

结合 HowNet 的可拓策略生成软件研制

李卫华^{1,2}, 杨春燕^{1,3}

1. 广东省创新方法与决策管理系统重点实验室, 广州 510006
2. 广东工业大学计算机学院, 广州 510006
3. 广东工业大学可拓学与创新方法研究所, 广州 510006

摘要 针对现有可拓策略生成软件自然语言理解能力不高的现状, 研制了一种结合 HowNet(知网)的可拓策略生成软件, 通过学习 HowNet 的知识组织, 改进可拓策略生成软件的基础知识库结构, 并通过人机界面的智能引导, 减轻计算机对自然语言理解的困难。以企业提高客户价值需求为案例研制的结合 HowNet 的可拓策略生成软件, 能利用人和计算机各自的优势, 为企业提高客户价值提供技术支持。以服装企业客户数据进行测试, 结果表明该可拓策略生成软件有效地满足了服装企业提高客户价值的目标。

关键词 可拓策略生成软件; HowNet; 客户价值

中图分类号 TP30

文献标志码 A

doi 10.3981/j.issn.1000-7857.2014.36.004

Develop an Extension Strategy Generating Software System Combined with HowNet

LI Weihua^{1,2}, YANG Chunyan^{1,3}

1. Key Laboratory of Innovation Method and Decision Management System of Guangdong Province, Guangzhou 510006, China
2. School of Computers, Guangdong University of Technology, Guangzhou 510006, China
3. Research Institute of Extenics and Innovation Methods, Guangdong University of Technology, Guangzhou 510006, China

Abstract The extension strategy generating software can help people generate strategies to solve contradiction problems. But the ability to understand natural language of the current extension strategy generating software is very low and its strategy generative capacity is limited. In view of this situation, a system of extension strategy generating software combined with HowNet is designed. By learning HowNet's knowledge organization the basic knowledge base structure of the extension strategy generating software is improved, and by intelligent guidance at the man-machine interface, its natural language understanding capacity is improved. As a case study, an extension strategy generating system combined with HowNet for enhanced customer value requirement is implemented. It can use both advantages of human and computer to help enterprises improve customer value. Tested by a clothing company, the extension strategy generating system can meet the requirement of enhance customer value for that clothing company.

Keywords the extension strategy generating system; HowNet; customer value

利用可拓学^[1]的理论与方法, 研究解决矛盾问题的策略生成的计算机实现系统, 简称可拓策略生成系统^[2], 可拓策略生成软件是其软件实现部分。研究可拓策略生成软件的意

义是利用计算机技术协助人们解决矛盾问题, 即在当前条件下无法实现人们所希望的目标的问题。目前, 笔者团队已研制了多个可拓策略生成软件^[3,4], 并获得国家颁发的软件著

收稿日期: 2014-10-23; 修回日期: 2014-11-29

基金项目: 国家自然科学基金项目(61273306); 广东省自然科学基金项目(815009001000041); 广东省创新方法与决策管理系统重点实验室开放课题(2011A060901001-13B)

作者简介: 李卫华, 教授, 研究方向为智能软件, 电子邮箱: lw@gdut.edu.cn

引用格式: 李卫华, 杨春燕. 结合 HowNet 的可拓策略生成软件研制[J]. 科技导报, 2014, 32(36): 32-36.

作权。

信息技术的飞速发展使得利用计算机技术尤其是利用可拓策略生成软件来实现更多的解决矛盾问题的辅助工具成为可能。然而计算机理解自然语言的能力很有限,无法深刻认识人们所希望解决的问题,因此,必须设法提高计算机的自然语言处理能力。

HowNet(知网)是一个描述概念与概念之间的关系,以及概念的属性与属性之间的关系的知识系统^[5]。在纵览 HowNet 应用的国内外现状时,发现基于 HowNet 进行语义相似度计算方面的研究文献最多,其次是基于 HowNet 的词义消歧方面的文献,接下来是 HowNet 和汉语研究方面的文献,还检索到语义情感倾向方面的文献、基于 HowNet 文本分类与聚类方面的文献,等等。这说明 HowNet 技术有利于提高计算机的自然语言处理能力,但如何应用于可拓策略生成仍需要深入研究^[6]。

