log c}) = O(\operatorname{poly}(\log(1/\epsilon))). \tag{9}$$

The standard approach gives $c = 3/2$ and $d = 5$, so the exponent is about 4, but this can be tightened.

□

One immediate consequence of the Solovay-Kitaev theorem is that if we have two different universal gates, we can convert a circuit written using one such gate set to the other one with minimal overhead. In particular, suppose $C_{\mathcal{G}}$ is a circuit of size T using gates from gate set $\mathcal{G}$. In order to rewrite this circuit using gates from gate set $\mathcal{H}$, we should replace each gate in $C_{\mathcal{G}}$ with an approximation from $\mathcal{H}$. We need an accuracy $\epsilon = O(1/T)$ so that the approximate circuit is close enough to the original. Therefore, each gate from $\mathcal{G}$ gets replaced with $O(\operatorname{poly}\log(T))$ gates from $\mathcal{H}$. Since the gates from $\mathcal{G}$ act on a bounded number of qubits, we don't have to worry about the dimension factors in this approximation. The new circuit thus has $O(T\operatorname{poly}\log T)$ gates, which is certainly polynomial in T. While the polylogarithmic scaling of Solovay-Kitaev is not necessary for defining BQP (polynomial in $1/\epsilon$ would have sufficed), it *is* necessary for getting meaningful polynomial speedups, such as Grover's algorithm, from a quantum computer.

5.2 Non-Universal Gate Sets

The standard example of a universal set of gates is $\mathcal{G} = \{\text{single-qubit gates}, CNOT\}$. Two standard examples of approximately universal sets of gates are $\mathcal{G} = \{H, R_{\pi/8}, CNOT\}$ and $\{H, R_{\pi/4}, Tof\}$. Here $CNOT$ is the controlled-NOT, H is the Hadamard transform, R_θ is the diagonal phase gate $R_\theta|0\rangle = e^{-i\theta}|0\rangle$, $R_\theta|1\rangle = e^{i\theta}|1\rangle$, and Tof is the 3-qubit Toffoli gate (also called the controlled-controlled-NOT). (Often, $R_{\pi/4}$ is called S and $R_{\pi/8}$ is called T.)

What if instead of a universal gate set, we take some non-universal gate set? There are a lot of different choices, and now they are not all equivalent. So let us think about polynomial-size quantum circuits built from some different sets $\mathcal{G}$ which are not universal. It turns out that there are a lot of different variations and even small changes in the rules can produce different results. For the moment, let us assume that we always initialize the circuit with states in the standard basis, measure at the end of the computation in the standard basis, and do not allow intermediate measurements.

As a first example, consider the gate set $\mathcal{G} = \{X, CNOT, Tof\}$. What complexity class do we get from polynomial circuits with this gate set?

Answer: P! These are all classical gates and indeed, they form a universal set of reversible classical gates. Since every classical computation can be converted to a reversible computation, this allows everything in P. Note that we don't even have any way to introduce randomness, so we have P, not BPP. But if we allowed some qubits to be initialized in the state $|+\rangle = |0\rangle + |1\rangle$, then we could use those as random bits and we would get BPP. (This is a simple example of what I mean when I say that small rule changes can give different results.)