Классические и квантовые вычисления

А. Китаев; написал М. Вялый

6 лекция (19/03/98)

1. Базисы для квантовых схем

Как выбрать базис для вычислений в квантовых схемах? Унитарных операторов бесконечно много, поэтому полный базис должен содержать бесконечное количество элементов, или же мы должны ослабить условие точной реализуемости оператора схемой на условие приближенной реализуемости. Мы рассмотрим обе возможности.

Точная реализация

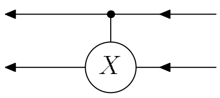
Теорема. *Базис, содержащий все унитарные операторы, действующие на парах кубитов, позволяет реализовать любой унитарный оператор в расширенном смысле (с дополнительными q-битами, как это объяснялось раньше).*

Для доказательства этой теоремы введем важный класс операторов: *операторы с квантовым управлением.*

Напомним, что рассматриваются унитарные операторы $U : \mathbb{B}^n \to \mathbb{B}^n$, где $\mathbb{B} = \mathbb{C}^2$ с выделенным базисом $\{|0\rangle, |1\rangle\}$.

Определение. *Определим по оператору U оператор $\Lambda(U) : \mathbb{B} \otimes \mathbb{B}^n \to \mathbb{B} \otimes \mathbb{B}^n$ с квантовым управляющим q-битом (первый сомножитель) следующими соотношениями:*

$$\Lambda(U)\,|0\rangle \otimes |\xi\rangle = |0\rangle \otimes |\xi\rangle$$
$$\Lambda(U)\,|1\rangle \otimes |\xi\rangle = |1\rangle \otimes U|\xi\rangle$$



Графически будем изображать оператор $\Lambda(X)$ с квантовым управлением как показано на врезке. Верхняя линия соответствует первому сомножителю, нижняя линия — второму. Направление стрелок соответствует направлению перемножения операторов (справа налево).

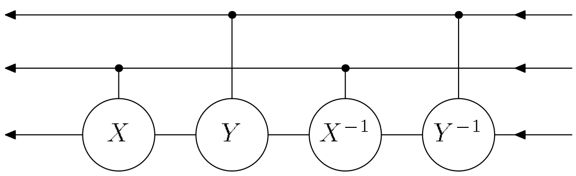
Теперь построим элемент Тоффоли, используя преобразования двух q-битов. Для начала найдем пару операторов, удовлетворяющих следующему соотношению $XYX^{-1}Y^{-1} = \sigma^x \stackrel{\text{def}}{=} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Например, годится такая пара: $X = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$; $\quad Y = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

Поясним геометрический смысл этой конструкции. Унитарная группа $U(2)$ действует на трехмерном евклидовом пространстве. Чтобы описать это действие, заметим, что эрмитовы матрицы 2×2 с нулевым следом образуют трехмерное евклидово пространство: скалярное произведение задается формулой $\mathrm{Tr}(XY)$, ортонормированный базис образуют *матрицы Паули* $\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\sigma^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, $\sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Унитарный оператор $U \in U(2)$ действует на этом пространстве так: $U : H \mapsto UHU^{-1}$.

При этом действии σ^x соответствует поворот вокруг оси x на 180°, X соответствует поворот вокруг x на 90°, а Y соответствует поворот вокруг z на 180°.

На врезке изображено графическое представление схемы, вычисляющей элемент Тоффоли с помощью операторов $\Lambda(X)$, $\Lambda(Y)$. Проверим это. Пусть на вход схемы подается вектор $|a\rangle\otimes|b\rangle\otimes|\xi\rangle$, где $a, b \in \{0,1\}$, $|\xi\rangle \in \mathbb{B}$. Если $a = b = 1$, то к $|\xi\rangle$ будет применен оператор $XYX^{-1}Y^{-1} = \sigma^x$, т. е. $|0\rangle$ и $|1\rangle$ переставляются. Если же хотя бы один из управляющих битов равен 0, то к ξ будет применен тождественный оператор. Это и есть действие элемента Тоффоли.



