

一种 DNA 计算机与电子计算机之间的通信模型^{*}

朱莹¹ 丁永生^{1,2} 李汪根¹

(东华大学信息科学与技术学院 上海 201620)¹

(数字化纺织服装技术教育部工程研究中心 上海 201620)²

摘要 提出了一种 DNA 计算机和电子计算机之间通信的层次模型。首先,为了使电子计算机能够对 DNA 计算进行控制,研究并设计了适合于 DNA 计算的反应器;然后给出了 DNA 计算机与电子计算机之间通信的层次模型;最后描述了选择操作在该层次模型上的求解方法以表明其可行性。本文工作为发展 DNA 计算机和电子计算机相集成的杂合计算机提供了一种通信模型和方法。

关键词 DNA 计算机,电子计算机,通信模型,探针选择操作

A Communication Model between DNA Computer and Electronic Computer

ZHU Ying¹ DING Yong-Sheng^{1,2} LI Wang-Gen¹

(College of Information Sciences and Technology¹, Engineering Research Center of Digitized

Textile & Fashion Technology, Ministry of Education², Donghua University, Shanghai 201620)

Abstract We propose a hierarchical model to describe the communication procedure between DNA computer and electronic computer. Firstly, we design a DNA computing reactor, inside which the electronic computer controls the DNA computer. Then, there is an illustration of the proposed hierarchical model, under which the Probe Selection is pictured to set out its feasibility. Looking ahead, a hybrid computer integrates DNA computer with electronic computer may come into the stage someday.

Keywords DNA computer, Electronic computer, Communication model, Probe selection operation

1 引言

以 DNA 分子作为计算载体的计算方法最早由 Adleman 博士于 1994 年在生物实验室实现^[1]。随后, DNA 计算取得了研究成果。1995 年^[2], Lipton 提出了用 DNA 实验解决布尔可满足性问题(SAT)的方法。1997 年^[3], Ouyang 等人用分子生物技术解决了最大团问题,并相应地建立了六个节点的 DNA 分子计算池。2000 年^[4], Liu 等人设计了一个 DNA 计算系统,将计算问题的所有可能候选解编码成一系列 DNA 分子,把这些分子综合起来贴在磁珠的表面。2001 年^[5], Wu 分析并改进了前者所采用的基于表面磁珠的方法。这种新颖的计算方式是建立在其高密度信息存储和大量并行计算基础上的,有望在求解 NP 问题、破解密码、疾病诊断、新材料等领域发挥重要作用。

微流控系统为快速化学反应和生物分析提供了基础。微流控系统作为 DNA 计算的一种平台,已经有了初步的研究成果。2004 年^[6], Ledesma 等人提出了一种用微流控系统解决 Hamilton 路径问题的线性 DNA 算法,实现了并行计算。1999 年^[7], Gehani 和 Reif 研究了用微流生物分子计算模型解决某个问题在理论上所需要的最少 DNA 序列和最少反应时间,并且提出了反应池之间有效地传送 DNA 序列的方法。2001 年^[8], McCaskill 采用枚举法用 DNA 序列对每个可能子图进行编码。该算法使用了所谓的选择模块(STM)来保留

图中所有可能团,然后用排序的方法确定了最大团。2001 年^[9], Chiu 等提出了一种新的方法,把子图和图的边分别编译成反应池和缓冲池。这些反应池和缓冲池带有荧光剂,由通道连接,输出以各个子图所发出的荧光强度不同来区分。2004 年^[10], Livstone 和 Landweber 提出一种微反应器用来解决布尔函数“与”和“或”的问题。

随着微机电(MEMS)技术的快速发展,在生物芯片上集成各种能与电子计算机之间相互通信的传感器是各种功能芯片研究的热点^[11]。DNA 计算的一个严重缺点是操作的不可控,严重影响了 DNA 计算的实用性。本文在微流控平台上,讨论基于生物芯片的 DNA 计算机和电子计算机之间相互通信的层次模型,从而为发展 DNA 计算机和电子计算机相集成的杂合计算机提供一种通信模型和方法。