本文提出在设计可拓策略生成软件时利用 HowNet 技术增强软件的语义理解能力,学习 HowNet 的概念组织关系改进可拓策略生成软件的基础知识库结构。另一方面,设计可拓策略生成软件时加入人机界面的智能引导,与基础知识库配合,利用人机交互方法,引导人输入计算机需要了解的信息,发挥人和计算机各自的优势,提高可拓策略生成的速度。

1 软件结构设计

现有的可拓策略生成系统^[2]的主要功能模块与流程如图 1 所示,主要功能模块包括:基础数据库、可拓规则库、问题库、可拓变换库和可拓策略库等。笔者认为,作为支持策略生成的基础知识库,其功能模块应该包括基础数据库、可拓规则库和可拓变换库,而问题库和可拓策略库则与具体问题有关,不属于基础知识。

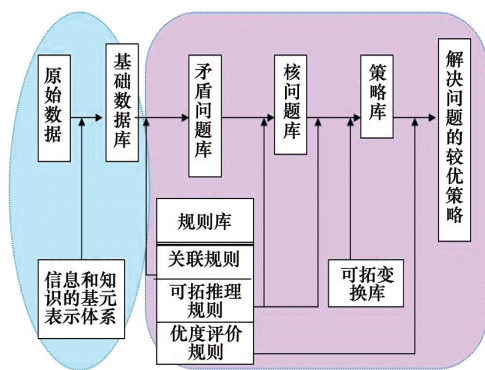


图 1 现有可拓策略生成系统模块与流程

Fig. 1 Current extension strategy generating system modules and flow diagram

由于计算机理解自然语言的能力不足,难以建立用户的矛盾问题模型。比如用户输入“我想去南沙”,计算机不清楚用户说的是“南沙群岛”还是“南沙区”。目前已研制的可拓

策略生成软件主要用人机界面分别直接输入矛盾问题的目标与条件,避免计算机理解错误而无法建立正确的问题模型。例如有关文献的自助游问题^[3],其输入界面直接要求用户输入可以旅游的时间和消费金额数值(条件),在 4 种指定的不同类型的旅游景点中选择一两种景点(目标),计算机并不涉及自然语言理解过程,这样自助游问题模型才可以顺利建立。然而这种限制性很大的输入问题方式使得可拓策略生成软件缺乏灵活性,不适合解决更广泛的矛盾问题。

为了改进现有可拓策略生成软件的不足,可借助 HowNet 的语义处理技术。从 HowNet 的知识结构可以看出,HowNet 描述了下列各种关系^[5]:

上下位关系:由概念的主要特征体现,参看“知网管理工具”。

同义关系,反义关系,对义关系:参看“同义、反义以及对义组的形成”。

部件-整体关系:由在整体前标注%体现,如“心”、“CPU”等。

属性-宿主关系:由在宿主前标注&体现,如“颜色”、“速度”等。

材料-成品关系:由在成品前标注?体现,如“布”、“面粉”等。

施事/经验者/关系主体-事件关系:由在事件前标注*体现,如“医生”、“雇主”等。

受事/内容/领属物等-事件关系:由在事件前标注\$体现,如“患者”、“雇员”等。

工具-事件关系:由在事件前标注*体现,如“手表”、“计算机”等。

场所-事件关系:由在事件前标注@体现,如“银行”、“医院”等。

时间-事件关系:由在事件前标注@体现,如“假日”、“孕期”等。

值-属性关系:直接标注无须借助标识符,如“蓝”、“慢”等。

实体-值关系:直接标注无须借助标识符,如“矮子”、“傻瓜”等。

事件-角色关系:由加角色名体现,如“购物”、“盗墓”等。

相关关系:由在相关概念前标注#体现,如“谷物”、“煤田”等。

利用 HowNet 作为基础资源,来建设本研究所需要的二次资源^[7]。比如在原有的可拓策略生成系统的信息和知识的基元表示体系中引入 HowNet 的各种关系,改进可拓策略生成软件基础数据库结构,以利于解决语义问题。例如对下面的物元 M_2 引入 HowNet 的上下位关系,使计算机很快理解“低价客户”是“客户”的子类。

