

doi:10.3969/j.issn.1001-2400.2013.05.027

云模型中正向云算法的修正

李 庆, 董庆宽, 赵 蕾

(西安电子科技大学 综合业务网理论及关键技术国家重点实验室, 陕西 西安 710071)

摘要: 云模型中现有正向云算法存在一些与人类认知相悖的图像特征, 使用正向云算法产生的点(云滴)也具有误差, 直接影响了应用的准确性. 针对这一问题对正向云算法做出修正, 对不同概念点处的认知歧义加以限制, 并给出了计算各点认知歧义值的函数. 对修正的方案对比和仿真, 分析结果表明, 修正后的算法更满足人们认知, 能实现对定性概念更准确的刻画.

关键词: 云模型; 正向云算法; 信任管理; 模糊理论; 人工智能

中图分类号: TN918.91 **文献标识码:** A **文章编号:** 1001-2400(2013)05-0169-06

Modified forward cloud generator in the cloud model

LI Qing, DONG Qingkuan, ZHAO Lei

(State Key Lab. of Integrated Service Networks, Xidian Univ., Xi'an 710071, China)

Abstract: Some image features of the cloud model created by the forward cloud generator available contradict human cognition, which influences the accuracy of applications directly. For these contradictions, modifications to the forward cloud generator are made in this paper. A function for calculating the cognitive ambiguity value of each sample point in concept is given. Comparing two algorithms, simulation results show that the cloud model created by the modified forward cloud generator is capable of satisfying human cognition and can describe the natural concept more precisely.

Key Words: cloud model; forward cloud generator; trust management; fuzzy logic; artificial Intelligence

人类语言不可避免具有不确定性. 不确定性主要是模糊性和随机性. 如何处理和描述不确定性是人工智能领域的研究热点之一. 概率统计很好地描述了事物的随机性; 模糊理论则针对了事物的模糊性. 云模型是在概率理论和模糊理论相结合的基础上, 实现对定性概念不确定性的描述, 形成了定性概念与定量模型之间转换的模型.

云模型^[1-2]于1995年由李德毅院士提出, 并提出了正向正态云发生器(后简称正向云算法)和逆向正态云发生器(后简称逆向云算法). 随后, 正态云模型的普适性^[3]被证明, 使得很多领域引入了正态云模型来描述概念的不确定性. 目前主要有正向云算法、逆向云算法和多维多规则云^[4]. 多维多规则云是在正向云算法的基础上延伸出来的.

云模型主要用于信任管理^[5-8]、算法改进^[9]和数据挖掘^[10]等领域, 尤其在信任管理中, 它能够兼顾“信任”的主观性、模糊性和随机性, 使得信任的描述更加准确, 是其他方案无法达到的. 现有基于云模型的信任模型多使用逆向云算法^[11-12]. 逆向云算法稳定性差, 同一组数据通过多次求出的云参数都不同且相差较大. 此外, 逆向云算法需要收集大量的数据, 实时性低, 复杂度高, 影响了云模型实际应用的效果和可操作性.

使用正向云算法设计新的信任模型, 可以避免逆向云算法的使用. 正向云算法的稳定性和复杂度都优于

收稿日期: 2013-03-15 网络出版时间: 2013-06-06

基金项目: 国家自然科学基金资助项目(61373172); 西安电子科技大学昆山创新研究院2013年度科技创新计划资助项目(201314); 中央高校基本科研业务费资助项目(K50511010009)

作者简介: 李庆(1989-), 女, 西安电子科技大学硕士研究生, E-mail: yake_y@aliyun.com.

网络出版地址: http://www.cnki.net/kcms/detail/61.1076.TN.20130606.0923.201305.208_024.html

逆向云算法. 但现有正向云算法实现的图像的一些特性与人类认知相悖. 笔者指出了这些问题, 并给出了出现这些问题的原因. 在此基础上对现有正向云算法做出修正, 通过对比, 证明修正算法实现的图像更满足人们认知, 对定性概念的刻画更准确.

1 云模型的概述

正态云模型(简称云模型或云), 是对定性概念的数学建模, 即通过分析定性概念的规律及特点, 实现对概念在数学上的描述, 具有随机性和模糊性. 使用云模型对概念“20 km 左右”的描述如图 1 所示.

横坐标是指定性概念的样本值, 也是论域空间的数值点; 纵坐标是该样本点属于概念的隶属度. 云模型的核心表现在 3 个数字特征上:

Ex 称作期望, 是论域空间中最能够代表定性概念的点, 是这个概念最典型样本点.

