

包含飞：生物医学知识整合论(V)，《医学信息》杂志，17（5）：244-250，2004

生物医学知识整合论（V）

——再论知识背景空间

包含飞

上海中医药大学中医学信息化-标准化研究室，200032

摘要 本文沿生物医学知识整合论研究方向进一步探讨生物医学知识复杂的背景空间。文章介绍了笔者对医学信息学及医学知识领域中物理性知识、人类智能和人工智能世界三维关系的认识；为了使该理论逐步清晰化，文中又讨论了一系列概念定义；接着分别以欧几里得几何、VB 编程语言和胆汁淤滞为例描述了人类智能、人工智能和物理性知识三种知识背景空间的特点；最后探讨了先哲们建立理性学科背景空间的技巧及对我们的启示。

关键词 生物医学知识整合论 医学信息学 知识背景空间 认知科学 医学人工智能

在某种意义上，生命是一切自然机理的总和，生命科学是一切科学的总和。生命系统的奥妙之处不仅仅在于“一切”二字，更在于“总和”二字，在于生物界的完美无缺的、和谐共济的“总和”！这是自然界的物理整合。关于科学知识的“总和”机制的研究就是知识整合科学。在生物医学知识的整合研究^[1-7]中，生物医学信息学将起到无可取代的作用。

1. 医学信息学，知识整合和三维知识世界

医学信息学拥有一个庞大的学者团体，其国内和国际学术会议的规模一般相当大。在有关的各种书刊杂志上发表的及各种会议上交流的，大都是关于医疗卫生领域中的应用系统。但有些学者指出，医学信息学并不等于建立和维护形形色色的与医疗卫生有关的应用系统^[8]。那么医学信息学究竟是什么？它仅仅是把现成的信息技术引进到医学领域中的努力？它的内容仅仅是开发、引进和维护一些信息系统？它有没有自身的概念系统，基本理论和体系结构？

笔者认为对前两个问题，答案是明确否定的，但对第3个深层次问题就不那么容易回答。首先，生物医学信息学涉及三大学问信息学、生物学和医学。信息一词外延极为广泛，常与物质、能量等被并列为构成世界的基本要素或维度，所以几乎是一个无处不在的带有哲学意义的“千面概念”。笔者认为信息这个概念很多方面类似于另一个无处不在的概念：“结构”^[4, 6]。我们并没有根据物质和能量概念提出“物质学”或“能量学”，但我们却根据信息概念提出了“信息学”。这也许是因为香农曾对关于通信的信息概念给出了定量公式和人类发明了诸如计算机和网络这些被称为信息技术的新事物并且对人类生活产生了空前的影响之故。我们还在生物学和医学领域中分别提出了生物信息学和医学信息学，并且把财务管理、

设备药物管理、医院局域网建设、电子病历、PACS、疾病诊断-治疗-护理、药理和药物动力学、DNA 的表达及调控和蛋白质功能-空间构型-蛋白质网络等的有关算法-查索-计算机处理、..... 统统地归到“生物医学信息学”这面大旗下。生物学和医学也是两个超繁领域，几乎囊括和整合了诸如机械、电、化学、大分子、随机、非线性..... 所有的方面，真可谓集“万千机制于一体”。因此三个超繁领域结合的“产儿”生物医学信息学的极端复杂性和内在千差万别的异质性就在所难免。再加上参与到生物医学信息学领域中的学者群体的知识背景、追求目标、思维方式又彼此相差甚远。对一些学者来说“理所当然”的事，对另一些学者也许会“莫名其妙”。所以，我们不得不承认目前生物医学信息学丰富的异质性，截止今日，它更象一个学科群体，尚未成为真正统一的科学。

划时代级的科学大师牛顿和爱因斯坦在追求统一结构和理论时都选择了万古不变的天体系统，没有选择生命系统。另一个大科学家量子力学的奠基人薛定谔研究过生命系统，留下了一句脍炙人口的名言：“生命是负熵”。讲得通俗一点就是“生命的结构可以越来越复杂”。他认为原因是生命系统是“开放系统”，任何复杂的生命系统都可“统一”到负熵。这一伟大的论断在哲学界和物理学界意义远大于生物学和医学界，因为即使一个老婆婆也拥有生物学常识知道她的小鸡崽与她的针线盒不同，“会长大下蛋”。由此可以想象探索医学信息学的内在统一的结构和理论决非轻而易举的事。相反，它是一件令人望而生畏的苦差事。

MA Musen 指出，“Blois 的工作的原理性贡献在于确认了认识论（*epistemology*）是令人信服地说明医学信息学是一门学科研究的核心元素。^[8]”这一点与笔者十多年来的努力方向可谓不期而遇。知识从其本质上来讲是客观物理世界和人类意识世界的对话，是物理世界的规则和意识世界的规则的交流。Weed 把病历数据分为主观类（病人主诉），客观类（临床发现，实验室结果），评估类（结论，诊断），计划类（治疗措施）^[14]。其实他的主观类，评估类，计划类均属主观意识类，客观类属客观物理类，所以生物医学知识系统反映了这种对话。讨论生物医学知识整合问题不可能离开物理世界，意识世界及人工智能世界的规律及其相互关系。根据唯物主义哲学观，人类大脑的知识和思维运动可看成是客观物理世界的抽象或模型，而人工智能可看成人类智能的一部分或“代理”并可作为人工智能或知识整合的具体手段，因此也可看成客观现实世界的一种抽象或模型（见图 1，2）。因为人工智能的载体与大脑神经系统不同，具有帮助人类智能大容量积累和大规模联合的潜能。而且正因为人工智能的载体不同，它们一经成立，便获得了自身的个性并按自身的规则运动，成为一种独立于人类智能的自在体。就象一辆汽车固然是人类智慧的产物，但一个人必须掌握汽车的“个性（性能）”，才能获得“驾驶执照”。

