

为什么计算机科学家们应该了解量子计算

刘宇攀

浙江大学计算机学院

2016 年 2 月 10 日

摘要

量子计算是验证困难的物理实验的正确性的绝佳方式, 但是直到真正的量子计算机出现之前, 计算机科学家们需要了解量子信息吗? 事实上, 量子计算不仅是一个关于新的计算设备的理论, 而且也是一种崭新而令人惊讶的理解世界的方式. 在这篇文章中, 我会介绍量子计算的来源, 量子计算是什么, 以及我们会从中知道什么.

译者按

本文译自 Aram Harrow 的 *Why now is the right time to study quantum computing*, 经 Aram 本人授权翻译. Aram 最广为人知的工作应该是求解线性方程组的 HHL 算法 (Scott 的介绍), 这几乎是现在绝大多数的量子机器学习 (统计学习) 算法的基础. 他现在是 MIT 理论物理中心 (Center of Theoretical Physics) 的助理教授.

在 *XRDS: Crossroads, The ACM Magazine for Students* 上, 关于 Aram Harrow 有段风趣的介绍:

Aram Harrow 并不能确定他究竟是个物理学家, 还是计算机科学家. 他早年在 MIT 获得了学士学位 (物理和数学) 和博士学位 (物理), 后来曾在 *University of Bristol* (数学和计算机科学) 和 *University of Washington* (计算机科学) 拿到过教职. 然而, 他现在仍然是量子计算的信徒.

1 量子计算的发展

在长达数个世纪的时间里, 我们以确定性的观点来审视世界, 这使得我们把它想象成一个非常复杂的机械装置. 但是, 当计算机变得无处不在的时候, 它提供了新的比喻手段, 并且改变了我们思考自然科学、数学甚至社会科学的方式. 计算机不仅帮助我们解决问题, 而且在建造它们和编程的过程中, 它也以全新的思考方式启发了我们.

比如说, 我们将 DNA、语言或者认知这些纷繁复杂的现象, 视为演化出对编码压缩和纠错的信息传输机制. 自博弈论和经济学始, 开始出现与计算的效率和计算能否实现相关的概念, 正如 Kamil Jain 关于纳什均衡的著名评论, “如果你在你的手机找不到它, 那么它就不可能在市场上出现.” 计算机科学也被这些领域的目标所改进. 而对数学, 其自身与计算

有效性也日益引发关注, 造就了与计算机相关的蓬勃发展的新兴领域: 信息论, 图论和统计学. $P \stackrel{?}{=} NP$ 问题是最新的 Clay 千禧年问题, 它的解决将会为求解数学中的古老谜题提供新的曙光: 寻找证明如此困难的原因是什么?

事后想来, 如此这般的计算观点似乎是非常自然的. 但当计算机第一次出现的时候, 即使是预言了其在商业上的巨大成功的寥寥数人, 也无法预见它掀起的思维方式革命. 譬如熵的概念 (熵是压缩和纠错编码中的核心概念) 在 Gauss 的时代, 或者中世纪的阿拉伯人, 甚至是古希腊人就能轻易提出. 但是它直到十九世纪才有了实践意义下的进一步发展, 当我们的热力学理论足够理解蒸汽机的时候. 当 Bell 实验室借研究密码学之名, 在战争时期雇佣了 Claude Shannon 之后, 同样的事情也发生在了二十世纪. 而这样的事情无独有偶, 也并非局限在计算机科学之中. Einstein 在专利局当小职员的经历帮助他想出了相对论, 比如高速铁路网中的时钟同步就曾是个重要的工程问题, 以及对时钟及火车的比喻为他著名的思想实验提供了素材. 总而言之, 自然科学的进展总是跟随着技术的发展, 因为发明总是带给我们全新的认知世界的方式, 以及亟待理论解释的新现象.

关于量子计算的故事大抵亦复如是. 量子力学的提出在二十世纪初叶, 而它现在广泛使用的形式则是建立于 1930 年. 但是量子力学潜在的计算上的优势却没有顺势被发现, 直到物理学家们试图在计算机上模拟量子力学. 为了这样的模拟, 他们尝试考虑实际问题: 当一个孤立系统 (譬如极化后的光子) 可能只需要用两个复数描述 (比如极化后的水平和竖直分量的概率幅), 那么完全描述 n 个这样的系统需要的不是 $2n$ 而是 2^n 个复数, 尽管在测量后我们只能提取出 n 个比特. 针对这个问题, 物理学家们发展了它的解析解以及物理模拟的近似解, 而这些解都与大量的实际应用相关.

