

《医学信息》杂志，2003，第16卷，第6期  
生物医学知识整合论(I)

包含飞  
上海中医药大学(200032)  
中医学信息化-标准化研究室

生物医学信息学除了本身具有完整的理论结构、系统的方法学及各个领域中的应用研究<sup>[1]</sup>以外，还担负着生物医学数据、信息、知识的大整合的新的历史使命，本研究把生物医学数据、信息、知识统称为生物医学知识，因此这一使命也可称为生物医学知识大整合。这一使命具有无与伦比的重要性，其艰难程度也超乎人们的想象。它要求生物医学信息学更深入地、更密切地、更具体地涉及生物医学各个方面的数据、信息、知识。目前人们很难想象生物医学知识大整合究竟涉及多大空间，也很难区分这种新研究究竟是属于新的信息学还是新的医学。

### (一) 生命系统——物理整合的大师

是我们人类的科学活动而不是机体本身把机体分解为量子生物学、分子生物学、生物物理学、生物化学、细胞学、组织学、解剖学、生理学、免疫学、遗传学.....各层次。相反，大自然把万千机制都有条不紊地融为一体，把无数元素、无数机制融合和联合成一个完美的动态网络。因此，笔者认为把所有的生物医学知识重新连接或整合起来的努力顺应了大自然的意志。

机体的自然整合能力令人赞叹。为了显示机体的天赋整合神力之点滴，笔者总结了心脏搏动周期中从分子生物学层次到心脏整体泵血层次各循环机制相互衔接的动力学定性模型，这些动力学循环分别由诸如离子通道、离子泵、膜电位、动作电位、内质网与钙离子、ATP酶、肌丝蛋白(粗肌丝、细肌丝)、神经递质-受体、突触、心肌细胞、心肌、流体力学、活塞原理等结构和生理机制所驱动和执行，神奇地相互协调地完成正常的泵血循环。一个正常的泵血动作依赖于这些异质驱动力循环各相态的相互准确衔接，最后整合为心脏整体的泵血行为。任何相态之间的不准确衔接均有导致疾病的可能(见图1)。

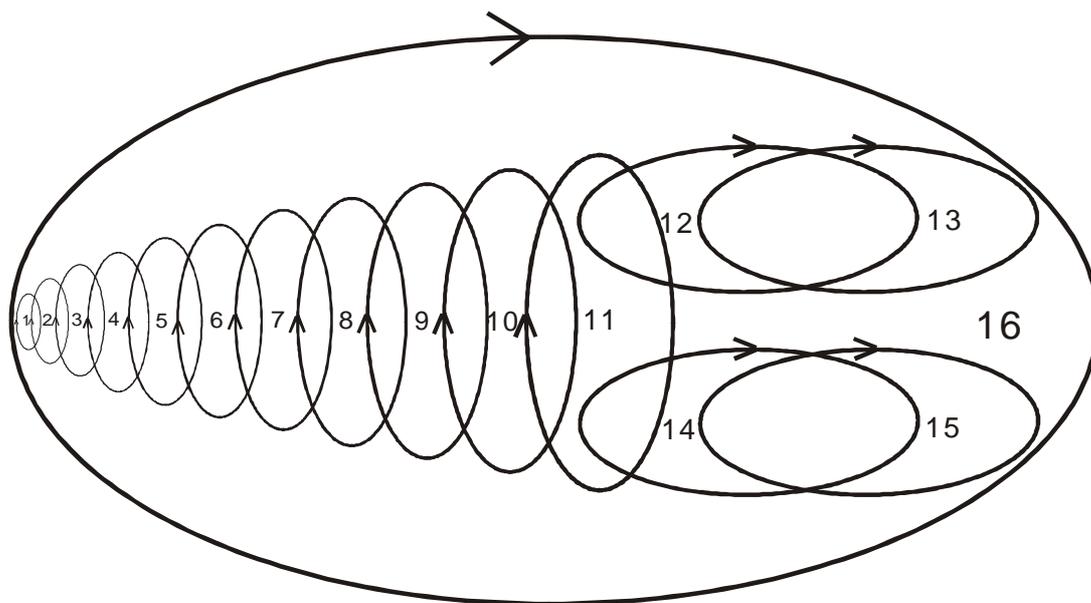


图 1 心脏泵血循环连接图：1.膜电位循环；2.钙离子从肌浆网收放循环(借助于钙泵即  $\text{Ca}^{++}$  依赖式 ATP 酶的循环)；3.肌钙蛋白与  $\text{Ca}^{++}$  结合与释放循环；4.肌钙蛋白分子构型循环；5.原肌凝蛋白的构型变化循环；6.横桥与肌纤蛋白结合-扭动-解离循环(依赖于肌球蛋白的头部即 ATP 酶激活-分解 ATP-释放能量-抑制循环)；7.粗肌丝与细肌丝的滑动循环；8.肌节的缩舒循环；9.肌纤维的缩舒循环；10.肌肉的缩舒循环；11.心室内血液压力循环；12.房室瓣开闭循环；13.心房内血液压力循环；14.主动脉瓣开闭循环；15.主动脉内血液压力循环；16.心脏泵血循环。