2 基于电子计算机的 DNA 计算反应器模型

为了使电子计算机能够对 DNA 计算进行控制,我们研究并设计了适合于 DNA 计算的反应器。如图 1 所示,该反应器通过 RS232 接口与电子计算机连接,人机交互界面的平台是 LabVIEW。

(1)数据流:在电子计算机软件客户端由用户选择某个 NP 问题,程序开始设计 DNA 计算机,包括:输入符号、终止符号、〈状态,符号〉、转移分子、扩增所需的引物和 DNA 编码;酶的选择;扩增、酶切、酶连的动作选择,针对上述酶设计

^{*}基金项目:国家自然科学基金重点项目(60534020),教育部新世纪优秀人才支持计划(NCET-04-415),教育部科技创新工程重大项目培育资金项目(706024),上海市国际科技合作基金项目(061307041)。朱莹 博士研究生,从事 DNA 计算机等研究。丁永生 博士,教授,博士生导师,从事智能系统、网络智能、DNA 计算、人工免疫系统、生物网络结构、生物信息学、数字化纺织服装、智能决策与分析等研究。

反应温度和阶段。

(2)控制流:设计具体问题的程序输入分子和采用的微流控芯片,芯片的通道数量和通道之间的连接以及通道的形状,

设计实验实现的详细步骤:每一步的动作、所需的时间、反应的底物、目标产物、代表的中间变量等。

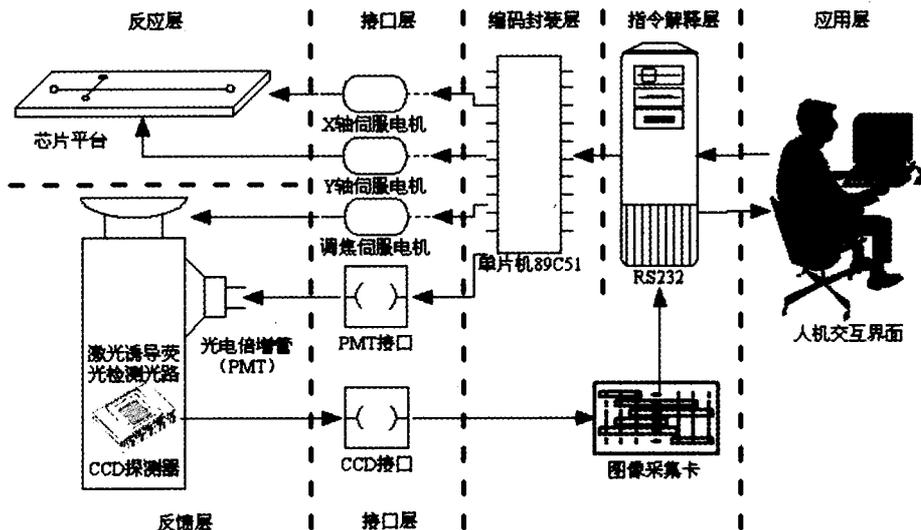


图1 DNA计算的反应器模型

电子计算机主机与89C51系列单片机相连,将数据流和控制流同时送到接口端的高精度全方位机械手:分配试剂,根据反馈的图像信号定位芯片反应平台。生化反应的动态结果反映在应用层,包括当前反应所在的通道、该反应所需的时间、已经消耗的时间、通道的切换、反应产物的解释。

DNA计算反应器与电子计算机通信的系统控制模块、光电检测模块、高压电源模块和温度加热模块与电子计算机之间的连接如图2所示。整个反应器能够实现激光诱导荧光检测、芯片电泳和仅与温度有关的生化反应,如PCR、退火、复性等。

