



基于“DNA 折纸术”设计哈密顿路径问题的解决方案

俞洋^①, 苏邵^②, 晁洁^{②*}

① 上海科技管理干部学院电子信息系, 上海 201800

② 有机电子与信息显示国家重点实验室培育基地&信息材料与纳米技术研究院; 先进生物与化学制造协同创新中心, 南京邮电大学, 南京 210023

*通讯作者, E-mail: iamjchao@njupt.edu.cn

收稿日期: 2015-03-11; 接受日期: 2015-03-18; 网络版发表日期: 2015-05-19

doi: 10.1360/N032015-00035

摘要 哈密顿路径问题是著名的 NP-完全问题. 本文基于“DNA 折纸术”提出了一个通过 DNA 纳米结构的自组装找出最短哈密顿路径的解决方案. 利用“DNA 折纸术”可以折叠出具有固定大小的长方形 DNA 纳米结构, 这些结构可用来编码哈密顿路径图中的顶点和路径. 这些折纸结构具有黏性末端, 可以在溶液中通过分子自组装直接连接起来, 从而产生有向无权的不同大小的纳米结构. 利用磁珠筛选和电泳等分子生物学手段, 可以找到对应于只经过图的顶点一次的最短有向哈密顿路径. 该解决方案具有高度并行性, 是一种很有潜力的哈密顿路径问题解决方案.

关键词

哈密顿路径
DNA 折纸术
磁珠筛选
原子力显微镜

1 引言

于 20 世纪 60 年代初, Feynman 首次提出分子计算的概念. 生物科学技术的快速发展, 特别是人类基因组计划的实施以及 DNA 纳米技术的发展, 极大地促进了分子计算的发展. 1994 年, Adleman^[1]首次创立 DNA 计算模型, 解决了有向哈密顿路径问题, 拉开了 DNA 计算的序幕. 近年来, 人们提出了诸如粘贴模型^[2,3]、剪接模型^[4]等多种 DNA 计算模型算法, 并且这些模型已在实验室中得到了证实^[5~7]. 研究表明, DNA 计算无论在理论上还是应用模型的建立上以及实验与检测技术上都取得了惊人的进展.

除了传统的 DNA 分子结构被用作计算材料外, 其他的 DNA 分子结构也被用于分子生物计算, 如质粒 DNA 分子^[8]、分子信标结构^[9,10]以及 DNA Tile^[11]等. 随着 DNA 分子结构研究的不断深入, 人们提出

了利用 DNA Tile 构建各种简单分子器件的思想^[12]. 通过对 DNA 分子结构自组装过程的研究, 人们不断挖掘了 DNA 计算在解决组合优化方面的潜力. 1998 年, Winfree^[13,14]证明了基于自组装技术的 DNA 分子计算模型具有和图灵机一样的计算能力. 2002 年, Adleman 等^[15]通过对自组装 DNA 计算模型的复杂度研究给出了有关树状自组装的程序复杂度的界限. 2007 年, Brun^[16]基于 DNA Tile 的相应设计从理论上给出了两种实现二进制加法和乘法运算的 Tile 自组装系统; 次年, 他又给出了路径寻找问题以及可满足性问题的自组装计算模型的解决方案, 该方案将所需要的 Tile 集大小由 $O(n^2)$ 降为常数个^[17]. 研究表明, DNA 计算无论在理论上、应用模型的建立上, 还是实验与检测技术上都取得了惊人的进展. 从已获的研究成果来看, DNA 计算模型在解决图与组合优化中的一些困难的 NP-完全问题上具有一种天然的优势^[18~20].

哈密顿路径问题是著名的 NP-完全问题. 在 Adleman 用 DNA 分子找出一个有向图中的哈密顿路径之后, 日本 Tokyo 大学的 Morimoto 等^[21]又给出了该问题的表面计算方式, 刘文斌等^[22]对于赋权图的哈密顿路径问题进行了研究, 设计了一个基于试管的赋权有向图中的问题的计算模型. 文中结合 DNA 纳米折纸技术给出了一个用 DNA 纳米结构组装找出最短哈密顿路径的方法. 首先用 DNA 长方形折纸编码图中的顶点和边, 然后将这些折纸结构放入缓冲溶液中自组装, 利用黏性末端将它们连接起来, 从而产生有向无权的不同 DNA 折纸结构. 利用磁珠筛选和电泳等常规分子生物学手段可以寻找到对应于只经过图的顶点一次的最短有向.

