流形学习方法中的若干问题分析

高小方
(山西大学计算智能与中文处理教育部实验室 太原 030006)

摘 要 流形学习是近年来机器学习与认知科学中的一个新的研究热点,其本质在于根据有限的离散样本学习和发现嵌入在高维空间中的低维光滑流形,从而揭示隐藏在高维数据中的内在低维结构,以实现非线性降维或者可视化。介绍了几种主要的流形学习算法,分析了它们的优势与不足,总结了流形学习方法中需要解决的若干问题及其研究现状,并展望了流形学习未来的研究前景。

关键词 流形学习,维数约简,等距映射算法,局部线性嵌入算法

Problems and Analysis in Manifold Learning

GAO Xiao-fang
(Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education, Shanxi University, Taiyuan 030006, China)

Abstract Manifold learning is a newer research direction of machine learning and congnitive science in recent years, its essence is to find out the low dimensional manifold hidden in high dimensional space though learning discrete samples, and get the hidden dimensional structure of the high dimensional data to realize non-linear dimension reduction. The paper introduced some manifold learning algorithms, summarized some problems of manifold learning and its research status, and discussed the prospect of manifold learning.

Keywords Manifold learning, Dimensionality reduction, ISOMAP, LLE

随着信息时代的到来,对高维数据的处理成为迫切需要解决的问题。如何对高维数据降维,以解决"维数灾难",成为人工智能和机器学习领域的研究热点。美国著名杂志《Science》在2000年连续发表了3篇论文[1-3],从认知的角度提出并讨论了流形学习的研究课题。不仅首次使用了术语"manifold learning(流形学习)",而且给出了可行的算法,为维数约简提出了一个新的研究方向,从而引起了"流形学习"的研究热潮。

在这3篇论文中,Seung等人认为人类很大程度上是通过低维流形来认知各种事物的,而高维信息一般情况下就嵌入在一个低维流形中[1]。Tenenbaum与Silva研究了图模型意义下的观测空间数据集与内在低维空间测地线距离之间的关系,提出了等距映射算法(ISOMAP)[2]。Roweis与Saul也提出了局部线性嵌入算法(LLE)[3]。

本文首先介绍了流形与流形学习的概念和数学描述,分析了几种主要的流形学习算法,总结了流形学习研究中的若干问题以及研究现状,最后展望了流形学习未来的研究前景。

1 流形与流形学习

流形是微分几何中的一个基本概念,最早由Riemann在1854年提出,其定义为:设M是一个Hausdorff拓扑空间,若M的每一点p都有一个开邻域$U\subset M$,使得U和n维欧氏空间R^n中的一个开子集同胚,则称M是一个n维拓扑流形,简称为n维流形。

从流形的定义可以看出,所谓流形就是一类特殊的连通Hausdorff仿射拓扑空间,其本质就是局部可坐标化的拓扑空间,可以看作是欧氏空间的非线性推广。"局部坐标"可以将问题分解为局部问题进行计算,而拓扑空间又能保证将局部计算结果合理、光滑地拼接起来,揭示问题的整体结构。流形研究的主要目标是经过坐标卡变换而保持不变的性质。

流形学习就是要根据有限的离散样本学习和发现嵌入在高维空间中的低维光滑流形,揭示隐藏在高维数据集中的内在低维结构,重构并进行非线性降维或者可视化。其定义为:设$Y\subset R^d$是一个低维流形,$f:Y\rightarrow R^D$是一个光滑嵌入,其中$d<<D$。数据集$\{y_i\}$是随机生成的,且经过f映射为观察空间的数据$\{x_i=f(y_i)\}$,流形学习就是在给定观察样本集$\{x_i\}$的条件下重构f和$\{y_i\}$。

有了对流形的定义,就可以形式化地给出流形学习问题的数学描述:

给定高维观测数据集$\{X=\{x_1,x_2,\cdots,x_N\}$,其中$x_i\in R^D$为独立同分布的随机样本,散布在光滑的$d$维流形$M\subset R^d$上,即$M$为嵌入在$D$维欧氏空间中$d$维流形,定义嵌入映射上$f:M\subset R^d\rightarrow R^D$,这里$d<<D$。流形学习是在没有任何关于$M$和$d$的先验知识的条件下,根据有限的观测数据集$X$