生物医学知识整合论（The Theory of Bio-Medical Knowledge Integration, BMKI）^[1-7]的宗旨是探讨整合生物医学知识的信息学方法，使较局部和较零碎的生物医学知识能更整体化地反映生物机体和医学过程。诗曰“会当凌绝顶，一览众山小。”目标不高难以“盖全”，笔者有意借此“高难”生物医学知识整合研究为目标，对生物医学知识“超复杂”世界的哲学-信息学细节作一番神游和苦旅，并且希望有所收获。MA Musen 说过意味深长的一句话，“如果我们相信医学信息学的研究对世界的某些方面贡献了自己独特的有内在价值的基本理解，那么阐明医学信息学的理论基础并使之区别于计算机科学和其他科学如信息理论，生物统计学和保健研究等，是我们的当务之急。然而，不幸的是我们学科中没有致力于提出基础理论的哲学家。^[8]”看来有责任性的学者之所虑，有很多共通之处。笔者生性偏爱学术的完美性，眼中的医学信息学犹如一位婷婷待妆的“美丽公主”，具有自身的美学潜质，不是人们心目中的生物学和医学的“美丽女仆”。更希望 BMKI 成为医学信息学的一个新学科，新成员。笔者将尽其所能为 MBKI 建立一套相对严密的基础概念和学术结构，期望开创一片蓝天碧地，湖影垂柳的学术境地。

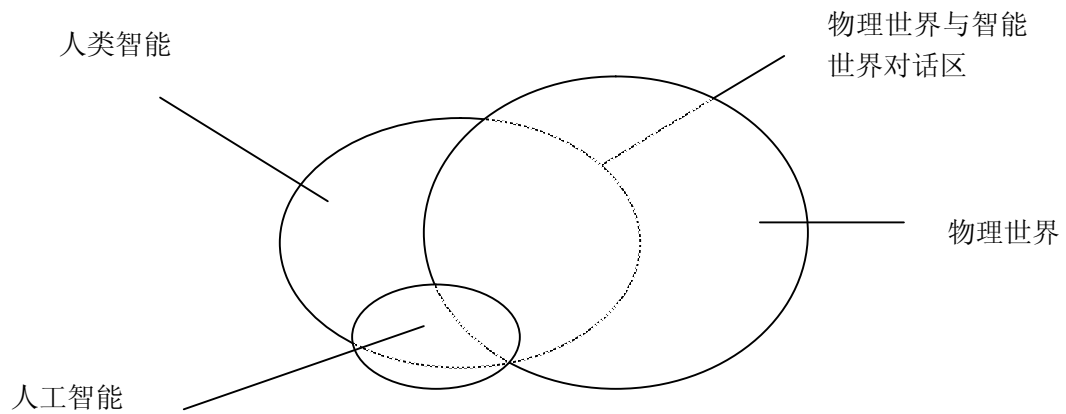


图 1 人类智能，人工智能和物理世界为各自具有自身运动规律的三种不同的领域，医学知识系统处于它们之间的对话或重叠区域，是三类运动规律的综合表现。

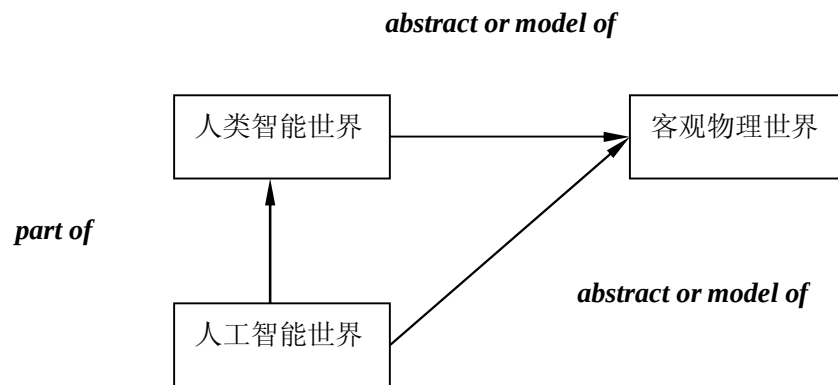


图 2 客观物理、人类智能、人工智能世界三者的关系。从人工智能这个全透明体，我们可以观察和探索客观现实世界的部分原理。

2. “可爱的孩子”与“温暖的摇篮”

人类的天性喜爱活泼可爱的孩子，但娇嫩的孩子离不开它温暖的摇篮。“温暖的摇篮”是“可爱的孩子”的背景空间。在我国住房紧张时期，有人感叹说“找一个姑娘容易，找一套住房难”。说明在一定的情况下，“背景”可以比一切都重要。

所谓知识背景空间即为某条知识的准确意义成立所依赖的条件、环境或前提的集合。就象人们容易记住孩子而忽略摇篮一样，知识的背景也最容易为人们所遗忘。这是因为生存原则作为生物系统的第一原则离不开另一个基本原则即经济原则或效率原则，而思维作为高等生物的功能之一也遵循这一原则，往往把思维活动的背景性或非直接事物放置在大脑的相对非活跃的区域即下意识或潜意识区域（就象电脑把不直接参与操作的资源放在成本相对较低的外存中一样），而把直接参与思维活动的诸要素放在明意识区域（相当于电脑的直接操作空间即成本相对较高的内存）。

把各级相对间接的资源放置在大脑各级下意识或潜意识区域以达到“最佳效率”堪称大自然“神来之笔”的杰作，但给人工智能研究者带来了诸多迷惑。因为人工智能要求知识成立的条件或背景“一个不能少”。有人总结所以在医学信息学中人工智能发展“不尽人意”的原因之一就是因为人们“本以为”医生能把他们的思维过程讲清楚^[13]，但事实证明事情远非如此。那么思维过程究竟那些部分“难以讲清楚”，笔者认为就是相应的“知识背景”部分。计算机科学史上所犯的关于“第五代”智能计算机神话的全球性估计错误也源于此。

我们知道，人们发现和寻找知识目的有二，一为满足求知欲望，二为达到某种意识的或物理的功利目标。难以想象有朝一日我们会穷尽世界的一切奥秘，人类面前不再存在任何无知领域，因此求知欲望是无穷的。但如果获得知识是为了达到某种功利目标，我们就可以此目标为中心，把直接参与的知识作为“当事者”知识，而把间接影响的知识作为“背景性”知识。在这种情况下，知识的背景空间带有“功利性”。这并不会冒犯理论的“纯洁性”，我们将在下文看到即使众所周知的被认为是“绝对精确”、“绝对理性”的欧几里得几何其背景也带有“拿来主义”或“经验主义”的意味，带有某种功利性。