2 哈密顿路径问题描述

对于一个由 n 个顶点和 m 个有向线组成的图 G , 存在若干经过顶点的路径. 在这些路径中是否能找到一条路径, 从顶点 0 出发到顶点 $n-1$ 结束, 途中经历每个顶点一次且只有一次. 这样的途径就称为哈密顿路径. 哈密顿路径问题是组合数学中一个典型的 NP-完全类问题, 即该问题的计算量极大, 超出了多项式算法的范围. 例如, 图 1 中就存在一条哈密顿路径, 其是 $0 \rightarrow 1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 6$. 如果将 $2 \rightarrow 3$ 这一路径取消掉, 则该图形无有效的哈密顿路径. 解决哈密顿路径一般步骤为: (1) 合成出所有可能的路径; (2) 筛选出包含始点和终点的路径; (3) 如果该图存在 n 个顶点, 从(2)中筛选出共 n 个顶点的路径; (4) 从(3)中筛选出包含每个顶点的路径; (5) 通过以上步骤得到的路径就是所谓哈密顿路径, 否则无解.

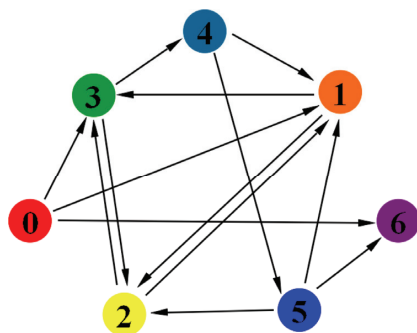


图 1 哈密顿路径问题选取的含有 6 个顶点以及顶点间具有有向路径的图

3 基于 DNA 纳米折纸结构的算法实现

“DNA 折纸术”是 Rothemund^[23]提出的一种全新的 DNA 自组装的思路. 他利用一条 DNA 长单链和几百根 DNA 辅助短链相互杂交形成相对复杂的二维纳米级形状和图案. 而 Shih^[24]、Yan^[25,26]和 Yin^[27] 3 个研究组将这种方法成功用于三维复杂图形的组装. DNA 折纸技术是 DNA 纳米技术发展的又一里程碑. 本文中, 我们利用 DNA 折纸技术对每个顶点进行设计编码, 由于 DNA 折纸图形足够大因而结果可从原子力显微镜或者电镜照片中得到验证. 利用 DNA 自组装来计算通常有 3 个步骤: (1) 用分子组件来编码问题; (2) 通过组装这些组件来生成问题的可能解; (3) 从组装的组件中来提取答案.

3.1 编码

文中使用的图如图 1 所示, 它包括 7 个顶点 0, 1, 2, 3, 4, 5, 6, 以及它们之间的关系 $0 \rightarrow 1, 0 \rightarrow 3, 0 \rightarrow 6, 2 \rightarrow 1, 4 \rightarrow 1, 5 \rightarrow 1, 1 \rightarrow 2, 1 \rightarrow 3, 3 \rightarrow 2, 5 \rightarrow 2, 2 \rightarrow 3, 3 \rightarrow 4, 4 \rightarrow 5$ 和 $5 \rightarrow 6$. 图 2(a)显示了我们如何用 DNA 纳米折纸结构编码这个图. 每个顶点对应一个带有标号的 DNA 纳米折纸结构 (图 2(b)). 该结构包括一个月牙形磁珠筛选位点和与其他点联系的手臂链. 手臂链之间通过黏性末端进行杂交匹配即可形成图 1 对应的各种路径 (具体序列见支持信息). 以 1 号点为例, 它与其他点结合形成的二聚路径为 $0 \rightarrow 1, 2 \rightarrow 1, 4 \rightarrow 1$,