到稿日期:2008-05-05 本文得到国家863计划项目(2007AA01Z165),国家自然科学基金(70471003,60773133),高等学校博士学科点专项科研基金(20050108604),教育部科学技术研究重点项目(206017),山西省重点实验室开放基金(200603023)资助。

高小方 博士研究生,研究方向为机器学习,E-mail: gxfhtp@mail.sxu.cn。

发现未知嵌入映射 $f(\cdot)$ 且找到与高维观测数据一一对应的低维嵌入 $Y=\{y_1,y_2,\cdots,y_N\}$，其中 $y_i\in R^d$。

可以看出，流形学习的本质是当样本空间为一个高维光滑流形时，要从样本数据中学习出这个高维流形的内在几何结构或者内在规律，得到对应的低维数据集，实际上也就是非线性降维。这种维数约简的方法比传统的线性降维方法(诸如 PCA，MDS)更能体现事物的本质，更利于对数据的理解和进一步处理。

2 流形学习的主要算法及其分析

流形学习方法大体可以分为两类：一类是基于全局的方法，即计算每一个数据点与所有其他数据点的关系，建立全连接图，例如等距映射(ISOMAP)。另一类是基于局部的方法，即考虑每一个数据点与它邻域内的点的关系，通常是 k 近邻或 ε 邻域方法定义图中的边，例如局部线形嵌入算法(LLE)和拉普拉斯特征映射(LE)[5]、海赛局部线性嵌入算法(HLLE)[6]、局部切空间排列算法(LTSA)[7]等。

2.1 等距映射(ISOMAP)

ISOMAP 算法是 Tenenbaum 与 Silva 于 2000 年在《Science》上提出的[2]。其基本思想是当数据集的分布具有低维嵌入流形结构时，可以通过保距映射获得样本数据集在低维空间的表示[2]。该算法建立在多维尺度变换(MDS)的基础上，先计算邻域图中的最短路径，得到近似的测地线距离，代替不能表示内在流形结构的欧氏距离，然后输入到 MDS 中处理，进而得到嵌入在高维空间的低维坐标。这种算法是一种全局的降维方法，力求保持数据点的内在几何性质(即测地距离)。它同 MDS 的最大区别在于，MDS 构造的距离矩阵反映的是样本点之间的欧氏距离，而 ISOMAP 构造的距离矩阵反映的是样本点之间的测地距离。

ISOMAP 的优点在于：(1)使用 ISOMAP 维数约简，不仅将流形上邻近的点映射到低维空间中的邻近点，同时保证将流形上距离远的点映射到低维空间中远距离的点；(2)能够更忠实地表达数据的全局结构，易于从理论角度理解度量的保持；(3)能够很好地处理具有单一流形结构的数据，在降维过程中可以产生“elbow”现象，并由此判断出流形的本征维数[2]。

ISOMAP 的缺点在于：(1)ISOMAP 没有定义样本空间到嵌入空间的映射，对于一个未知点不能直接投影到嵌入空间；(2)ISOMAP 要求流形所对应的低维空间的子集是凸的，仅适用于内部平坦的流形，而不适用于有较大内在曲率的流形；(3)在噪声干扰下，ISOMAP 用于可视化会有不稳定现象，取较大的邻域会产生“短路”现象，而选取较小的邻域，虽然能够保证整体结构的稳定，但低维投影结果会产生大量“空洞”，使得最短路径算法重构的图不连通[4]；(4)ISOMAP 的本征维数通常要经过多次实验绘制残差曲线才能得到，而计算残差需要运行整个 ISOMAP 算法，这使得整个算法不仅极其耗时而且不能保证结果的有效性。

2.2 局部线性嵌入算法(LLE)