## (二) 生物医学知识整合——逐步逼近对人体的整体理解

提出生物医学知识整合问题的理由，一方面在于我们对世界的整体性本质的理解，在于对人类认知活动的自然归结的理解，即对人类认知活动的本义的理解。令人担忧的是这种整体性本质在当今崇尚“拆零”的主流科学家眼中已十分漠然。另一方面出于一种责任，试图提醒人们在生物医学知识中就事论事地或满足于在越来越狭窄的领域中理解和考虑问题是近利短视的，是偏狭的。人们应该时刻以整合的眼光来看生命系统，提醒并记住这么一个事实，即在复杂的生命系统面前我们也许仅仅相当于一个一知半解的小学生。从整合的观点来看，即使是两条看来逻辑上十分密切的知识，要在物理上联系起来是相当困难的，有多少医学知识盲区和裂隙需要探索、开发和贯通。整合性研究也将激励一种无穷无尽的整体探求的欲望，激励我们不断努力把已获得的知识相互连接起来。

说到底，整合问题是个局整问题。局整问题称得上是一个基本宇宙矛盾。它与另一个基本原理即所谓经济原理（或称最低能量通道原理）有内在的联系。个体的本能行为总是“急功近利”的，但群体的前途（实际上最终也是个体的前途）却并非完全决定于个体的“急功近利”行为，很多情况下往往适得其反，群体的前途决定于个体行为的“总和”或整合。再往深处探究也许要追溯到更基本的原理即场或势的空间不均匀原理（场或势的空间距离原则），即任何事物在广义和狭义空间中都有近强远弱的规律。

生物进化或人类行为常常为局部-整体问题“伤透脑筋”。有科学家指出,生物进化的基本机制是小尺度的、急功近利的、短视的,不是大尺度的(长远的、全生物种群的)、理性的。生物进化常常会出现这样的情况,从目前讲这一突变有利于生存的,但久而久之,许多对当前有利的小突变积聚成一个大突变,却在总体上却不利于生存,甚至造成物种灭绝。例如物种的数目极度增加就是一例<sup>[4]</sup>。古人说:“民以食为天”,现代人很多已不用为温饱发愁,但可悲的是很多人为了为一时口福或仅仅是为了摆摆阔气以满足虚荣而大办筵席,大吃珍稀动物,这可能把地球上的物种斩尽杀绝,从而把人类自身推入绝境。

人类在科学研究上也有明显的急功近利倾向。科学探索行为一般都是先从获得简单的、直接的知识着手,通过经验的(相对地不精确)和实验的(相对地精确)手段获得二二关系,其简单的数学表达形式为: $X=F(Y)$ ,至今我们的医学图书馆或知识仓库中已经堆满了各式各样广义的、狭义的、概括的、具体的“二二关系”,可形式化表达为 $X_i=F_i(Y_i)$ , ( $i=1,2,3, \dots$ )。但这些知识所展示的是在限定范围内或仅仅是在实验条件下的简单的、离散的“生命关系”,真正的“生命世界”要复杂得多,真正的“生命世界”决非仅仅是二二关系,而是三三关系,四四关系,……以及它们的多重复合关系。

长期以来,知识整合就是很多科学家特别是系统学家的追求和理想,但往往被放在哲学或数学等理论象牙塔中进行讨论,或把它当成一种智慧追求或思维享受,并不实际地论及具体的知识整合问题。无庸讳言,生物医学知识系统是个令人生畏的超级复杂系统。一旦开始实际的医学知识的整合问题,我们就可能立刻掉入万劫不复的云天雾海中,难有出头之日。以此看来,生物医学知识整合似乎是非常非常遥远的事情,现在来具体谈论这个问题是否为时尚早。但笔者理想的知识体系是完美无缺的,结构清晰的。即便我们一时不能达到全部目的,我们也不能放弃这种追求,更不应该盲目等待。

