量子计算机的回顾与展望

Look Back and Prospect of Quantum Computer

苏运霖

(广州暨南大学计算机科学系　广州 510632)

Abstract This paper reviews the development of the quantum computers in the last twenty years. Ranging from the beginning of the idea until the deep investigation of its construction, it is said that they are more powerful than their present counterparts; even the first modest-sized quantum computer may make many encrypted data files insecure; with only a few hundred qubits it is possible to represent simultaneously more numbers than all atoms in the universe. The explications of these facts are made in the paper. After summarizing the important and major events happening in the last twenty years, this paper briefly explains the working principles of quantum computing and then looks forward the future.

Keywords Quantum, Quantum location, Eddy revolve

一、量子计算机的兴起

今天，当大规模集成电路的计算机正按照所谓的摩尔定律，以走马灯般的速度向着越来越高的档次发展时，大概谁都只能说，它仍然处于"如日中天"的鼎盛发展时期。然而，明天会是怎样？具体地说，十年或二十年后，计算机的发展会是怎么样？这肯定是人们普遍关心的问题。

二十年前，正当计算机工业正在为微芯片的未来而烦恼，人们思考着计算机最终每秒钟可以达到多高的速度。而这要产生更多的热量，硅片能否承受这样持续的烘烤？等等问题时，就已经有人在考虑替代当前计算机的新途径了。

1981 年，在奥斯汀的德克萨斯大学，著名的理论物理学家，被誉为黑洞之父的 John A. Wheeler 教授举办了一次聚会。在这个聚会上，IBM 公司的物理学家 Charles Bennett 和英国牛津大学的研究人员 Davis Deutsch 相遇并就他们共同感兴趣的计算的基础问题进行了深入交谈。从这个交谈中，Davis Deutsch 产生了这样一个想法，即今天的计算机的理论，是以牛顿的定律为基础的，而不是以量子理论所提供的对宇宙更基本的描述为基础的。从那时开始，Davis Deutsch 便投入这方面的研究工作。1985 年，他发表了关于量子计算机的头一篇论文。现在一般人都把它当作是这个领域的经典之作。在这篇论文中，他论述了计算机如何可以使用量子力学奇异的规则来运行，以及为什么这样的计算机根本上不同于今日通常的计算机，等等。

理论是实践的先导，在这个问题上再次证实这一论断的正确性。从那时开始的十五年后，量子计算机不再被看成是胡思乱想的空中楼阁，而被认为是计算机工业强有力的未来。争论的焦点也从它们是否真的能够实现，变成为什么时候能够实现。它们令人振奋的不仅仅是量子计算机要比通常的计算机更强有力，而且在于，它们能够求解以及模拟在通常的计算机上基本不可能的问题。

在巨大的潜在经济利益的驱动下，世界上最大的远程通讯和计算机企业都争先赞助这方面的研究计划。这些企业包括 IBM、HP、Lucent 技术、AT&T 以及微软等。更有一个总部设在纽约的 Magi Q 技术应运而生并把量子计算机作为它的主攻目标。

世界各国的科学家也都争先恐后，纷至沓来。由于 Deutsch 是在牛津大学，他们号称在那里有最强大的研究班子。但是美国自然不卖帐，因为美国的研究班子可以说遍布全国，而且他们不仅力量雄厚，而且还具有充裕的研究经费。在美国，实质性地开展这方面研究并取得进展的单位包括在科罗拉多州的博尔德(Boulder)的美国国家标准和技术研究所、在新墨西哥州的洛斯·阿拉莫斯国家实验室、麻省理工学院、加州理工学院、哈佛大学、马里兰大学，IBM、AT&T 等等。英国除牛津大学外，主要有在默尔文的国防评估和研究机构，等。欧洲包括瑞士、法国、荷兰、奥地利等，再有澳大利亚和以色列、日本也在尽力迎头赶上。

关于这方面的研究，主要集中在两个方面。一是有关它的工作原理的进一步研究。另一个是这样的计算机真正的物理实现。前者，由于 Deutsch 指出，量子计算机表示和对数进行运算的能力是随着它的位数而指数地增加的，因此 1994 年在新泽西州的 AT&T 贝尔实验室的 Peter Shor 证明，量子计算机在进行因子分解方面比通常的计算机快。对于通常的计算机来说，实现对于很大的数的因子分解是如此困难，因此密码的设计正是利用了这一弱点，设计出供政府部门和军队使用的密码。下面我们引用 D. E. Knuth 在介绍著名的 RSA 加密方案时关于这个问题的一个精彩描述[1]。