量子力学中以指数增长的态空间, 应该自然地联系着更多的计算资源, 甚至远远超过我们的想象. 除此之外, 它和量子力学的新奇特性看起来更像是某种限制, 甚至是”造化弄人”. 譬如 Heisenberg 不确定性关系被认为是对测量的限制. 而诸如量子纠缠的现象则被认为是量子力学基础中的一部分, 或者是关于量子力学的哲学, 但是并没有什么具有可操作性的关联. 直到量子计算和量子密码学, 在上世纪七十年代和八十年代被各自独立提出.

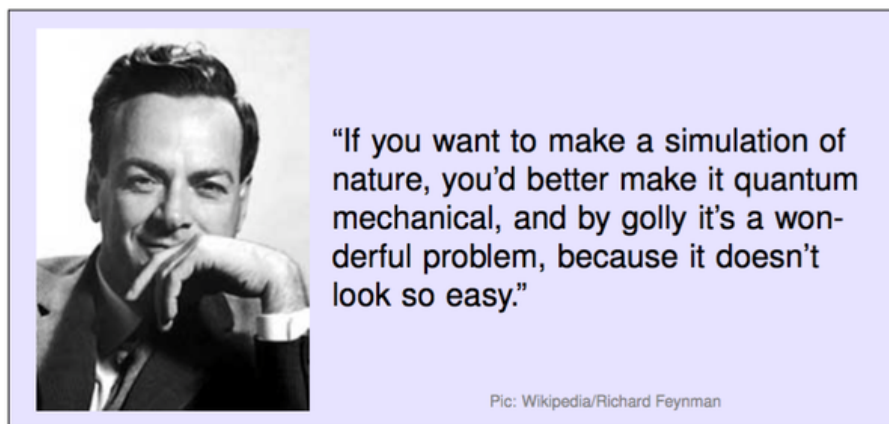


图 1: Richard Feynman

量子计算 (或者更准确的说, 利用量子力学优势的计算) 来自 Richard Feynman 在 1982 年的建议: 既然传统计算机需要用指数级别的资源来模拟量子力学, 也许一台基于量子力学的计算机能够更为有效的完成任务. David Deutsch 在 1985 年对这一想法进行了形式化, 他令人惊讶地展示了一个量子计算机与经典计算机相比存在优势的例子 (计算两比特的异或),

当然, 这个问题在表面上似乎和量子力学并没有什么关系. 很快, 一系列能有效加速算法被提出, 但都是针对人为设计的问题. 直到 1994 年, Peter Shor 提出了多项式时间的素因子分解算法.

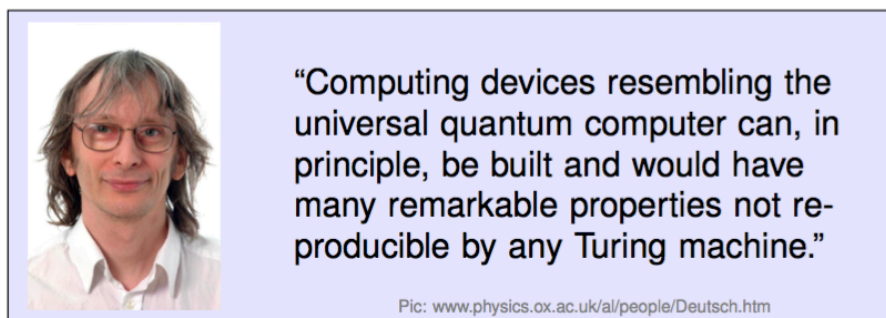


图 2: David Deutsch

而在更早些的时候, 即 1970 年, 还是研究生的 Stephen Wiesner 提出了利用 Heisenberg 约束测量, 来防止对方窃取秘密信息的方法. 因而, 量子密码学应运而生, 尽管 Wiesner 的论文几乎被所有期刊拒收, 并且这样的想法一直并不为人所知. 直到 Charles Bennett 和 Giller Brassard 在 1984 年发表了一篇关于量子密钥分发 (Quantum Key Distribution) 的工作, 甚至这个工作也一直没有受到关注, 直到他们的想法在 1991 年获得实验证实.

这一关键的概念上的演进, 使得量子计算和量子密码学, 开始被用于思考量子力学在操作意义上与信息相关的影响, 而不是被作为某种极限、好奇心或者是悖论的来源. 一旦这样的变化完成, 技术层面的量子信息会变得比早先发展的量子力学更为简单, 就像是上世纪五十年代完成的对量子力学和狭义相对论的统一.