En 称作熵, 表现了定性概念的范围, 衡量定性概念的模糊度与随机性. 满足概率论中“ $3En$ 准则”.

He 称作超熵, 表示熵的不确定性度量, 反应云厚度变化规律的方差的最大值 σ_{\max} , 是表示人类认知歧义的物理量, 表现在图中就是云图中点的凝聚度, 也就是云层的厚度.

云模型服从的 3 个正态分布^[2]是它能够满足拟合人类认知的根本:

(1) 反映模糊概念亦此亦彼性的云期望曲线(正态分布的曲线 $N(Ex, En)$)服从正态分布规律; 其物理意义在于人们对事物的认知一般符合正态分布或者近似于正态分布.

(2) 反映隶属度随机性的云上的点服从以云期望曲线上相应点为期望的正态分布. 即不同人对同一点的认识程度不同, 某点 x 的不同隶属度满足以隶属云期望曲线上的相应点为期望, 以相应点的认知歧义值为方差的正态分布.

(3) 反映云厚度变化规律的方差, 即各点的认知歧义值近似服从两个半正态分布规律. 这两个半正态分布有 5 个特殊点: $(Ex, 0)$ 、 $(Ex \pm 3En, 0)$ 、 $(Ex \pm (\ln 8)^{1/2} En, He)$. 特殊点表示在均值处和远离概念点处, 人们的认知并没有歧义, 云层的厚度为 0; 在云覆盖范围下的面积的形心 ($u = 2^{1/2}/4$) 所对应的点 $x = Ex \pm (\ln 8)^{1/2} En$ 处的认知歧义最大^[3], 云层的厚度最大.

云参数的物理意义和云模型满足的 3 个正态分布是正向云算法实现的主要理论基础, 也是本文对正向云算法做出修正的依据.

尽管 He 是实现云模型的最主要的参数, 但以往关于云模型的论文中对 He 并不重视. 且认为 He 是垂直于云期望曲线切线方向上的距离. 这个定义使得 He 在物理意义上难以相符, 在数值上难于度量.

He 是一个具有实际物理意义的参数, 其大小理应可以计算和度量. 依据 He 的物理意义以及其满足的第三个正态分布, 将 He 重新定义. 定义 He 的方向在 y 轴上, 数值为各点隶属度最大值和最小值的差值的最大值. 这样, 不同点处差值不同, 表现每个 x 点概念的认知歧义不同, 并且实现可以度量.

云模型主要有正向云算法和逆向云算法, 本文主要关注正向云算法. 正向云算法是已知概念的云参数 (Ex, En, He) 生成 N 个云滴. 在已知概念的云参数的基础上, 多次模拟实现人们对这一参数的认知, 形成云模型的过程.

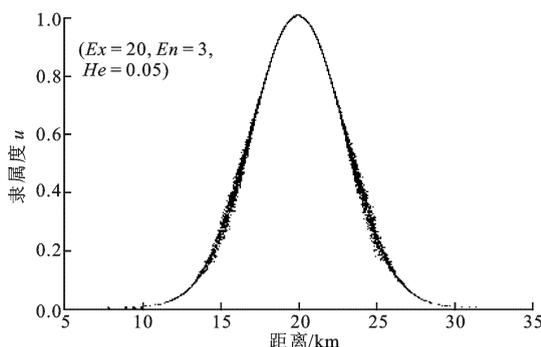


图 1 表示“20 km 左右”的云模型

2 正向云算法及其问题分析

2.1 正向云算法概述

正向云算法是已知的概念云参数, 重复实现多次, 得到云模型的过程. 每次实现都是对现实概念样本点

的一次认知. 目前常使用的正向云算法只有一种, 描述如下:

输入: (Ex, En, He) 和 N ;

输出: N 个云滴 (x, u) ;

Step1 以 En 为均值, He 为方差, 随机产生 En' 的值;

Step2 以 Ex 为均值, En' 为方差, 随机产生 x 的值;

Step3 以 Ex 为均值, En' 为方差, 代入 x , 计算 x 对应的 u ;

Step4 重复上述步骤, 直到产生 N 个点 (x, u) .