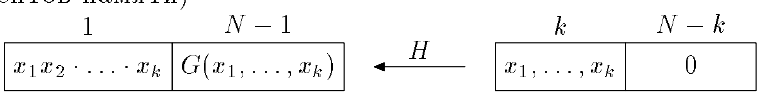
С помощью элемента Тоффоли можно реализовать любую перестановку базисных векторов (с использованием дополнительной памяти).

Введем еще один оператор

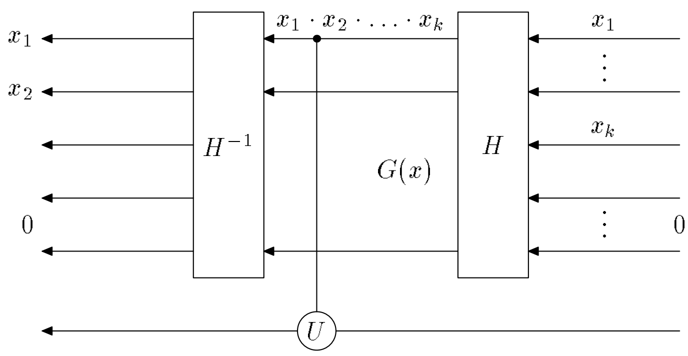
$$\Lambda^k(U)\,|x_1,\ldots,x_k\rangle\otimes|\xi\rangle \;=\; \begin{cases} |x_1,\ldots,x_k\rangle\otimes|\xi\rangle & \text{если } x_1\cdot\ldots\cdot x_k = 0 \\ |x_1,\ldots,x_k\rangle\otimes U|\xi\rangle & \text{если } x_1\cdot\ldots\cdot x_k = 1 \end{cases}.$$

Для его реализации также потребуется дополнительная память. Будем строить оператор W, действующий в пространстве N q-битов $\mathbb{B}^{\otimes N}$ и удовлетворяющий условию $W\left(|\eta\rangle\otimes|0^{N-k-1}\rangle\right) = \Lambda(U)|\eta\rangle\otimes|0^{N-k-1}\rangle$. (Предостережение: это условие ***не означает***, что $W = \Lambda(U)\otimes I$.)

Существует обратимая схема H размера $O(k)$, вычисляющая произведение входных битов (с мусором); графически ее можно представить как (сверху обозначено число битов в каждом из выделенных фрагментов памяти)



На врезке слева показано, как с помощью схемы H и оператора с одним управляющим q-битом построить оператор $\Lambda^k(U)$. Мы применяем схему H, а затем — обратную схему H^{-1}, после чего все вспомогательные биты возвращаются в исходное состояние. В промежутке самой верхней линии соответствует бит со значением $x_1\cdot\ldots\cdot x_k$. Его мы и используем для управления оператором U, действующим на самой нижней линии. Другим способом $\Lambda^k(U)$ можно записать как $H^{-1}\Lambda(U)H$.



Действие $\Lambda^k(U)$ можно описать так: на подпространстве, порожденном векторами $|1,\ldots 1,\,0\rangle$ и $|1,\ldots 1,\,1\rangle$, действует оператор U, а на ортогональном дополнении к этому подпространству — тождественный оператор. Наша следующая задача: реализовать оператор, который устроен так же, но нетривиальное действие осуществляется на подпространстве, натянутом на произвольную пару базисных векторов. Пусть мы хотим реализовать произвольный оператор в подпространстве, натянутом на базисные вектора $|x\rangle$ и $|y\rangle$, где $x = x_1\ldots x_n$, $y = y_1\ldots y_n$, $x_j,\,y_j \in \{0,1\}$. Пусть f — такая перестановка, что $f(x) = (1,\,\ldots,\,1,\,0)$, $f(y) = (1,\,\ldots,\,1,\,1)$. Тогда нужный нам оператор представляется в виде $\hat{f}^{-1}\Lambda^{n-1}(U)\hat{f}$. (Напомним, что $\hat{f}$ — оператор, соответствующий перестановке f.)

Итак, на парах базисных векторов мы можем действовать произвольно. Пока все использованные схемы имели размер $O(n)$, так что построенные действия реализуются эффективно. Следующая часть неэффективна. Имеет место следующая лемма.