2 用量子比特计算

量子比特 (qubit) 是计算观点对量子力学的基础贡献之一. 一般而言, 对于可区分 d 维量子系统的量子态, 我们可以用 \mathbb{C}^d 上的一个单位矢量来描述它. 最简单而有趣的情形即是 $d = 2$, 这样的系统就叫量子比特 (qubit). 考虑对态矢量 $x = \begin{pmatrix} x_0 \\ x_1 \end{pmatrix}$ 的测量, 为了保证得到结果 0 的概率是 $|x_0|^2$, 并且得到结果 1 的概率是 $|x_1|^2$, 那么 x 必须是单位矢量. 如此说来, 任何一个线性动力系统都是可能的备选方案, 只要能够保证态矢量的模长不变. 换句话说, 这样的演化把态矢量从 x 映射到 Ux , 这里的 U 是酉矩阵 (unitary), 即它能够保证态矢量的模长不变. 在数学上, 这等价于 $U^\dagger U = I$, 即对于每个矩阵元有 $(U^\dagger)_{ij} = \bar{U}_{ji}$. 这个解释的美妙之处, 在于它是与实现系统无关的, 这样的系统可能是光子的极化, 也可能是原子中某个电子的能级, 还可能是原子核的自旋, 亦或是超导线圈中的电流方向. 因此, 量子比特是一种设备无关的量子信息描述方式. 它就像是经典信息中的比特一样, 我们能够从中推理出信息, 却不必要知道它究竟是编码在何处, 不管是在 RAM¹上, 在硬盘上, 还是在算盘上.

¹Random-Access Memory, 随机访问存储器, 用于实现现代计算机中的内存.

Figure 1. This molecule was used in a 5-qubit nuclear magnetic resonance (NMR) quantum computer. While NMR was useful in the early QC implementations, it has limitations that make it unlikely to lead to a large-scale implementation.

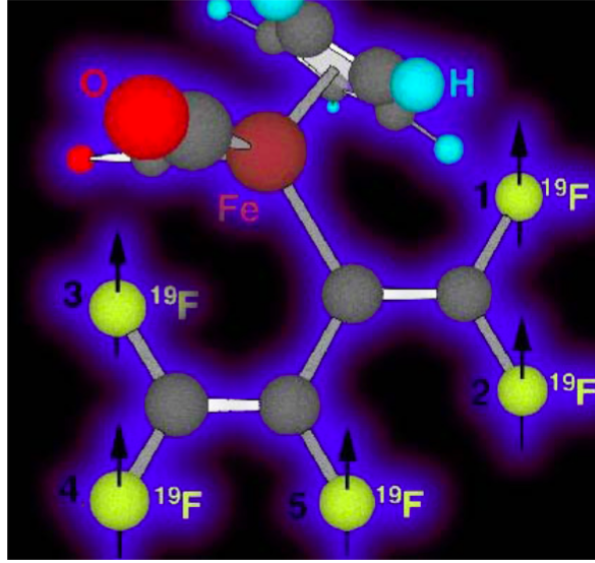


图 3: 量子计算机的 NMR 实现

就物理实验而言, 单量子比特的情形已经足够有趣了. 但从计算观点来看, n 量子比特情形下发生的事情更吸引我们. 此时, 一个量子态 x 就是 \mathbb{C}^{2^n} 中的一个单位矢量. 那么, 对于矢量空间单位正交基中的每个元素, 我们都用一个 n 比特字符串标记. 大多数 n 量子比特的量子态是纠缠的, 这意味着它们的振幅在某种意义上, 对 n 比特串的各处都是关联的. 酉矩阵给出了这样的系统的动力学描述, 它们由一系列两比特量子门构成². 为了表示作用在具体的量子比特上的量子门, 比如第 3 个或者第 7 个量子比特, 我们考虑一个矩阵元 U_{ij} 非零当且仅当 n 比特字符串中 i 和 j 等于 3 或者 7 的酉矩阵 U . 进一步地, U_{ij} 的值应该只取决于四个比特, 即 i_3, j_3, i_7 和 j_7 .

这样的线性代数形式看起来似乎很抽象, 但它还可以被用于描述经典的确定性计算和随机计算. 对于经典的确定性计算, 虽然有些浪费, 但确实用 n 比特字符串来表示. 譬如说对于长度为 2^n 的矢量, 其中只有一个分量为 1, 其余均为 0 的情形. 那么考虑一个 0/1 矩阵 M , 且其中每列只有一个 1, 这样的动力学描述即表示了从 x 到 Mx 的映射. 事实上, 随机计算³也是类似的. 只不过此时的状态矢量的各个分量均为非负实数, 且它们的和为 1. 对于任何仅有非负矩阵元, 且每一列矩阵元和为 1 的矩阵 M , $x \mapsto Mx$ 都是有效的状态转移. 在一般意义上, 这样的图象忽视了一个事实: 一些 n 比特的变换比另一些更容易. 那么为了表示单个操作, 即两比特操作, 我们也需要用一个矩阵 M . 它满足对于所有未作用比特 i, j , 有 $M_{i,j} = 0$, 即 $M_{i,j}$ 的值仅仅取决于其对应的比特 i 和 j 是否被作用. 这启示我们, 如果我们不对一个比特作用, 那么它不应该发生变化.