以下生成 p 和 q 的过程几乎是无懈可击的。比如说从 10^{80}和 10^{81}之间的一个真正随机的数 p_0 开始，寻找大于 p_0 的头一个质数 p_1；这将要求对大约 $1/2 \ln p_0 \approx 90$ 个奇数进行测试。在 50 次运算 P 之后，只要以 $>1-2^{-100}$的概率证实 p_1"大概是质数"就够了。然后在 10^{39}和 10^{40}之间选择另一个真正随机的数 p_2。寻找形如 kp_1+1 的头一个质数 p，其中 $k \geqslant p_2$，k 为偶数，而且 $k \equiv p_1 (\text{modulo } 3)$。在发现 p 是一个质数之前，这将要求测试大约 $1/3\ln p_1p_2 \approx 90$ 个数。质数 p 大约有 120 位。类似的构造可以用来找出大约 130 位长的另一个数 q。为更加安全起见，大概还应检验，无论是 p+1 还是 q+1，都不完全由小的质因子组成。现在乘积 N=pq 的数量级将是 10^{250}次方，它满足所有要求。就现时而言，要对这样一个 N 进行因子分解是不可想象的。

比如说，假如我们知道有一种方法可以在 $N^{0.1}$微秒的时间内对于一个 250 位数进行因子分解，这就等于 10^{25}微秒了，而一年之内仅有 31,556,952,000,000 微秒。因此为了完成这个因式分解，我们将需要超过 3×10^{11}年的 CPU 时间。即使一

个政府部门购买100亿台计算机并把它们全都投入到这个问题中来，那在把N完全分解成质因子之前，还需要花上超过31年的时间。当政府采购如此大量的计算机用于这一目的的消息一经走漏出去，人们又立即使用300位的N，而使上述努力前功尽弃。

Peter Shor的工作的意义，在于指出，一旦头一个中等规模的量子计算机开通，政府和它们的军队将被迫承认，它们的许多加了密的数据统统都变成不安全了。

另外一个很好的例子是由新泽西州默里山(Murray Hill)的Lucent技术贝尔实验室的Lov Grover给出的。他提出了使用量子计算机对于表进行查找的一个算法。问题是这样的：给定某个人的电话号码，要从一本电话目录中找出这个人的名字来。如果这个目录包含有N个条目，则根据线性查找的算法，在把这个人的名字找出之前，平均说来，你将需要查找N/2个条目。但是通过使用Gover的量子算法，平均说来，你将只需要查找N个条目。因此，如果这个电话目录有10000个名字，用量子算法将查找10000＝100次，而不是5000次。

对于量子计算机来说，一个必须解决的问题是探测和校正误差。不解决这个问题，量子计算机便无法发展。有些科学家曾经认为，这是不可能的，因为探测和校正误差意味着要度量一个量子系统的状态，而这会破坏它所含的信息。但是在20世纪90年代初，Deutsch证明，未必会这样。1994年，牛津大学的Andrew Steane和AT&T贝尔实验室的Peter Shor各自发现了误差校正算法。

在研制真正的量子计算机方面，第一个大的突破是在20世纪90年代中期。当时科学家们发现了使用核磁共振(NMR)技术进行计算的方法。

这里关键的思想是一个分子可以像一个微小的计算机那样运作。信息存储在分子中核旋转的方向上，每个核都持有一个量子位。核旋转之间的交互，称为旋转耦合，起作传递逻辑操作的作用。在一个强磁场之下，这些核以依赖于它们的化学环境的频率环绕磁场的方向进动。

核磁共振像是一个棘手问题的梦幻般的解。核被自然地同外部世界的噪音隔绝开来。因此可维持一致性(见下面解释)许多秒的时间，从而有足够的时间来实现几百个逻辑操作。此外，核磁共振是一个成熟的技术，已经被用于图像和化学分析许多年了。