在 Matlab 7.0 环境下对上述算法仿真, 实现参数为 $Ex = 20, En = 3, He = 0.1$ 和 $Ex = 20, En = 3, He = 0.5$ 的云模型, 得到的图形如图 2 所示. 绘制不同 He 的两个云图是为了后文中的对比分析. 其中两条点划线分别是均值为 Ex 、方差分别为 $En + 3He$ 和 $En - 3He$ 的正态曲线. 实线是均值为 Ex 、方差为 En 的云期望曲线. 这 3 条曲线是为了后文对比分析而绘制的.

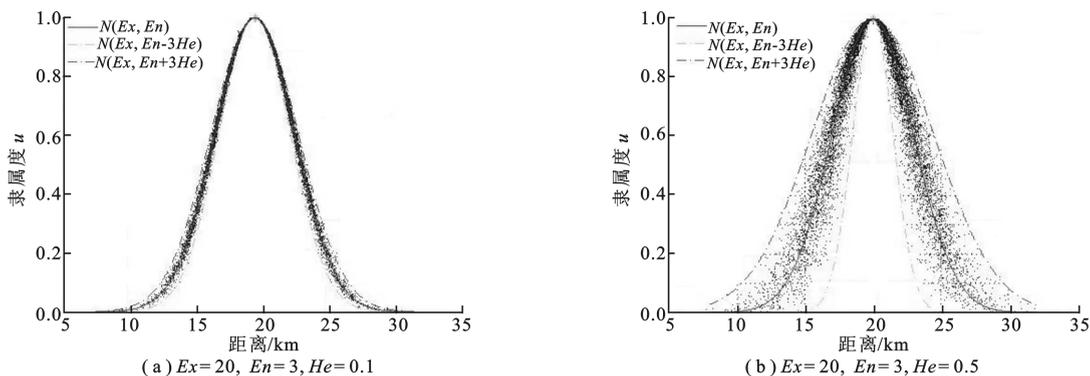


图 2 现有的正向云算法形成的云模型

2.2 现有正向云算法存在的问题

对正向云算法的分析从现有正向云算法的仿真入手. 由图 2 可得, 现有正向云算法存在的问题主要表现为云层厚度(各点的云层厚度为该点处隶属度最大值与最小值之差, 形象起见, 称为云层厚度)在远离期望值处没有收敛(收敛是指云滴紧密分布在期望曲线附近, 是针对认知歧义而提出的), 在物理意义上表示该处仍然存在认知歧义. 一般情况下, 远离概念典型值处, 人们都会没有歧义地认为这些值不属于这个概念. 同时, 尽管云层厚度在 Ex 处为 0, 但在 Ex 的邻域内, 云层厚度较大. 在物理意义上表示 Ex 的邻域内认知歧义较大. 人类认知中, 在一个概念的典型值附近, 一般没有歧义地认为这个样本点属于该概念. He 较大时这个问题较显著, 如图 2(b) 所示.

绘制云层厚度曲线如图 3 所示, 分别对应图 2 中的两图. 图中五角星形点标记了 He 的位置, 即理论上的歧义最大值的位置, 菱形点标记了理论上认知歧义收敛为零的位置.

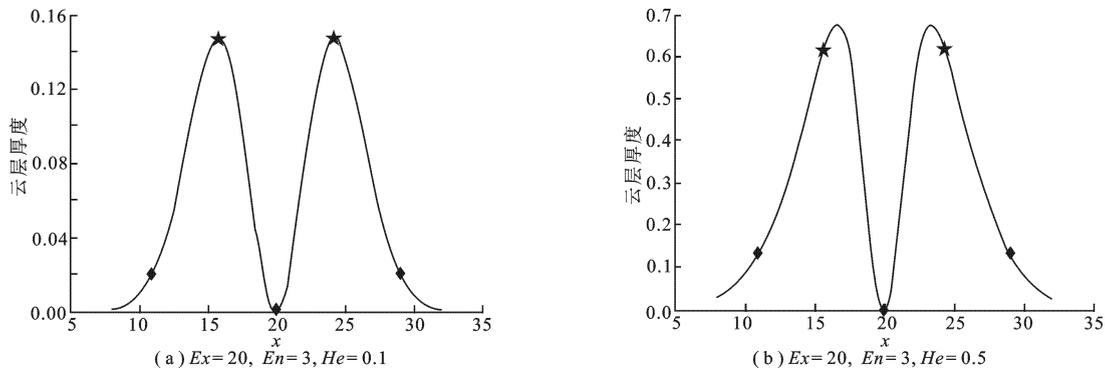


图 3 现有的正向云算法形成的云模型云层厚度曲线

云层厚度曲线上, 纵坐标表示概念点的认知歧义值. 菱形点处的 y 值证实了云图在远离期望点处无法收敛. 当 He 较小时, 该值接近 0, 随着 He 的增大, 该值也逐渐偏离 0. $Ex = 20$ 附近的云层厚度曲线增加幅度剧

烈,证实了 E_x 附近认知歧义没有收敛.增加的幅度随着 He 增加而增加.