自然和生物系统对人类认识来讲是无穷的,人类对它们的认识具有相对性,人类对生命的知识和理解总是取决于当代的科学手段和认知水平。对于人类的认知能力来讲,它们永远是可望而不可及的迷,人类对它们的理解永远只能是一种模拟和近似。所以知识整合的任务是不断地、逐步地逼近对人体系统的整体理解,不断地、逐步地逼近对疾病及其控制过程的整体理解。

### (三) 人类认知过程的关系拓扑图

人类获得对世界包括人类本身的知识 and 理解的过程是认知过程,因此知识整合与认知科学有着内在的联系,同时知识整合也离不开丰富多彩的知识表达方式。

先谈谈人类认知过程的最基本的拓扑图。

人类的认知对象是客观世界,当然也包括我们人类自身和认知过程本身。作为人类的认知对象的客观世界原型(或实例),总是“犹抱琵琶半遮面”,我们永远只能观测到或理解它的一部分,即表型部分;它们总是把另一部分,即潜型部分,藏匿起来。此为大自然(包括人类自身和认知过程本身)的神秘莫测部分,尚未被观测和理解部分,它永远是人类及其科学家们所毕生追求的对象。

我们知道,人类从客观世界所获得的知识包括二类即叙述性知识和过程性(即干涉性)知识。前者是对世界的描述,回答“是怎样的”,“是什么”;后者与如何改造世界有关(包括为了科学研究的认知性干涉),回答“怎么办”。在生物医学知识中,二者往往不能截然分开而反反复复地组合在一起。有些医学过程性知识如外科手术为技艺性知识,其实施对人的依赖性很大。根据成熟程度、详尽程度或理解的深度,知识可分为经验性或实验性知识和科学理论性知识,前者一般只知其输入与输出关系,而不知道为什么会有这种关系,其中间过程为“黑箱”;科学理论性知识是较为成熟的知识,不仅知道输入与输出关系,而且还知道为什么

会发生这种关系，中间过程是明的，为“白箱”。但大部分知识，特别是生物医学知识，是半经验性半科学性知识，输入-输出关系的中间过程半明半暗，被称为“灰箱”。作为认知主体的大脑，其思维或认知方式可以分形象思维和逻辑思维二种。形象作为认识对象是一种未经加工分析的原型。很幼小的孩子就有能力识别妈妈的音容笑貌，就是一种形象思维，是对形象或场景的认识，往往是一种直观的无须训练的本能的思维能力。逻辑思维一般是经过逐步训练的抽象和分析思维能力。

本文希望指出的一个事实是知识的储存和表达有显性和潜性二种属性，二者与生物医学知识工程有着密切的关系。人类大脑在思维过程中大量应用潜型知识，而人工智能处理的一般只是知识的显性部分，此问题与人工智能的成败极为关键，但其重要性往往认识和估计不足。更为重要的是人们往往忽略了显性知识和潜性知识在储存媒体方面的本质差异。显性知识的存储媒体为我们所熟悉的话语、文字、信号、数字数据、字符数据、文本、图形、图象、录音、录象等媒体之中的知识；而潜性知识的存储媒体形式为神经系统，其本质和特点我们至今了解甚少。人们往往盲目地认为把潜型知识转化成显型知识是轻而易举的事，但问题并非如此。相反，事实证明潜型知识并不能轻易地转化成显型知识。由于对此类问题认识不足，致使人工智能研究在其发展中走过很多弯路，吃过很大的亏。

我们不可能离开人类当时的认知水平、认知方法手段或认知语境（context）来谈论世界是什么样的。大脑作为认知主体从认知对象原型（或实例）获得数据必须通过某种观测方法或手段，或视（或摄像）、或听（录音）、或触、或电，或磁、或电磁、或化学、或动（动态、动力等）、或静……观测方法决定观测数据，认知参量或广义变量及其性质。任何知识都或明或暗地依赖于保证其成立的语境或条件，无一例外，离开了这种具体语境或条件，它们就不一定能成立。这一原理可称为知识的语境原理。牛顿三大定律的背景是常观常速世界，在量子领域或宇观领域就不一定成立。又例如一条最简单的生理学知识，“成年人心率平均每分钟75次，每个心动周期持续0.8秒。”我们知道这条知识成立的语境为一般正常人群（不是病人和运动员人群）、安静状态（非运动或激动状态）、男女性成人年龄段的心率平均值，因此这条知识只能应用于如此定义的“标准人体”的整合上。即使标准的、经典的或被称为本体的知识也有其相应的语境。知识对其语境的依赖性往往远远超过人们的想象。因为知识的语境虽为知识的固有组件，但基于经济原则，人类在自然语言或其他信息交流中对双方不言自明的部分常常予以省略，固在表达上它们往往并非必需部分。这一重要原理也就容易为人们所忽视。