但是这个技术也有其严重的局限性，那就是，单分子不能产生出足于被观察的信号。核磁共振的实验必须涉及极大数量级(为10^{23})的分子，才能使得磁感应信号大到足于被挑选出来。而这些分子通常分布在溶剂当中，因此最初的量子计算机实际上都有流体的心。又有一个问题提了出来，为了开始一个计算，必须知道计算机的初始状态。但是在室温之下的一个材料中，上旋转状态和下旋转状态几乎同样随机地分布。换言之，这许多计算机每一个状态在溶液中无法得知，也就表示任何随后的计算毫无意义。然而，1997年加利福尼亚圣·约瑟(San Jose)的IBM Almaden实验室的Isaac Chuang和麻省理工学院的Neil Gershenfeld发现了解决这个问题的方法。他们用这个方法建立了一类人工的基础状态(在一个双量子位系统中表示00)。于是计算就可由此开始。与此同时，麻省理工学院的David Cory和哈佛大学的Amr Fahmy和Timothy Havel发现，通过用无线电脉冲碰撞样品，可以从除了基础状态之外的所有其它状态把信号"阻塞"掉。

然而，基于液体的核磁共振的量子计算机注定没有生命力，因为它们工作在室温之下，又工作在液体中，因此其能力受到诸多限制；如随着计算所涉及的量子位的增加，它们所产生的读出信号按指数律垂直下降等等。因此科学家们开始研究在固体状态下单个原子上核磁共振式的操作。在这方面，马里兰大学Bruce Kane的方案特别受到注意。他的思想是把磷原子的一个阵列埋入硅中，上面铺上一个绝缘层。在这个绝缘层上面又放置电子极阵列，这每个电子极都可施加电压到它下面的原子上。这个设置的巧妙方面是Kane提出的控制每个原子旋转的方案。

问题在于如何把这个装置真正地构造出来。Kane的合作者们已经在进行着这方面的研究。而与此同时，在澳大利亚新南威尔士大学的量子计算机技术中心，以Robert Clark为首的一个研究班子也在积极开展着研究工作。他们希望克服Kane的装置所面对的许多障碍。Clark相信，在一段适当长时间内，有少量量子位的量子计算机就可能研制出来。

一个更先进的技术是量子点阵。实质上就是保持离散个数的电子，一个半导体陷波器。自20世纪90年代初人们就已经开始对它的研究。到1998年，IBM的David DiVincenzo和瑞士伯塞尔大学的Daniel Loss提出使用量子点阵作为一个量子计算机的结构部件。自那以后，利用点阵的量子性质来进行计算的各种各样方案就陆续地被提了出来。

以上概述的是以核磁共振为基础的量子计算机技术。还有另一个技术，在公众眼里，它没有核磁共振那么受注意。这是使用离子陷波器来构造量子逻辑门的技术。它是由奥地利的Innsbrucik大学的Iqnacio和Peter Zollar于1995年首先提出的，在离子陷波器中隐含的技术已被用于光谱学上，并且用于改进时间和频率标准。但要用于量子计算，还需要有巨大的进展才行。目前，全世界大概至少有五个研究小组在进行离子陷波器量子计算机的研究，其中国家标准和技术研究所的David Wineland所领导的研究班子被认为处于领先地位。

二、量子计算机的工作原理

人们自然关心量子计算机是怎么工作的？它的工作原理如何？以下对于这个问题作一简要的概述。

数字信息看起来像是无所不在的东西。二进制代码的0和1可以很容易地被测量，被复制和到处传送。但若把一组信息赋给量子的粒子，则它就具有了量子世界的奇特的性质。量子信息的基本单位是量子位，英文为qubit，它同其经典对立面二进位十分不同。

一个直流电磁场中有一个粒子的放置类似于围绕一个场的进动的一个旋转陀螺。在这样一个场中，粒子具有两个状态之一，即上旋转或下旋转，它们分别代表数字逻辑中的0和1。在一个旋转状态下的一个粒子可以通过垂直于这个磁场的一个无线电频率脉冲推向另一个状态。一个正确频率和持续时间的脉冲将使该旋转完全触发(往上)。而一个较短的正确频率的脉冲将把该旋转翻转成上和下状态的迭加(往下)。也就是说，把电子放在鬼使神差的对偶位置，称作状态的迭加，其中粒子同时具有向上旋转和向下旋转的状态，即同时为0和1。因此利用量子位进行计算，你可以同时对0和1进行计算，即以1的代价，获得2个计算的结果。