$$M_1 = \begin{bmatrix} \text{客户}A, & \text{当前价值}, & v_1 \\ & \text{潜在价值}, & v_2 \end{bmatrix}$$

$$M_2 = \begin{bmatrix} \text{低价值客户,} & \text{上位,} & \text{客户} \\ & \text{当前价值,} & \text{低} \\ & \text{潜在价值,} & \text{低} \end{bmatrix}$$

对客户价值相关概念的基元表示体系及其结合 HowNet 的改进,可参看有关文献^[6]。

另一方面,增加智能引导软件模块,在 HowNet 知识组织的支持下,通过人机交互克服计算机理解自然语言的困难。一般性的人机对话是决策支持系统的基本功能^[8],但计算机的决策支持能力与系统中的知识有很大关系。比如用户输入“我想去南沙”,在 HowNet 的义原中南沙和南沙群岛的意义是相同的:

```
NO.=061494
W_C=南沙
G_C=N
E_C=
W_E=the Nansha Islands
G_E=N
E_E=
DEF=placel地方,#waters|水域,ProperName|专,(Chinal中国)
```

```
NO.=061495
W_C=南沙群岛
G_C=N
E_C=
W_E=the Nansha Islands
G_E=N
E_E=
DEF=placel地方,#waters|水域,ProperName|专,(Chinal中国)
```

而本研究的基础数据库则采用不同上下位属性的基元来区分“南沙”,有上位属性是广州市的南沙记录,还有上位属性是三沙市的南沙记录,智能引导软件模块根据基础数据库会引导用户回答“您是去三沙市南沙(群岛)还是广州市南沙(区)”,当用户确定是广州市南沙区后,计算机很快理解了用户的目标。否则的话,计算机只能认为用户去南沙群岛。

结合 HowNet 的可拓策略生成软件的主要功能模块与流程如图 2 所示。

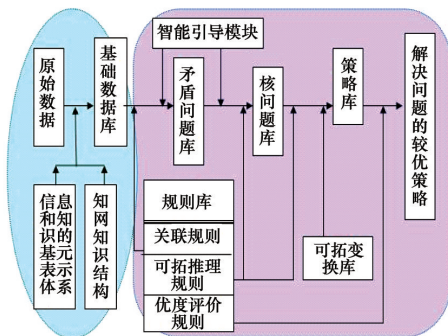


图 2 结合 HowNet 的可拓策略生成系统模块与流程

Fig. 2 Modules and flow diagram of the extension strategy generating system combined with HowNet

基础数据库按 HowNet 知识结构整合后更容易支持自然语言理解,智能引导模块可以协助建立矛盾问题模型以及核问题模型,而后者一直是可拓策略生成过程的难点。

2 软件功能设计

企业很关心如何提高客户价值的问题,但没有很好地提高客户价值支持工具。客户价值即客户给企业所带来的收益^[9],客户价值的高低决定着品牌企业的持续竞争力^[10],客户价值的相关概念可参看有关文献^[11]。为了帮助企业提高客户价值,使低价值客户、潜价值客户和次价值客户向价值客户转变,在客户信息和客户数据的采集基础上设计软件功能。

2.1 知识存储

为了存储客户价值领域知识^[6]和查询领域知识库,以利于生成提高客户价值的方案,采用关系数据库 MySQL 对客户价值领域知识进行存储。部分存储结构如表 1 所示。

表 1 客户价值知识存储结构

Table 1 Storage structure of customer value knowledge

id	customer	CCV	PCV	Upper
1	High value customer1	high	high	customer
2	Low value customer2	low	low	customer
3	potential value customer3	low	high	customer
4	inferior value customer4	high	low	customer