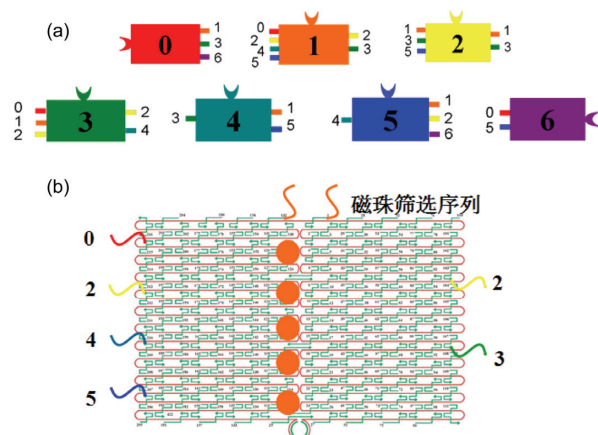


图 2 (a) 针对图 1 路径问题的 DNA 纳米折纸结构编码图; (b) DNA 纳米折纸结构设计图 (以顶点 1 为例): 长方形折纸上方的磁珠筛选用的两条链, 折纸左右两边与其他结构相互配对的手臂链, 折纸中间的修饰数字 1

5→1, 1→2, 1→3. 而三聚路径就是上述二聚路径的排列组合, 分别为 0→1→2, 0→1→3, 2→1→2, 2→1→3, 4→1→2, 4→1→3, 5→1→2 和 5→1→3. 每个点都可以形成多个路径, 只要黏性末端裸露在外就会与其匹配的点进行结合, 因此对应图 1 的路径就由 DNA 折纸结构的 n 聚体体现出来.

3.2 解决方案

对应一般解决哈密顿路径步骤, 利用 DNA 纳米折纸术的解决方法分以下几个步骤进行(图 3):

(1) 将每个顶点对应的带有黏性末端的 DNA 纳米折纸结构进行退火, 这是一种随机的自由结合的反应, 当它们充分反应后形成对应图 1 的所有路径, 即随机生成问题的所有可能解. DNA 分子的大量冗余的特点保证了对于一定的问题一定可以找到所有可能的组合, 当然也包括所有的可行解.

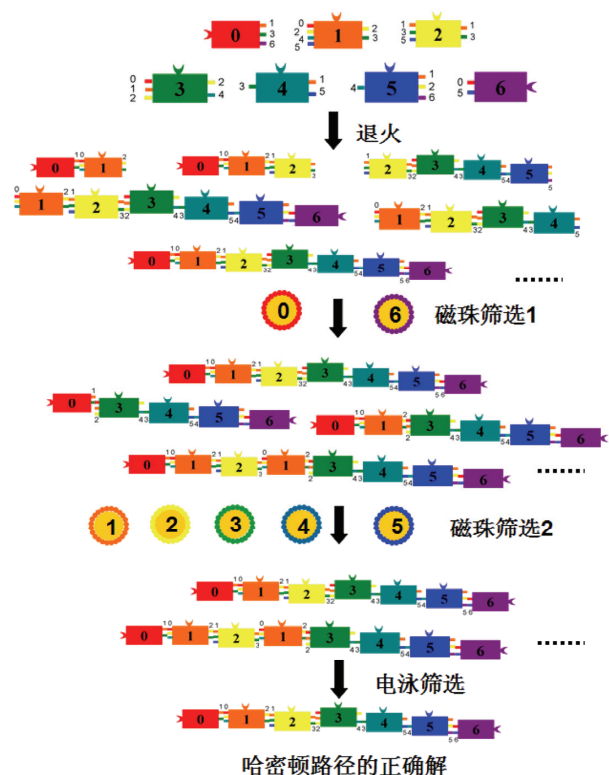


图 3 利用 DNA 纳米折纸术解决哈密顿路径问题的详细步骤示意图

补充材料

本文的补充材料见网络版 chem.scichina.com. 补充材料为作者提供的原始数据, 作者对其学术质量和内容负责.

(2) 寻找那些从 0 开始到 6 结束的路径. 将每个顶点对应的磁珠进行特异性互补 DNA 链修饰, 通过该 DNA 与顶点对应 DNA 纳米折纸结构定位杂交组装, 将路径中带有 0 和 6 顶点的路径筛选出来得到初步纯化的结果.

