

# 社会性群体协作学习环境中的认知聚合研究\*

王蓉, 刘妍妍, 陈禹, 方美琪

(中国人民大学信息学院, 北京 100872)

**摘要** 首先提出社会性群体协作学习和知识共享(SGCL)系统构想,介绍了SGCL系统的框架及需要研究的五个子任务,重点介绍群体认知聚合研究。又提出群体聚合中的节点聚合、关系聚合及用户声望计算模型,然后针对这些模型进行实验验证,一是利用互联网上的分类目录生成个体认知,模拟用户行为,进行聚合,然后与原分类目录进行对比;二是利用Repast多主体建模工具,对系统进行建模仿真。实验结果证明聚合模型与算法符合系统目标的结论。

**关键词** 认知聚合,协作学习,社会计算,集体智慧

**中图分类号** TP391, G434

目前整个社会正在发生着学习的变迁。这种变迁一方面来自人类对于自身认知的深入研究;另一方面来自技术对学习过程的支撑。从认知角度来看,知识的建构性、社会性、情境性、复杂性和默会性<sup>[1]</sup>等新的认知观得到重视,这些认知观强调学习者的学习过程应该是在一定的社会环境下与他人交流、协作,不断建构自己的认知或知识体系的过程,即强调社会性群体协作过程。计算机支持的协作学习(CSCL)领域也在适应着这种变化趋势。Koschmann认为CSCL的核心研究领域所关心的是在共同活动环境中意义形成的实践,以及实现这些实践活动的途径<sup>[2]</sup>。从这个阐释来看,“共同活动环境”强调了社会性;“意义的形成”强调人们对客观世界的认知是一个动态、变化、不断建构的过程;“实现这些实践活动的途径”则强调为意义的形成提供一种环境和工具。此外,全球大脑<sup>[3-5]</sup>和社会计算<sup>[6-7]</sup>的思想促使信息系统的研究向集成深度、群体广度和历史经验方向发展。在这种背景下,需要研究一种能够充分发挥数字化学习特点,支持社会性群体协作学习和知识共享(Social Group Collaborative Learning, SGCL)的机制与技术,这种机制与技术能够存储和展现微观上的个体认知和宏观上的群体认知,以及个体和群体的认知历史轨迹<sup>[8-9]</sup>。

在上述背景与前提下,本文首先介绍了社会性群体协作学习与知识共享的系统目标及框架,以及该系统所支持的个体认知模型与行为,然后在此基础上重点阐述系统将个体认知聚合为群体认知的模型与方法,并通过实验的方法验证聚合模型的有效性。

## 1 SGCL 系统目标及框架

从微观来看,每个个体从出生起就生活在社会中,从事各种社会活动,通过与其他个体或环境的交互、沟通,不断建构自己对客观世界的认知,整个人的生命过程就是一个不断学习的过程。每个个体的建构过程在宏观上又体现出人类对客观世界的认识,这种认知反过来又会影响个体知识的建构。

\* 基金项目: 国家社会科学基金(06BTQ016), 信息管理与信息经济学教育部重点实验室开放基金(F0607-16)

通信作者: 王蓉, 中国人民大学信息学院, 讲师, E-mail: rong@ruc.edu.cn

社会作为一个整体,在不断地发现和积累知识,因此人类的历史也是一个不断探索、发现的过程。微观个体认知和宏观人类认知是不断发展、变化的,有其建构的历史轨迹(变迁)。SGCL系统的目标就是能够提供这样一个环境,使得个体能够在其中建构自己的认知(结构),系统则将个体建构的认知聚合成群体认知,聚合后的群体认知反过来为个体建构认知提供帮助。个体与群体的交互关系如图1所示。

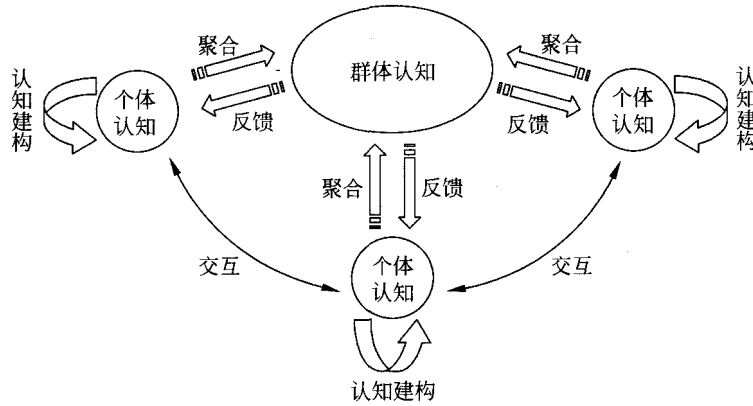


图1 SGCL系统中个体与群体的交互关系

根据系统目标设计出SGCL系统的系统框架(图2)。SGCL包括五个主要的子课题:(1)个体认知建构研究;(2)群体认知聚合研究;(3)促进系统向良性方向演化的协作机制研究;(4)知识群、学友群和资源群的发现算法研究;(5)支持系统功能与特点的可视化研究。本文重点介绍SGCL系统中的群体认知聚合的相关研究。