在存储了领域知识库后,如果希望将低价值客户向价值客户转变,则从知识库检索得知,可以提高其当前价值和潜在价值。检查当前价值的特征发现有毛利润、购买量、服务成本等,检查潜在价值的特征发现有信用度、忠诚度、信任度等。因此企业可以发掘该类客户的购买潜力,提升客户购买量,而且提高其忠诚度等,从而提高客户价值。

2.2 功能模块

用于提高客户价值的可拓策略生成软件包含了图 2 中的主要功能模块。

基础数据库是策略生成的数据准备,本文的基础数据库是企业提供的客户信息,包括购买数量、品牌忠诚度等,如图 3 所示。

客户名称	会员号	性别	年龄	职业	支付额度	购买数量	单位产品支付	品牌忠诚
客户43	43	女	23	学生	800	7	200	5
客户44	44	男	28	会计	1500	5	280	4
客户45	45	女	25	公务员	1800	3	400	4
客户46	46	男	21	职员	950	2	215	2
客户47	47	女	23	医生	1200	6	210	5
客户48	48	男	23	教师	1200	8	410	5
客户49	49	女	24	经理	2500	10	210	5
客户50	50	男	61	退休人员	360	5	180	4
客户51	51	女	25	教师	2200	2	255	2
客户52	52	男	29	秘书	1300	1	300	1
客户53	53	女	22	服务员	800	2	180	1
客户54	54	男	23	司机	1500	3	200	2
客户55	55	女	22	医生	1800	6	300	4
客户56	56	男	32	教师	350	1	185	1
客户57	57	女	21	经理	1200	10	228	5
客户58	58	男	63	退休人员	1200	5	188	4
客户59	59	女	25	教师	2500	7	300	5
客户60	60	男	23	秘书	360	10	278	5
客户61	61	女	23	服务员	2200	5	180	4
客户62	62	男	22	司机	1300	2	185	1
客户63	63	女	23	学生	800	2	180	1

图 3 基础数据库

Fig. 3 Basic database

可拓规则库包括可拓推理规则、关联函数公式与优度评价规则等,本文采用Java程序Relationfunction类提供关联函数公式,部分代码如下:

```

public class Relationfunction{
    //初始化所有变量
    public Relationfunction(){
    public float relationfunction_K3(float x3){
        float temp;
        if (x3==5) temp=1;
        else if (x3==4) temp=(float) 0.5;
        else if (x3==3) temp=0;
        else if (x3==2) temp= (float)-0.5;
        else temp=-1;
        return temp;
    }
    //建立客户建筑的综合关联函数
    //x1,x2为每个客户“单位产品支付d1”、“购买数量d2”的值,其他为最大最小值
    public float comprehensiverelationfunction(float x1, float x2, float rs, int l1, int b1, int l2, int b2){
        float k=0;
        float k1=0;
        float k2=0;
        float rs1=rs/100;//变为0-1的实数
        int low1=l1;
        int big1=b1;
        if ((big1-low1)!=0) k2=(x2-3)/(big1-low1);
        else k2=(x2-3)/9;//3是企业定的值,
        int low2=l2;
        int big2=b2;
        if ((big2-low2)!=0) k1=(x1-300)/(big2-low2);
        else k1=(x1-300)/570;//300是企业定的值
        k=(float) (rs1*k1+(1-rs1)*k2);//这是用户输入的比例
        return k;
    }
}

```

可拓变换是生成策略的基本技术,本文采用可拓变换库存放可以提高客户价值的营销手段,如图4所示。

序号	矛盾问题	策略	预计提升效果	见效时间
1	品牌忠诚度	提供良好的咨询服务	5%	6个月
2	品牌忠诚度	提供包装送货服务	5%	3个月
3	品牌忠诚度	建立客户导向的企业文化	20%	12个月
4	品牌忠诚度	主动回访与沟通	10%	6个月
5	品牌忠诚度	妥善处理投诉	10%	6个月
6	单位产品支付	提供分期付款	10%	3个月
7	购买数量	买2送1	10%	1个月