LLE 算法是 Roweis 和 Saul 于 2000 年在《Science》上提出的[3]。该算法侧重于保证局部邻域结构。也就是说，该算法仅仅保证将流形上邻近的点映射到低维空间中的邻近点，并不保证将流形上距离远的点映射到低维空间中远距离的点。在算法中，LLE 通过用重构权 w_{ji} 表示样本点 x_i 对它的邻域点 x_j 的重构的贡献，反映出每个样本点 x_i 同它的邻域点之间的局部几何性质：在样本点和它的邻域点做平移、旋转和缩放时，重构权保持不变，由此将高维输入数据映射到统一的全局低维坐标系，同时保留了邻接点之间的关系，保留了固有的几何结构。

与 ISOMAP 算法相比，LLE 是一种局部方法，通过局部线性拟合来获得内在的全局线形结构，其优点在于：(1)只考虑近距离点，因此无需要求流形对应的低维空间的子集为凸，因此应用更广；(2)计算对象是多项式数量级的稀疏矩阵，计算复杂度远远小于全局方法。但是由于 LLE 的低维嵌入所保持的并不是一个距离关系，对于等距流形，LLE 不一定能很好地恢复出同它等距的低维嵌入；在计算重构权的时候，最小二乘问题面临着参数 γ 的选择问题，对于不同的 γ 会求出不同的重构权，从而影响最终的嵌入结果。另外，对于有噪音、样本稀疏或者相互关联较弱的数据集，在维数约简时很可能会将远距离点映射到邻近点的位置。

2.3 拉普拉斯特征映射算法(LE)

拉普拉斯特征映射(LE)是 2003 年 Belkin 等基于局部保序的思想提出的[5]，有着很直观的降维目标，即高维空间中近距离点映射到低维空间之后距离应该很近。该算法利用 Laplace-Beltrami 算子(定义为流形切空间上梯度向量的负散度函数)的特性，通过计算该算子的特征函数来实现流形的最优嵌入。

LE 在降维时将邻近点映射到低维空间上仍然是邻近的，因此能够很好地处理分类问题；但是算法的参数 σ 对嵌入结果有着重要的影响，如何选择最佳的或者最合适的 σ 是 LE 的一个难题。另外，作为一种保持局部特征的算法，LE 对于噪音比较敏感。

2.4 海赛局部线性嵌入算法(HLLE)

海赛局部线性嵌入算法(HLLE)是 Donoho 和 Grimes 于 2003 年提出的[6]。该算法的出发点就是针对局部等距于低维欧氏空间中开连通的子集的流形。对于这种流形，HLLE 不要求其开子集为凸，而且能够恢复出与流形等距的低维嵌入。但是 HLLE 在获取切空间坐标时就需要知道流形的本征维数。当流形噪音分布不一致或流形局部低维特征不明显时，获取的切空间坐标可能偏差较大，从而影响嵌入结果；由于 HLLE 选取零空间中一组在某个领域内保持正交性的基作为嵌入结果，对于不同领域，嵌入结果有可能不同；另外，HLLE 对样本个数和近邻参数敏感，选取不当会导致算法失败。

2.5 局部切空间排列算法(LTSA)

局部切空间排列算法(LTSA)是浙江大学张振跃等人于 2004 年提出的[7]。其基本思想是利用样本点邻域的切空间来表示局部的几何性质，然后将这些局部切空间排列起来构造流形的全局坐标。

LTSA 的计算步骤简单，同 LLE 和 LE 相比能够很好地恢复出等距流形的低维嵌入，而又无需进行海赛矩阵的估计和构造二次项。但是它对样本点的密度和曲率的变化的影响比较敏感，在每个局部邻域计算投影坐标时，会较大程度地受到流形的曲率和样本点密度的影响，可能导致嵌入结果出现偏差，影响整体结果。

3 流形学习方法中若干重要问题分析

通过对算法的描述和分析可以看出，上述算法具有一些共同的特性：首先，它们都要构造流形上样本点的局部邻域，确定邻域大小(k 或 ϵ)，然后用这些邻域将样本点映射到一个低维空间；其次，它们都需要确定本征维数 d；再次，它们都是一些非线性的方法，都可以将问题简单地转化为特征值的求解问题，不需要用迭代算法，避免了局部极值问题。而它们之间的不同之处也只是在于局部邻域结构不同以及利用这些邻域构造的低维嵌入方式不同。