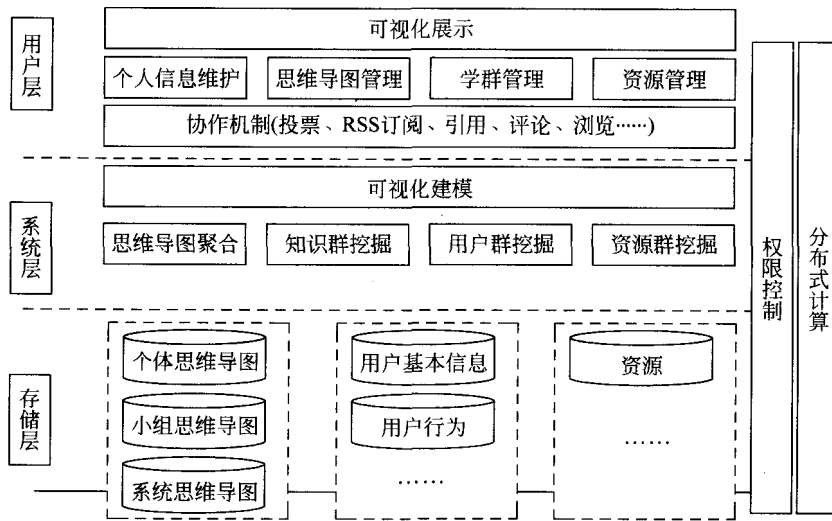


图2 SGCL系统框架

## 2 个体认知模型和行为

为了更清楚地说明SGCL系统中的认知聚合研究,先简单介绍一下系统如何支持个体的认知建构。

## 2.1 个体认知模型

个体认知建构部分的实质是个体认知的外化。这种外化可以通过语言、肢体动作、文字和作图等方式进行展现。对于基于 Web 的环境而言,考虑到计算机处理数据的特点,并遵循“左右脑分工理论”(split-brain)<sup>[10]</sup>及“双重编码理论”(dual-coding theory)<sup>[11]</sup>,SGCL 系统为个人提供的认知建构工具需要采用图文的方式,以同时激活语言系统和非语言系统。

通过对多种现有的知识可视化工具进行分析研究<sup>[12,13]</sup>,SGCL 系统选择思维导图作为个体建构自己认知体系的工具。思维导图(Mind Map)也称为心智图,是英国人托尼·巴赞(Tony Buzan)于 20 世纪 60 年代创造的一种笔记方法。思维导图从一个中心点开始,以一种无穷无尽的分支链的形式从中心向四周放射。这种方法遵循大脑放射性思维的方式,全面调动左脑的逻辑、顺序、条例、文字、数字以及右脑的图像、想象、颜色、空间和整体思维,使大脑潜能得到充分的开发,从而极大地挖掘人的记忆、创造、身体、语言、精神和社交等各方面的潜能<sup>[14]</sup>。

用户在自己的空间以类似于画思维导图的方式建构自己的认知体系,为描述方便,本文称这种建构出来的图形为认知图。认知图以一个节点为中心,向四周扩散到一级节点,然后一级节点再扩散到二级节点,以此类推,如图 3 所示。虽然节点与节点之间连接主要呈树形关系,但是也会存在图中虚线所示的那种连接,因此个体认知可以表示为图形拓扑结构。认知图可以抽象为节点及节点与节点之间的连接。这样,第  $i$  个用户的个体认知体系的拓扑结构可以表示为  $G_i(U_i, R_i)$ ,  $G_i$  表示个体  $i$  建构的知识体系,其中  $U_i$  表示个体  $i$  建构的知识点集合,  $R_i \subseteq U_i \times U_i$  表示知识点之间的映射关系。另外,用户可以为每个节点添加与之相关的资源。

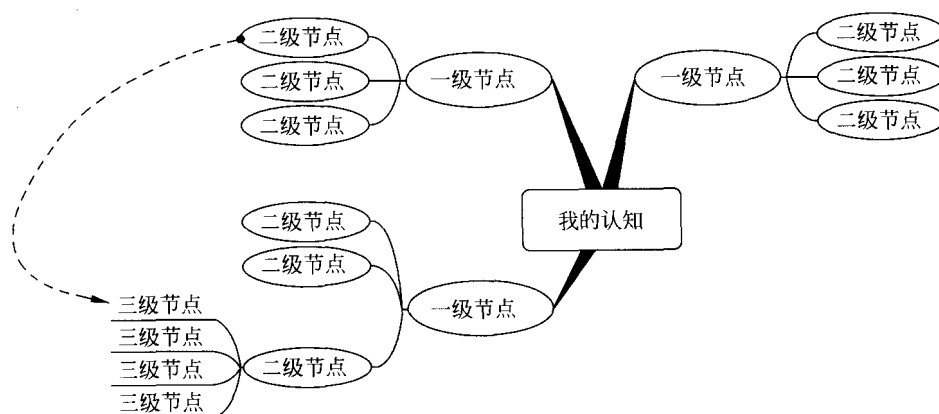


图 3 个体建构认知图示意

## 2.2 个体行为

SGCL 系统中个体参与将作为聚合参数,影响聚合结果,即用户群体行为决定着群体认知演化。用户在建构过程中可以参照系统聚合后的群体认知,以此实现个体与个体、个体与环境之间的相互协作。

系统中个体行为可以分为两类:一类个体行为作用于自己的个体认知图;另一类行为作用于系统聚合后的认知图。为了表述方便,我们把前者称为建构行为,后者称为交互行为。建构行为包括个体对自己的认知图的编辑行为(认知图节点和关系的增、删、改,资源的添加、删除与修改)。交互行为体现为个体对群体认知图的浏览、引用、投票、导入和订阅等行为。用户建构自己的认知体系时,系统

