

地理信息科学

M. F. 古德柴尔德

1. 引言 在过去的十年中,地理信息系统(GIS)已发展到了一个较为成熟的阶段。很多国家建立了相关的研究和培训项目,新的应用领域不断被发现,新产品不断涌现,遥感平台技术的发展使得新的重要数据的获得成为可能。可以说,GIS的研究前景,具有相当的吸引力。

然而,所有这些活动的目的是什么?象“空间数据处理”这样的名词可以描述我们做了些什么,但描述不了为什么我们这样做。这就是1984年苏黎世召开的第一届国际空间数据交换系统间数据交换标准,后来的新系统应当使用这些数据标准格式。

3. 制图综合问题作为一直存在的瓶颈,现在更为突出地表现了攻克这个难关的重要性。因为地图数据库不可能存储每一种比例尺的地图,这就需要解决一种比例尺数字地图产品如何综合以满足其它比例尺的分析与制图要求。这个问题已经得到了越来越多的重视,在1993年科隆举行的国际制图大会上成了中心议题之一,根据新技术专业委员会的三次讨论会的讨论结果,已将制图综合问题归纳为三类:①选取,②单个图形综合,③多个图形相关处理。前二个问题已基本解决,最后一个问题正在组织围攻之中,可望在几年内得到解决。

4. 空间分析精度应当有一个明确的回答。许多用户都希望GIS的每一步操作过程所产生的结果应有一个可信度。对于最终结果的误差积累过程和大小有一个正确的答案。例如,GIS执行快速图形叠置技术实现3—4层以上要素的叠合分析时,将能产生一种多要素索引,它能带来许多方便之处。但是对合成后的地图进行地面校核时,即选择一个确实具备各层次属性的地点进行实地对照,其结果往往令人失望,其根源主要来自于资料图的不精确和不协调。因为在通常情况下,土壤专家划分的土壤类型界线只是一种研究结果,而在实地并不存在这种确切的界线。同样的问题也可出现在地貌类型图上或土地类型图、土地利用图、岩性图等等。可想而知,采用这些不精确的地图叠合处理后只能偏离和误差,甚至使你无法利用这种叠合结果。目前,解决这类问题的一个有效方法是通过预先编制地理单元图并进行多要素描述和编码,而后让其进入计算机,从而使上述不协调问题克服于预先由人工编制地理单元图的过程中,但是这种方法并没有解决“不精确”的问题。而且需要有多种知识和经验的人才能胜任这类费时的工作,可见应进一步研究以找出更为有效的办法。

除此之外,还应研究GIS操作过程中的误差传播和积累过程,包括简单的逻辑运算中的误差传播,加权叠合处理对于输入要素和权重误差的敏感性,以及栅格化或矢量化过程中的几何误差产生的原因和大小等等。

5. 加速大区域地理数据库的查询速度,需要从多方面的努力才能得到有效的解决,首先是依赖于计算机性能的提高。目前内存达到10GB以上的计算机已问世,使得查询工作又可能在内存进行,其查询速度将有大幅度提高。另外,改进数据结构和查询方法依然成为今后研究的重要课题。

总之,GIS已经成为社会进步的一个标志,促进经济发展的先进工具,它必然会随着社会经济文化的发展和科学技术水平的不断提高而得到相应的发展和自身的不断完善。

理会议上 Tomlinson 主要发言中涉及到的问题之一。会议名称意味着空间数据的处理有一定的难度, 但是否永远如此? 它还暗示了某种程度上的数据无关性, 正如美国地质调查局在分发出去的磁带上都标注有“处理困难”这样的字句。这使人们联想起上一届国际地理联合会召开的会议名称: 地理数据的感知与处理会议。这次及以后的会议论文名称足以说明这些论文的作者所关心的不仅只是数据处理——从美国这一方面看: 这个团体不仅仅是 GIS 联合服务部。

地理信息系统有时被指责为由技术所推动, 仅是一种正在寻找应用领域的技术。GIS 25 年的发展历史似乎也恰好证明了这一点。例如, 如果把 Tomlinson 及加拿大地理信息系统的开发者们在 1965 年硬件那种原始初级的水平上做的工作认作为技术推动还比较困难的话, 那么在以后, 技术的迅猛发展使得 GIS 后面这种推动力自然也强大起来。386 微机上的菜单驱动, 全颜色、下拉式菜单栅格 GIS 功能使得过去几年中系统销售量大增。也许正是这种原因驱使人们去进行空间数据处理, 但这仍然不能解释会议和文献中所报告的多种多样的研究工作。