Лемма. *Любая унитарная матрица U в пространстве $\mathbb{C}^M$ может быть представлена в виде произведения $M(M-1)/2$ матриц вида*

$$\begin{pmatrix} 1 & 0 & \ldots\ldots\ldots\ldots\ldots\ldots\ldots & & & & \\ \vdots & \ddots & 0 & \ldots\ldots\ldots\ldots\ldots\ldots & & & \\ 0 & \ldots & 1 & 0 & \ldots\ldots\ldots & & \\ 0 & \ldots\ldots & \begin{pmatrix} a & b \\ c & d \end{pmatrix} & 0 & \ldots\ldots & & \\ 0 & \ldots\ldots\ldots\ldots\ldots & & 1 & 0 & 0 & \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots & & & & \ddots & 0 & \\ 0 & \ldots\ldots\ldots\ldots\ldots\ldots\ldots & & & & 1 & \end{pmatrix},$$

где $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *— унитарная матрица.*

Заметим, что в нашем случае $M = 2^n$, так что получаем представление в виде произведения экспоненциального большого числа базисных операторов.

Приближенная реализация

Теперь перейдем к рассмотрению конечных базисов. В этом случае возможно только приближенное представление операторов произведениями базисных. Чтобы определить приближенную реализацию, нам потребуется норма на пространстве операторов.

Во-первых, отметим, что на пространстве состояний определена норма $\big\||\xi\rangle\big\| = \sqrt{\langle\xi|\xi\rangle}$. Она, как и любая норма, удовлетворяет следующим условиям:

$$\big\||\xi\rangle\big\| = \begin{cases} 0, & \text{если } |\xi\rangle = 0; \\ >0, & \text{если } |\xi\rangle \neq 0; \end{cases} \tag{1}$$

$$\big\||\xi\rangle + |\eta\rangle\big\| \le \big\||\xi\rangle\big\| + \big\||\eta\rangle\big\|; \tag{2}$$

$$\big\|c|\xi\rangle\big\| = |c|\big\||\xi\rangle\big\|. \tag{3}$$

Эти условия и есть определение нормы.

Определим теперь норму на пространстве операторов. Пусть $\mathcal{N}$ — пространство с нормой. Пространство операторов, действующих на нем можно представить как $\mathbf{L}(\mathcal{N}) = \mathcal{N} \otimes \mathcal{N}^*$ (изоморфизм задается матричным представлением $\sum a_{ij}|i\rangle\langle j|$).

Определение. *Норма оператора X (так называемая операторная норма, вообще говоря, есть и другие) равна*

$$\|X\| \;=\; \sup_{|\xi\rangle\neq 0} \frac{X\big\||\xi\rangle\big\|}{\big\||\xi\rangle\big\|}$$

Эта норма обладает всеми перечисленными выше свойствами нормы, а кроме того, еще несколькими специфическими

$$\|XY\| \le \|X\|\,\|Y\|; \tag{4}$$

$$\|X^\dagger\| = \|X\|; \tag{5}$$

$$\|X \otimes Y\| = \|X\|\,\|Y\|, \qquad \text{где } X \in \mathbf{L}(\mathcal{N}),\ Y \in \mathbf{L}(\mathcal{M}). \tag{6}$$

Доказательства этих свойств нормы получаются непосредственно из определения и оставляются в качестве упражнения.

Дадим теперь определение приближенной реализуемости. Если искомый оператор — U, то его приближенная реализация будет обозначаться $\tilde{U}$.