虽然这些流形学习算法能够更好地挖掘隐藏在高维数据中的流形分布，但是还存在很多问题，诸如 ISOMAP 的领域选择不当会产生“短路”或者“空洞”现象、LLE 不能很好地处理等距流形等，因此还需要不断完善流形学习的理论并改进其算法。在算法改进中遇到的一些具体问题同样是研究者们关注的热点，诸如邻域的选取、本征维数的确定、噪音的处理、稀疏样本的处理、算法的评价、算法的泛化问题等等。这些问题在“流形学习若干问题研究”[8]、“非线性流形学习方法的分析和应用”[9]、“流形学习中非线性维数约简方法概述”[10]等文献中已做了较好的分析。

3.1 理论完善与算法改进

近几年来，研究人员针对流形学习理论和算法的不足做了大量的工作，取得了一些有意义的研究结果。张长水等人在 LLE 基础上提出了一种从低维嵌入空间向高维空间映射的方法[11]，并在多姿态人脸图像的重构实验中得到有效的验证；Saul 等人在 LLE 方法的基础上，结合流形的微分几何特征，发展出基于半正定规划的流形方法[12]，适用于不同类型的流形；谭璐等人结合 LLE 和 LE，提出了局部不变投影方法[13]，在维数约简时也能够保持数据集的几何结构和拓扑结构不变。

3.2 邻域的选择

邻域问题是各种流形算法首先要考虑的问题。很显然，邻域越小可以认为邻域的线性结构越明显，但是邻域之间需要有足够的交叠以保证较远的点之间有足够的联系，这又使得邻域不能过小。而且在流形上曲率大的样本点的邻域应该小一些，而流形上曲率小的样本点处的邻域可以大一些，因此流形上的曲率以及样本点密度的变化也会使邻域结构产生偏差。

Kouropteva 等人[14]在 2002 年提出为 LLE 自动选取最优邻域因子的算法，Samko 等人[15]借鉴 Kouropteva 的思想也提出为 ISOMAP 自动选取最优领域因子的算法，但是这种算法得到的 k 值结果并不理想；王靖等人[16,17]提出了一种邻域的收缩与扩张算法，用于在每个样本点上自适应地选择邻域参数 k。这种方法在可视化意义下取得了较好的效果，但是增加了计算复杂性；邵超等人[18-20]通过二阶最小生成树等方法去解决 ISOMAP 算法中的“短路”/“空洞”问题，取得了较好的效果。但是通过在小邻域上的学习得到一个全局的坐标是非常困难的，如何将全局与局部数据的学习结合起来是一个非常有意义的研究课题[21]。

3.3 本征维数的确定

本征维数的确定是流形学习算法中的一个研究难点，赵连伟等人[22]区分了嵌入空间维数、高维数据的本征维数与流形维数这些容易混淆的概念，完善了流形学习的研究理论，并且证明如果高维数据空间存在环状流形，流形维数则要小于嵌入空间维数，还给出一种有效的环状流形发现算法，以得到正确的低维参数空间。

本征维数的设定值对维数约简的结果有很大影响。设定值过大，降维结果中的噪音就比较多；设置值过小，又会使本来不同的点降维后在低维空间重叠。目前，对本征维数的研究已经存在一些算法，诸如分形维[23]、测地线最小生成树[24]、Packing Number[25]、最大似然估计[26]等。

3.4 噪声处理

当观测数据是对一个光滑流形较好的采样时，使用非线性降维可以找出其内在本质的流形分布。在实际高维采样数据中，由于各种因素经常存在噪声，使得映射到低维空间后会出现对原始数据结构的扭曲和变形。非线性降维方法对噪声敏感是一个非常普遍的现象，因此解决噪声问题对非线性降维的结果是非常重要的。

詹德川、周志华针对 ISOMAP 算法，通过引入集成学习技术，扩大可以产生有效可视化结果的输入参数范围，以此降低该算法对噪音的敏感[27]。张振跃等人[28]也提出一种局部线形平滑的思想来解决噪声问题，先采用加权 PCA 来构建局部小块，然后用迭代方法优化权值。这种方法可以用作其它非线性降维方法的预处理，平滑噪声，得到对原始数据较好的重构。但是这种改进方法也同样存在一些问题，如迭代算法容易陷入局部极小值、鲁棒性不够强等问题。