也有过由 GIS 发展应用来推动的阶段。CGIS 本身正是这样一个例子, 这种强大的推动力导致了第一台地图扫描仪样机的问世和其它技术的发展。McHarg 研究出了地图叠合技术原理, 而很长时间以后 Berry 及其它人才在 MAP 及其它同类产品中完成了自动操作技术; 学校公车路线软件也是问题提出在先, 标准 GIS 中的实现在后。尽管如此, GIS 研究中很多问题的产生还是和应用无关的。

随着 Landsat 和 SPOT 影象的广泛使用, 数字高程模型的提供、街道文件的建立, 很多国家的数据应用已经大大超出了纠正、编辑范围。例如, TIGER, 尽管其主要目的是为 1990 年美国人口普查服务, 但它已逐步发展了自己的更新、调整、应用开发软件。

尽管推动 GIS 发展的因素很多, 本文要描述的是另外一种推动力, 这就是“S”, 即“科学”这个词。我认为迫切需要认识和发展“科学”在 GIS 中的作用。这有两方面的含义。第一, 作为一个领域, GIS 包含哪些一般的科学性问题的, 这些问题范围有多大, 普遍性多强? 在多大程度上 GIS 研究是由 GIS 本身引发的一些问题所推动而深入的? 再有, 如果 GIS 可作为一门科学, 那么它可以划分成哪些子学科? 它们都研究些什么问题? 它们是怎样产生发展的? 第二个内容与 GIS 在科学中的工具箱作用有关——是为科学而产生的 GIS 而非 GIS 科学。我们需要做些什么来保证 GIS 和空间数据处理技术能取得一种合法地位。在此, 地理学是关键, 或者说是分析、解释和理解的起点。

要做到这一点, 首先必须清楚, 确切地说, 空间数据, 是地理数据, 独具特点, 它们的一些问题不可能归入到另一些领域中。我们还必须清楚, 所有地理数据都存在一些普遍问题; 至少应清楚, 区分地理数据和非地理数据是可能的。

2 空间数据的特点 在很多基础设施管理系统中, GIS 是提供数据的一种方法, 它提供一种基于地理位置的数据形式。实质上, 一个空间数据库具有双重作用, 使得记录可以按属性方式或地理位置方式使用。空间信息的关键特性是区分, 它使得可以定义标准查询语言中没有的操作。例如, 可以查询任意的, 用户自定义的一个多边形上的所有点的记录, 这是 SQL 等标准查询语言中没有的操作。从本质上说, 空间方式是多维的, 但反过来, 其它领域中也存在多维方式。

空间数据的特点在于它在二维空间的连续性。任一点的位置对应的真实地理空间坐标在理论上的定义都可能达到无限高的精度, 并返回一个变量值。例如, 设高程为 Z , 则地形可

用一个无限数量的元组 (X, Y, Z) 来描述。在用网格表示时, Z 只是定义为网格中的一个点的高程, 但集合中元组的个数仍是无限的, 只要变量在链和结点的一维结构中是连续的。时间维也是连续的, 其在测量或表示时也用连续的形式, 但在时间数据处理问题上感兴趣的人并不多。与此相比, 空间数据处理方面做的工作则很多。

我们的很多数据模型, 特别是多边形网络或不规则三角网 (TIN_s), 反映了一种公认连续空间概念, 以及用户在任意的精确的 (X, Y) 坐标处确定 Z 的需要。就是说, 对于空间数据来说, 存在着多种可能的概念性数据模型。针对一种特定的地理现象, 选择那一种数据模型, 这是空间数据处理中更基本的东西。

空间数据的另一显著特点是 Anselin (1989) 提到过的空间依赖性, 邻近区域互相影响并具有相似特性。如果不存在空间依赖性, 用一种有限的离散模型来近似表示连续变化的空间就会是不可能的。