²这里说的是 n 比特量子门到 2 -unitary 的分解算法, 2 -unitary 和下面讨论和 2 比特量子门其实并不相同, 因为前者的对角线除了 i, j 所在行列都是 1.

³这里的随机计算指的就是 Markov 过程

Figure 2. When light, or any other wave, is split into two beams and recombined, the relative phase of the two beams determines whether they interfere constructively or destructively. Since the phase of light depends on how far it's traveled, this is a very sensitive way to measure distance. Similarly, quantum computers can split computations into multiple branches and recombine them to obtain either constructive or destructive interference.

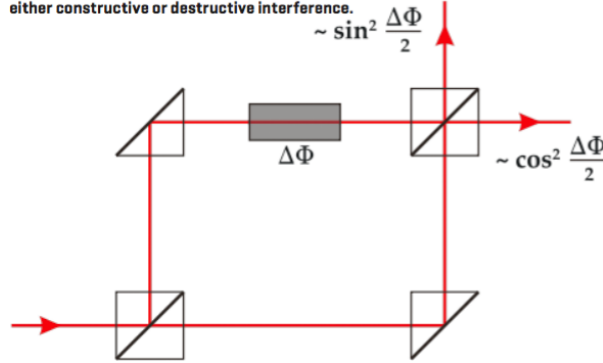


图 4: 可叠加性: 光的反射和折射

因此, 说到随机计算和量子计算最关键的不同, 它“仅仅”是把实数换成了复数 (甚至允许负实数), 以及用于度量态矢量模长的范数从 ℓ_1 换成了 ℓ_2 . 也就是说, 于量子态而言, 其态矢量各分量的概率幅平方和等于 1; 而概率矢量的分量的和为 1, 并没有平方. 然而, 在计算中, 可以用不同的相位作为不同分支这一事实, 意味着当我们重新组合它们的概率幅的时候, 概率幅可能得到增强 (即相长干涉) 或者削弱 (即相消干涉). 如果“0”用矢量 $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ 表示, 而“1”用矢量 $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ 表示, 那么非 (NOT) 操作能够表示 (不论量子或经典情形) 为 $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. 论及它的几何意义, 这是一个角度为 $\pi/2$ 的旋转. 然而, 只有量子计算机才能作用“非门的平方根”⁴, 这样的量子门即 $\pi/4$ 旋转:

$$\sqrt{\text{NOT}} = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$$

如果我们从量子态“0”开始, 然后作用一个 $\sqrt{\text{NOT}}$ 门, 那么我们会得到态 $\begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$. 要是我们测量这个态, 那么得到结果 0 或者 1 的概率都是 1/2. 然而, 如果我们对初态“0”作用两次 $\sqrt{\text{NOT}}$ 门, 那么我们得到的结果是“1”. 这展示了量子叠加和经典随机性之间的核心差异: 将一个量子态进行叠加, 那么它不会有任何不可逆的信息丢失. 当我们有 n 个量子比特的时候, 叠加 (superposition) 和干涉 (interference) 使得我们能得到更大的计算优势. 一个著名的例子就是 Grover 算法: 给定一个 n 比特二进制函数 f , 我们要找到一个满足 $f(x) = 1$ 的输入 $x \in \{0, 1\}^n$, 这里仅仅需要计算 f 大约 $2^{n/2}$ 次, 即平方级加速. 正是概率是概率幅的平方这一事实, 使得 Grover 算法得以可行. 因此, 在 2^n 个基矢量上的均匀叠加的量子态, 对应着每个基矢量的概率幅都是 $1/\sqrt{2^n}$. 更进一步地, 我们能够演示计算 f 造成的影响, 并且这样的影响是可比较的: 每个目标量子态 x (即 $f(x) = 1$) 的概率幅大致以 $1/\sqrt{2^n}$ 的速度渐进增长, 而这样的破坏仅仅出现在总概率幅非常大的时候. 因此, 总影响 (渐进时间复杂度) 的

⁴这里给出的就是 Hadamard 门, 再加上 Phase Gate S 就构成了 Clifford 门, 而 Clifford+T 就能够实现通用量子计算. 还有个有趣的事实, 实际上 Hadamard 门就是单量子比特的量子 Fourier 变换, 这或许从另一个角度说明了为什么 QFT 如此基本.

阶是 $2^{n/2}$; 或者更一般地, 对于 M 个解有 $\sqrt{2^n/M}$.