3.5 稀疏样本的处理

流形学习算法在样本稠密的情况会取得很好的结果，但是当样本稀疏时就不能得到理想的结果，甚至还会失败。曹顺茂等人对样本稀疏问题进行了研究[29-32]，不仅改进了 LLE，使得其在样本稀疏时具有较强的处理能力和实用能力，而且将 LLE 与核方法相结合，提出了 SDLE 算法，有效地解决了样本稀疏问题。

3.6 监督与半监督

现有的非线性维数约简方法多数是无监督的，如何将其推广到半监督以及有监督的情况，着眼提高方法的泛化能力，是一个有价值的研究课题。Ridder 等人针对 LLE 提出了监督局部线形嵌入方法(SLLE)[34]，Geng 等人则提出了监督流形学习方法[33]，Belkin 等人则成功地将 LE 应用于半监督的 Riemannian 流形学习[35]和流形规则化[36]。

3.7 算法的评价方法

对于高维数据集的内在维数如何影响高维空间的流形结构，现有的流形学习没有一般性的结论，这使得要进一步分析流形学习的有效性相对困难。张军平等人采用了局部放大因子和主要延伸方向对现有的流形学习方法进行了研究和分析[8,37]，显示出观测空间的高维数据与降维后的低维数据之间的联系，并实验比较了 ISOMAP 算法和 LLE 算法的性能。这种研究可以作为评价流形学习算法的方法。

3.8 算法的泛化

无论哪种流形学习算法总是可以得到训练样本的低维嵌入。但是一旦引入新样本，就需要重新构造权值矩阵，并计算相应矩阵的特征向量，求得低维嵌入值。也就是说，现有的非线性降维方法基本上无法处理实际的大批量数据，只能应用于样本空间。

杨剑、王珏等人在 VQ PCA 和 LTSA 算法的基础上提出一种改进的局部切空间排列算法(PLTSA)[38],该算法解决了 LTSA 中大规模矩阵的特征值分解问题,而且在存储少量数据的条件下能对新样本进行处理,但是这种算法对新样本的处理不具有可推广性。正如 Bengio 指出,现有的基于特征分解的非监督算法,不管是对已知训练集的嵌套还是聚类,在无需重新计算特征向量而直接推广至未知样本上都存在一定的困难[39]。基于对 LLE,ISOMAP 等 5 种不同的流形学习算法进行了分析,Bengio 提出了对这些算法的 out-of-sample 学习能力扩展,试用一个训练好的模型用于 out-of-sample 数据的学习,而无需重新计算特征向量。这种推广形成了一种统一的框架。也就是说,在这种统一的框架下,关系矩阵被看作一个独立于数据的核,这里的核与训练数据相关。根据这个核学习其特征函数,可以不必重新计算特征向量,直接通过 Nystrom 公式得到新样本的低维嵌入。Ham 等人同样考虑到要将流形学习统一到对核函数的研究中,并且把不同关系矩阵看作不同的核矩阵,然后用核方法对这些学习算法进行解释[40]。这些方法在一定程度上对流形学习算法进行了推广,为实质性研究开辟了新的途径。

结束语 流形学习的理论基础问题研究涉及拓扑学、微分几何学等多个数学分支,比较复杂,而要研究鲁棒性更强、应用范围更广的算法又必须有充分的理论支持。因此流形学习的研究和发展已经将机器学习、数学、生物学、信息学等各个学科紧密结合在一起,成为一个具有基础性和前瞻性的研究方向。

谱方法意义下的流形学习算法主要用于维数约简,对图像分析和计算机视觉、信息检索的文本分析、人类基因分布和特征认证等方面的研究具有很重要的意义。目前,非线性降维已经成为高维数据分析必需的预处理步骤。采用传统的降维方法可能并不会取得很好的结果,而通过流形学习方法进行维数约简,一方面可以有效地解决这些高维数据的维数灾问题,另一方面易于直接理解并发现数据集内在的规律。