3 DNA计算机与电子计算机之间通信的层次模型

DNA计算本质上是以DNA分子及生化酶作为物质基础,施以适当的生化操作来解决数学问题的一种新型的计算模式。由于DNA计算的处理对象是DNA片段,因此运用DNA计算求解数学问题时首先需要将实际问题用(A, T, C, G)四个碱基来编码,原理类似电子计算机求解这些问题时需要用二进制编码。然后需要为求解过程设计合适的生化操作,这个完成运算的生化操作序列我们称之为生物算法。

DNA计算是在分子尺度内进行的,完成计算过程的生化操作的不可控一直制约着DNA计算的进一步应用。随着电子计算机技术和传感器技术的发展,二者在DNA计算中的结合可实现对生化操作的精确控制,提高DNA计算的可靠性,为DNA计算进一步走向实用化发挥重要作用。图3给出了电子计算机和DNA计算机之间通信的层次模型。

为了便于描述通信过程,将该模型分为六个部分,每一部分的组成和完成功能描述如下。

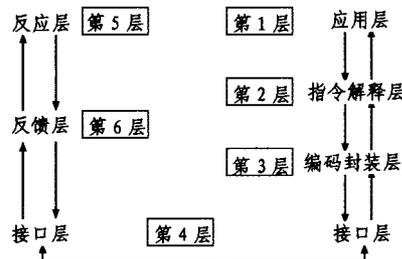


图3 DNA计算机和电子计算机的层次通信模型

(1)应用层

提供用户与DNA计算机之间交互的接口。应用层主要完成两个功能:一是提供用户操纵DNA计算机的界面。在这个界面上用户可以完成原始问题到DNA碱基域的映射以及完成生物算法的设计。通过这个界面,用户可以像使用Office办公软件一样方便地使用DNA计算机。另一个功能是接收指令解释层传送的DNA计算结果,并将结果可视化。指令解释层传送的结果也是用电子计算机语言来描述的。运算结果的可视化可以帮助直观地对这些结果进行合理的解释。应用层由安装在Windows操作系统的电子计算机上的应用程序组成。

(2)指令解释层

由于应用层的指令是用户所熟悉的电子计算机语言描述的,而DNA计算机的基本指令是具体的生物操作,所处理的对象是DNA分子,因此需要将应用层的电子计算机指令解释成DNA计算机上具体的生物操作指令(这些生物操作指令是DNA计算机的最小执行单位——基本指令),确定这些生物操作的执行顺序,并依次将这些生物操作指令单个传送给编码封装层。另一方面,还需要将编码封装层反馈的DNA计算结果解释成计算机语言。指令解释层也是由安装在电子计算机内部的代理程序构成。

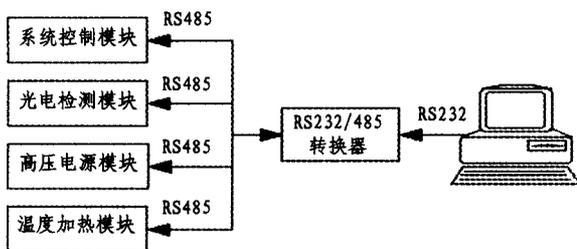


图2 DNA计算反应器与电子计算机的通信

(3) 编码封装层

将指令解释层传送的单个生物操作指令封装成 DNA 计算机能直接执行的指令。这里需要考虑每个原子生物操作的实验室实现方法。然后将这个操作的步骤映射成控制传感器和生化仪器的一系列指令, 包括对生物芯片上发生该反应的位置信息。编码封装层由操作传感器和生化仪器的接口程序构成。

(4) 接口层

接口层是传感器和生化仪器的各种信号接口。一方面, 将编码封装层中的控制指令转换成控制 DNA 计算机执行生化操作的指令; 另一方面, 也将 DNA 计算机上的反馈信号转换成电子计算机中的控制指令。接口层之间的通信采用电子计算机的串口通信方式, 也可以设计成并口通信方式。