将提示聚合后的群体认知图中的相似和相关节点,用户可以浏览群体(或个体)认知图中节点,若认可其节点或关系,可对其实施浏览、引用、投票、导入、订阅和评论等行为。系统将记录这些行为,并将它们作为参数参与群体认知聚合计算。

### 3 群体认知聚合

将个体认知聚合为群体认知是 SGCL 系统的核心,是聚合个体智慧为群体智慧的关键。在 SGCL 系统提供的群体协作环境中,虽然每个个体行为是主观、多样的,但是群体行为在整体上呈现出的结果代表了群体的认知结构,而且这种结构会随着系统贡献内容的变化而不断演化。

SGCL 系统中的聚合是指将个体建构的个体认知图通过聚合算法进行计算,生成群体认知图,是个体(主体)行为的宏观呈现。这种聚合不是个体认知图中的节点和关系的简单累加,它还反映了用户之间交流与协作后对群体和个体认知图的影响。因此,SGCL 系统中的群体认知是以个体认知为基础,以用户行为为依据,采用聚合算法提取出的群体认知。

聚合的基本原则为将个体认知图中完全相同的节点聚合为一个节点,尊重用户的输入,依靠人本身强大的语义理解与处理能力对聚合后的群体认知图进行评判与选择,而不是通过计算处理判断语义来干涉聚合结果。其步骤为:

- (1) 完全相同的节点则合并为一个节点;
- (2) 增量计算聚合后的节点权重;
- (3) 增量计算聚合后关系的权重;
- (4) 根据用户的行为,增量计算用户声望;
- (5) 聚合后的结果将反馈给个体,对个体下次建构产生影响。

“增量”的含义是指上次聚合至本次聚合之间所发生的变化,包括个体对本人认知图的增、删、改,个体对群体认知图的交互行为。

SGCL 系统中的认知聚合分为三个层次,系统整体聚合、学群聚合和用户自定义聚合。学群聚合又分为显性学群聚合和隐性学群聚合。如图 4 所示。系统整体聚合将所有用户的认知图进行聚合,它体现系统中所有个体的共同认知。系统整体认知体现的是系统内所有个体的意见,多样化的视角会帮助该个体从不同角度思考问题,突破某个领域的局限,培养发散性思维,而发散性思维对于知识学习和知识创新是极为重要的。

学群聚合分为显性学群聚合和隐性学群聚合。显性学群是用户自己以显性方式建立的学群,显性学群聚合就是将显性学群中的个体认知进行聚合。隐性学群是系统根据个体建构的认知图以及个体行为通过算法挖掘出的认知结构相似或兴趣相似的学习群体,隐性学群聚合就是将隐性学群中的个体认知进行聚合。因为同一群体中的成员的认知结构或兴趣爱好相似,所关注的问题也比较集中,所以,这种小群体的认知聚合会形成更收敛的聚合结果,当个体需要聚焦某个学习领域时,这种聚合能提供更大的帮助。

用户自定义的群体聚合是由用户指定将哪些个体的认知图进行聚合。这种自定义的功能允许用户专注于某些“专家”的意见,而这种专家是用户通过自己的分析与思考进行判断的。

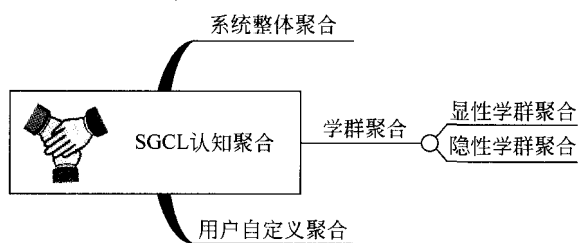


图 4 SGCL 系统中的认知聚合分类

不论哪种聚合方式均采用相同的聚合模型,只是聚合对象的范围和用户行为因素上存在差异。聚合时首先在个体认知模型的基础上建构群体的聚合模型,然后再考虑用户的交互行为对聚合结果的影响。下面分别阐述系统中节点聚合、关系聚合和用户声望计算模型。

### 3.1 节点聚合模型

个体认知模型可以抽象成节点与节点之间的连线组成的图。聚合首先是将个体认知图中“相同”节点合并成一个节点,并记录源节点的相关信息,这样用户可以通过聚合的系统认知图追溯到每个节点的历史。这里的关键问题就是如何判断不同个体认知图中哪些节点为“相同”节点。可以有两种处理方法:一是节点内容完全相同,则判断为相同节点,如“Web2.0”和“web2.0”虽然表示同一意思,但不被认为是相同节点;二是对节点内容进行语义分析,将语义相同的节点判断为相同节点。在此我们采用第一种处理方法,理由是SGCL系统的目的是体现集体智慧,系统尊重用户的输入,不通过计算处理干涉人的语义理解,而是依靠人本身强大的语义理解与处理能力来进行判断和筛选,通过众人的力量来促成系统中共同语义的达成。但是系统提供一定机制促进相同语义的收敛,如个体在编辑认知图节点时,系统提示相似的节点名称,允许用户在聚合的认知图中指示两个节点表达同一语义等。