我们把背景性知识的全体称为“知识背景空间”。一般来说，

（i）人类思维的“知识背景空间”往往位于潜意识或下意识区域，包括常识，共识和本能性知识。其基础部分多来自先天本能，上层部分多来自后天；

（ii）对电脑“思维”或人工智能来说，无背景的知识是“不完整的知识”。一个基于规则的系统如果只包含人类的明意识规则，缺乏背景知识空间的支撑，那么系统就不够健壮。它们往往只能在高度限定的空间中才有效。

因此在人工智能和任何知识工程中，应重视、分析和挖掘相关的潜在的知识背景空间。

3. 概念与定义

有鉴于生物医学知识领域的高度的结构性和复杂性，过于抽象地泛泛而谈“熵与信息互为倒数”恐怕无济于事。笔者认为首要任务是先把我们所讨论的对象尽可能限定或确定，在本系列文章之四^[4]的基础上，这里再给出一组定义。

定义 1 知识单元的基本组分 一般说来，一个知识单元由四类基本组分，直接因子、间接因子、直接规则和间接规则。因子又称操作子或元素，规则又称关系。

定义 2 认知动机(cognitive motive) 激发人类获得知识努力的动力，包括本能的求知欲望（instinctive or inherent longing for acquiring knowledge, ILAK）或对世界的好奇性和对具体的物理性或意识性功利（particular physical or mental “profits”, PPP or PMP）的追求。

定义 3 认知目标(cognitive goal) 某种认知动机下获取知识的努力所希望达到的目的集合。

定义 4 ILAK 的认知目标 (cognitive goal driven by ILAK) 在人类求知本能和探索欲望驱使下的无穷目标。在这一目标下,下文所述所谓知识的规则粒度、因子粒度、理解粒度、操作粒度和知识背景空间粒度等都是无穷尽概念。但在某一历史阶段,人们常常把这个无限目标限定为一个该历史时期的具体的有限目标。

定义 5 PPP 或 PMP 的认知目标(cognitive goal driven by PPP or PMP) 为 PPP or PMP 驱动下获取知识的努力希望达到的具体的物理的或意识的目的(事物和关系)的集合。意识性功利目标如建立一种理论,构思一首诗,物理性功利目标如诊断疾病、治疗疾病、种植一亩水稻、喂养一头奶牛等。此目标决定相关的知识规则粒度、因子粒度、理解粒度、操作粒度和知识背景空间粒度等。

定义 6 认知量子 (cognitive quantum, CQ) 指在某种认知目标下,无法或无须再细分或深究的量子因子(或操作子)和量子规则(或关系)。

定义 7 ILAK 认知量子 (CQ by ILAK) 当代科学技术测出或发现的独立存在的稳定的最小粒子或基础力。也称为时代量子粒度。

定义 8 PPP or PMP 认知量子 (CQ by PPP or PMP) 为 PPP or PMP 目标下的最小认知因子或规则。如外婆的老母鸡和鸡蛋,欧几里得的点,病理学的细胞等。

定义 9 认知粒度 (cognitive granularity, CG) 认知粒度为在某一认知目标之下对某事物的细节的了解程度的量度。它可以用公式 $CG = 1/d$ 表示,式中 d 为某一事物相关细节的计数。一个知识的细节可以包括直接或间接因子(或操作子)和规则(或关系)等方面。

认知综度(cognitive synthesis degree, CSD) 为一个事物的细节合成的程度,可视为认知粒度(CG)的反概念。如一个合成算法 $F=A \wedge B \wedge C$,其中 A, B, C 为子规则或子因子, F 则为它们的合成规则或因子。

定义 10 ILAK 目标认知粒度(CG for cognitive goal driven by ILAK) 为在某 ILAK 认知目标下的某事物的认知粒度。

定义 11 PPP 或 PMP 目标认知粒度(CG for cognitive goal driven by PPP or PMP) 为在某 PPP or PMP 认知目标下某事物的认知粒度。

定义 12 认知量子粒度 (cognitive quantum granularity, CQG) 为某一认知目标下的认知量子的粒度,因为量子是不能再分的最小粒子,因此量子的细节是量子本身,即 $d = 1$,所以 $CQG = 1$ 。

定义 13 知识表示粒度(granularity of knowledge expression, GKE) 知识表示粒度为在某一认知目标之下对某事物的细节的描述程度的量度。为某一知识的直接/间接因子或规则的表达粒度的总计。可以比作一种“鸟瞰粒度”,比如一个人在地面,高楼,飞机,卫星等不同高度来观察或描述地球上物体的粒度。

定义 14 知识表达的完全性 (completeness of knowledge expression, CKE) 为在某一认知目标下知识表达的粒度与该目标的要求粒度之比。

定义 15 知识操作粒度(operational granularity of knowledge, OGK) 由某一认知目标下操作因子或规则所需要的认知粒度。例如水果商只须了解水果的质量和价格,无需知道果树的栽种技术。又例如可以从水分和肥料而无须从生物结构来控制庄稼的生长。可从发展教育和交通大环节来支持贫困地区。

定义 16 知识背景空间 (background space of knowledge, BSK) 为“在知识表示形式以外的但可影响该知识表示含义的事物的总称。换言之,知识背景空间为一个知识表示的精确含义所依赖的除知识表示本身以外的所有事物的总称”^[4, 6]。因此 BSK 由两部分组成:(1) 认知目标 (集合);(2) 由认知目标决定的认知量子 (集合),后者决定了知识背景空

间的广义维度。

有时知识的背景空间可以是“零空间”，即该条知识“无条件成立”。例如“人的生命是有限的”这条知识就是“无条件成立”知识。但即使如此，也只是相对地“无条件”。因为首先这是信仰科学的人士的看法，秦始皇就不承认，所以才有派徐福出海寻找“长生不老药”之说。又例如我们限定这条知识所指的是物理生命，而不是指精神方面，因为例如一个人的学说，如欧几里得的学说是“不朽”的。