(5) 反应层

DNA 计算的生化操作在这一层得以物理实现, 以完成解释层下达的任务。此层包含有完成生化反应的生物芯片以及控制这些生化反应的各种传感器和生化仪器。

(6) 反馈层

这一层由监控生化反应的传感器构成。反应层的生化操作的执行情况由这些传感器收集, 以便反馈给指令解释层。

4 通信模型的实例

为了更直观地理解本文提出的层次模型, 我们以选择操作为例, 解释在层次模型下 DNA 计算机上实现选择操作的过程。

DNA 计算机中的选择操作是用来选择编码特定信息的 DNA 片段。一般是将包含特定模式的探针固定于芯片上, 被选择的片段在芯片上与其发生复性(Renaturation)操作, 最后固定在芯片上的片段就是选择的片段。复性是变性过程的逆过程, 即两条完全互补的单链在适当的条件下恢复到天然双螺旋结构的过程。热变性的 DNA 一般经过冷却后即可复性。因此, 此过程有时也称退火(Annealing)。复性温度一般应该比该 DNA 的解链温度值低 20℃~25℃。

对于选择操作, 需要将其解释成以下几个独立的生物操作:

操作 1: 选择固定有特定模式探针的生物芯片;

操作 2: 将被选择的片段在该生物芯片上发生复性反应;

操作 3: 在激光诱导荧光检测仪上通过荧光检测判断选择的结果。

图 4、图 5、图 6 分别为层次模型下以上操作的流程图。

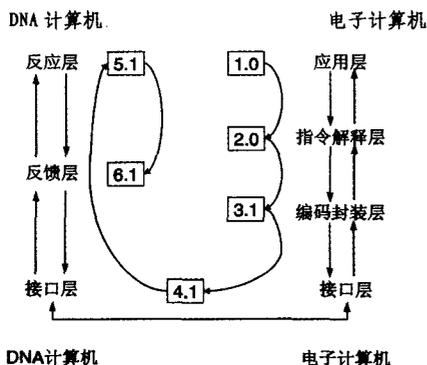


图 4 操作 1 流程图

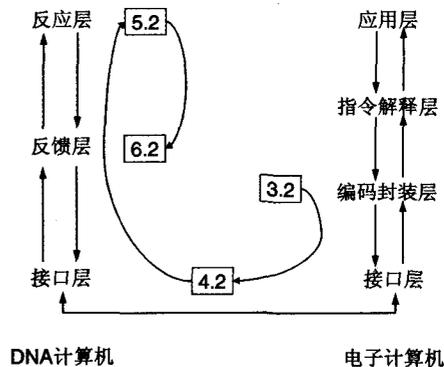


图 5 操作 2 流程图

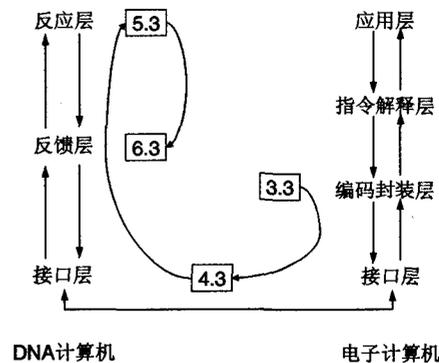


图 6 操作 3 流程图

动作 1.0 应用层: 用户在接口程序发出选择操作, 例如: 选择操作符号 1。

动作 2.0 指令解释层: 逐条解释生物算法中的每一个描述, 将其解释成分子实验室中具体的生物操作, 包括反应名称和反应条件, 然后将单个的生物操作依次发给编码封装层, 待编码封装层返回操作的执行状态后再发下一个生物操作, 类似于电子计算机中的指令寄存器。

动作 3.1 编码封装层(第 3 层操作 1): 发出选择探针芯片的指令。通过选择可寻址的生物芯片来实现。

动作 4.1 接口层(第 4 层操作 1): 电子计算机通过 RS232 与 DNA 计算机进行通信。

动作 5.1 反应层(第 5 层操作 1): 无生化反应。

动作 6.1 反馈层(第 6 层操作 1): 相应的传感器接收到确认信息, 确认该操作 1 完成, 并将反馈信息返回到接口层, 接口层反馈到编码封装层, 再发下一个操作。