最后, 地理数据是分布于地球表面上的, 这一点在许多GIS项目的有限区域研究中经常被忽略。怎样把地球表面描绘在一张平面的纸上, 这个问题困扰了我们几个世纪, 由此已发展了成熟的地图投影技术。正因为如此, 我们分析球面数据的方法很少, 对于表面上的过程建模知之甚少。再者, 我们在对待GIS的显示时也把它们当作是显示在实际的一些纸上, 并一直把地理数据当作已被投影到平面上的, 而未利用电子显示的能力来显示这个球体。如果我们想要在球体尺度上理解地理过程, 并有效地开展地球研究工作, 我们需要开发适当的技术来研究这个球体, 要利用真实的模型而不是传统的二维图形。我们必须恢复正射投影的作用。

3. 地理信息科学的内容 一旦意识到地理信息具有自身的特点和自身的问题, 我们可以看看构成一门地理信息科学的一系列基本问题。可把这些问题以一个线性流程来表示: 从数据收集到数据分析, 尽管中间有些专题可能会和这条简单主线交叉。这里先作一个声明: 因为这个领域的多样性和多变性, 仅靠一个人尝试一种根本性的突破是困难的, 甚至是不可能的。这篇文章中很多地方仅是我个人观点, 不可避免地存在着不完备之处。

研究工作经常被区分为纯理论性或应用性。它们的区别于是被人类最基本的求知欲推动, 还是被人类社会现实的正常需要推动。许多GIS的产生是应人们对信息管理与分析的需要, 从这个方面说, GIS研究应用性多于纯理论性。还有一种看法把纯理论研究视为还未找到应用领域研究; 纯理论研究是长时效投资而应用研究为短时效投资。从学术方面来看, 纯理论研究常常联系于高荣誉, 而应用研究联系于高经费。我考虑了纯理论和应用研究的所有方面, 觉得两者对GIS来说都是重要的。比如, 美国地理信息与分析中心的宗旨是“基础研究”, 而同时这个中心也很幸运, 获得了很多经费支持, 而其应用则是未来几年甚至几十年以后的事情了。

六十年代CGIS还处于设计阶段时, 有一点就很清楚, 要把这么多的地图都输入, 唯一可行的办法是通过某种形式的扫描设备。而那时, 还没有地图幅面大小的扫描仪, 这就产生了发明这样的扫描仪的需要。于是加拿大IBM造出了滚筒扫描仪的样机, 从现代标准来看, 会被认为是很大的花费。

经过将近三十年的发展, “怎样做”这个问题的研究产生了大量的算法、数据结构、空间索引方法以及其它技术手段。其中一些是GIS特有的, 但很多在一些相关学科中也被用到。例如, 确定 Thiessen 多边形算法的描述在大量的期刊中都可查到。同时, 当GIS中越来越

多的技术问题得到解决以后,人们越来越感觉到研究的重点已经由原始的算法和数据结构转移到了更复杂的数据库设计,以及围绕GIS技术在实际中的应用方面。下面论述一些重要问题。

3.1 数据收集和测量 如果真实空间具有连续性和空间依赖性,那么应该怎样来描述和测量它,更通俗一点说,人们是怎样来观察、了解地理空间的?尽管这些问题中很多是属于遥感、摄影测量、大地测量、认识心理学的研究范围的,它们的分界并不十分清晰。是GIS还是遥感研究把信息从一种技术传递到另外一种技术,并保持其合理性和有效性?用辅助地理信息来提高分类精度,用遥感影象来更新一个GIS数据层,这类问题是属于GIS还是遥感?总的来说,我们把问题归入哪个学科并无太大关系。但毫无疑问,这里存在着很多科学性问題,需要从深度来理解空间变化的本质。一个人认为是“遥感”很可能另外一个人认为是“地理信息科学”。

离散的,以及伴随其中的总结、抽象、近似,作用于数据收集、解译或编辑过程,此阶段所采用的方法影响其后数据的使用。正如在很多GIS应用中都会遇到一样,当应用目的变化以后,重新考查数据收集过程是很有必要也是很有益处的。如果最终目标是地下三维数字形式,传统的地质野外制图是否还仍最合适?如利用全球定位系统(GPS),地貌制图该怎样调整?地理数据收集这一类问题要回答好还需要很多年。