乍一看去，这并没有什么了不起的，但若加上更多的量子位和更多的数，则它就变得十分有吸引力了。如上所述，一个量子位可以是两个状态0和1的迭加，两个量子位就成为四

个状态 00、01、10 和 11 的迭加，也即同时表示四个数。这个增加是指数地增长的。因此当有 m 个量子位时，就有可能并行地实现对于 2^m 个数的计算。上边谈到过的 10^{250} 这么大数的因式分解，如果我们有 250 位量子位，那它就可以并行地进行对 2^{250} 个数的计算。要进行 10^{250} 这么大的数的计算，仅仅需要使用 $<(2^{250})^4$ 的计算就够了。由于对于 2^{250} 个数的计算是在很短时间完成的，因此对于 $(2^{250})^4$ 个数的计算，也不会有多长时间，决不会象上边 D·E·Knuth 所计算的用 100 亿台计算机仍然花 31 年以上的时间。这就是为什么我们说，用几百位的量子位，就有可能同时表示比宇宙中所有原子个数还大的数；以及头一个中等规模的量子计算机，就能使许多加了密的文件变成不安全了，因为对它们的破译，那不过是几秒钟的事，顶多也不过几分钟罢了。

然而，由于这种并行计算，同时计算出许多个数运算的结果，就存在如何得到所需要的答案的问题。一个简单的度量，可以破坏迭加，使系统保持在一种状态或另一种状态之下。因此，我们的目标就是去设计出这个简单的度量，使系统保持的状态就是我们所希望的答案的状态。然而事先很少有可能确定出这将是哪一个状态。通过利用量子干扰现象，可以解决这个问题。每个迭加状态都有同它相关联的概率，而这个概率有着类似于波的特性——它能破坏性地或建设性地干扰其它状态的概率。因此得到一个计算所希望的答案意味着以这样一种方式处理信息，即对不想要的解进行破坏性的干扰。最后，仅仅保留所需要的状态，或者一些或多或少地需要的状态。这个过程称为量子算法。这个算法的设计自然是对于物理家、数学家以及计算机科学家的尖锐挑战。一个最终的度量于是给出所想要的答案。而在有几个最终状态的情况下，一系列的度量给出它们的概率分布。从这些分布，也就可以计算出所想要的答案来。

这里还要介绍量子信息所引起的疑惑问题。早在 20 世纪 30 年代，科学家们对于量子力学所预测的结论是否有其现实存在性，或者说量子力学的奇怪结论是否由于理论上的某些缺陷引起的，这样一些问题展开了激烈争论。艾尔伯特·爱因斯坦就不相信，宇宙是如量子力学所断言的那样构成的。因此，他和他的同事 Bovis Podolsky 及 Nathan Rosen 一起，设计了一个思维实验，试图找出新理论的漏洞。

这个思维实验专注于粒子的行为特性，按照量子力学，它们是以经典的世界没有类似物的一种深奥的方式缠绕在一起的。如果刺激其中一个，则似乎是，另一个立即感受到影响。不论它们可能相距多远。但这样一来，就可能涉及到这个过程中的两个粒子之间存在超光速的信号传输——这是不可能的。他们的这个结论称为 EPR（爱因斯坦-博多尔斯基-罗森）谬论，而缠绕的粒子称为 EPR 粒子偶。

这个争论最后是由靠近日内瓦的欧洲粒子物理实验室 CERN 的 John Bell 和法国物理学家 Alarn Aspect 平息下来的。他们证明，量子世界的邋罗双胎，即 EPR 偶，确实有量子力学所预测到的行为特性。然而实验也证实，没有超过光速的信号，缠绕不能用于超光速的通讯。EPR 偶之间不是通讯，而是共享相同的存在，相同的密度。缠绕是在量子信息处理中所利用的关键现象之一。今天环绕全球，几乎每天都在进行着 EPR 的实验。