Определение. *$\tilde U$ представляет U с точностью δ, если $\|\tilde U - U\| \le \delta$.*

У этого определения есть два замечательных свойства. Во-первых, если мы имеем произведение нескольких операторов $U = U_L \dots U_2 U_1$, каждый из которых имеет свое приближение $\tilde U_k$ с точностью δ_k ($\|\tilde U_k - U_k\| < \delta_k$), то произведение этих приближений $\tilde U = \tilde U_L \dots \tilde U_2 \tilde U_1$ приближает U с точностью $\sum \delta_k$ (***ошибки накапливаются линейно***):

$$\left\|\tilde U_L \dots \tilde U_1 - U_L \dots U_1\right\| \le \sum_j \delta_j.$$

Достаточно рассмотреть пример с двумя операторами

$$\|\tilde U_2 \tilde U_1 - U_2 U_1\| = \|\tilde U_2(\tilde U_1 - U_1) + (\tilde U_2 - U_2)U_1\| \le \|\tilde U_2(\tilde U_1 - U_1)\| + \|(\tilde U_2 - U_2)U_1\| \le$$
$$\le \|\tilde U_2\|\,\|(\tilde U_1 - U_1)\| + \|(\tilde U_2 - U_2)\|\,\|U_1\| = \|(\tilde U_1 - U_1)\| + \|(\tilde U_2 - U_2)\|.$$

В этой выкладке последнее равенство справедливо благодаря унитарности операторов. Если рассматривать неунитарные операторы, то ошибки приближения могут накапливаться гораздо быстрее, например, экспоненциально. Поэтому всякая модель, претендующая на решение сложных задач какими-то реальными физическими процессами, должна обязательно изучаться на предмет устойчивости к ошибкам приближения. (В реальной жизни параметры любого физического процесса можно задать лишь с некоторой точностью). В частности, вычисление с экспоненциальным накоплением ошибок почти заведомо бесполезно с практической точки зрения.

Второе свойство понятия “$\tilde U$ представляет U с точностью δ” мы сформулируем в более общем контексте.

Определение. ***Приближенная реализация в расширенном смысле.*** *Оператор $U : \mathbb{B}^{\otimes n} \to \mathbb{B}^{\otimes n}$ приближается в расширенном смысле оператором $\tilde U : \mathbb{B}^{\otimes N} \to \mathbb{B}^{\otimes N}$, если для любого $|\xi\rangle$ из $\mathbb{B}^{\otimes n}$ выполнено*

$$\left\|\tilde U\left(|\xi\rangle \otimes |0^{N-n}\rangle\right) - U|\xi\rangle \otimes |0^{N-n}\rangle\right\| < \delta\left\||\xi\rangle\right\|.$$

Сформулируем это определение еще одним способом. Введем оператор $V : \mathbb{B}^{\otimes n} \to \mathbb{B}^{\otimes N}$, который действует по правилу $V : |\xi\rangle \mapsto |\xi\rangle \otimes |0^{N-n}\rangle$. Оператор V не унитарный, но изометричный. Условие из последнего определения можно переписать так

$$\|\tilde U V - V U\| \le \delta. \tag{7}$$

Рассуждение про накопление ошибок проходит и в этом случае (что, конечно, следует проверить).

Во-вторых, справедливо следующее утверждение: если $\tilde U$ приближает (в расширенном смысле) U с точностью δ, то $\tilde U^{-1}$ приближает U^{-1} с той же точностью δ. Это следует из того, что $\|W_1 X W_2\| = \|X\|$ для унитарных операторов W_1, W_2. Умножая выражение под нормой в (7) слева на $\tilde U^{-1}$, а справа — на U^{-1}, получим следствие из неравенства (7) $\|\tilde U^{-1} V - V U^{-1}\| \le \delta$.

Определение. *Будем называть базис $\mathcal{A}$ полным, если любой унитарный оператор U можно с любой точностью представить в расширенном смысле квантовой схемой в базисе $\mathcal{A}$.*

Теорема (см. УМН, №6, 1997). *Базис $\mathcal{Q} = \{S, K, \widehat{\oplus}, \widehat{\wedge_\oplus}\}$, где*

$$S = \frac{1}{\sqrt 2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \qquad K = \begin{pmatrix} 1 & 0 \\ 0 & \mathrm{i} \end{pmatrix},$$

$$\widehat{\wedge_\oplus} \text{ — квантовый элемент Тоффоли}, \qquad \Lambda(\sigma^x) = \widehat{\oplus} : |a, b\rangle \mapsto |a, a \oplus b\rangle,$$

является полным.