3.2 数据采集 过去十年,地理数据采集技术获得较大发展,现行的一些系统有较强的扫描输入能力。其问题主要是文件的低质量和地图设计时的模糊性。因此,手工数字化仍是广泛使用的方法,尽管它花费高、工作枯燥、效率低。在未来几年内有两种趋势将改变这种情形。一是输入及数据编辑过程中将越来越少用到地图文件。遥感和摄影测量使编辑越来越少地使用纸张地图,一些解译性更强的领域如土地利用、植被或土壤制图也将如此。数字总站将为数字平面板(Digital plane table),也许甚至于数字地质野外记录本所代替。另外一个则是长期以来就认识到的地图设计方面的一些小改进的可能性,这种改进使得扫描和解译大大地方便起来。

3.3 空间统计学 由于空间数据常常是现实世界的近似或概括,它们是十分不确定和不准确的。数据模型或比例尺的改变可能导致信息的损失,如数字化或扫描过程。在限定的机器上的处理同样也会引入它自身的不确定形式,尽管比较起数据本身的误差来,这些是微不足道的。许多人文地理概念更是不确定的,包括空间物体(“印度洋”、“欧洲”)以及它们之间的联系(“包含于”、“穿过”)。如果我们用模糊数学或统计学术语通过概率微分来研究这种不确定性以及它的测量和建模,它通过空间数据处理过程的传送,那么这些研究无疑都是地理信息科学的一部分。应该使用怎样的模型来代表地理变化才算准确?在一数字模型中怎样表示不确定性或不准确性?这种不确定性该怎样从数据库传送到GIS产品?

地理数据给空间统计学提出了独特的问题。如果说处理医学影像的目的是从“内脏”图象上确定某目标的真实位置,那么在地理影象上甚至经常找不到真实的明确的概念,因为目标常常是解译或概括的产物。我们需要测量和描述不确定性的更好的方法,特别是研究GIS中常见的复杂的空间物体时。世界是一个错综复杂的连续体,我们不应硬性把它变成一个生硬联系在一起的对象。要解决这些问题,主要还得依靠空间统计学,但是地理信息专家们必须提供动机和事例,并且定义整体目标和限制条件。

尽管所有的地理数据在一定程度上都是不确定的,现行的GIS系统都是沿循制图学的惯

例，地理物体的位置和属性也已完整清楚地知道，数据质量的单独说明已无关紧要。GIS产品不确定性的影响从来没有作过总结评价。最近作了一些数据质量方面的研究。其中一个为用精度来表示。在一个地区范围内，这意味着在数据表示和处理中使用有限精度，绝大多数通过使用网格大小由数据精度确定的栅格表示。另有用很多形式的四叉树来表示以达到已知精度级的方法。最近还有GIS方面几篇关于有限分辨率处理的文章，有限分辨率几何是数学上一个很活跃的研究领域。

另一个很有用的方法是采用了地理统计学的一些技术，这些技术的统计学基础使得不确定性清晰起来。现在我们已有几种有用的数字误差模型。以及用其结果来量算面积等的模型。最后，还有几个把地理数据作为随机区域或随机区域的较为成功的模型，以及用这种方法来模拟GIS中地物不确定性的模型。在所有这些方法中，我们可以得到一系列精度模型，用它们可以建立起一个误差跟踪的GIS。但是空间统计学并非一个简单领域，在概念复杂性方面，用到的很多技术远远超出了初级统计学。

3.4 数据模型及空间数据理论 数据模型是我们用来表示数据库中地理过程的逻辑框架。因为每一模型都是一近似，模型选择不仅要考虑已有的功能，还要考虑产品的精度。过去十年GIS取得的进步中，最引人注目的是数据模型的激增，以及关于它们的评价文献大量涌现。有关栅格数据与矢量数据的争论是最早期期的事了，但是现在这种争论中又加入了关于对象、层面、面向对象的概念及复合物体的层次模型、时间和三维空间的固有可能性。除此以外，我们对于地理数据模型，甚至静态的二维情形都未形成一个完整清晰的框架，而没有这个框架，要弄清楚GIS是怎样摆脱特定系统实现限制这个问题是困难的。多系统采用的多层栅格模型或者一点一线一面模型的使用，究竟使当前应用受到了多大限制？这即是一个纯理论又是一个应用性问题。一方面，我们必须结合相关术语，开发出一个地理数据模型的综合框架提供一个标准基础和一个理想模式，以此来衡量特定系统。另一方面，应该清楚对实际工作不起作用的抽象框架是没有意义的。但这种影响将会是怎样的呢？是否可能扩大或“重整”已有产品中的数据模型，而不致于导致不一致？