图4 可拓变换库

Fig. 4 Extension transformation base

2.3 人机界面智能引导

为了减轻计算机处理自然语言的困难,设计了智能引导模块与用户进行交互,获取用户信息语义,在出现语义歧义时引导用户进一步说明含义,不必艰难地进行自然语言理解,同时尽量引导用户输入数值而不是文字。例如,询问用户希望提高客户价值的百分比,用户只要输入0~100的数字即可,不必输入文字。同理,确定关联函数、选择希望提升的客户群、优度评价、询问比重因子等,都是只需输入数字。

3 测试结果

采用某服装企业的客户信息数据^[12]来测试所研制的可拓策略生成软件。

首先软件导入客户数据(Excel表)并导入数据库,如图5所示。



图5 导入客户数据

Fig. 5 Software to introduce customer data

然后确定关联函数,软件自动查出关联函数的几个上下限,用户只需输入比例因子。根据关联函数计算出客户的分类,如图6所示。

客户ID	购买数量	单位产品支付	品牌忠诚度	当前价值	潜在价值	客户分类
客户1	3	267	3	-0.02847368	0.0	提升型客户
客户2	2	750	2	0.3391813	-0.5	次升级客户
客户3	10	180	5	0.2832672	1.0	升级客户
客户4	1	350	1	-0.04726146	-1.0	低升级客户
客户5	3	400	3	0.04771893	0.0	升级客户
客户6	2	450	3	0.0760224	0.0	升级客户
客户7	5	200	4	0.023891813	0.5	升级客户
客户8	2	180	1	-0.16881871	-1.0	低升级客户
客户9	4	700	5	0.40643275	1.0	升级客户
客户10	2	500	3	0.13883046	0.0	升级客户
客户11	2	190	2	-0.16204678	-0.5	低升级客户
客户12	2	490	3	0.11111111	0.0	升级客户
客户13	2	214	2	-0.1309416	-0.5	低升级客户
客户14	1	185	1	-0.2118983	-1.0	低升级客户
客户15	1	190	1	-0.20780294	-1.0	低升级客户

图6 确定客户分类

Fig. 6 Customer classification

当用户在智能引导模块提示下输入希望提高的客户价值比例、选择希望提升的客户群后,软件根据可拓变换给出满足目标的策略,如图7所示。



检查结果:	
客户ID	变换方法
客户1	将其单位产品支付从267提高到300,就提升为价值客户
客户16	将其单位产品支付从222提高到300,就提升为价值客户
客户21	将其单位产品支付从267提高到300,就提升为价值客户
客户36	将其单位产品支付从222提高到300,就提升为价值客户
客户2	将其品牌忠诚度提高到3,就提升为价值客户
客户22	将其品牌忠诚度提高到3,就提升为价值客户
客户4	将其品牌忠诚度提高到3,且将其购买数量提高到3和将其单位产品支付提高到300,就提升为价值客户

选择评价因子

图7 变换结果

Fig. 7 Transformation results

最后是优度评价,软件按优度大小从高到低提供给用户,如图8所示。



评价结果:		
客户名称	可拓变换	优度
客户1	将其单位产品支付从267提高到300,就提升为价值客户	0.839
客户21	将其单位产品支付从267提高到300,就提升为价值客户	0.839
客户16	将其单位产品支付从222提高到300,就提升为价值客户	0.8243333
客户36	将其单位产品支付从222提高到300,就提升为价值客户	0.8243333
客户2	将其品牌忠诚度提高到3,就提升为价值客户	0.75
客户22	将其品牌忠诚度提高到3,就提升为价值客户	0.75
客户4	将其品牌忠诚度提高到3,且将其购买数量提高到3和将其单位产品支付提高到300,就提升为价值客户	0.3666667

首页

图8 评价结果

Fig. 8 Evaluation results

从测试结果可以看出,可拓策略生成软件有效地满足了服装企业提高客户价值的目标。

4 结论

设计了一种结合 HowNet 的可拓策略生成软件,通过学习 HowNet 的概念组织关系改进可拓策略生成系统的基础知识库结构,提高了系统生成解决策略的能力。另一方面,该软件的研制实践为利用计算机技术协助生成可拓策略提供了参考。