图2是关于认知关系的潜性和显性的拓扑图。从此图我们也许可以领会到医学人工智能和医学知识整合的高度复杂性。

#### （四）生物医学知识整合论的基本概念

整合一词为多义词，在很多场合并不是严格定义的概念，有时仅仅是“综合”，“一起考虑”，“联起来”的代用词。本文讨论的生物医学知识整合是指“严格科学意义下的物理性生物医学知识整合”或“严格逻辑学意义下的定性生物医学知识整合”，对其所依赖的整合机制来说，它们最终的目标应该是“无缝的”（seamless）整合，是“滴水不漏”的整合。

事物或对象的全部称为整体。整体相对于部分而言，整体依赖于我们观测视角和操作目标，例如个人、家庭、公司、国家、地球在不同的情况下，可以被视作整体，也可以被视作部分，因而整体是一个相对概念。如果两个事物不是相互独立，我们就称二者之间存在关系。非连续的即离散的关系由四要素组成：关系作用子、关系主体元素（集合）（ $\tau$ ）、关系客体元素（集合）（ $\sigma$ ）、关系条件元素（集合）（ $\gamma$ ）。关系作用子反映关系本质的逻辑要素，为关系构成的逻辑驱动力，驱动关系的成立。关系作用子的类型取决于我们的观测视角，因此带

有某种任意性。逻辑作用的直接施与者为主体元素（集合），直接受施者为客体元素（集合），与关系实现间接相关的是条件元素（集合），三者都是关系成立的物理要素。在由笔者提出的整合性关系化医学电子书（IREMB）研究中的人机共解互通语言（HCSL）按逻辑属性和操作属性关系作用子被分为泛化的相生（CREAT）、相协（INCRS）、相克（或相制）（DECRS）、相涵（CONTN）、相并（COMPN）、转化（TRANS）、等价（或赋值）（EQUAL）、泛序（PORDR）、泛移（PASTO）、泛离（PNULL）十大关系<sup>[6]</sup>。由美国国立医学图书馆主持的 UMLS 项目<sup>[7]</sup> 的关系作用子可分为两大类型：①主观（理性）类，如“is\_a”，“conceptually\_related\_to”；②客观类，如“physically\_related\_to”、“spatially\_related\_to”、“functionally\_related\_to”、“temporally\_related\_to”。UMLS 的关系元素被分为实体（entity，指相对静态、固定事物）和事件（event，指相对运动、变化和过程类事物）两大类。二者又可根据主观（理性、概念）-客观（实在、物理）、正常-异常、自然-人工、定量-定性、宏微组成（局整关系）、发展历史、结构-功能等轴向或这些轴向的“包含”、“与”、“或”关系进一步细分。由于关系元素和作用子的抽象程度不同，其可确定性和可操作性也不相同。保证关系成立的必要的元素集合称为关系成立的必要条件，它由必要的关系主体元素（集合）（ $\tau$ ）和必要的关系条件元素（集合）（ $\gamma$ ）组成。只要关系实施的必要条件得到充分满足，关系即告成立。

对两个有序关系，如果第一个关系实施的结果或输出能充分满足下一个关系实施的必要条件，下一个关系即告成立。这种两个或两个以上的关系的联立称为关系复合。所以关系复合成立的基本机制是一种求供索交机制<sup>[8]</sup>。若求供索交结果可以满足供求二关系实施的必要条件，供求二种关系的复合即告成立。整体意义下的关系的复合称为整合。若整体意义下所有的供求二二关系之间的求供索交的结果可满足所有的二二关系实施的必要条件，整合即告成立。

人类的科学研究活动把世界分析得越来越细，从中我们获得了大量分析意义下的知识和细节，整合的任务是根据这些分析性知识和细节重新组装成一个模拟世界或世界模型。

关系的选择带有某种任意性，因此整合的选择也带有某种任意性。决定一种整合是否有效的最后认可权属于生物医学系统。其中经济性是一个关键。如果整合是经济的（有优良的广义性价比），整合后事物的适应度或生存度是增加的或最佳的<sup>[4]</sup>，那么该整合是有效的并可能被接受及留存下来。

与生物医学知识整合有关的操作和运算可以是数学的、普通逻辑的、医学逻辑的（概念的抽象过程就是一种逻辑整合行为）或物理的（即有生物医学操作性的）。一般来说，战略性整合是一种逻辑性整合，战术或战法性整合是物理性整合。如一部《三国演义》就是逻辑整合，而具体的赤壁之战就是物理整合。