如果说相对于 10 年前，建立缠绕和迭加已经成为普通的事情，但量子信息仍然是脆弱的东西，和环境通常的交互就会破坏它们所包含的量子位和信息，这称为去一致性（与此相对应的一致性，是一个量子位保持诸如迭加这样量子特征的能力）。如果要把量子信息传送到计算机科学的世界，那就需要针对去一致性的误差校正过程。这个问题，如前边所说，已经由牛津大学的 Andrew Steane 和 AT&T 实验室的 Peter Shor 解决了，他们设计了实用的量子误差校正算法。

计算的初始化问题，如前所述，也已经得到了解决方法。现在我们简单地列出量子计算机有关的重要术语：

量子位：用于量子计算的信息单位，它不同于通常的二进位，在于它可进行值的迭加。

旋转：粒子的量子力学的性质。在某些情况下它仅能取相互排斥的值，它广泛用于核磁共振中。

迭加：如果像一个粒子这样的一个物理系统被发现处于一个以上的状态之中，而且它的状态是未知的，则说这个系统处于这些状态的迭加中。也就是说，如果有两个可能的状态，则说系统同时处于这两个状态中直到它的状态真正被度量为止。这样的度量使系统瓦解成一种状态或另一种状态。

缠绕：处于不确定状态的两个量子系统链接在一起的状态，使得对于一个系统的操作同时也对另一个系统进行操作。

一致性/去一致性：前者是一个量子系统保持状态迭加的能力。去一致性是同环境进行交互的过程，以破坏迭加，迫使系统进入一种状态或另一种状态。

振动状态：在一个线性离子陷波器中离子的集体运动的量子化状态。振动状态可对一个量子位进行编码，并且在计算过程中用于把诸离子链接在一起。

远程移动：使用缠绕的粒子在两个部分之间的通讯。通过缠绕，一个粒子的状态可以被传输到它缠绕的另一个远距离的粒子，以下将对它再作描述。

此外，为了进行有用的计算，量子计算机还必须能够实现任何逻辑运算，对于量子计算机来说，有两个逻辑运算。从这两种逻辑运算出发，就可导出所有其它的运算来，它们稍微类似于经典计算机中的与门与非门。一个运算涉及转动一个量子位，另一个运算称为受控的非门，它类似于可控制的反转器线路。在这样一个门元件中，一个量子位，即控制量子位的状态，确定第二个量子位，即输入量子位的最终状态是否将为一系列的正确频率脉冲所翻转。

三、量子计算机未来的展望

在我们展望量子计算机的未来之前，有必要先来谈谈量子互联网的概念。在科学家们研究量子计算机的各种各样问题和各种各样想法的过程中，他们终于又认识到，如果要使在不久之后量子计算机变成为有用的，就有必要把一些小的量子计算机联成网络，然而从一个位置向另一个位置传送量子信息是十分复杂的，如果实际地移动量子位，就可能受到一致性的影响。1993 年，新泽西州约镇海特斯（Heights）IBM 托马斯·J·沃森实验室的 Charles Bennett 和他的一些同事提出一个不同的选择：远程移动。

远程移动利用了在宇宙中一点和另一点之间建立起来的缠绕的深入链接。Bennett 论证，这个缠绕起着一类电话线的作用，沿着它可以发送量子信息。换言之，就是要建立粒子的缠绕偶，并且在保持一个的同时把另一个传送给接收者，这个过程以这样一种方式链接两个点，即它允许从一个量子位向另一个量子位进行量子信息的交换。

Bennett 和他的同事们等待了四年才看到他们的预言被证实，1997 年在奥地利 Innsbruck 大学的一个小房间里，由 Anton Zeillinger 领导的一个小组实现了第一个远程移动实

基于概念空间的文本语义索引*)

Text Semantic Indexing Based on Concept Space

李 源[1] 郑 毅[2] 何 清[2] 史忠植[2]

(中国科技大学研究生院计算机学部 北京 100039)[1]

(中科院计算技术研究所智能信息处理开放实验室 北京 100080)[2]

Abstract A new method of using cluster analysis to establish a text semantic indexing based on concept space is presented in order to get the needed information promptly from a great number of text documents. The keywords with high semantic proximity are direct clustered, an indexing based on the semantic proximity between keywords is set up on text concept space, and a strategy of indexing expand is presented. Therefore corresponding text files can be obtained after users input keywords. According to the results of experimentation, it is clear that the indexing efficiency and accuracy have been improved using this method.