这些观点引起了激烈的讨论，我们研究、分析、表示建模时是把整个世界当作一个离散的还是连续的，是一些物体的集合还是一系列区域？我们考虑的变量在空间各处都有确定的值，或只是一个杂乱堆放着的一些物体的空的空间？从本质上说，这些观点使GIS的争论从内部数据结构这种比较模糊的概念转向了更普遍的我们对地理过程的理解上。日常人们所见的是一个物质的世界，但是关于自然过程的科学关注的更多是连续的变化。

3.5 数据结构、算法和程序 这个领域在过去25年内的很多研究成果属于内部数据表示，以及操作这些数据的算法设计方面。四叉树，叠加的分带扫描算法（Band Sweep Algorithm），计算复杂性分析以及弧——结点数据结构，这些当年的重大突破现在仍在使用。但仍遗留了许多挑战性的问题，例如，制图设计中的一些问题，各种地形数据模型之间转换的更好实现方法。通过数据库管理系统来处理数据的很多系统，以及数据结构方法越来越偏向计算机科学领域。尽管如此，我们似乎只是达到了这样一个阶段，所有的简单的，更基本一些的问题都解决了，而全部难题都留了下来。例如，地形数据模型转换的进一步发展（例如，从等高线到TIN）要求对地形本质的深入了解，且不同的地形可能需要不同的转换模型（例如何流地貌区别于冰川地貌）。对于未来十年将获得的大量数据将怎样有效地存储和处理，这一研究也将进一步深入。

3.6 显示 地理信息系统因为没有对制图设计原理给予足够的注意而常常受到批评,换句话说,它把地图视为信息的简单存贮,而忽视了其作为交流信息工具的功能。如果我们把数据库视为真实的,那么一张地图就不只是存贮体,在数据库和地图之间常常存在着一种直接明显的联系。尽管如此,如果数据库只是被看成地理事实的近似,那么,输出显示的设计就十分关键,因为它可以影响用户的观察角度。一些简单的手段如背景颜色选择,或邻近多边形的对比都有很大的影响。

电子显示的能力大大超过了传统制图学。我们需要研究动画显示、三维显示、用户界面中的图标和象征物,颜色和色调的变化序列,放大、漫游、多媒体,允许同时使用多种具有空间和时间性数据的窗口。我们需要使用电子手段来达到分区地图设计的进一步提高。这些都属于地理信息科学的一些最基本的问题。

3.7 分析工具 GIS是支持一系列空间分析技术的工具,包括产生空间地物的新分类、地物属性及位置分析,以及利用地物多类别和它们之间关系的建模过程。从一些基本的几何操作如计算多边形的重心,线的缓冲区,到更复杂一些的操作如在网络中找到最短路径。一些主要产品的功能一直不断增加。

尽管广泛认为分析功能是GIS的中心,GIS和空间分析仍未有机地结合起来,很多系统的分析功能仍较简单。统计软件包SAS在其早期随着技术和试验的完成发展很快,但GIS并未如此,其在把已知空间分析技术应用到现有产品中的进展很慢。

这种情况有很多原因。最明显的一个原因是GIS市场更强调其信息管理功能。GIS市场虽然繁荣,但其需求还处于一种相对简单的阶段,只需要简单的查询和制表。另一原因是空间分析的相对模糊性,这一种在很多不同学科上发展起来的技术,缺乏清晰的准则和概念、理论框架。尽管有人希望GIS给空间分析提供一个标准系统,却局限于现今较低的地理数据模型理解水平。Tomlin (1990)已经做了一些努力在GIS多种多样的功能中加入一些框架。我们急需一空间分析的分类学,也许可从很多种数据模型中总结出来。

在这个阶段,GIS和空间分析的结合进展很慢,至少有三种不同的方式。一些分析功能直接加入到GIS里,例如网络分析中几种模块新扩展的功能。一些进展是在松散连结的分析中,这里有一个独立分析模块依赖于GIS输入数据,显示等。尽管如此,紧密连结仍缺少一有效形式,在紧密连结中,数据可以在GIS和空间分析模块间传递,而不损失较高结构关系,例如拓扑关系、目标特征属性以及其它联系。现在这些仍不可能达到,在很大程度上是因为缺少数据模型标准。这样,连结的水平较低,而高结构的建立缺乏统一基础。