我们可以对某些整合进行运算，如果关系类型  $A, B, C, D, \dots \in F$ ，那么整合  $(A, \tau, o, \gamma)$ ， $(B, \tau, o, \gamma)$   $(C, \tau, o, \gamma)$   $(D, \tau, o, \gamma)$ ， $\dots \in (F, \tau, o, \gamma)$ 。例如有二杂交技术确立的蛋白质之间的二二关系，设  $\Lambda$  为蛋白质之间的相互关系， $\lambda_1$  为与细胞周期控制有关的蛋白质之间的相互关系， $\lambda_2$  为与信号传导有关的蛋白质之间的相互关系， $\lambda_3$  为与小泡运输有关的蛋白质之间的相互关系， $\lambda_4$  为与细胞极性有关的蛋白质之间的相互关系.....因为与细胞周期控制，信号传导，小泡运输，细胞极性，RNA 处理/修饰，RNA 接合，有丝分裂，蛋白质合成，碳水化合物新陈代谢，细胞张力，染色质/染色体结构，蛋白质衰变，氨基酸新陈代谢有关的蛋白质相互关系均属于蛋白质相互关系，也即

$$\lambda_1, \lambda_2, \lambda_3, \lambda_4 \dots \in \Lambda,$$

那么这些关系整合的网络都从属于“蛋白质相互关系网络”<sup>[9]</sup>。即

$$(\lambda_1, \tau, o, \gamma), (\lambda_2, \tau, o, \gamma) (\lambda_3, \tau, o, \gamma) (\lambda_4, \tau, o, \gamma), \dots \in (\Lambda, \tau, o, \gamma)。$$