Keywords Text indexing, Concept space, Cluster analysis, Direct clustering, Semantic proximity

1 引言

据统计,在现今的联机存储信息中,80%以上的信息以文本的形式存在。信息的多元化、复杂化,致使信息的自动索引成为急需解决的问题。本文研究的内容是建立一个基于概念空间的文本语义索引。目前的文本索引都是建立在文本空间,或关键词空间上的,而建立在概念空间上的索引具有条理清晰、人机界面友好、符合通常检索习惯等许多优势,这也是文本语义索引发展的方向。另外,在建立文本索引的过程中,国内外大多使用 Hopfield 神经网络联想的方法[5,6,8],本文首次使用直接聚类法代替了 Hopfield 神经网络联想功能,这样使得索引具有很好的可扩展性。基于语义关联度的文本索引可以广泛应用于 Internet 搜索引擎、数字图书馆、电子商务等众多领域中。建立文本索引的过程主要有以下几部分:

1)对文档分类,建立文档的概念空间,在概念空间的层次上组织文档并确定文档中出现的关键词。

2)分析概念空间中的文档,计算关键词的语义关联矩阵,标识出任意两个单词的语义关联度。

3)对关键词根据语义相近的原则聚类,建立分层索引。采用直接聚类法对关键词聚类,返回同类关键词共现率较高的文档。

为了对文本索引的准确性以及算法的可行性进行验证,选取 1061 篇足球类文档,对文档切词、分词后得到 876 个关键词,在此基础上建立起一个完整的基于概念空间的文本语义索引,并进行了检索试验。实验证明,这种索引能够大大提高用户的检索效率和准确度。

2 文档概念空间

文档的概念空间是指对文档自然聚类后所得的一种概念的层次结构,简单地说是一个树型的结构。从一个根节点依次地打开,逐层得到最终的叶子节点,形成一个空间的结构。它的叶子节点是一篇篇文档,各层的根节点是一个个从文档中提取出的概念,也就是一个个能概括其子节点共同特征的关键词。概念空间与目录的性质有些相同,但它是机器通过学习,由下而上自动生成,而不是由人工逐层分类、自上而下构成的。生成概念空间的方法如下:

1)采用传统的向量空间模型的方法来实现文本的数学抽象。

*)本文得到国家自然科学基金资助(课题号 60073019、69803010)。李 源 研究生,主要研究方向:人工智能。

验。Zeillinger 使用光子作为他的研究对象,并使光子在他们实验室的一边移动到另一边,但距离仅为一米。然而三年多以后,Zeillinger 等已经工作在第二步了,他们把光子传送了超过一公里之遥。

在 Zeillinger 等人的突破之后,Cirac 和 Zoller 提出,远程移动可以变成量子互联网的一个基础。2000 年 3 月,麻省理工学院的 Seth Lloyd 和 Selim Shahriar 以及马萨诸塞州林肯市的美国空军实验室的 Phillip Hemmer 提议,在光纤上传送缠绕的光子到包含冻原子的节点上,这些原子将吸收光子,因此将保存缠绕。这个缠绕然后可用于误差校正、远程移动以及许多其它有价值的应用。一些研究小组正在对这个思想开展研究,其中包括加州理工学院的 Jeff Kimble 和在洛杉矶的加州大学的 Eli Yablonovitch。他们希望,在 10 年内有包含三个节点的量子计算机网络投入运行。因此,我们作如下展望:

10 年内,肯定将会有真正意义下的量子计算机问世,人们将不会像现在这样对于量子计算机感到陌生。

然而,由于对于量子计算机的使用要求全新的程序设计技术,如前边所说的量子算法,误差校正算法,特别是真正用于解决确定问题的算法,因此它的普及应用将需要更长得多的时间。

量子互联网和分布式量子计算机系统,是使量子计算机系统走向普及达到当前通常计算机发展水平的必经之路。

参 考 文 献

1 Knuth D E. The Art. of Computer Programming. Addison-Wesley. 1998,2

2 Mullins J. The topsy turvy world of quantum computing. IEEE Spectrum,2001(2):42~49