4. 人类智能的背景空间

(i) 欧几里得几何的背景空间

欧几里得几何 (Euclidean geometry) 是理性演绎而成的象牙之塔，是用来研究意识世界的运动规律及其所依赖的背景很好的实例。

有人认为，欧几里得 (Euclid) 是为了堵住怀疑派与诡辩派之口而建立几何学。怀疑派与诡辩派声辩：“为何知道甲？因为乙；为何知道乙？因为丙；……”，如此没完没了地回溯下去，最终还是无法知道甲^[10]。为了回答这些诡辩，欧几里得在物理与意识的“共识区域”，设立了一道回溯屏障，告诉诡辩派人士不能越过他所确定的“量子层”或底层背景空间。

欧氏利用 10 条公设和公理，配合“优多诸斯检定法则、反证法（归谬法）与尺规作图”，推导出一系列几何定理。欧氏以“朝生暮死”之躯，成就了“永恒之事”。他成就的不是“单独一个命题的前提与结论之间的连结”，而是一个“逻辑网络”，一个逻辑体系^[10]。用笔者的说法，欧氏几何的“前提”是个“知识背景空间”，其“结论”是一个“知识领域”。

发现和证明点、线、面、体、三角形、四边形、圆、弧等规则图形之间的关系是欧氏几何的意识性功利目标（不包括球面或曲面上的图形）。为此，欧几里得首先设定 5 个公设，5 个公理和 23 个定义，在此基础上证明了 286 个命题^[9]。定义是对一个领域的量子因子的阐明和规定，公设和公理是“直观自明”的真理，亦即量子规则。欧几里得这套体系必须在物理世界与意识世界均可获得“认可”，无须也无法证明（也即量子化的），方可作为知识的背景空间或数学的源头。

如按人类无穷的求知目标，欧几里得几何的公设、公理和定义下面应还有其元概念，还有更深层次的知识背景空间（见表 1，2，3）（也许是无穷的）。但对于欧几里得几何为自己所规定的意识性功利目标来说，这一知识背景空间可以作为底层背景，足以建立欧氏几何大厦。

(ii) 欧氏几何如何摆脱毕氏几何的困境

笔者所以一直努力试图阐明科学知识中的意识成分和物理成分^[7]，是因为物理世界和意识世界的基本规则在某些情况下大相径庭。芝诺悖论如“运动是不可能的”，“阿基里斯(《荷马史诗》中的善跑的英雄)追不上乌龟”，“飞矢不动”等清楚地告诉我们意识世界和物理世界的运动规律之间的严重分歧。

生物医学知识是意识和物理成分的混合体。讨论医学知识领域的人工智能或知识工程，也包括我们的 BMKI，不可避免地会遇到这两种既相互区别又难以割舍的异质成分。但令人不解的是在人工智能或知识工程界，笔者从未见过关于二者区别的认真讨论。因为这一问题十分复杂，本文拟先从欧氏几何到毕氏几何的历史转化中的相关问题谈起。

毕氏 (Pythagoras, 公元前 582-497) 几何采用了离散派的观点，主张线段只能作有

限的步骤的分割，线段经过（很大的）有穷步骤分割后，得到一个点，其长度 d 虽然很小很小，但是不等于 0，并且 $d > 0$ ^[10]。我们知道对于物理世界的一个线段，毕氏的主张无疑是天经地义的。因为世界上任何人，用任何方法去分割任何实际的线段，那么他总有一天会穷尽其技，再也无法分割下去。所以物理世界的分割永远是有限的。但毕氏希望建立的几何毕竟是个典型的意识系统，它的每个结论特别是量子规则势必受到意识性世界严格审查。

毕氏学派另一个奠基性或量子性观点是“任何两条线段皆可共度（总是存在公因子）”。在物理世界，毕氏的这一观点无疑也是千真万确的，因为只要分到足够细小，没有线段不可“共度”。毕氏的观点的致命打击来自意识世界的严密性。当毕氏的学生希伯斯发现边长为 1 的正方形的对角线的长度不是既约分数，并用 $\sqrt{2}$ 表示对角线的长度，因此标志了无理数概念的诞生时，毕氏学派非常紧张，对这一发现实行“保密”，不让外传。从几何的意识性功利目标来说，毕氏这一原则被粉碎了。

鉴于毕氏学派的教训，欧几里得巧妙地在意识世界和物理世界容易达到“共识”的区域借助或物理的“直观自明”或理性的常识概念的一套公设，公理和定义作为量子因子和量子规则，来演绎他的几何体系。公理的是“直观自明”的真理的提法本身就隐含着以生活，经验等物理内容为根据。所以从某种意义上可以说由严密推理建立起来的“纯理性”的欧几里得几何的基础却是“非理性”的经验和物理。虽然欧氏定义的点是无高、无长、无宽的“三无”之点，是典型的只有在意识世界才存在的“意识之点”（物理世界不存在这种点）。但对人们用严格理性审查的眼光加以批评的老问题“没有长度之点怎么能累积成具有长度的线”，欧氏却不予理睬。显然这里他又不得不“借用”了毕氏的物理世界的 $d > 0$ 的点。也许欧氏清楚，在这个问题上意识世界和物理世界是无法调和的。在物理世界不费吹灰之力的事，在意识世界却是无法超越的鸿沟。阿基里斯“一不小心”就可赶上和超过乌龟，但从“严格的理性思维”来说他永远无法超越。

(iii) 欧氏几何的启示

根据上述分析，欧氏在建立其“理性演绎”系统时，至少在下述几点没有拘泥于理性的“阳春白雪”式的严密性，而是借助于“下里巴人”的物理经验：（1）他没有理会怀疑派无穷无尽地问“为什么”的习惯，而是在物理和意识两个世界的“并轨”处选择了一套公设，公理和定义作为建立其几何系统的背景空间或底线。他以此告诉人们，不用无穷无尽地问下去，否则将自进“地狱”；（2）他虽然确立了无高、无长、无宽的“三无”意识之点，但没有理会“没有长度之点怎么能累积成具有长度的线”的责难，因为物理空间的点都可均匀累积成线。