GIS和空间分析之间的结合还可能以语言的形式出现。这种语言一开始在很多GIS现行版本中就以宏语言的形式存在,并且试图扩展SQL到空间操作。尽管如此,这些都是针对并紧密依赖于有限数据模型的,而此时它们之间几乎无任何相似之处。在加州大学圣巴巴拉分校,我们尝试从现今GIS使用的多种语言中,找出了一种共同的语言来,但是要达到较满意的结果,必须从综合数据模型的概念性框架入手。

结合GIS和空间分析的另一个问题是,前者的空间是离散的,然而在很多空间分析中这一点常常是不清楚或未指明的。很多形式的空间分析是在连续空间中的,且未能解决好离散化过程带来的不可避免的不确定性。例如,在GIS中,独立于离散形式的坡度无法测量,同样,一个面状地物的边界依赖于它的数字表示方式。尽管如此,坡度和长度在空间模型中常常是一种不合格的量数。这样说来,在GIS和空间分析的结合问题上,GIS和空间分析两方

面的问题都必须解决。

绝大多数现行GIS版本提供一些宏指令或手写体设施,允许用户从复杂操作系列中定义产品,只需用一条指令来激活它们。尽管这些经常包括构造习惯性环境和界面的能力,它们却未提供空间分析所需要的工具。仅有的一个例外是 Prime/Wild 的 ATB,在 System/9 上建立了一套工具,可使用户做较复杂的分析、显示和管理中间结果。这样的工具会越来越需要,以使GIS进入到一个更复杂的分析和决策支撑阶段。我们需要研究跟踪数据谱系和误差传递的方法,研究恢复中间结果的方法以防止用户错误或无意义地进行信息复合。我们还必须研究把这种分析和GIS数据获取和规划过程结合的方法。

把重点放在多阶段复合分析以及来自多层数据库的产品上似乎和基于知识的系统、空间推理和空间查询的研究大相径庭, GIS 领域的一个诱人之处在于其广泛的应用,及相应的用户界面设计的多种多样环境。在数据建模中,最重要的问题不是扩展关系模型或面向对象的模型之间,那一个更适于地理数据,而是哪一种模型方法更适合于表示特定的一类地理数据。同样,用户界面设计中一个很重要的问题是决定每一种GIS应用各自的最优环境。最适合于交通导航系统的环境可能会和最适合森林资源管理系统的环境完全不同。

3.8 机构的、管理的、伦理的问题 关于完善和管理GIS,特别是在大型机构中的一些问题的研究才刚刚开始。这是一研究难点。尽管如此,美国一些成功的大项目,联邦机构关于几个大型数据获取的讨论,给研究提供了很好的例子。需要深入了解GIS技术被采纳的过程以及它在机构中的影响;了解地理信息的价值和GIS的益处,了解使用GIS进行决策的过程。一些有关的社会科学方面的书籍已提到过这些问题的理论框架,要解决GIS特有的问题,我们需要更好地利用已有的这些材料。

尽管在接受任一新技术的过程中都会遇到一些问题, GIS 在地方政府部门、公共设施及资源管理机构中已被广泛运用。事实上, GIS 的引入为社会上地理信息的管理起了重要的作用。同时,人们日益关注GIS对个人信息的监视和侵犯方面的能力。研究机关有责任监察和研究GIS现象中这些更实在的方面,包括它对社会的整体影响。GIS对社会力量的平衡将起什么样的作用?它们是否只是为当权者提供工具?或者它们可起到平衡权力分布的作用?这些都是GIS伦理学中的问题。

4. 普及性研究 前面的部分讲述了作为一门地理信息科学可能包括的一些领域。很清楚,在每一领域内都有挑战性的科学性问题的提出并被研究。提到的内容可能是不完整的。在每一领域中空间联系都显得很显著,虽然显著程度各领域中不同。例如,对于决策理论,其空间联系不是很显著,但在数据模型中空间联系便是很明显的。

在 NGGIA 研究计划中,我们坚持这些问题如得不到解决将会给GIS技术的应用带来障碍。GIS研究中的其它讨论虽然重点不同,但结论也是同样的,其中许多都是旧观点,在GIS出现以前地图学、大地测量学和地理学中就意识到过。一些可能是GIS独有的。例如, GIS 的被采纳方式和其它技术截然不同,这一点没有被很快意识到。GIS 益处的衡量是一独特问题,还只是信息技术益处衡量问题中的一般例子?