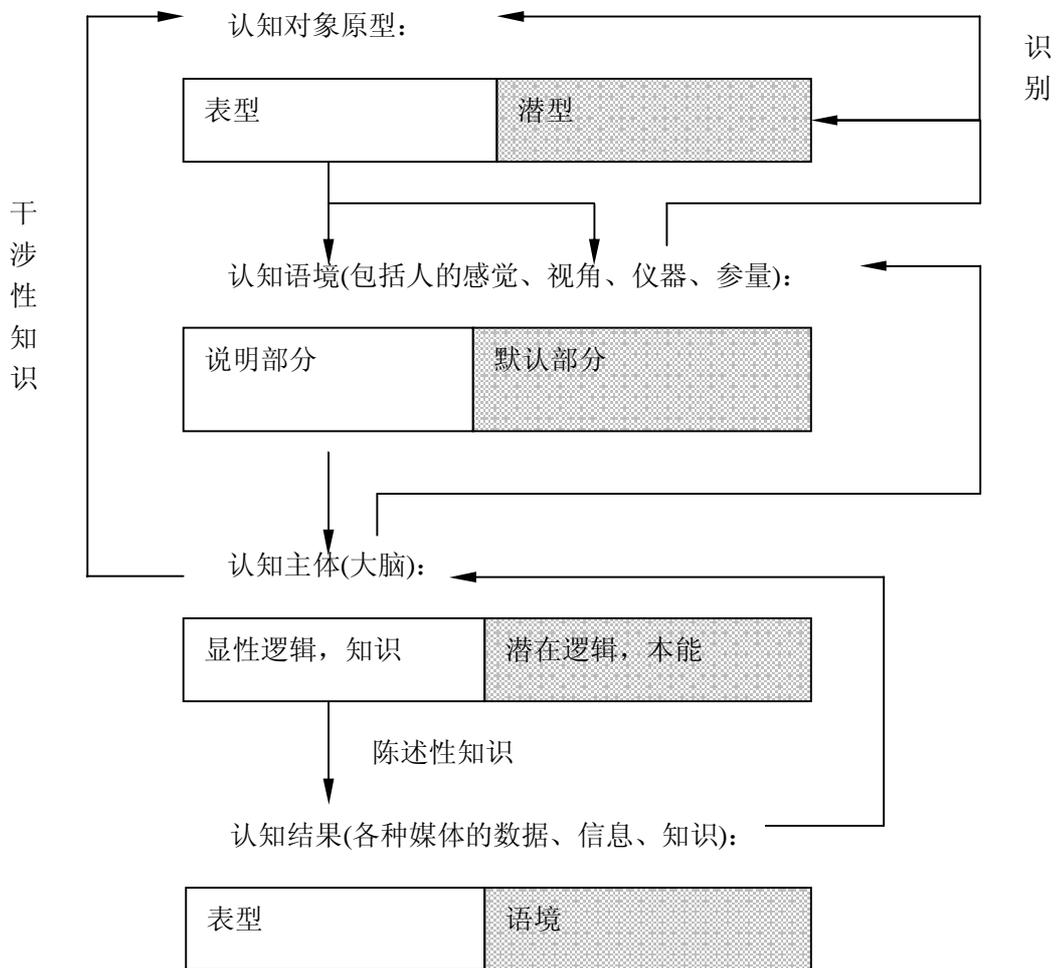


图2 认知对象、语境、主体、结果的关系拓扑图。认知对象的表型部分表示已知部分，潜型表示未知部分；认知语境说明部分表示明确说明的知识获取的手段、仪器设备、观察指标、眼睛视觉、耳朵听觉、背景等，默认部分是其公认的、不言自明的部分；认知主体的潜在逻辑指大脑的潜在的知识模势；认知结果表型为存储或表达的数据、信息和知识部分，语境指上下文、语势等。

### (五) 生物医学知识整合研究举例

当今生物医学科坛，大规模的数据、信息和知识的整合性研究业已兴起。科学家在不同领域和不同的层次、从不同的角度、根据不同的机制、面向不同的目标来讨论、探索和研究整合问题。但截止今日，笔者尚未见到关于生物医学整合的基础的、严谨的和系统的理论研

究。如上所述，整合的驱动力是关系，所以不同的关系或关系的复合可产生不同的整合。很多所谓整合研究是指非生物医学意义下的整合，如信息系统之间物理上的连通或数据之间在通讯协议基础上的相互访问，包括医院信息和信息系统整合（或称集成），多媒体信息整合等。有关生物学和医学知识的整合可进一步分成几个层次。自古以来就有哲学意义下的朴素的整合，如中医学的整体观。但对生物医学发展有着革命性影响的是逻辑意义下和物理意义下的生物医学知识整合。前者如美国国立医学图书馆主持的 UMLS 的语义网络及方兴未艾的医学知识本体研究，二者都可发展为逻辑意义下的生物医学知识整合。最有实质性意义的整合研究包括数据或数据库层次的整合，分子生物学（基因组、功能性基因组或调控网络、蛋白质组、蛋白质相互作用网络、蛋白质空间结构网络）等生物医学整合，生理学整合，生物化学整合，分子-细胞-器官的信息的“垂直”整合，基因-基础科学-临床科学整合，解剖学（生理学、药理学）虚拟人体等。这些是严格科学意义下面的物理操作性整合，甚至定量的或无缝的整合。

较为全面地介绍所有领域的这方面的发展并非易事，本文仅举几例最有影响的新的整合研究团体及其“宣言”，生物医学整合研究及其研究者的学术气势和胆魄从中可见一斑。

(1) 美国西亚图：系统生物学研究室<sup>[10]</sup>

该研究所创始人 L. Hood 教授在关于该所的任务的讲话中指出，“21 世纪对生物学和医学最最主要的挑战是对生物系统的理解，研究所的任务是实现生物学是一门信息科学，完成生物学和医学革命。”该所的宗旨被定为：“对生物系统在系统层次上的理解”。追求各层次（分子层次、细胞层次、组织层次、器官层次、整体层次）与各系统（各种生理和病理系统、器官系统）协调一致。

(2) 美国 Entelos 虚拟病人项目<sup>[11]</sup>

该项目应用其专利技术 PhysioLab 开发动态的大规模人类疾病的计算机模型。每个 PhysioLab 整合在某种疾病背景下包括基因组、蛋白质组、生理学和环境在内的各种数据，重点是理解和测定某种新疗法的临床反应。对经验的和科研的各种生物学探索加以整合，进行数学的和系统工程学的研究，以此促进药物学研究。其成员由生命科学家及系统工程科学家组成。

华盛顿大学放射学、生物工程教授 J. B. Bassingthwaight 说：“Entelos 模型的疾病状态模型开发是一种崭新的和富有科学趣味的研究...它预示整合生物学的一种新的冲击波，为理解健康和疾病中的生物学复杂行为创造系统性基础。”

(3) 英国牛顿研究所的一个专业组的研究：从分子到整体生物学的垂直整合<sup>[12]</sup>。

这里是有关该专业组宗旨的一段论述：生物学系统在个体和群体层次都表现了不同的行为，在各层次的经验的和理论的研究中我们都获得了大量知识，但我们对从个体层次的行为如何整合和协调并进而导致群体层次行为的内在机制很少了解。专业组将把最前沿的实验研究者和理论研究者及其交缘领域的研究者结合起来，确定和探索新的关于垂直整合 (vertical integration) 问题的交缘学科方法。具体研究将涉及细胞、发生生物学到生理学的广阔领域，如细胞信号，形态发生，基因网络，神经元网络和神经生理学，心脏病学等。