而用于分解整数的 Shor 算法的加速, 于经典计算而言更为戏剧性. Shor 算法有更为本质的经典部分, 即把素因子分解 (Factoring) 问题转化为更抽象的寻阶问题 (Period-finding). 这个问题输入是一个在集合 $\{0, 1, \dots, 2^n - 1\}$ 上函数 f , 满足性质 $f(x) = f(y)$ 当且仅当⁵ $x - y$ 能够整除某些隐含周期 r . 它的目标是寻找合适的 r . 既然 r 相对于 n 来说可以指数大, 如果我们把 f 视作一个黑盒, 那么经典计算机用于寻找它的时间自然是指数的.

然而, 量子计算机却能够概率幅近似表示为一系列 $\sqrt{r/2^n}$ 的量子态的叠加, 即考虑每个 $z, z + r, z + 2r, \dots$, 这里的 z 是某个随机选择的 $z \in \{0, 1, \dots, r - 1\}$. 迄今为止, 这很像一个概率算法, 选择随机的 x 用于计算 $f(x)$. 此外, 根据对 $f(x)$ 的, 选择考虑 x 的分布. 但算法的下一步是量子的, 并且应用了一类称为量子 Fourier 变换 (Quantum Fourier Transform, QFT) 的酉矩阵. 这个矩阵中的矩阵元为 $(y, z) = e^{2\pi i y z / 2^n} / \sqrt{2^n}$. 为了有效的执行算法, 我们需要经典快速 Fourier 变换 (Fast Fourier Transform, FFT) 的量子版本. 但是它们的相异之处同样是戏剧性的: 因为它是对量子态的概率幅进行变换, 而不是像经典的 FFT 一样作用在一串数上. 在这种情形下应用 QFT, 得到的叠加态 y 的概率幅约为 $\frac{r}{2^n}(1 + e^{\pi i \frac{y r}{2^n}} + e^{\pi i \frac{2 y r}{2^n}} + e^{\pi i \frac{3 y r}{2^n}} + \dots)$. 如果 $y r$ 能够近似地被 2^n 整除, 那么它的和会非常大; 反之, 如果 $y r$ 并不能近似地被 2^n 整除, 它会涉及到很多不同相位的复数, 不难验证, 此时它们倾向于互相抵消. 因此, 测量 y 会返回一个接近 $2^n/r$ 的倍数的答案. 最终, 通过经典连分数展开算法, 我们重新得到了 r .

Figure 3. Any function can be written as a sum of waves of different frequencies. The coefficients of this sum are called Fourier coefficients. The quantum Fourier transform (QFT) transforms a quantum state into one whose amplitudes are the Fourier coefficients of the original state. Measuring this state can be an incredibly efficient way to detect periodic structure.

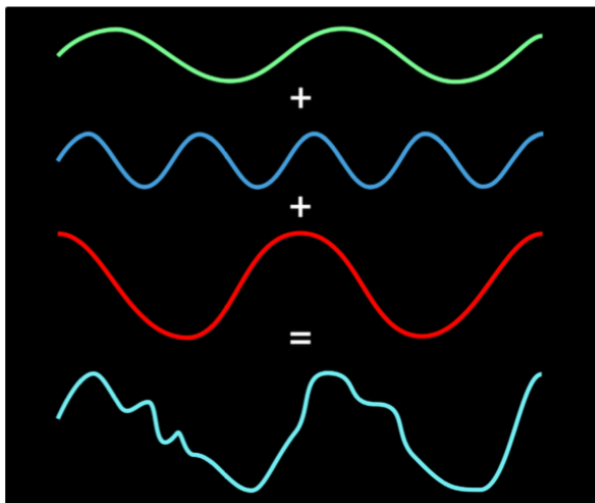


图 5: 波的叠加

一旦我们开发出了新的公钥密码系统来抵抗量子计算机的攻击, 量子模拟就在实践上逐渐变得更加重要. 而 Shor 算法所展示的量子计算的力量, 甚至远不能用令人惊讶来形容, 因为这一问题的解决和量子力学并没有明显的联系. 在 Shor 算法之后, 人们越来越难以相信

⁵熟悉群论的读者不难发现, 其实这里 $f(x) = f(y)$ 刻画了一个群, 所以 QFT 框架下更一般的就是隐含子群问题 (Hidden Subgroup Problem, HSP). Shor 算法对应的是 \mathbb{Z}_n 的情形, 而最著名的非交换群的例子就是 S_n 对应的图同构问题.