同时,区分GIS与别的学科不同的特点是很重要的。在GIS初期,可以说这种技术填补了一空白,使原来存在的一些任务得到完成。通过叠加层面 (Tomlin 1990) 用GIS来进行合理的分析,反映了被McHarg所推广的这一技术,尽管其中明显增加了一些新的有趣的功能。CGIS说明计算机比起手工叠加区域测量来花费—效率比更好。而且GIS可以完成连一些数

据收集者都未设想过的事情。GIS在地理数据收集、处理和分析方面起了革命性的变革。

下面是一些GIS的特点：怎样建造时间性地理数据模型，怎样获取、处理三维地理数据；怎样使用数据；和更传统的观念相比，怎样评价地理学在信息和过程方面的前景？

GIS中有很多重要问题，要取得重大进展还有很多工作要做。因为这些问题是在GIS中产生的，所以它们在别的学科中并不是主要关注点。尽管如此，GIS可从交叉学科中得到很多有益的启发。统计学对解决GIS中的误差问题很有价值，认知心理学对GIS中用户界面的设计很有帮助。

这种论述最后自然给GIS研究提出了一个初步定义，研究围绕GIS技术应用的普遍问题，研究仍阻碍GIS完善的普遍问题，研究开发其潜在功能的普遍问题。这是“关于GIS的研究”还是“利用GIS的研究”？从某种意义上来说二者兼有，因为它们是GIS技术中的基本问题，也是在这种技术被成功应用以前必须解决的问题。如果使用GIS进行研究的问题是普遍的。那么可以把它当作GIS研究的一部分来对待。尽管如此，在一具体领域中特定的GIS应用问题需要结合特定领域来考虑，并且利用其专门知识。精确性问题提供了一个有用实例。很多类型的地理数据都存在精确性问题，需要用普遍的模糊模型来解决，与测量误差理论中的高斯分布相似。尽管如此，利用GIS来做一犯罪数据的分析也会带来此应用特有的精度问题，这需要对犯罪学及犯罪数据的收集的过程有一个深入的理解。

尽管如此，只有科学性问题的存在还远远不能为使其成为一门科学提供有力的基础。人们对此普遍存在兴趣吗？这些子领域是否可以找到足够的相互作用的基础来使它们维持作为一门科学需要的装备，如期刊、社团、书籍和学者？这些子领域的研究者是否表现得象一群学者？在系统与地理信息科学的关系是否可类比于统计软件包和统计学？统计学是一门高度规模化的学科，但是也可找到一些更趋向于技术性的领域如考察性数据分析、统计数据显示以及应用统计学。自然，纯科学与工具之间的关系经常有争议，但是这并不妨碍这两者的成功。关于统计软件在统计学教学中价值的存在争论，相同的问题在GIS中也存在。

首先从原理上，然后从实践上简短回顾地理信息科学中那些关于共同兴趣的论述。这个领域较小——除了产业方面的发展外，没人承认GIS这个领域是一个主要的学科。它是独特的，有其合理、特殊的问题。很自然地它是富有挑战性的，自身充满吸引力。从另一方面说，它是多学科的，有很多分化出来的竞争对象。它缺少一个核心原则，不象统计学那样，在过去几十年中学院系科的数目和规模都在稳步上升。也不象地理学那样，传统上就是一个非技术领域，特别在社会地理学的一些领域内，技术方法十分不受欢迎。

在实际中，频繁召开的GIS会议是其受到普遍重视的一个证明，且我们开始有了一些书籍和期刊。尽管如此，GIS会议上表现出的科学性还不具规模。参加GIS会议的人需要不停的创新，不论是在科学性研究还是在推销产品方面，否则这些活动就会因为来源不足进行不下去了。