参考文献(References)

- [1] 杨春燕,蔡文.可拓学[M].北京:科学出版社,2014.
Yang Chunyan, Cai Wen. Extenics[M]. Beijing: Science Press, 2014.
- [2] 李立希,杨春燕,李锐汶.可拓策略生成系统[M].北京:科学出版社,2006.
Li Lixi, Yang Chunyan, Li Huawen. Extension strategy generating system[M]. Beijing: Science Press, 2006.
- [3] 方卓君,李卫华,李承晓.自助游可拓策略生成系统的研究与实现[J].广东工业大学学报,2009,26(2):83-89.
Fang Zhuojun, Li Weihua, Li Chengxiao. Research and realization of extension strategy generating system for independent travel[J]. Journal of Guangdong University of Technology, 2009, 26(2): 83-89.

- [4] 李承晓,李卫华.租房可拓策略生成系统[J].智能系统学报,2011,6(3):272-278.
Li Chengxiao, Li Weihua. Research on tenement extension strategy generating system[J]. CAAI Transactions on Intelligent Systems, 2011, 6(3): 272-278.
- [5] 董振东,董强.知网[EB/OL].[2014-08-25].http://www.keenage.com/zhiwang/c_zh_iwang.html.
Dong Zhendong, Dong Qiang. HowNet[EB/OL].[2014-08-25].http://www.keenage.com/zhiwang/c_zh_iwang.html.
- [6] 李卫华.利用知网增强可拓策略生成机制研究[J].广东工业大学学报,2013,30(2):1-6.
Li Weihua. Research on taking advantage of the HowNet to enhance mechanisms of extension strategy generation[J]. Journal of Guangdong University of Technology, 2013, 30(2): 1-6.
- [7] 董振东,董强,郝长伶.知网的理论发现[J].中文信息学报,2007,21(4):3-9.
Dong Zhendong, Dong Qiang, Hao Changling. Theoretical findings of HowNet[J]. Journal of Chinese Information Processing, 2007, 21(4): 3-9.
- [8] 陈文伟,廖建文.决策支持系统及其开发[M].第3版.北京:清华大学出版社,2008.
Chen Wehwei, Liao Jianwen. Decision support system and development [M]. 3rd ed. Beijing: Tsinghua University Press, 2008.
- [9] 陈静宇.基于客户价值的中间商客户细分与客户管理策略研究[D].重庆:重庆大学,2004.
Chen Jingyu. Study on customer segmentation and customer management strategy of distributor based on customer value[D]. Chongqing: Chongqing University, 2004.
- [10] 陈桂玲.品牌企业客户价值的保持和提升研究[J].商场现代化,2009(10):62-63.
Chen Guiling. Research on maintain and improve research of brand enterprise customer value[J]. Market Modernization, 2009, (10): 62-63.
- [11] 权明富,齐佳音,舒华英.客户价值评价指标体系设计[J].南开管理评论,2004,7(3):17-23.
Quan Mingfu, Qi Jiayin, Shu Huaying. An evaluation index system to assess customer value[J]. Nankai Business Review, 2004, 7(3): 17-23.
- [12] 杨春燕,李小妹,陈文伟,等.可拓数据挖掘方法及其计算机实现[M].广州:广东高等教育出版社,2010.
Yang Chunyan, Li Xiaomei, Chen Wenwei, et al. Extension data-mining method and computer implement[M]. Guangzhou: Guangdong Higher Education Press, 2010.

(责任编辑 韩星明)

《科技导报》征集“封面文章”

为快速反映中国最新科技研究成果,《科技导报》拟利用刊物最显著位置——封面将最新科研成果第一时间予以突出报道。来稿要求:研究成果具创新性或新颖性;反映该领域中国乃至世界前沿研究水平;可以图片形式予以反映,图片美观、清晰、分辨率超过300dpi;文章篇幅不限,要说明研究的背景、方法、取得的结果,以及结论。在线投稿:www.kjdb.org。