Доказательство этой теоремы здесь не приводится, заинтересованные слушатели адресуются к упомянутому обзору в УМН.

Замечание. Если убрать из базиса $\mathcal{Q}$ квантовый элемент Тоффоли, он перестает быть полным. Однако многие важные вычисления можно делать и в таком усеченном базисе. В частности, как будет видно в дальнейшем, схемы, исправляющие ошибки, можно реализовать без элемента Тоффоли.

Можно оценить сложность реализации оператора U в этом базисе. Если $U : \mathbb{B}^{\otimes n} \to \mathbb{B}^{\otimes n}$, то можно реализовать этот оператор с точностью δ квантовой схемой в базисе $\mathcal{Q}$ размера $L = \exp(O(n) \cdot poly(\log(\frac{1}{\delta})))$. Если U задан матрицей с элементами в двоичном представлении, то эта схема строится по U с помощью некоторого алгоритма примерно за то же время (множители и степени полинома могут отличаться).

2. Квантовые вероятности

Перейдем теперь к обсуждению некоторых "физических" аспектов квантовых вычислений. Пусть система из n q-битов находится в состоянии $|\psi\rangle = \sum_x c_x|x\rangle$. Коэффициенты разложения по выделенному базису (классических состояний) называются *амплитудами*. Квадрат амплитуды $|c_x|^2$ интерпретируется как вероятность обнаружить систему состоянии x. Т. е. если сделать *измерение* этой квантовой системы, то будут получаться классические состояния, распределенные как $|c_x|^2$. Не вдаваясь в обсуждение физических объяснений этого факта, положим по определению, что вероятность получения состояния x при измерении состояния $|\psi\rangle$ равна

$$\mathbf{P}(|\psi\rangle, x) = |c_x|^2.$$

Определенная таким образом величина обладает основными свойствами обычной вероятности. Тот факт, что квадрат амплитуды является вероятностью наблюдения системы в состоянии $|x\rangle$, согласуется с тем, что физические состояния в квантовой механике соответствуют векторам единичной длины, а преобразования этих состояний не меняют длины, т. е. унитарны. Действительно, $\langle\psi|\psi\rangle = \sum_x |c_x|^2 = 1$ (сумма вероятностей равна 1), а применение физически реализуемого оператора должно сохранять это соотношение, т. е. должно быть унитарным.

Нам потребуется более общее определение квантовой вероятности, к которому можно прийти, рассматривая ряд примеров. Для начала перепишем уже полученное выражение для вероятности в виде

$$|c_x|^2 = |\langle\psi|x\rangle|^2 = \langle\psi \underbrace{|x\rangle\langle x|}_{\Pi_x} \psi\rangle,$$

где через Π_x обозначен проектор на подпространство, порожденное $|x\rangle$.

Чтобы сделать следующий шаг к общему определению квантовой вероятности, подсчитаем вероятность того, что первые m бит имеют заданное значение $y = (y_1, \ldots, y_m)$. Для этого представим состояние в виде двух блоков $x = \overset{m}{\boxed{y}}\overset{n-m}{\boxed{z}}$. Имеем

$$\mathbf{P}(|\psi\rangle, y) = \sum_z \mathbf{P}(|\psi\rangle, (y,z)) = \sum_z \langle\psi|\,|y,z\rangle\langle y,z|\,|\psi\rangle = \langle\psi|\,|y\rangle\langle y| \otimes I\,|\psi\rangle = \langle\psi|\Pi_{\mathcal{M}}|\psi\rangle.$$

Здесь через $\Pi_{\mathcal{M}}$ обозначен проектор на подпространство $\mathcal{M} = |y\rangle \otimes \mathbb{B}^{\otimes(n-m)}$. Последняя формула и задает определение квантовой вероятности в общем случае, когда $\mathcal{M}$ — произвольное подпространство.

В общем случае проектор на подпространство $\mathcal{M} \subseteq \mathcal{N}$ определяется как $\Pi_{\mathcal{M}} = \sum_j |e_j\rangle\langle e_j|$, где e_j пробегают любой ортонормированный базис $\mathcal{M}$.