(3) 同理利用每个顶点对应的磁珠, 将路径中含有从 1 到 5 的每个顶点的路径逐一筛选, 得到二次纯化的结果.

(4) 对于形成了所有可能路径的初始溶液进行寻找指定长度路径的操作. 该问题具有 7 个节点, 通过在 1% 琼脂糖凝胶电泳层析分离出长度为 7 个 DNA 纳米折纸结构的路径, 即那些经过的顶点个数为 7 的那些路径.

3.3 识别答案

对于图 1 问题的哈密顿路径只有 1 条, 即 0→1→2→3→4→5→6. 由于 DNA 纳米折纸结构的大小为 100 nm×70 nm, 原子力显微镜的分辨率足够进行观测, 因此对以上操作所得到的分子将使用原子力显微镜进行识别, 可以通过其长度进行判断, 以确定最终得到的路径是否为正确解. 图 3 显示了正确解的结果可由原子力显微镜下测试出来.

4 结论

哈密顿路径问题有着广泛的理论与应用研究意义, 各种 DNA 计算方法都将其作为研究目标. DNA 纳米自组装技术为 DNA 计算的发展提供了新思路, 各种 DNA 自组装计算模型也被用于解决不同的 NP 问题. 这里, 采用 DNA 纳米折纸结构编码信息, 借助于纳米结构之间的黏性末端进行自组装, 给出了一种非确定性的哈密顿路径模型. 通过创建数以亿计的参与计算的 DNA 纳米折纸结构, 该算法可以并行的测试每种可能的路径. 对于给定 n 个顶点和它们之间相互联系图形, 哈密顿路径的存在与否以及存在几条哈密顿路径都可以通过相应的 DNA 纳米结构表达出来. 这些结构可以直接由原子力显微镜读出. 该方法具有高度的并行性和解决哈密顿路径问题的潜力.

致谢 本工作得到江苏高校优势学科建设工程资助项目(YX03001)、教育部创新团队(长江学者和创新团队发展计划)(IRT1148)和南京邮电大学引进人才科研启动基金(20140175)资助, 特此一并致谢.