其实后来的极限、微分和积分概念的建立都借用了类似的技巧。据说哥德尔也认为，数学家应该不断从现实世界中发现新的数学问题，然后在此基础上构筑新的数学体系。可见象哥德尔这样的理性思维大师也是主张借助“非理性的直觉”的。他的著名的不完备性定理更反映了意识系统对物理世界的依赖性。

笔者敬请读者注意，这些启示来自被称为纯理性的“科学的皇冠”数学领域。面对现象和物理结构主宰一切的生物医学世界，我们就更有理由借鉴这些科学巨擘做法而无须顾虑来自理性世界苛求和责难：

- （1） 确定要建立问题解决的功利目标；
- （2） 在物理和意识的两可领域获得我们的学术背景空间基石（充分假借经验的或实验的即物理的事实）；
- （3） 在此基础上建立相应的生物医学信息学。

表 1 举例说明欧几里得几何公设的元概念与经验

欧几里得几何公设	元概念与经验
过两个不同的点，能作且只能作一条直线 线段可以任意延长 以任意一点为圆心及任意长距离为半径可作一圆 凡直角都相等 两直线被第三条直线所截，如果同侧两内角之和小于两个直角之和，则两直线延长时在此侧相交	两点之间用尺划线的经验与空间的同质性 事物无限均匀伸展概念和局部空间经验 环绕一点作等距运动的经验和概念 “直角的不变性或空间的同质性” ^[10] 直线延伸，平行，斜行等概念和局部空间经验

表 2 举例说明欧几里得几何公理的元概念与经验

欧几里得几何公理	元概念与经验
等于同量的量彼此相等 等量加等量，其和仍相等 等量减等量，其差仍相等 彼此能重合的物体是全等的 整体大于部分	三个事物的度量的比较、等价概念和经验 事物的度量相加、等价概念和经验 事物的度量相减、等价概念和经验 事物的形态比较、等价概念和经验 整体和部分定义的另一叙述

表 3 举例说明欧几里得几何基本定义的元概念与经验

基本定义	元概念与经验
点是没有部分的 线段只有长度而没有宽度 线的两端是点 线由均匀直放的点组成 面只有长度和宽度 面的端部是线	整体和部分概念的抽象 细长概念的抽象 细长物的截面经验和抽象 点惯性重复运动中连续与离散互转概念 薄而广事物的经验和抽象 薄而广事物切断经验及抽象

5. 人工智能的背景空间

计算机程序是典型的“全透明体”人工智能的实例，用它来研究人工智能原理有其独特的明晰性。任何计算机程序的基础是编程语言。任何编程语言都由各个等级要素及规则集合组成，任何等级的要素及规则集都以其下层的要素及规则集为背景空间，其本身也作为其上层要素及规则集的背景空间。现以 Visual Basic 语言为例说明如下^[14]。

(i) 数据类型及其规则：数据种类，占用字节数，取值范围，操作范围及其行为规律，相互转化规律等。为数据的抽象体系。

(ii) 变量：为数据“实例”，关于变量的规则为数据实例化规则，包括变量的声明方法，初始值（缺省值，标准实例）的确定，变量赋值规则及控制，覆盖范围（有效范围）的确定，固定值变量和可变值变量确定。数组为批量对象的数据变量。

(iii) 运算符：类型包括算术、比较、连接、逻辑运算符（类）等。运算符规则包括：运算功能，操作数要求，表达式组成规则，结果值类型规定，精度不同的混合类操作数的精度转换规定，违规处理（如超出精度的正确范围）等。

(iv) 表达式：为关键词、运算符、数据类型转化函数、变量、字符串常数、数字或对象等的组合。基本表达式类型包括算术、关系、字符串、逻辑表达式。表达式规则：操作数的数据类型要求规定，结果数据类型的规定，异类数据共存时输出数据类型规定等。

(v) 基本语句：包括执行语句，非执行语句。语句规则包括语句结构规定，赋值时数据类型规定，给变量、属性、自定义类型变量（一组频繁共存的变量如学生的一般性数据）赋值的操作类型。执行语句规则还包括赋值，注释，程序退出或暂停，错误处理，跳转，判断，循环语句，语句的嵌套结构等。

(vi) 子程序和函数：语句的复合，包括子过程或函数等。子过程为完成一些经常重复或共享任务的代码体。VB 的子程序包括不返回值的 Sub 过程，返回值的 Function 过程和设置对象引用的 Property 过程。

子过程总体规则包括代码编写机制：组织，编辑和格式化代码规则，语句执行顺序规定和控制等。

Sub 子过程及 Function 过程的规则还包括语法或结构，被访问的范围，可见性，名称或标识符，输入参量要求，变量值的变化性（稳定性），调用形式等，Function 过程还包括返回值的方法和类型等。

(vii) 程序：为子程序，函数，变量，表达式等的复合。规则包括整个程序的组织规则。

各级因子集合和规则集合组成的背景空间的层次结构如图 3。

6. 物理性知识的背景空间

与上述人类智能和人工智能例子相比，关于生物医学世界的知识要复杂的多。生物医学系统以其令人敬畏的、神秘莫测的物理性傲慢地嘲弄着人类思维、理性和数学。由于生物医学知识及其形形色色的背景空间极度复杂，对它们的探讨是一件非常困难的工作。限于篇幅，本文只涉其一二。为易于理解起见，仍以一些生物医学知识实例展开讨论。

(i) 生物医学知识背景空间的学科异质性

知识举例 1：“肝细胞主动分泌胆汁酸盐，使毛细胆管内渗透压升高，使水和电解质进入毛细胆管，因此胆汁酸盐是决定毛细胆管胆汁流量的重要因素，这部分胆汁称为‘胆盐依赖性胆流’（Bile salt dependent bile flow, BSDF）。”^[12]