(4) 美国斯坦福大学 Bio-X 项目<sup>[13]</sup>

Bio-X 孵化基础科学、应用临床科学中的方兴未艾的各种前沿研究，促进横跨从分子到机体所有科学的未来的发现和技术进步。1998 年，斯坦福大学正式开展生物科学和工程科学之间学科间合作，被命名为 Bio-X，斯坦福大学为此作出了认真的承诺。Bio-X 项目把工程、物理、化学、信息科学与生物学和医学的结合以孵化新的发现和创新，创造机会促进传统的分散学科相互有机地联结以获得新的重大发现。

(5) 英国剑桥生物学整合团体 (Cambridge Biology Integration Group, BIG) [14]

生物界越来越多的全基因组完成测序将引航新纪元的生物学研究并对各学科分枝产生深刻的影响。从细菌的基础基因学到发育生物学,从进化生物学到与人类疾病有关的复杂性细胞生物学,一系列的尖端的后基因时代技术正在被开发,对生物学过程的研究将达到空前详尽程度。探索新生物学并充分理解新技术的潜力的关键之一是驾驭不同领域的专业知识,为探索生物学问题提供整合性的多学科研究方法。

(6) 衰老生物学 e-科学整合和模拟系统 (Biology of Ageing e-Science Integration and Simulation system, BASIS) [15]:

衰老是个复杂的生物学过程,涉及一系列影响分子、细胞、组织和整个机体层次的生物化学机制之间的相互关系。在衰老的分子生物学和基础基因组学 (genomics) 方面已取得了巨大进展,但当务之急是开发能整合各种生物学资源的数据和假设的生物数学和生物信息系统。这一设想导致了衰老生物学 e-科学整合和模拟系统的开发,旨在鼓励和支持生物科学、生物信息学、IT、计算机科学、数学和统计学在前沿领域中的相互结合。

## (六) 十年追求“弹指间”

由上文可知,本系列文章希望在医学信息学领域中开拓一种新的研究,探索在多大广度、多大程度、多大精度上我们可以把当今的医学数据或知识整合起来。我们可把这一研究方向称为生物医学知识整合 (Biomedical Knowledge Integration, BMKI) 研究。其始创灵感不仅来自生物医学处理对象生命系统的高度整体性这一本质特点,而且来自倡导一种关注医学知识全体的新科学精神的愿望,希望尽可能减少科学家终身钻在各自狭窄的领域中所造成的局限性和盲目性。再者,生物医学知识整合研究毫无疑问地将有力地支持包括医学人工智能在内的其他医学知识工程的发展,因为很显然这些医学知识工程研究无法回避生命系统各概念、各事件在整体意义下的复杂的相互关系。

笔者涉及医学知识模拟和整合研究始于 1989 年<sup>[16-19]</sup>, 1991 年首次正式提出“大规模医学知识整合”研究方向<sup>[20]</sup>。历经医学的计算机定量或定性模拟与运算 (QMSOC) 的理论探索<sup>[16-19]</sup>, 全息编码编辑技术 (HICE)<sup>[21]</sup>、整合性关系化医学电子书 (IREMB)<sup>[22]</sup>、人机共解互通医学语言 (HCSL)<sup>[23]</sup> 探索及 IREMB 国际互联网网站建设等多个发展阶段。常常不无感慨地谈起那些不为人们理解的岁月,十多年来笔者集理论探讨、算法设计、数据采集、程序开发、网站维护等工作于一身,常自嘲新研究为“单枪匹马巨型工程”。医学知识整合新研究自提出后,很长一段时间形单影只,孤独无援。“整合”一词在当时国内科学领域中还是个冷僻单词,笔者为寻找其英文对应词“integration”曾化了些功夫。由于当时 Internet 尚未普及,笔者对国外相关研究所知甚少。令人感叹不已的是形势发展竟如此迅速,仅隔短短的十多年,弹指一挥间,生物医学知识整合性研究的势态已蔚为壮观,Internet 的普及更使其如虎添翼。因此更应该庆幸的是 BMKI 研究的“生逢其时”。

我们知道,人类基因组全体基因的测序工作刚完成不久。作为人类科学史上一个伟大工程,人类基因组工程对生命科学作出了莫大贡献。但有人借此而提出的“一个基因一个病”这种极端还原论观点,这种观点最后受到严重挫折完全是意料中的事。但挫折却使科学家们更清醒地认识到,由一个一个基因、一个一个蛋白质(或酶)提供给我们信息和对生命的理解是有限的,我们必须从众多基因众多蛋白质之间的复杂的关系(网)中寻找关于生命和疾病的真谛。于是后基因组时代应时而至。上文已经指出,诸如“蛋白质网络整合”、“药物本体整合”、“生物化反应网络”、“多个独立研究的医学证据统计学整合”(指循证医学的 meta-分析)等新的生物医学整合研究不断涌现,层出不穷。有道是“有朋自远方来,不亦