参考文献

- Adleman L. Molecular computation of solutions to combinatorial problems. *Science*, 1994, 66: 1021–1024
- Roweis S, Winfree E, Burgoyne R, Chelyapov N, Goodman M, Rothmund P, Adleman L. A sticker based model for DNA computation. In: Landweber LF, Baum EB, Eds. *DNA Based Computers II*. Princeton: American Mathematical Society and DIMACS, 1999, 44: 1–27
- 许进, 董亚非, 魏小鹏. 粘贴 DNA 计算机模型(I): 理论. 科学通报, 2004, 3: 205–212
- Boneh D, Dtmworth C, Lipton RJ. Breaking DES using a molecular computer. In: Lipton RJ, Baum EB, Eds. *DNA Based Computers*. Princeton: American Mathematical Society and DIMACS, 1996. 37–65
- 许进, 李三平, 董亚非, 魏小鹏. 粘贴 DNA 计算机模型(I): 应用. 科学通报, 2004, 4: 299–307
- 王延峰, 崔光照. 粘贴 DNA 计算模型的几种分子逻辑门的实现. 计算机工程与应用, 2006, 1: 31–33
- Csuhaj VE, Kari L, Paun G. Test tube distributed systems based on splicing. *Comput Artif Intell*, 1996, 15: 21–41
- Head T, Kaolan PD, Bladergroen RS, Breck KD, Lommersed PHM, Spaink HP. Computing with DNA by operating on plasmids. *Biosystem*, 2000, 57: 87–93
- Sakamoto K, Gouzu H, Komiya K, Kiga D, Yokoyama S, Yokomori T, Hagiya M. Molecular computation by DNA hairpin formation. *Science*, 2000, 283: 1223–1227
- Tyagi S, Kramer FR. Molecular beacons: probes that fluoresce upon hybridization. *Nat Biotech*, 1996, 14: 303–308
- Seeman NC. Nucleic-acid junctions and lattices. *J Theor Biol*, 1982, 99: 237–247
- Mao C, LaBean TH, Reif JH, Seeman NC. Logical computation using algorithmic self-assembly of DNA triple crossover molecules. *Nature*, 2000, 407: 493–496
- Winfree E. Algorithmic self-assembly of DNA. Doctoral Dissertation. Pasadena: California Institute of Technology, 1998. 12–24
- Winfree E, Liu FR, Wenzler LA, Seeman NC. Design and self-assembly of two-dimensional DNA crystals. *Nature*, 1998, 394: 1223–1226
- Adleman L, Cheng Q, Goel A, Huang MD, Kempe D, Espanés PM, Rothmund PWK. Combinatorial optimization problems in self-assembly. In: *Thirty-Fourth Annual ACM Amposium on Theory of Computing*. Montreal, Canada, 2002. 19–21
- Brun Y. Arithmetic computation in the tile assembly model: addition and multiplication. *Theor Comput Sci*, 2006, 378: 17–31
- Brun Y. Solving satisfiability in the tile assembly model with a constant-size tileset. *J Algorithm*, 2008, 63: 151–166
- Ouyang Q, Kaplan P D, Liu S, Libchaber A. DNA solution of the maximal clique problem. *Science*, 1997, 278: 446–449
- Benenson Y, Paz-Elizur T, Adar R, Keinan E, Livneh Z, Shapiro E. Programmable and autonomous computing machine made of biomolecules. *Nature*, 2001, 414: 430–434
- Braich RS, Chelyapov N, Johnson C, Rothmund PWK, Adleman L. Solution of a 20-variable 3-SAT problem on a DNA computer. *Science*, 2002, 296: 499–502
- Morimoto N, Arita M, Suyamay A. Solid phase DNA solution to the Hamiltonian path problem. In: Rubin H, Wood DH, Eds. *DNA Based Computers III*. Philadelphia: American Mathematical Society and DIMACS, 1999, 48: 193–216
- 刘文斌, 许进. 赋权 Hamilton 路的 DNA 计算模型. 系统工程与电子技术, 2002, 24: 99–102
- Rothmund P. Folding DNA to create nanoscale shapes and patterns. *Nature*, 2006, 440: 297–302
- Dietz H, Douglas S, Shih W. Folding DNA into twisted and curved nanoscale shapes. *Science*, 2009, 325: 725–730
- Han D, Pal S, Nangreave J, Deng Z, Liu Y, Yan H. DNA origami with complex curvatures in three-dimensional space. *Science*, 2011, 332: 342–346
- Han DR, Pal S, Yang Y, Jiang SX, Nangreave J, Liu Y, Yan H. DNA gridiron nanostructures based on four-arm junctions. *Science*, 2013, 339: 1412–1415
- Ke YG, Ong LL, Shih W, Yin P. Three-dimensional structures self-assembled from DNA bricks. *Science*, 2012, 338: 1177–1183

A “DNA origami”-based approach to the solution of Hamilton path problem

Yang Yu¹, Shao Su², Jie Chao^{2*}

¹ Information Technology Department, Shanghai Institute of Science & Technology Management, Shanghai 201800, China

² Key Laboratory for Organic Electronics and Information Displays & Institute of Advanced Materials (IAM); Jiangsu National Synergetic Innovation Center for Advanced Materials (SICAM), Nanjing University of Posts & Telecommunications, Nanjing 210023, China

*Corresponding author (email: iamjchao@njupt.edu.cn)

Abstract: Hamiltonian path problem (HPP) is known to be in the problem class NP-complete. In this paper, a DNA origami-based method for solving the HPP was described. Among the various predefined 2D nanostructures created by DNA origami technology, a square shaped DNA origami body with decoration of numbers was selected for coding the vertexes while their single-stranded arms coded the paths. Because these arms could hybridize with their corresponding ones in the solution, all the paths with possibility would exist after annealing. Take advantage of magnetic beads and gel electrophoresis, the shortest path that pass each vertex once for HPP could easily found under the atomic force microscope. This method not only simplified the experimental procedures for solving the HPP but obtained a direct and readable result, which could also be applied to other problems requiring a massive parallelism in computation.

Keywords: Hamiltonian path, DNA origami, magnetic beads, atomic force microscope