5. 未来的选择 回顾GIS研究将近三十年的历史，可以清楚看到，最大的进步是在那些定义最好的、最容易的问题上取得的，这些问题的解决很大程度上依赖于技术本身。70年代和80年代，进展最快的是算法和数据结构，但很多难题，如数据模型、误差模型、空间分析和机构的管理的问题的结合等仍遗留下来。其中的一些可能难以解决：例如，政府机构采纳GIS的过程无简单规律可循，尽管提出这个研究问题是容易的。

另外一些问题在纯理论意义上已经解决，但其实现却仍是应用研究中的一大难题。例如，

大量技术已被定义,但仍无法把它们实际应用起来。为进行重要领域的研究的学术研究环境已经建立起来了,但在实现手段提供方面仍较弱。因此,我们需要一个与研究团体紧密联系的软件工业。更重要的是,我们需要一个教育系统,对新的研究作出及时反映,并且能把新概念很快地用在其教育项目中。很遗憾,高教部门经常是保守的,要过很多年才会把新思想纳入教学计划中去。

GIS研究很象地理数据——观察的距离越近,出现的问题越有趣。GIS才刚刚开始研究出现的重要问题。我们处在一个可大显身手的地位。我希望在这篇文章中能讲清这一点:用GIS技术来处理空间数据比“空间数据处理”这个词表现得更富于科学性和挑战性——实际上,这是一门地理信息科学。“地理”这个词很必要——很多GIS研究是关于地理世界以及人类与其联系的,另外,这个词的含义也比“空间”更丰富。“S”这个字的改变——从“系统”到“科学”——显得很正确,一系列空间数据处理会议的成功、《地理信息系统国际期刊》的征订,期刊上收到的文章如“地理分析”、“计算机与地球科学”、“计算机视觉”、“图形和图象处理”以及区域科学协会和IEEE的出版物等等都可证明这一点。

我希望我还说明:一强大的科学项目不仅为其自身服务,而且还可满足产业和GIS用户的需要。如果要使其脱离纯粹的商业现象,在技术领域并非只是昙花一现,GIS需要补充很多科学的成分。现在的GIS太易于被看成为寻找应用领域的一种软硬件技术,或被视为由其主流流行的几种产品的有限功能定义之物。我们需要从“系统”发展到“科学”,把GIS建成为人们普遍感兴趣的交叉学科,其本身为一工具箱支撑,反过来以其基础研究支持技术发展。有时,GIS和科学的关系,很象Fortran和代数的关系。

近年来GIS在两种传统之间的分歧越来越大,一是空间信息传统,另一是空间分析传统。空间信息传统强调大型数据库,供给地理学使用。空间分析传统强调多功能以及大量数据模型。两种传统共享数据结构和算法,依赖于同一种数据来源和硬件。尽管如此,这不足以说明这一科学领域的存在。要说明这一点,我们需要看得更广,包括数据模型精度,认知,推理,人机界面(HCI)和视觉化,并指出这些是怎样成为两种传统的统一部分的。

如果没有这些证据,GIS领域将会分化,直至消亡。推销商们将会在数据输入工作站,空间分析工作站或设施管理方面专门化,而无交互或统一的可能。这将是可悲的。

怎样保证地理信息系统及其科学的长远未来?学科领域就好比一个部落,有其自身的图腾、族规、语言和社会关系。现在的GIS部落联系非常紧密:资金来源充足,问题富于挑战性,很多有用的研究正在开展。尽管如此,从长远来说,作为一门学科,它表现得并不太好,学术界至今还在怀疑是否需要严肃认真地对待它。科学是严谨的,其执行者负有重大责任。我们的工作过于匆忙,而技术又发展太快,以至于很多文献都是会议记录,虽然出版快,但读者有限,并缺乏足够的质量控制。很少有人有时间来编写教材,识别其真正的学术核心,或出版较好的研究实例。

我相信,把GIS作为一门科学而不是一个系统,并且研究清楚此领域的主要问题及其领域范围,有利于其发展。地理信息系统是地理信息科学的一个工具,学科的发展最终又会促进工具的完善。我们应该告诉学术界,可以直接地告诉它,也可通过介绍此领域思想、方法、基础的文章、教材,并在有影响力的期刊上发表关于GIS的文章。用户,销售商和研究者——三种人都有各自重要的工作要做。

章燕燕译自《IJGIS》1992,第6卷,第1期