各层次的作用因子

各层次的规则集合

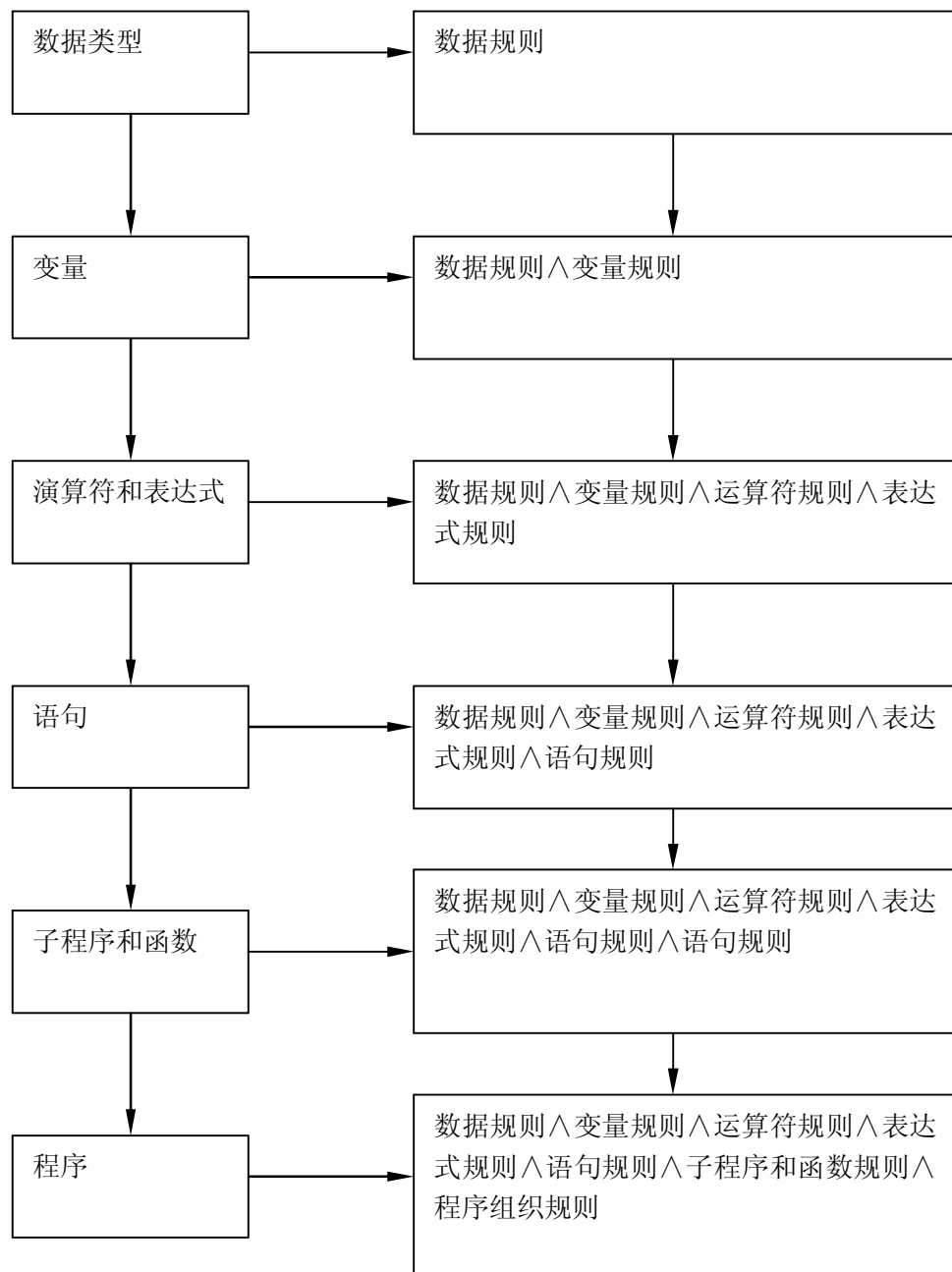


图 3 计算机程序语言的各级作用因子及其规则集合组成的知识

知识背景 1: 包括: (1) “肝细胞主动分泌胆汁酸盐”由细胞生理学规则决定; (2) “毛细胆管内渗透压升高, 使水和电解质进入毛细胆管”为非生理胶体物理学规则决定。前者引发了后者。

知识举例 2: “胆道始于相邻肝细胞形成的胆小管, 这些胆小管逐渐汇集成汇管, 小叶间胆管及较大的肝管出肝门后, 肝总管与胆囊管合成胆总管并开口于十二指肠。胆道中任何一处的胆流受阻均可引起胆汁淤积的特殊临床表现和生化改变。”^[11]

知识背景 2: 对于防病治病功利目标来说, 其基本作用规则至少包括: (1) 一般流体力学规则, 包括管径, 压力梯度, 液体黏度; (2) 生物化学规则: 包括胆结石形成机理和堵塞等。

知识举例 3: “胆汁酸为两亲性化合物, 一端为极性基团 (-OH 等), 另一端为非极性基团 (CH 等), 胆酸和鹅脱氧胆酸含羟基较多, 极性较高, 与胆红素和胆固醇形成的乳粒亲水性强, 较易通过毛细胆管微绒毛分泌入毛细胆管。”^[12]

知识背景 3: 包括: (1) 胶体物理规则; (2) 生物膜规则; (3) 细胞生物学 (微绒毛) 规则。

知识举例 4: “滑面内质网质受损, 羟化胆固醇功能降低, 此时胆固醇直接经侧链氧化而生成大量的石胆酸。”“石胆酸的羟基少, 极性低, 与胆红素和胆固醇形成的乳粒亲水性差, 在毛细胆管微绒毛内发生淤积, 渗透活性降低, 从而使胆盐依赖性胆流减少, 引起肝内胆汁淤积。”^[12]

知识背景 4: 包括: (1) 生物化学规则: “胆固醇直接经侧链氧化而生成大量的石胆酸”; (2) 胶体物理规则: 石胆酸“渗透活性降低”; (3) 细胞生物学规则: 微绒毛在胆汁分泌中的重要作用。

(ii) 知识的描述粒度:

(1) “直接操作子明确-间接操作子不完全明确-关系不完全明确性类”描述: 如不论是物质转化还是池室间转移关系, 一律用箭头表示关系或过程, 如图 4 物质因果链^[12]。

(2) “直接操作子或规则不完全-间接操作子或规则不完全-关系类型明确”知识描述: 以下以“肝细胞对胆红素的处理”三个过程^[12]为例说明之 (其中黑体-斜体字为关系动词, 黑体字与关系动词发生直接相关的一些因子和规则):

摄取: 未结合胆红素**随血一流至肝脏**, 很快就被**肝细胞摄取**, 与肝细胞载体蛋白**Y 蛋白和 Z 蛋白结合** (这两种载体蛋白, 以 Y 蛋白为主, 能够特异地结合包括胆红素在内的有机阴离子) **被动送至滑面内质网**。

结合: Y 蛋白-胆红素和 Z 蛋白-胆红素在滑面内质网内, 未结合胆红素通过微粒体的 UDP-葡萄糖醛酸基转移酶 (UDPGA) 的作用, 与葡萄糖醛酸**结合**, **转变为结合胆红素**。结合胆红素主要的是胆红素双葡萄糖醛酸酯, 另外有一部分结合胆红素为胆红素硫酸酯。这种胆红素的特点是**水溶性大**, 能从肾脏**排出**, 胆红素定性试验呈直接阳性

反应。故称这种胆红素为结合胆红素，也称直应胆红素。

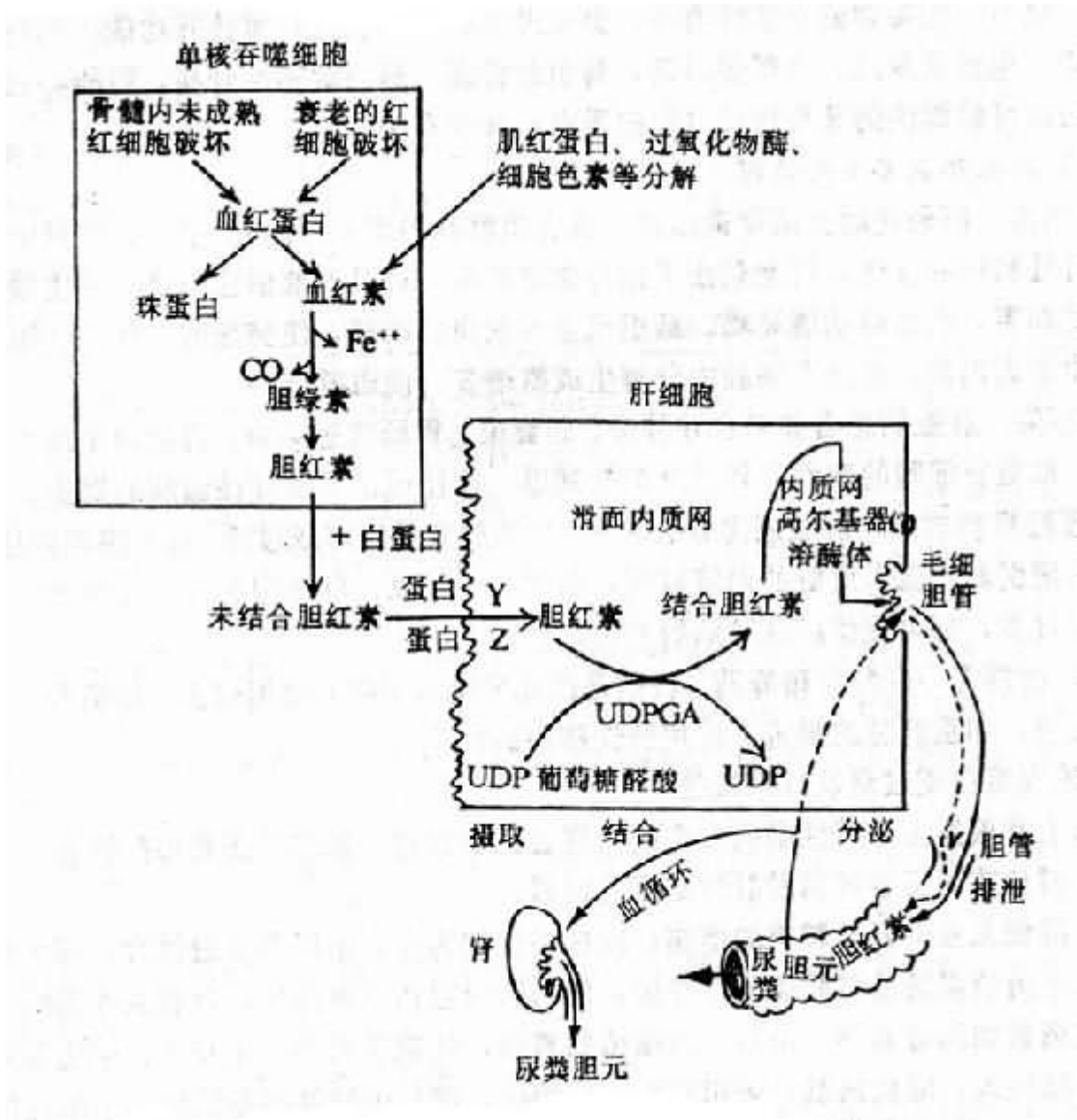


图 4 正常胆色素代谢途径 (UDP: 尿嘧啶核苷二磷酸, UDPGA: 尿嘧啶核苷二磷酸-葡萄糖醛酸基转移酶)

分泌: 结合胆红素在肝细胞浆内, 与胆汁酸盐一起, 经胆汁分泌器 (高尔基复合体在细胞分泌过程中有重要作用), 被分泌入毛细胆管, 随胆汁排出。由于毛细胆管内胆红素浓度很高, 故胆红素由肝细胞内分泌入毛细胆管是一个较复杂的耗能过程。

(iii) 知识背景底层空间: 正如上文所述, 不同的意识或物理的功利目标决定其知识背景空间及其量子化因子和规则, 也决定“底层背景空间”。从临床医学的功利目标来讲, 致病、