本文介绍了几种主要的流形学习算法,并对其优缺点做了分析。针对流形学习近几年取得的成果总结了现存的问题,包括邻域的选取、本征维数的确定、噪音的影响、稀疏样本的处理等。这些问题也正是流形学习目前面临的主要研究课题,一旦在理论上取得突破性的进展,必然会推动整个机器学习领域的发展。

参 考 文 献

[1] Seung H S,Lee D D. The manifold ways of perception[J]. Science,2000,290(5500):2268-2269

[2] Tenenbaum J,Silva D D,Langford J. A global geometric framework for nonlinear dimensionality reduction[J]. Science,2000,290(5500):2319-2323

[3] Roweis S,Saul L. Nonlinear dimensionality reduction by locally linear embedding[J]. Science,2000,290(5500):2323-2326

[4] Balasubramamanian M,Schwartz E L. The Isomap Algorithm and Topological Stability[J]. Science,2002,295(5552):7

[5] Belkin M,Niyogi P. Laplacian eigenmaps for dimensionality reduction and data representation[J]. Neural Computation,2003,15(6):1373-1396

[6] Donoho D,Grimes C. Hessian eigenmaps:Locally linear embedding techniques for high-dimensioinal data[J]. PNAS,2003,100(10):5591-5596

[7] Zhang Z Y,Zha H Y. Principal Manifolds and Nonlinear Dimensionality Reduction via Tangent Space Alignmnet[J]. SIAM Journal of Scientific Computing,2004,26(1):313-338

[8] 张军平. 流形学习若干问题研究. 王珏,周志华,周傲英. 机器学习及其应用[M]. 北京:清华大学出版社,2006:135-169

[9] 尹峻松,肖健,周宗潭,等. 非线性流形学习方法的分析和应用[J]. 自然科学进展,2007,17(8):1015-1025

[10] 黄启宏,刘钊. 流形学习中非线性维数约简方法概述[J]. 计算机应用研究,2007,24(11):19-25

[11] Zhang C S,Wang J,Zhao N Y,et al. Reconstruction and analysis of muti-pose face images based on nonlinear dimensionality reduction. Pattern Recognition,2004,37(1):325-336

[12] Weinberger K Q, Saul L K. Unsupervised learning of image manifolds by semidefinite Progralnming. International Journal of Computer Vision,2006,70(1):77-90

[13] 谭璐,易东云,冯国柱,等. 局部不变投影[J]. 自然科学进展,2004,14(3):282-287

[14] Kouropteva O,Okun O,Pietikäinen M. Selection of the Optimal Parameter Value for the Locally Linear Embedding Algorithm// FSKD'02. Singapore,2002

[15] Samko O,Marshall A D,Rosin P L. Selection of the Optimal Parameter Value for the Isomap Algorithm. Pattern Recognition Letters,2006,27(9):968-979

[16] Wang J,Zhang Z Y,Zha H Y. Adaptive Manifold Learning // NIPS 17. Cambridge,2004

[17] 王靖. 流形学习的理论与方法研究[D]. 杭州:浙江大学,2006

[18] 邵超,黄厚宽,赵连伟. P_ISOMAP_一种新的对邻域大小不甚敏感的数据可视化算法[J]. 电子学报,2006,1497-1501

[19] 邵超,黄厚宽,赵连伟. 一种更具拓扑稳定性的 ISOMAP 算法[J]. 软件学报,2007,18(4),869-877

[20] 邵超,黄厚宽. 一种新的基于 ISOMAP 的数据可视化算法[J]. 计算机研究与发展,2007,44(7),1137-1143

[21] Bengio Y,Monperrus M. Nonlocal manifold tangent learning // NIPS. 2005

[22] 赵连伟,罗四维,赵艳敞,等. 高维数据的低维嵌入及嵌入维数研究[J]. 软件学报, 2005,16(8):1423-1430

[23] Camastra F,Vinciarelli A. Estimating the intrinsic dimension of data with a fractal-based method. IEEE Transactions on Pattern Analysis and Machine Intelligence,2002,24(10):1404-1407

[24] Costa J,Hero A O. Geodesic entropic graphs for dimension and entropy estimation in manifold learning. IEEE Transactions on Signal Processing,2004,25(8):2210-2221