聚合认知图中的每个节点由于众人支持程度不一样,而具有不同的权重。众人支持程度由两部分因素衡量,一是个体建构的认知图叠加因素,如两个人同时都在自己的认知图中创建了某节点或关系;二是群体对聚合后的认知图施与行为所体现出来的认可程度,如某节点获得更多的浏览次数、投票数和引用次数等。本文建立了式(1)来计算聚合后的节点权重。

$$W_{i_t} = dW_{i_{t-1}} + pR + xQ + yV + zL \quad (1)$$

式中:

$W_{i_t}$ —— $t$ 时刻节点 $i$ 的权重

$W_{i_{t-1}}$ —— $t-1$ 时刻节点 $i$ 的权重

$d$ ——时间衰减因子

$p$ ——声望权重系数

$R$ ——创建该节点的用户声望之和减去删除该节点的用户声望之和

$x$ ——引用权重系数

$Q$ ——引用该节点的人数减去引用后删除该节点的人数

$y$ ——投票权重系数

$V$ ——对该节点投票的人数

$z$ ——浏览权重系数

$L$ ——浏览过该节点的人数

从式(1)可以看到,聚合节点的权重根据用户参与及认可程度动态变化。每个节点的当前权重均与上一次聚合后的该节点权重,对该节点施加的行为及施加行为的人有关。因为老的节点有较长的生存期来积累权重,这样会抑制新产生认知内容,所以用时间衰减因子来降低老节点在系统中的影响力。各种行为对聚合节点的权值影响程度是不一样的,因此,用户声望、引用、投票、浏览权重系数的设置决定着聚合结果的价值取向,也体现着系统演化机制。

### 3.2 关系聚合模型

节点与节点之间关系(连线)代表着个体的认知路径,若多个用户拥有相同路径,在聚合后的认知

图中会有较大的权重。相同节点聚合为一个节点的原则确定了聚合认知图的拓扑结构。关系聚合不会改变聚合认知图的拓扑结构,只是计算节点与节点间关系(连线)的权值,通过权值来标识群体对认知路径达成共识的程度。关系聚合可用 $(U, R, W)$ 表示,其中 $U$ 是用户感兴趣的话题构成的节点集合, $R \subseteq U \times U$ 是 $U$ 和 $U$ 之间的关系集合, $W: R \rightarrow R$ 是关系的权重。对节点关系的聚合,就是要确定节点与节点之间的关系和每个关系的权重 $W$ 。式(2)为关系权重的计算公式。

$$W_{i_t} = dW_{i_{t-1}} + pR + xQ + yV + zL \quad (2)$$

式中:

$W_{i_t}$ —— $t$ 时刻关系 $i$ 的权重

$W_{i_{t-1}}$ —— $t-1$ 时刻关系 $i$ 的权重

$d$ ——时间衰减因子

$p$ ——声望权重系数

$R$ ——创建该关系的用户声望之和减去删除该关系的用户声望之和

$x$ ——引用权重系数

$Q$ ——引用该关系的人数减去引用后删除该关系的人数

$y$ ——投票权重系数

$V$ ——对该关系投票的人数

$z$ ——浏览权重系数

$L$ ——浏览过该关系的个体数

由式(2)可以看到,关系创建权值取决于用户的声望高低和创建该关系的用户的数量。越是常被一起使用的话题,越是被权威人士认为相关,创建权值就会越高。引用代表着用户在浏览他人认知图或者浏览群体聚合认知图的时候对比较赞同的节点关系(也就是认知路径)引用到自己的认知图中,一个关系被引用的越多,表明赞同这两个话题相关的用户越多,引用权值也就会越高。投票结果也代表着用户对话题相关度的赞同度,所有人投票结果之和,构成了投票权值。浏览行为也反映这种多用户对相关认知路径的使用频繁程度,浏览权值越高,代表着认知路径对用户的导向作用越明显,对认知的贡献程度越高。

### 3.3 用户声望计算模型

用户声望反映的是用户对群体建构贡献的程度,是用户个人的认知建构被群体所支持的程度,在SGCL系统中用户创建的节点和关系被他人引用、投票支持、浏览得越多,他建构的认知对集体智慧的贡献也就越大,声望也就越高。用户声望的计算公式如式(3)。

$$R_{i_t} = R_{i_{t-1}} + xQ + yV + zL \quad (3)$$

式中:

$R_{i_t}$ ——当前时刻 $t$ 个体 $i$ 的声望

$R_{i_{t-1}}$ ——前一时刻 $t-1$ 个体 $i$ 的声望

$x$ ——引用权重系数

$Q$ ——个体 $i$ 的认知图被其他个体引用的次数

$y$ ——投票权重系数

$V$ ——个体 $i$ 的认知图被其他个体投票的次数

$z$ ——浏览权重系数

$L$ ——个体  $i$  的认知图被其他个体浏览的次数

用户声望可以通过建构用户模型,针对用户专长的领域计算得更加细致。

## 4 系统建模与仿真

为了验证系统聚合模型的有效性,观察聚合模型中的参数对聚合的影响,我们从两方面做了实验,一是分类目录模拟实验;二是多主体建模与仿真实验。下面分别介绍。

### 4.1 分类目录模拟实验

实验一是基于聚合模型,将模拟生成的个体认知建构数据和用户行为数据通过聚合算法计算群体认知,并反馈给用户,为用户下一步的认知建构提供帮助。本文考虑采用当前流行的分类目录作为认知的数据来源。分类目录是通过人工的方式收集网站,并把这些具有一定价值的网站资源按照主题进行整理和组织,之后存放到相应的目录下,从而形成网站的分类目录体系,因为它的分类结构源于现实世界,与人认知结构类似,能从一定程度体现认知语义。实验的基本思路为:

- (1) 首先在系统中模拟生成一些学习者;
- (2) 利用互联网上的新浪爱问分类目录<sup>①</sup>开放分类目录,根据一定的抽取原则,为用户生成初始的个体建构的认知图;
- (3) 针对每一个学习者,基于一定概率及当前群体认知情况,模拟生成一些可能产生的行为(浏览、投票、引用、创建、删除、无任何行为);
- (4) 利用聚合算法将用户新产生的行为结果利用增量方式进行聚合,形成新的群体认知;
- (5) 重复执行第 3、第 4 步,整个系统不断地运行演化。

本实验新浪爱问分类目录下抽取 10 个大的分类,总共包含 2 156 个节点,再以这 2 156 个节点作为数据源,生成模拟数据。如果将新浪爱问下抽取的 10 个分类目录作为系统中所有个体的认知全集,那么每一次聚合后得到的群体认知就相当于这个全集下面的一个子集,那么随着时间的推移,如果子集的结构逐渐趋同于全集,则说明群体认知聚合模型能够有效地把集体智慧逐渐呈现给个体,从而使得个体在集体智慧的指导下更加有效地建构自己的认知。

为此,本文采用图的拟合度作为衡量聚合模型的指标。本文所指的图的拟合度是:假设认知全集  $T$  中边的总数为  $T$ ,群体认知图中边的总数为  $G$ ,既在群体认知图中出现同时又在全集中出现的边的数目为  $m$ ,则该认知图的拟合度为  $m/T$ 。

另外,用群体认知图中有用边的比率( $m/G$ )来表达群体认知中集体智慧的含量。因为个体在建构自己的认知时,可能存在一些观点不能被大众所接受,这样的观点不应该代表集体的智慧。好的聚合模型应该能保证聚合后的群体认知中代表集体智慧的成分保持在较高水平,并且不会随时间推移而衰减。

我们对式(1)、式(2)和式(3)的参数进行调整,每个参数集运行 10 次实验,每次实验模拟系统运行 30 个时间步,在每个时间步里,首先模拟在上一轮聚合结果的反馈作用下用户的行为,然后进行本次时间步的群体认知聚合,最后搜集统计信息。图 5~图 9 显示的每个数据点是 10 次实验相应时间步采集数据的平均值。