量子力学能够被经典计算机有效模拟。在此之前，很多科学家们寄希望于基于量子力学的模型，希望它能够控制那些反直觉特征，比如指数大的态矢量空间，并且要保持它们与我们的观测相容。但是现在，我们知道这些模型并不能简化量子力学自身（或者说，至少并不容易模拟），除非素因子分解问题（Factoring）和寻阶问题（Period-finding）在经典计算机上有更快的算法。这样的情形就像是 NP 完全性（NP-Completeness）的发展，即昭示了不同领域在求解 NP – Complete 问题的努力，在实际上是等价的。

Shor 算法和 Grover 算法是最著名的量子算法，但并不是唯一的两个。比如最近提出的一个量子算法，即用于求解大的线性方程组的 HHL 算法 [1]：给定一个矩阵 A 和一个矢量 b ，找到满足 $Ax = b$ 的 x 。然而，不像这一问题的经典算法，量子版本的 x 和 b 都是量子态（因而我们可以只用 n 个量子比特表示 2^n 维的矢量）。更进一步地， A 必须是稀疏的，并且对于给定的隐含表示满足给定任意行指标 i ，能够有效的找到所有的非零 A_{ij} 。这一限制使得它能够在理论上在多项式时间里解决指数大小的方程组⁶。然而，需要说明的是：算法的运行时间还取决矩阵 A 的条件数的大小，条件数是一个关系到经典计算机上求解线性方程组的数值稳定性的参数。找一个能够使用这一算法的自然情形，是一个引人注目的公开问题（Open Problem）⁷。




Shor's Algorithm (1994)	<ul style="list-style-type: none"> • Breaks RSA, elliptic curve signatures, DSA, El-Gamal • Exponential speedups 	
Solving Linear Systems of Equations (2010)	<ul style="list-style-type: none"> • Applications shown for electromagnetic wave scattering • Exponential speedups 	
Quantum simulation (1982)	<ul style="list-style-type: none"> • Simulate physical systems in a quantum mechanical device • Exponential speedups 	

图 6: 三个著名的量子算法。

另一个最近的算法上的进展是 Metropolis 采样算法（Metropolis sampling）的量子对应 [2]。对于经典情形，Metropolis 法是一种在指数大小的状态空间上，从难于分析的分布中采样的方法。事实上，这是一个通过使用随机性来得到指数级加速的例子，而刻画对计算模型的改变会使得它变得多么强大也并不令人着迷。它应用范围很是出乎意料，包括统计推断和关于非负矩阵的积和式（the permanent of a nonnegative matrix）的近似算法，但是它最初的发展，是对系统的热力学分布采样所引起的。如果一个状态 x 有能量 $E(x)$ ，那么 x 在温度 T 下的热力学分布与 $e^{-E(x)/T}$ 成比例，所以更低的温度会使得系统更难以转化到低能态情形。类似地，量子 Metropolis 算法也从热力学分布中产生量子态。就像它的经典版本一样，量子 Metropolis 算法花费更多时间去产生低温态，正如其往往解决更困难的优化问题一样。但是除了少数情形，证明其运行时间的严格的界（bound），在一般意义上是困难的。尽管

⁶熟悉 Sparse Hamiltonian 的读者可能会觉得，这里的限制就是 Dorit Aharonov 等人提出的 Sparse Hamiltonian 的条件。所以不难看出 HHL 算法和绝热量子计算间的渊源。

⁷下图右侧自上而下的人物依次是：Peter Shor, Aram Harrow, Richard Feynman。图片来自 Krysta Svore 在 QIP 2015 的 Slides。

没有形式化的证明, 我们仍然能够运行经典 Metropolis 算法, 并且经验地观察到它在很多情形下执行的很快. 而对于量子 Metropolis 算法的经验观察, 当然需要等到第一台真正的大规模量子计算机出现. 但是我们已经知道了足够的工具, 如果我们能够指出如何组合它们, 以此使用量子 Metropolis 算法作为子过程设计新的量子算法, 就像非负矩阵的积和式算法使用经典 Metropolis 算法一样.

人们仍然在不断地提出着使用量子计算机的新想法, 而一个令人兴奋地进展是, 我们找到了越来越多的只需要 10 或稍多量子比特的应用. 量子比特的数量足够小, 意味着它能够很容易被经典计算机有效模拟, 特别是当我们并不能展示量子比特间足够大的相互作用之前. 这样的量子计算设备能够被用于提高量子测量的精度 (比如原子钟, 或者测量引力波), 在网络中作为“量子中继器 (quantum repeater)”分发用于密码学协议的纠缠, 或者甚至用于构建能够组成任意尺寸孔径的望远镜阵列. 正如经典计算机使用, 远远超过了作为计算模型的 Turing 机看起来的那样, 量子计算设备的使用也会变得比我们现在的想象更加广泛.