动作 3.2 编码封装层(第 3 层操作 2): 发出复性反应的指令。控制相应的温浴仪器先将温度升高到 65℃, 然后慢慢冷却, 降温到 20℃。

动作 4.2 接口层(第 4 层操作 2): 电子计算机通过 RS232 与 DNA 计算机进行通信。

动作 5.2 反应层(第 5 层操作 2): 相应的生化仪器先加热芯片, 然后慢慢冷却。

动作 6.2 反馈层(第 6 层操作 2): 相应的传感器接收到确认信息, 确认该操作 2 完成, 并将确认信息返回到接口层, 接口层反馈到编码封装层, 再发下一个操作。

动作 3.3 编码封装层(第 3 层操作 3): 指示激光诱导荧光生化检测仪工作。

(下转第 207 页)

结果优先返回给用户。

```

client request resource
check local LDAP Resource-info
if (valid)
  return result
else
  {send QueryAgent to VO Global LDAP Server
  query the position if satisfies condition
  if (valid)
  {QueryAgent copy itself and move to resource of multi-node
  query local resource of node
  return resultAgent to local LDAP server}
  else
  {send QueryAgent to Global LDAP Server of central domain
  central domain Global LDAP query and confirms the VO of resource
  central domain send QueryAgent to correlative VO
  return resultAgent to central domain}

```

基于移动 Agent 的资源发现系统本质上是分布式的,但是与传统的基于 RPC 方式的分布式应用相比,移动 Agent 的迁移不需要保持网络的长时间稳定连接,这可大大减轻网络负荷,避免了大量数据的网络传送,从而显著提高系统执行效率和可靠性。

4 性能分析

4.1 资源发现效率和资源质量

教育资源中存在着大量的非文本资源。传统的搜索引擎和主题指南无法对视频、音频等多媒体教育资源进行标引和发现,而元数据通过对资源类型的描述,只要用户在资源检索时对资源类型进行限定即可发现它们。

因特网上存在大量教育资源,这些资源的质量参差不齐,传统的搜索引擎和主题指南没法提供一个质量评估机制。而这正是教育元数据的一个特征,它通过提供教育资源使用者的评估来反映不同资源之间的质量差异,有助于用户发现高质量的教育资源。

4.2 动态性和扩展性

在网络系统中,资源节点的信息是动态变化的,为了保证服务的可用性,进行有效资源查找时必须考虑系统中节点的

动态变化。

网络层采用 Globus MDS 来组织和管理底层资源,通过调用虚拟组织 Agent 中间件部署的 Register Agent 和 Unregister Agent,进行虚拟组织的动态建立并且虚拟组织成员可以随时加入和离开,具有很大的灵活性,实现了动态注册和注销。通过 Upgrade Agent,社区节点周期性地向社区管理节点发送更新信息,将变化动态地反馈给虚拟组织中的 LDAP 服务器。同时,各个虚拟组织 VO 之间定期将目录服务器上的资源信息通过 LDAP 的异步远程复制机制更新到虚拟组织中心域中的 LDAP 中,这样就实现资源信息的动态收集和资源变化的实时反馈。

结束语 该模型采用教育元数据 LOM 规范来描述教育资源,使用分类元数据的方法,结合移动 Agent 的自治性、智能性等特点,具有较高的资源发现效率和资源质量、良好的动态性和扩展性,基于类别的虚拟社区划分,使得资源发现限定在所属社区范围内,缩小了搜索空间,减少了资源发现的响应时间,提高了资源发现的效率。