乐乎”，众多的志同道合的新研究对 BMKI 研究是一个莫大鼓舞。可以相信，BMKI 研究终将成为未来生物医学研究的主调。（待续）

#### 参考文献

1. 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
2. 张镜如主编, 乔健天副主编:《生理学》, 北京: 人民卫生出版社, 86-88, 39-45, 1996
3. 包含飞: 续议中医学是复杂性科学——中医标准化预备研究之三 (待发表)
4. 赵南元:《认知科学揭秘》, 北京: 清华大学出版社, 2002
5. 刘沁, 包含飞: 人机共享语言 HCSL (医学) 的第二版的计算机实现, 中外名医杂志, (1): 78-80, 2001
6. 包含飞: 创造一种人机共解互通的医学语言 (西医学部分), 第八届全国医药信息学大会论文集, CMIA' 99 (电子工程师增刊), 1999, 98: 18-24
7. Web site: UMLS Knowledge Sources 14 Edition, January Release, [http://www.nlm.nih.gov/research/umls/UMLSDOC\\_2003AA.pdf](http://www.nlm.nih.gov/research/umls/UMLSDOC_2003AA.pdf)
8. 吴学谋: <泛系——不合上帝模子的哲学>, 武汉: 武汉出版社, 1996
9. Benno Schwikowski, Peter Uetz, and Stanley Fields: A network of protein-protein interactions in yeast, <http://itgmvl.fzk.de/www/itg/uetz/publications/Schwikowski2000.pdf>
10. Klaus Prank : Data mining and mathematical modeling in systems biology, [www.systemsbiology.org](http://www.systemsbiology.org)
11. Web site: **Entelos is the leader in predictive biosimulation.** 美国 Entelos 虚拟病人项目, <http://www.entelos.com/science/index.html>
12. Nick Monk: Gene Networks Bridging the Gap Between Bioinformatics and Mathematical Modelling, <http://www.newton.cam.ac.uk/programs/ICB/icbw01.html>
13. Web site: <http://cmgm.stanford.edu/biochem/biox/intro.html>
14. Web site: Cambridge Biology Integration Group (BIG), <http://www.gen.cam.ac.uk/~big/>
15. Web site: **Biology of Ageing e-Science Integration and Simulation system (BASIS):** [http://umbriel.dcs.gla.ac.uk/NeSC/action/projects/project\\_action.cfm?title=91](http://umbriel.dcs.gla.ac.uk/NeSC/action/projects/project_action.cfm?title=91) 或 <http://www.basis.ncl.ac.uk/index.html>
16. Bao H.F.: The structure characteristics of the new research QMSOC and its relevant operators, J Tongji Med Univ, 1989, 9(4): 235-238
17. Bao H.F.: Quantitative and Computerized Medicine--New research QMSOC(I), The proceeding of the First Conference of the Frontier for Life Sciences on Central-south China, 1989, 211-216
18. Bao H.F.: The quantitative integration of biomedical information by computer--New research QMSOC(II), The Proceeding of the First Conference of the Frontier for Life Sciences in Central-south China, 1989, 216-222
19. Bao H.F.: The new functions of Quantitatively Medicine Simulating and Operating by Computer--New research QMSOC(III), 1990, J Tongji Med. Univ, 10(1): 52-56

20. Bao H.F., Geng J.H. and Su Z.F.: Pansystems Methodology(PM) and a new research on large-scale integration of biomedicine—An introduction of QMSOC and its recent progresses , Acta of Jiansu Industrial College, Journal of Jiangsu Institute of Technology, 1991, 4 (2): 69-75
21. 包含飞, 刘庚妹: 计算机编码编辑技术(HICE)及其初级信息开发功能——新研究 QMSOC(V), 计算机应用研究 (1991年专辑第2号), 1992, 4-9
22. Bao H.F., Ni X.W., Lou S.: Integratable Relationized Medical Electronic Book (IRMEB)--- An Exploration of An New Type of Intellegent Knowledge Medium Under the Influence of Pansystems Theory. Advances in Systems Science and Applications (Inauguration Issue) 1995, p304-309
23. Bao H.F.: HCSL: A Human-Computer Commonly Understandable and Communicatable Medical Language, 《Proceedings of The First China-Japan-Korea Joint Symposium on Medical Informatics(CJKMI'99)》, p177-181, 1999