3 (量子) 算法棱镜折射的科学

毋庸置疑, 大规模量子计算机的出现, 将会引起我们思考自然科学的方式的变化, 这样的变化是巨大且无法预测的. 但是, 即使没有真正的量子计算机, 我们到目前为止在理论上的进展, 仍然能够引起不少观念上的进步. 比如这样的想法就是一个主要的进步, 从物理学中分离量子力学的信息部分. 就像现在做的, 我们总是一起教授它们, 这样一来糅合了反直觉的概念和数学完备的图像, 比如说我们用偏微分方程中的 Schrodinger 方程来描述测量和纠缠, 就像生活在 \mathbb{R}^3 中的无穷维函数空间一样. 而概率似乎仅仅出现在教授统计力学的时候, 并且学生们学到的第一个分布往往就是理想气体的热力学分布. 相反, 如果我们先学到了量子力学关于信息的意义, 再用这样的框架来解释原子, 光子和其他物理学现象, 那么我相信量子力学会起到更大的作用.

刚刚起步的学生们, 并不是唯一能够从量子信息观点中获益的量子力学的实践者. 很多涉及量子力学的现象, 也与这样的一些问题相关, 比如熵、纠缠或者有物理上相关性的关联, 但它们最好以信息方式来描述. 寻找量子系统 (比如逐对相互作用的 n 个量子比特) 的最低能态就是个早年的成功例子. 对于经典系统, 这样的问题是 NP-Complete 的, 除非在特殊情况下, 比如这样的系统构成直线或者树的情形. 而对于量子系统, 能量最小化问题 (Energy-minimization problem) 是 QMA – Complete 的. QMA 代表了“量子 Merlin-Arthur 博弈 (Quantum Merlin-Arthur games)”⁸, 并且我们相信量子计算机很难解决它, 就像经典计算机难以解决 NP 中的某些问题一样.

⁸简单地说, 这里用到了交互式证明系统的概念. Arthur 具有有限的计算能力, 而 Merlin 拥有无所不能的计算能力. 对于具体的问题, Merlin 总是能想出证明, 并且费尽心机让 Arthur 觉得他的证明是对的. 类似 NP, QMA 也是让 Merlin 给出一个问题的证明, 再让 Arthur 检查证明是否正确. 区别包括两部分, 一是两人传递的信息不再是字符串 (比特串), 而是态矢量 (有叠加性); 二是 Arthur 的计算能力是量子 Turing 机的计算能力, 而不是经典 Turing 机的计算能力.

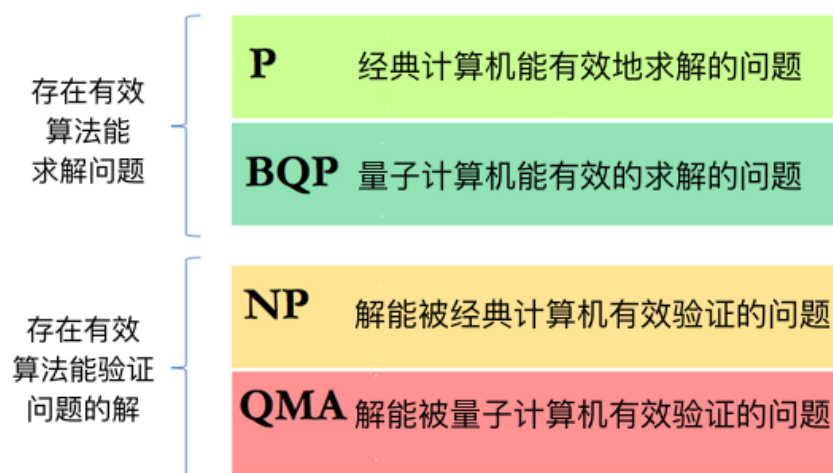


图 7: 量子计算的复杂性理论分层

这给出了很多根据经验观察, 总结而出的物理现象的理论判据, 比如说找玻璃的基态总是很困难的. 令人惊讶的是, 即使对于只有相邻点对有相互作用的直线情形, 能量最小化问题仍然是 QMA – Complete 的. 这与物理学家们的直觉相违背, 即一维的情形不总是容易的.