致残、致死机制均属“底层背景空间”，组成“底层背景空间”的是量子性因子和量子性规则，其余则以与“底层背景空间”的距离远近分层。在此功利目标下，底层背景空间以下的因子和规则均毋须深究。图 5 以“肝内胆汁淤积的发病机制”为例说明之。

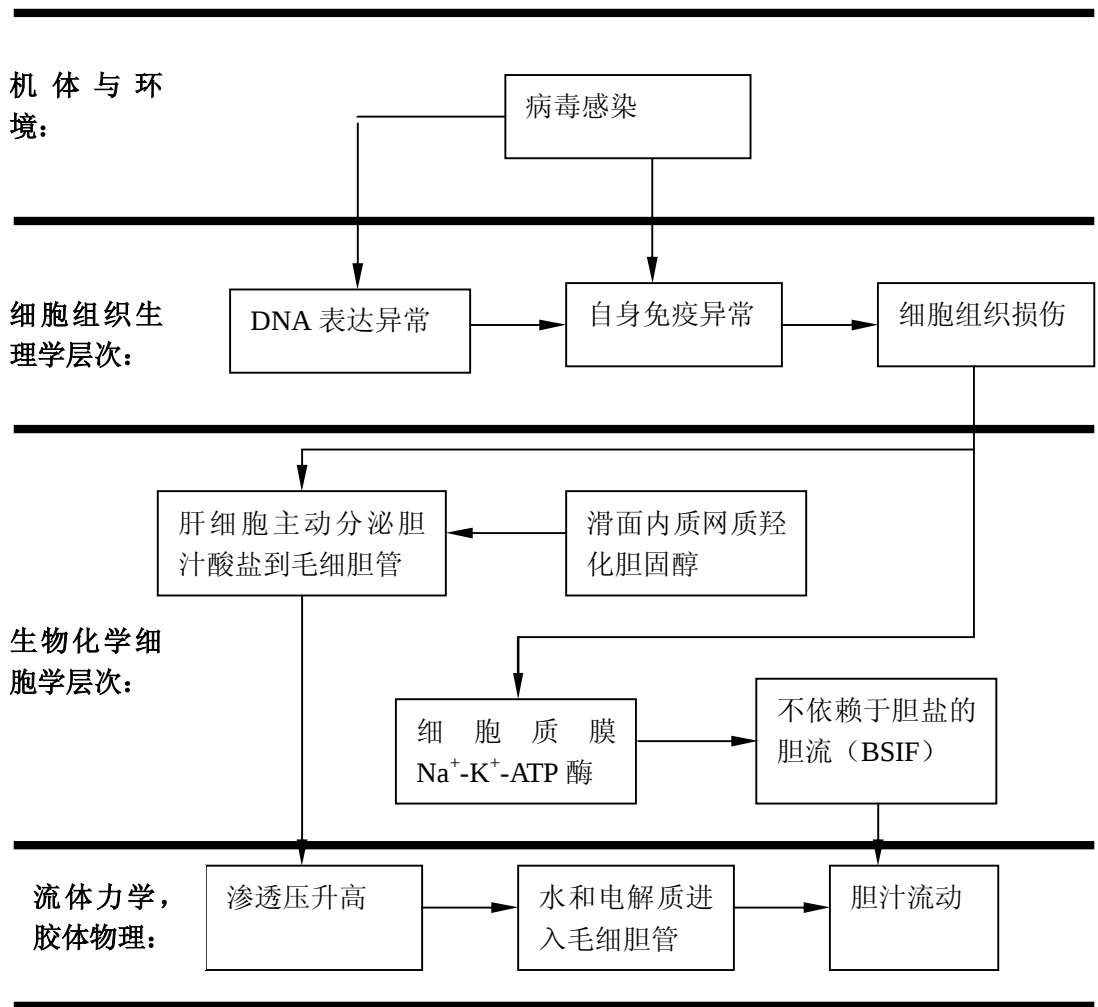


图 5 关于肝内胆汁淤积的发病机制的知识的动力学的层次，箭头表示正或负的作用，流体力学和胶体物理机制组成“底层背景空间”。

(待续)

参考文献

1. 包含飞：生物医学知识整合论(一)，《医学信息》杂志，16（6）：174-279，2003

2. 包含飞: 生物医学知识整合论(二),《医学信息》杂志, 16 (8): 410-415, 2003
3. 包含飞: 生物医学知识整合论(三),《医学信息》杂志, 16 (11) :602-605, 2003
4. 包含飞: 生物医学知识整合论(四),《医学信息》杂志, 16 (12): 666-670, 2003
5. Bao H.F.: The Theory of Biomedical Knowledge Integration(I), Chinese Journal of Medical Treatment, 2(13): 1-6, 2003
6. Bao H.F.:The Theory of Biomedical Knowledge Integration(II), Chinese Journal of Medical Treatment, 2(21), 1-6, 2003
7. 包含飞: 生物医学知识整合科学思维的成分异源性 (I),《计算机应用与软件》: 上海信息技术应用文集, 20 (增刊): 171-173, 2003,
8. Mark A. Musen :Medical Informatics: Searching for Underlying Components,
http://www.smi.stanford.edu/pubs/SMI_Reports/SMI-2002-0916.pdf
9. 兰纪正、朱恩宽译:《几何原本》第 I 卷, 台湾: 九章出版社, 1992
<http://www.i-mikekong.net/Maths/mathsf-frame.php>
10. 蔡聪明: 从毕氏学派到欧氏学派的诞生,
http://episte.math.ntu.edu.tw/articles/sm/sm_26_02_1/index.html
11. 默克诊疗手册: http://www.msdchina.com.cn/manual/section_4/4-36/a1.html
12. 大众医药网/文献资料/医学书籍/病理生理学 :
<http://www.windrug.com/pic/30/11/11/072.htm>
13. JH van Bommel, MA Musen 主编, 包含飞, 郑学侃主译:《医学信息学》, 上海: 上海科技出版社, 2002
14. 赵筠编著: Visual Basic 中文版基础培训教程, 北京: 人民邮电出版社, 2003