参考文献

- 1 虞惠达,朱庆华.元数据在教育资源发现中的应用[J].中国图书馆学报,2005(2)
- 2 Fukuda M, Tanaka Y, Suzuk N. A Mobile-Agent-Based PC Grid [A]. In: Autonomic Computing Workshop [C]. Seattle, USA; [s. n.], 2003. 142~150
- 3 Jun K, Boloni L, Palacz K, et al. Agent-based resource discover [J]. In: 19th IEEE Heterogeneous Computing Workshop (HCW'00), 1999(10); 157~268
- 4 Czajkowski K, Fitzgerald S, Foster I, et al. Grid information services for distributed resource sharing [A]. In: Proc. of the 10th IEEE HPDC [C]. Washington, DC; IEEE Computer Society Press, 2001. 181~194
- 5 李红梅,贺小扬,等.网络环境中教育资源发现机制[J].微电子学与计算机,2006(9)
- 6 Yu Jian, Zhao Chen, Pan Yunhe. Grid Resource Management Based on Mobile Agent. International Conference on Computational Intelligence for Modelling Control and Automation, and International Conference on Intelligent Agents, Web Technologies and Internet Commerce(CIMCA-IAWTIC'06)
- 7 rial problems. Science, 1994, 266; 1021~1024
- 8 Lipton R J. DNA solution of hard computational problems, Science, 1995, 268 (28); 542~545
- 9 Ouyang Q, Kaplan PD, Liu S, Libchaber A. DNA solution of the maximal clique problem. Science, 1997, 278; 446~449
- 10 Liu Q, et al. DNA computing on surfaces, Nature, 2000, 403; 175~179
- 11 Wu H Y. An improved surface-based method for DNA computation. Biosystems, 2001, 59; 1~5
- 12 Ledesma L, Pazos J, Rodríguez-Patón A. A DNA Algorithm for the Hamiltonian Path Problem Using Microfluidic Systems. Aspects of Molecular Computing, 2004. 289~296
- 13 Gehani A, Reif J H. Micro Flow Bio-Molecular Computation. In: 4th DIMACS Workshop on DNA Based Computers, University of Pennsylvania, USA, June 1998
- 14 McCaskill J S. Optically programming DNA computing in micro-flow reactors. BioSystems, 2001, 59; 125~138
- 15 Chiu D T, Pezzoli E, Wu H, Stroock A D, Whitesides G M. Using three-dimensional microfluidic networks for solving computationally hard problems. In: Proceedings of the National Academy of Sciences of the United States of America, 2001, 98 (6); 2961~2966
- 16 Livstone M S, van Noort D, Landweber L F. Molecular computing revisited: a Moore's Law? Trends Biotechnol., 2003, 21; 98~101
- 17 司士辉.生物传感器.北京:化学工业出版社,2003

(上接第 173 页)

动作 4.3 接口层(第 4 层操作 3):电子计算机通过 RS232 与 DNA 计算机进行通信。

动作 5.3 反应层(第 5 层操作 3):进行激光诱导荧光检测。

动作 6.3 反馈层(第 6 层操作 3):相应的传感器接收到确认信息,确认该操作 3 完成,并将确认信息返回到接口层,接口层反馈到编码封装层,再发下一个操作。

在编码封装层完成这 3 个操作后,将执行结果反馈到指令解释层。指令解释层再解释下一条电子计算机指令。所有电子计算机指令执行完毕,解释层将结果反馈到应用层。

结束语 本文提出了一种基于生物芯片的 DNA 计算机与电子计算机通信的层次模型,并介绍了选择操作在该层次模型上生物算法的求解过程。在进行通信时,每一个层都有独立的工作需要完成,在信息流动的方向上为上一层提供相应的服务,并负责将数据传送到下一层。

目前在 DNA 计算的研究中,DNA 计算机与电子计算机相结合的研究还鲜有涉及,将电子计算机算法自动地转化为 DNA 计算机上的生物算法是我们进一步的研究方向。

参考文献

- 1 Adleman L M. Molecular computation of solutions to combinato-