[25] Hein M,Audibert Y. Intrinsic dimensionality estimation of submanifolds in Rd// ICML 2005. Bonn,Germany,2005

[26] Levina E,Bickel P J. Maximum likelihood estimation of intrinsic dimension// NIPS 17. Cambridge,2004

[27] 詹德川,周志华. 基于集成的流形学习可视化[J]. 计算机研究与发展,2005,42 (9):1533-1537

[28] Park J H,Zhang Z Y,Zha H Y,et al. Local linear smoothing for nonlinear manifold learning // IEEE Computer Society Conference on CVPR. Washington DC,2004,2:452-459

[29] 郑守志,叶世伟. 局部线性嵌入算法改进研究[J]. 计算机仿真,2007,24(4):78-81

(下转第 59 页)

这些形式化刻画了看，听和缺省推理的人的感觉活动[12]。

(3) 遗忘与记忆：遗忘与记忆的逻辑刻画大概是比较成功的心理过程，它最初是描写机器人在动态环境中如何使数据库更新而使内存不变的逻辑条件，从而导致 Forgetting 理论的提出[13]。它是 Agent 遗忘一个事实这个过程的逻辑抽象。给定一个命题逻辑公式 ϕ，以及一个原子命题 P，在 ϕ 中遗忘 P 写成 forget(ϕ, P)，是一个公式：$\phi(p/\text{true}) \vee \phi(p/\text{false})$，$\phi(p/\text{true})$ 是用 true 处处替换在 ϕ 中的 P 的结果，$\phi(p/\text{false})$ 也以类似的方式定义。Agent 遗忘一个关系 P 可以这样定义：$\text{forget}(\phi, P) = \exists R\phi(P/R)$，$R$ 是一个二元关系谓词变元，$\phi(P/R)$ 是 ϕ 中 P 的每一出现用 R 替换的结果。与遗忘对立的概念是记忆。设 T 是一个理论，$P_1, \cdots, P_k$ 是出现在 T 中的关系谓词，$P_{k+1}, \cdots, P_n$ 是余下的出现在 T 中的关系谓词，则定义对 T 中的关系谓词 $P_1, \cdots, P_k$ 的记忆 $\text{remember}(T, P_1, \cdots, P_k) = \text{forget}(T, P_{k+1}, \cdots, P_n)$。林方真在文献[14]中提出一个计算遗忘的两个算法，这里就不详细讨论了。Forgetting 理论在 Abduction 推理和规划理论中有重要的应用[15]。

(4)社会 Agent 与 MAS(Multi-agent System)：对于单个 Agent，我们有 BDI 模型，即一个刻画 Agent 的信念(Belief)，愿望(Desire)和目标(Goal)等心理活动的认识逻辑，它也是 Kripke 式的模型。对于 MAS(Multi-agent System)，社会承诺(Social Commit)，合作(Cooperation)等社会性心理也得到不同方面的形式化[16,17]。

结束语 我们的论证既有来自心灵哲学的物理主义，也有来自量子计算的 Church-Turing-Deutsch 原理。心灵哲学的物理主义有很多版本，我们采用最小的物理主义，即以依附为基本出发点。我们的论证首次利用量子计算的 Church-Turing-Deutsch 原理，所以我们的计算概念是与通常的采用经典计算概念的功能主义(Functionalism)是不同的，我们的计算概念与实际的物理系统是密不可分的，且与量子物理学保持一致，而这些恰恰是功能主义或心灵的计算理论(Computational Theory of Mind，CTM)所没有的，因为它们的计算概念是基于 Church-Turing 论题的经典概念。所以我们的论证更可靠、更有力。另外，心灵的计算理论尽管承认心灵是可计算的，但这只局限与特定的心理过程，而我们则认为所有的心理过程是可计算的。值得一提的是，我们不仅仅论证所有的心理过程是量子可计算的，而且我们还有大量的计算实践工作，从个人感觉心理到社会心理，都有不同形式的设计和计算实践，其中有些算法已在知识表达和推理，机器人学和分步式智能等研究领域中有重要的应用。心灵既计算已得到哲学、逻辑和实践等不同研究学科的有意义的成果的有力支持。