Проекторы не являются физически реализуемыми операторами, точнее говоря, они не описывают переход от одного состояния системы к другому за определенный промежуток времени. Такой переход описывается унитарными операторами. Все же, допуская некоторую вольность, проекторам можно придать физический смысл. Проектор как бы выделяет часть состояний системы из всех возможных, поэтому он не сохраняет полную вероятность. Суммарная вероятность, равная $\langle\xi|\xi\rangle$, где $|\xi\rangle = \Pi_{\mathcal{M}}|\psi\rangle$, становится меньше единицы. (Заметим, что $\langle\xi|\xi\rangle = \langle\psi|\Pi_{\mathcal{M}}|\psi\rangle$.) Наглядно это можно представить так. Пусть в систему вмонтирована бомба. Если в результате наблюдения над системой она оказывается в состоянии $|\psi\rangle \in \mathcal{M}$, ничего не происходит. А если оказывается, что $|\psi\rangle \perp \mathcal{M}$, то бомба взрывается. Вектор $|\xi\rangle$ — это то, что осталось после возможного взрыва. (Впрочем, не нужно относиться к этой интерпретации слишком серьезно).

Сравним свойства классической и квантовой вероятности.

Классическая вероятность	Квантовая вероятность
Определение	
Событие $\mathcal{M} \subseteq \mathcal{N}$ — подмножество некоторого конечного множества.	Событие $\mathcal{M}$ — подпространство некоторого конечномерного унитарного пространства $\mathcal{N}$.
Распределение вероятностей задается функцией $w : \mathcal{N} \to \mathbb{R}$ со свойствами а) $\sum w_j = 1$; б) $w_j \geq 0$.	Распределение вероятностей задается вектором состояния $|\psi\rangle$, $\langle\psi|\psi\rangle = 1$.
Вероятность: $\mathbf{P}(w, \mathcal{M}) = \sum_{j \in \mathcal{M}} w_j$.	Вероятность: $\mathbf{P}(|\psi\rangle, \mathcal{M}) = \langle\psi|\Pi_{\mathcal{M}}|\psi\rangle$.
Свойства	
1. Если $\mathcal{M}_1 \cap \mathcal{M}_2 = \varnothing$, то $\mathbf{P}(w, \mathcal{M}_1 \cup \mathcal{M}_2) = \mathbf{P}(w, \mathcal{M}_1) + \mathbf{P}(w, \mathcal{M}_2)$	1^{q}. Если $\mathcal{M}_1 \perp \mathcal{M}_2$, то $\mathbf{P}(|\psi\rangle, \mathcal{M}_1 \oplus \mathcal{M}_2) = \mathbf{P}(|\psi\rangle, \mathcal{M}_1) + \mathbf{P}(|\psi\rangle, \mathcal{M}_2)$
2. (в общем случае) $\mathbf{P}(w, \mathcal{M}_1 \cup \mathcal{M}_2) =$ $= \mathbf{P}(w, \mathcal{M}_1) + \mathbf{P}(w, \mathcal{M}_2) - \mathbf{P}(w, \mathcal{M}_1 \cap \mathcal{M}_2)$	2^{q}. Если $\Pi_{\mathcal{M}_1}\Pi_{\mathcal{M}_2} = \Pi_{\mathcal{M}_2}\Pi_{\mathcal{M}_1}$, то $\mathbf{P}(|\psi\rangle, \mathcal{M}_1 + \mathcal{M}_2) =$ $= \mathbf{P}(|\psi\rangle, \mathcal{M}_1) + \mathbf{P}(|\psi\rangle, \mathcal{M}_2) - \mathbf{P}(|\psi\rangle, \mathcal{M}_1 \cap \mathcal{M}_2)$

Заметим, что условие $\mathcal{M}_1 \perp \mathcal{M}_2$ эквивалентно условию $\Pi_{\mathcal{M}_1}\Pi_{\mathcal{M}_2} = \Pi_{\mathcal{M}_2}\Pi_{\mathcal{M}_1} = 0$.