至于其他情形, 量子信息观点也给出了些正面的结果, 甚至是新的经典算法. 一个相关的例子是能隙 (Energy gap), 即系统的最低能级和系统的第二低能级之间的间隙. 从物理上说, 能隙大小对应于一次激发的难易程度. 对于一些物质的激发, 像是光子或者固体的振动, 它们不会产生物质变化; 而在半导体中移动的外部电子, 则会得到有效的质量. 从这样的图象中, 我们希望小的能隙会对应于长程关联 (Long-range correlation), 而大的能隙则意味着关联会迅速的减弱. 可现实中的结果却是更加微妙的. 对于一个可观测量, 我们考虑特定点的磁场强度, 当能隙很大的时候, 关联确实会迅速的减弱. 但是如果将系统状态作为整体考虑, 它仍然表现了长程关联. 下面我们来给出一个这样的情形的恰当类比, 考虑一百万比特的空间, 在其上固定一千个随机的一一映射函数. 如果这样的函数 f 是随机选择的, 那么有序对 $(x, f(x))$ 的互信息会接近极大值, 这意味着知道 x 会把 $f(x)$ 可能性的数量从 $2^{1,000,000}$ 降到 1000. 而在另一方面, x 中的一个独立的比特, 将会和 $f(x)$ 中的任意一个特定的比特几乎毫无关联. 即使我们把这样的随机函数被换成了有合适参数的扩展图 (Expander Graph), 这样的论断仍然有效. 在这样的直觉的指引下, 研究者们想出了扩展图的量子对应, 即用大的能隙和任意独立可观测量之间迅速减弱的关联来表示系统, 但是系统的不同部分之间的互信息, 在数值上仍然十分巨大. 为什么我们应该关系这些不同种类的关联, 除了对用理论预测真实世界中的可观测量的渴求? 在经典计算机上模拟量子系统, 就是这样一个令人兴奋的应用. 如果一个 n 量子比特系统中没有纠缠, 那么描述它所用的参数数量将不再是 2^n , 而是 $2n$. 如果这些量子比特被排成一条直线, 并且它们之间的关联被大致限制在某个很短的距离 k , 那么参数的数量的规模将会是 $n \cdot \exp(k)$. 因此, 控制量子系统中的关联, 也能够帮助我们有效的模拟它.

这条研究主线可以被视为一个大型项目的一部分, 即把量子系统划分能被经典计算机有效模拟的部分 (这么说是因为它们还是会一定程度上纠缠), 并且这也能够实现通用量子计算. 在另一方面, 对于一些系统来说, 它们的计算复杂性明显在经典模拟和通用量子计算之间. 这些系统包括不发生相互作用的光子, 也包括噪音比率过高的量子纠错码 (Quantum Error-correcting Code) 的作用, 但是太少以致于难以消除大规模纠缠的可能性. 分析这些边界情形的复杂性是公开问题 (Open Problem) 的一个令人着迷的来源.

量子信息观点在科学上的益处也不仅仅局限于量子力学。一些重要的问题表面上看起来和量子力学无关, 但和多维阵列的线性代数相关。比如说, 给定一个由数 A_{ijk} 组成的三维集合 (collection), 这里的 i, j, k 在 $\{1, \dots, n\}$ 中取值, 那么类似计算下面的最大奇异值问题有多难? 下式遍历了所有单位矢量 x, y, z .

$$\sum_{i,j,k} A_{i,j,k} x_i y_j z_k$$

对于这样的问题, 如果要求的计算精度在 $1/\text{poly}(n)$ 的话, 那么它是 NP-Hard 的。而对于大多数更真实的情况, 即只需要常数精度近似的话, 已知的唯一结果用到了量子技术, 就像大多数有希望的经典算法一样。量子信息观点的有效性是一个可能的理由, 这来自于此处的多维阵列的自然对应就是纠缠态; 并且随着我们对纠缠的量化理解日益深入, 它所导出的有关线性代数的结果的应用会愈加广泛。线性代数的重要性似乎是获益于理论计算机科学。这样的例子包括在布尔立方体 (Boolean cube) 上的 Fourier 分析, 和用图和矩阵的观点来看待相互作用, 它们因而混合了组合和代数的图像。在未来, 我希望我们关于线性代数和概率的观点的形成, 会越来越受到来自量子信息中的工具的影响。

最后, 量子信息对于计算机科学的绝大多数贡献都是理论意义上的, 因为现在并没有大规模量子计算机。但是一旦我们拥有了它, 我们希望理论科学家们能从中得到一些启示, 并加以解释, 就像经典计算机中那些启发式的成功例子, 如单纯形法 (Simplex)。举个例子, 我们在 Shor 算法中发现了周期结构, 并且它能够被用于获得一系列指数级加速。而在将来, 我们希望把类似寻阶 (Period-finding) 的工具用于数据分析, 这就像今天广泛使用的线性回归一样。绝大多数新工具在一开始都被认为是完全反直觉的, 但当我们深入理解了它们之后, 它们就昭示着科学而系统的看待世界的全新方式。

参考文献

- [1] A. W. Harrow, A. Hassidim, and S. Lloyd. Quantum algorithm for solving linear systems of equations. *Phys. Rev. Lett.*, 15(103):150502, 2009, [arXiv:0811.3171](#).
- [2] K. Temme, T. J. Osborne, K. G. Vollbrecht, D. Poulin, and F. Verstraete. Quantum Metropolis sampling. *Nature*, 471(7336):87–90, 2011, [arXiv:0911.3635](#).

延伸阅读

- M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2001.
- J. Preskill. Lecture notes for Ph219. <http://theory.caltech.edu/preskill/ph229/>
- D. Mermin. Lecture notes for CS483. <http://people.ccmr.cornell.edu/mermin/qcomp/CS483.html>