参 考 文 献

[1] Ravenscroft, Ian. Philosophy of mind: A beginner's guide. Oxford University Press, 2005

[2] [美] 司马贺. 人类的认知：思维的信息加工理论[M]. 荆其诚，张厚粲，译. 北京：科学出版社，1986

[3] Poland J. Physicalism: The Philosophical Foundations. Oxford: Clarendon, 1994

[4] Papineau D. Philosophical Naturalism. Oxford: Blackwell, 1996

[5] Chalmers D. The Conscious Mind. New York: Oxford University Press, 1996

[6] Ryle G. The Concept of Mind. London: Routledge, 1950

[7] Cutland N. Computability. An Introduction to Recursive Function Theory. Cambridge University Press, 1980: 48-67

[8] Deutsch D. Quantum Theory, the Church-Turing Principle and the Universal Quantum Computer // Proceedings of the Royal Society. Series A, 400: 97-117

[9] Fredkin E, Toffoli T. Conservative Logic. International Journal of Theoretical Physics, 1982, 21: 219-253

[10] 董荣胜，计算机科学与技术方法论(ppt)，2008

[11] Gershenson C. Modelling emotions with multidimensional logic // Fuzzy Information Processing Society, 1999. NAFIPS. 18th International Conference of the North American. Vol. July 1999: 42-46

[12] van Linder B, van der Hoek W, Meyer J-J C. Seeing is believing (and so are hearing and jumping). Journal of Logic, language, and Information, 1997, 6: 33-61

[13] Lin F, Reiter R, Forget it! // Proc. of the AAAI Fall Symposium on Relevance. 1994: 154-159

[14] Lin F. On strongest necessary and weakest sufficient conditions. Artificial Intelligence, 2001, 128: 143-159

[15] Lin F, Reiter R. How to progress a database. Artificial Intelligence, 1992: 131-167

[16] Rao A S, Georgeff M P. Modeling rational agents within a BDI-architecture // J. Allen, et al., eds. Proceedings of KR'91. Morgan Kaufman Publishers, 1991: 473-484

[17] Dignum F, van Linder B. Modelling social agents: Towards deliberate communication. Handbook of Defeasible Reasoning and Uncertainty Management Systems. Kluwer, 2002: 357-380

(上接第 28 页)

[30] 曹顺茂，叶世伟. 一种在源数据稀疏情况下的流形学习算法研究[J]. 计算机仿真，2007，24(3)：104-106

[31] 曹顺茂，叶世伟. 一种改进的局部线性嵌入算法[J]. 计算机仿真，2007，24(5)：87-90

[32] 宋欣，叶世伟. 一种在源数据稀疏情况下的数据降维算法[J]. 计算机工程与应用，2007，43(28)：181-186

[33] Geng X, Zhan D C, Zhou Z H. Supervised nonlinear dimensionality reduction for visualization and classification. IEEE Transactions on Systems, Man, and Cyberneties, 2005, 35(6): 1098-1107

[34] de Ridder D, Kouropteva O, Okun O, et al. Supervised locally linear embedding // Artificial Neural Networks and Neural Information Processing, ICANN/ICONIP 2003 Proceedings. Lecture Notes in Computer Science 2714, Springer, 2003: 333-341

[35] Belkin M, Niyogi P. Semi - supervised learning on Riemannian manifolds. Machine Learning, 2004, 56(1): 209-239

[36] Belkin M, Niyogi P, Sindhwani V. On manifold regularization // Proceedings of the International Conference on AI and Statistics. 2005: 17-24

[37] 何力，张军平，周志华. 基于放大因子和延伸方向研究流形学习算法[J]. 计算机学报，2005，28(12)：2000-2009

[38] 杨剑，李伏欣，王珏. 一种改进的局部切空间排列算法[J]. 软件学报，2005，16(9)：1584-1589

[39] Bengio Y, Paiemant J F, Vincent P. Out-of-sample extensions for LLE, Isomap, MDS, Eigenmaps and Spectral Clustering // NIPS 16. Cambridge, 2003

[40] Ham J, Lee D D, Mika S, et al. A kernel view of the dimensionality reduction of manifolds. Technical Report. Tübingen, 2003