<sup>①</sup> 新浪爱问分类目录是由新浪搜索专业编辑挑选而形成的, <http://dir.iask.com/>

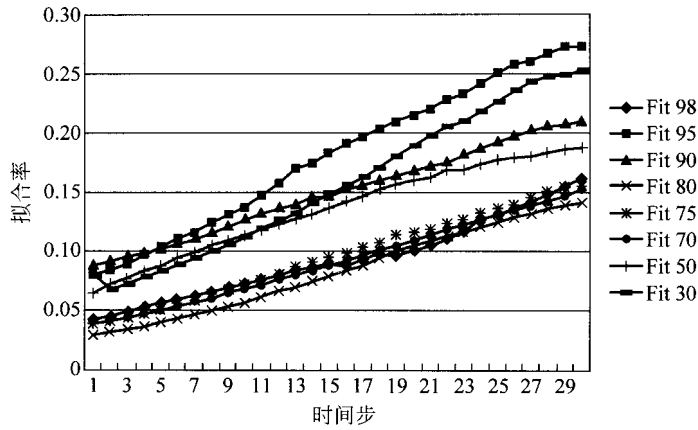


图5 不同衰减因子下的拟合率

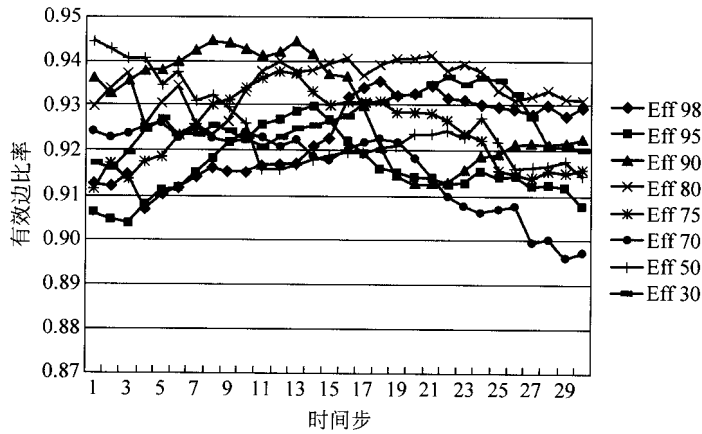


图6 不同衰减因子下的有效边比率

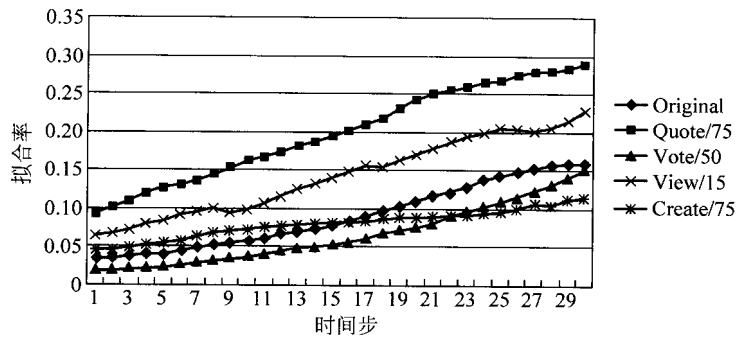


图7 不同权重系数下的拟合率



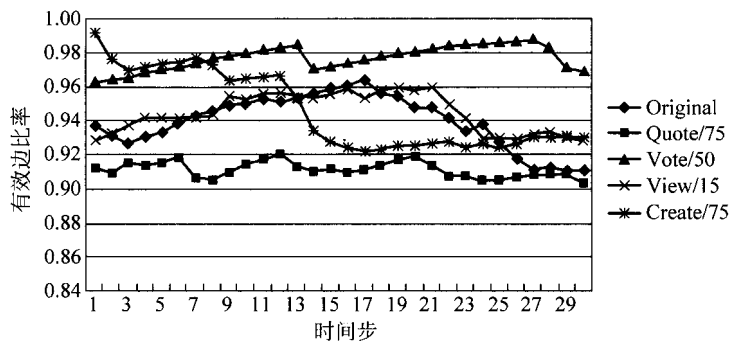


图 8 不同权重系数下的有效边比率

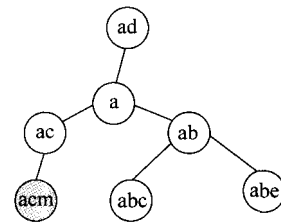


图 9 模拟中某用户可能生成的认知图

图 5 显示的是不同衰减因子下实验运行结果计算出来的聚合认知图拟合率(FitRate)。衰减因子越大,对系统热门节点和关系的衰减作用越小。从图中可以看到,不论衰减因子大小,拟合率的整体走势都是呈上升趋势,即随着时间的推移,聚合后的认知图中与原集中边的拟合程度越来越高。这说明系统的聚合模型反映了促进系统向良性方向演化的目标。

图 6 显示的是不同衰减因子下实验运行结果计算出来的聚合认知图中有效边比率(Effective Rate)。从图中可以看到,衰减因子大小与曲线走势之间的关系很难发现规律,随着时间的推移,每个衰减因子对应的曲线都在上下波动。因为每次实验产生的初始数据量都是随机生成的,所以 G 值也是随机的,但从曲线的变化来看,系统中每个个体都以自己的方式在建构认知图,所以体现一定的随机波动性。

图 7 显示了不同的权重系数发生变化时拟合率的变化情况。可以看到,权重系数的变化会影响拟合率的走势。

图 8 显示了不同的权重系数发生变化时有效边比率的变化情况。可以看到,权重系数的变化会影响有效边比率,但其走势也在发生波动。权重系数的确定代表了系统对不同用户行为(创建、浏览、投票、引用等)的重视程度。在系统运行后,需要收集实际数据进行分析,定出合理的权重系数。

### 4.2 多主体建模与仿真

实验二是利用 Repast(Recursive Porous Agent Simulation Toolkit),对系统进行建模仿真。Repast 是理想的多主体建模工具,提供了用于主体仿真建模的各种基本类。建模者可以通过直接使用这些类或从这些类继承子类来构造自己的模型<sup>[15,16]</sup>。

在 SGCL 系统中,将每个个体看做一个独立的主体,每个主体通过浏览、投票、引用其他主体认知图等方式与其他个体建立起协作学习关系。由于每个个体建构自己的认知图,个体之间的关系也是通过对认知图的相关操作而发生的,因此,为了模型的表达力更强,本文在模型实现时将每个个体的认知图看做行为主体,将聚合后的群体认知图作为环境主体。在每个时间步里,行为主体根据环境主体对其产生的反馈,并根据自己的行为规则采取相应的行为。这里的行为规则与 4.1 节的实验相同。另外,为了简化模拟,模拟模型中只是考虑了用户对其认知图的行为,及认知图聚合后对用户行为的影响情况,而没有考虑用户对资源的操作情况。

为了实现模拟,不失一般性,我们引入概念群(concept group),概念群中的概念与概念之间相关系数较大。设系统中有  $n$  个概念群  $G_i(1 \leq i \leq n)$ ,每个概念群包含  $m$  个基本的概念元,基本概念元用字母表示,每个概念群的基本概念元的交集为空。概念元是用户认知图节点上的基本组成单位,这些

概念元按一定规律组合,构成了认知图中的节点。这里还假设每个概念群包含  $p$  个用户,每个用户拥有自己唯一的认知图,每个用户的兴趣属于或专长于一个或多个概念群。

例如概念群数  $n$  为 3,每个概念群的基本概念元  $m$  为 6,则  $\{a,b,c,d,e,f\}$  为  $G_1$  概念群的基本概念元, $\{g,h,i,j,k,l\}$  为  $G_2$  概念群的基本概念元, $\{m,n,o,p,q,r\}$  为  $G_3$  概念群的基本概念元。假设某用户的兴趣为  $G_1$  概念群,则他的认知图主要围绕  $G_1$  中的概念元  $\{a,b,c,d,e,f\}$  生成,但也以小概率生成其他概念群中的节点。图 9 所示的是一个以  $a$  扩展的认知图,可以看到,除了节点  $acm$ ,其他节点中的概念均是由概念群  $G_1$  中的基本概念元  $\{a,b,c,d,e,f\}$  组合而成。

实验中模拟用户可以采取的行为有:创建、浏览、投票、引用,实验将采取一定概率决定将采取哪种行为。一般情况下,我们认为比较贴近实际的规律是:学习者浏览的行为要多于其他的几种行为,因为多数情况下,学习者首先会去浏览他感兴趣那部分系统整体的认知或者其他学习者的认知,然后确定如果他发现有他认同的认知,他才可能去创建这个认知点,或者将该认知点引入自己的认知图中。所以,我们在模拟用户行为时,赋予“浏览”这一行为的概率要比采取其他行为的概率大一些。

另外,聚合成群体认知以后,该群体认知可以认为是当时环境下所有学习者的集体智慧,学习者可以从中得到一些启示,从而采取相应的行为。在模拟群体认知对每个个体产生的反馈时,我们认为聚合后的群体认知图中,具备较大权值的那些认知点以及节点关系对学习者的影响更大,因为它们受到群体的关注度是相对较高的。

在 Repast 中,调度类负责仿真时钟的推进,行为的安排与调度。仿真时钟的单位是“tick”,可称为时间步。在 Repast 中,系统状态的改变是在离散时间点上被改变的,分为离散事件调度和离散时间调度。事件与事件的处理子程序被封装在一个行为类中,事件的安排与调度变成了行为的安排与调度。

本文实验的调度中,第一步,将每个用户的当前认知图进行聚合,生成群体的认知图;第二步,获得当前群体认知图中的热门知识点;第三步,每个用户发生一次行为(创建、浏览、投票、引用或无行为),用户的行为会改变自己的认知图;第四步,显示本轮结果;第五步,准备下一轮运行。

图 10~图 12 显示了时间步(tick)为 1、15 和 35 时的系统仿真的聚合后群体认知图。刚开始时,系统中的认知图相对分散[用不同形状(方块、圆点和三角)表示不同类别的概念,空心圆表示关注度较高的节点],经过一段时间的运行后,随着用户创建、浏览、投票、引用等行为的发生,各概念群发生聚合。

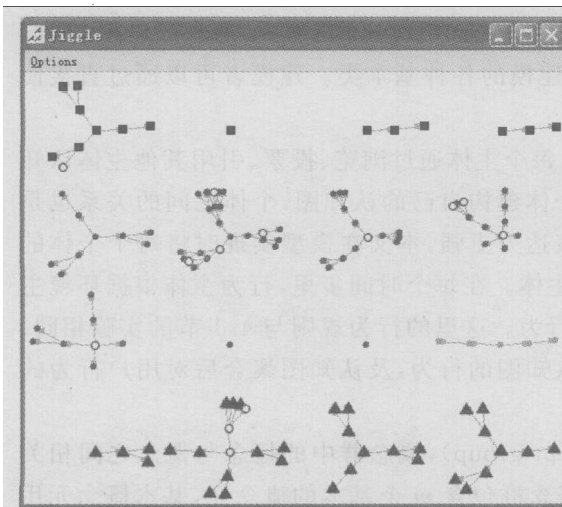


图 10 tick=1 的结果图

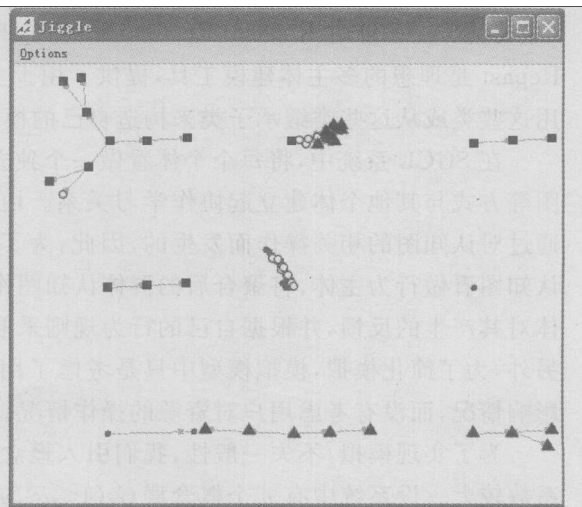


图 11 tick=15 的结果图

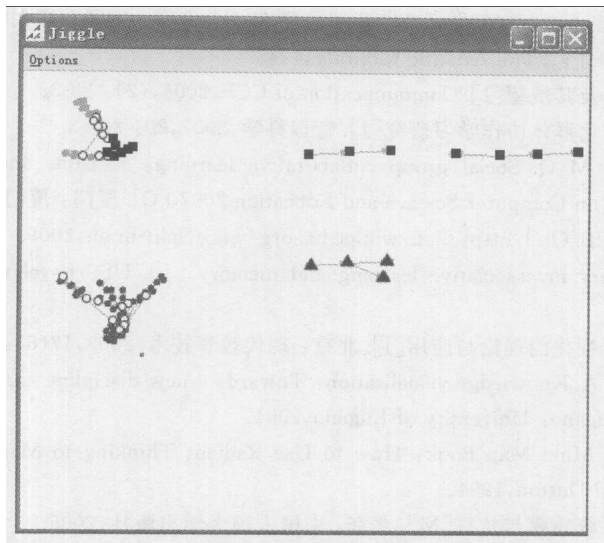


图 12 tick=35 的结果图

在时间步为 35 时(如图 12 所示),三个不同颜色概念群的群体认知聚合在一起,表明经过足够长的时间演化,各个概念群已经打破了原有的领域限制,形成了一个更丰富、更完善的认知体系。从认知角度来看,多数情况下,每个领域内的个体认知都集中分布在其所属的领域内,领域之间联系不是很紧密,在 SGCL 系统内,可以打破领域之间的壁垒,多样性个体的学习过程就可以在不同领域之间建立联系,从群体角度体现出来的就是一个相对完整的认知体系。

## 5 总结

随着认知科学、教育理论和计算机及网络的发展,人们的学习方式正在发生变化,在此前提下本论文首先提出社会性群体协作学习与知识共享的思想、系统框架以及在此框架下应该研究的五个子课题。然后重点介绍群体认知聚合研究,设计了节点聚合模型、节点关系聚合模型以及用户声望计算模型。通过实验对模型进行验证,证明聚合模型符合 SGCL 系统的设计目标。后续研究将进一步细化模型中参数之间的关系,将 SGCL 原型系统投入使用,收集实证数据,研究分析用户的行为,然后对群体认知聚合算法进行完善。

## 参考文献

- [1] John D Bransford. 人是如何学习的——大脑、心理、经验及学校[M]. 程可拉等译. 上海: 华东师范大学出版社, 2002.
- [2] Koschmann T. Dewey's contribution to the foundations of CSCL research [A]. Computer Supported for Collaborative Learning: Foundations for a CSCL Community, the Proceedings of CSCL2002[C]. Hillsdale, NJ: Erlbaum, 2002: 17-22.
- [3] Russell P. The Global Brain: Speculations on the Evolutionary Leap to Planetary Consciousness [M]. Boston, MA: Houghton Mifflin, 1983.
- [4] O'Reilly T. What is Web 2.0—Design patterns and business models for the next generation of software [EB/OL]. <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-2.0.html>, 2005.

- [5] 张树人. 从社会性软件、Web 2.0 到复杂适应信息系统研究[D]. 北京: 中国人民大学, 2006.
- [6] Bush V. As we may think [J]. The Atlantic Monthly, 1945.
- [7] 王飞跃. 社会计算的意义及其展望[J]. Communication of CCF, 2006, (2): 28-35.
- [8] 王蓉, 方美琪, 陈禹. 社会化群体协作学习研究[J]. 管理科学, 2007, 20: 80-83.
- [9] Wang R, Feng X, Fang M Q. Social group collaborative learning: Reasons and features [A]. Proc of 3rd International Conference on Computer Science and Education 2007 [C]. 厦门: 厦门大学出版社, 2007.
- [10] Wikipedia. Split-brain[EB/OL], <http://en.wikipedia.org/wiki/Split-brain>, 2009.
- [11] Paivio A. Mental imagery in associative learning and memory [J]. US: Psychological Review, 1969, 76 (3): 241-263.
- [12] 王朝云, 刘玉龙. 知识可视化的理论与应用[J]. 北京: 现代教育技术, 2007, 17(6): 18-20.
- [13] Eppler M J, Burkard R A. Knowledge visualization: Towards a new discipline and its fields of application [A]. ICA Working Paper. Lugano; University of Lugano, 2004.
- [14] Buzan T, Buzan B. The Mind Map Book: How to Use Radiant Thinking to Maximize Your Brain's Untapped Potential [M]. US: E P Dutton, 1994.
- [15] 方美琪, 张树人. 复杂系统建模与仿真[M]. 北京: 中国人民大学出版社, 2005.
- [16] Tobias R, Hofmann C. Evaluation of free Java-libraries for social-scientific agent based simulation[J]. Journal of Artificial Societies and Social Simulation, 2004, 7(1): visited at <http://jasss.soc.surrey.ac.uk/7/1/6.html>.

## Aggregation Model of Group Cognition in Social Group Collaborative Learning Environment

WANG Rong, LIU Yanyan, CHEN Yu & FANG Meiqi  
(School of Information, Renmin University of China, Beijing 100872)

**Abstract** At first, this paper introduces the idea of social group collaborative learning and knowledge sharing (SGCL), framework of SGCL system and its six research subtasks. Then this paper pays more attention to one of the subtasks, group cognition aggregation model research. As to group cognition aggregation, this paper suggests nodes aggregation model, relations aggregation model and users' reputation computing model. To verify these models, the paper introduces two experiments. In one experiment, we first draw individual cognition maps from directory on the Internet, then simulate user behaviors and aggregate individual cognition maps into group cognition map, at last compare the aggregation results with the original directory. In another experiment, we simulate the SGCL system with Repast. The results of the experiments show that the model and algorithm accord with the system's goal.

**Key words** Cognition aggregation, Collaborative learning, Social computing, Collective intelligence

### 作者简介:

王蓉, 中国人民大学信息学院, 博士, 讲师, e-mail: rong@ruc.edu.cn

刘妍妍, 中国人民大学信息学院, 硕士, 现任职于 IBM, e-mail: liuyy\_ruc@126.com

陈禹, 中国人民大学信息学院, 博导, 教授, e-mail: yuchen318@gmail.com

方美琪, 中国人民大学信息学院, 博导, 教授, e-mail: meiqif@gmail.com