此外,还可得到两个结论:一是由最大值的实际位置和五角星形点可见,现有算法歧义最大点的位置存在偏移.这个偏移随着 He 的增大而增大.二是云模型的认知歧义的最大值并不是 He .随着 He 的增加,认知歧义最大值与 He 的偏差增加.

这些问题存在的原因是现有正向云算法事实上是一系列以 E_x 为均值,以 $(E_n - 3He, E_n + 3He)$ 上的值为方差的正态曲线上的点,如图 4 所示.图 4 中, Enn 指不同正态曲线的方差, Enn 的变化体现了方差的变化.

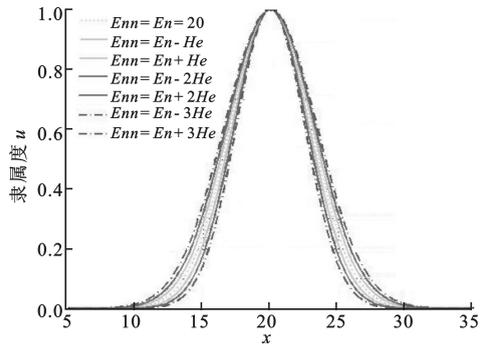


图 4 正态曲线组图

在 E_x 点的收敛为 0 是因为这组曲线都以 E_x 为期望,在 E_x 点处的取值都是最大值 1. He 较小时,在期望值附近和远离期望值处收敛较好是因为曲线 $N(E_x, E_n - 3He)$ 和 $N(E_x, E_n + 3He)$ 距离较近; He 较大时,曲线 $N(E_x, E_n - 3He)$ 和 $N(E_x, E_n + 3He)$ 距离较远,收敛变差.可见,云层厚度由 He 决定. He 不再是歧义最大值,而是影响认知歧义最大值的因素. He 发生变化时,歧义最大值的位置和大小都发生变化.而人类社会,语言概念的认知歧义只与概念本身有关,与认知歧义的最大值没有关系.

这些问题的本质原因是在设计算法的过程中,只考虑了云模型所满足的 3 个正态分布中的前两个而没有考虑第 3 个.这导致算法中的 He 成为了每个样本点认知歧义的通用值而非定义中认知歧义的最大值,使得云图出现了在远离期望值处和期望较小邻域内的无法收敛、特殊点的位置发生变化、特殊点的值和云层厚度曲线随着 He 变化而变化等现象,与人们认知相矛盾.这会导致在应用中随机产生的点的隶属度有较大的偏差.

3 正向云修正算法的描述及算法分析

3.1 正向云修正算法的描述

根据上述分析,对现有算法做出修正,主要考虑了认知歧义满足的第 3 个正态分布和 He 的物理意义,算法如下:

输入: (E_x, E_n, He) 和 N ;

输出: N 个云滴 (x, u) ;

Step1 以 E_x 为均值, E_n 为方差,随机产生样本点 x ;

Step2 构造云层厚度变化规律中满足的两个半正态分布函数

$$f_1(x) = He \exp\left(-\frac{(x - E_{x_1})^2}{2E_{n_1}^2}\right), \quad f_2(x) = He \exp\left(-\frac{(x - E_{x_2})^2}{2E_{n_2}^2}\right),$$

其中, $E_{x_1} = 0, E_{n_1} = (3 - (\ln 8)^{1/2})E_n/4; E_{x_2} = 0, E_{n_2} = (\ln 8)^{1/2}E_n/4$.

Step3 计算云层厚度 He' 的函数,该函数根据认知歧义的变化规律划分为 4 段,分别满足两个半正态分布函数, x 点对应的 He'

$$He' = He'(x) = \begin{cases} f_1((E_x - (\ln 8)^{1/2}E_n) - x), & x < (E_x - (\ln 8)^{1/2}E_n) \\ f_2(x - (E_x - (\ln 8)^{1/2}E_n)), & (E_x - (\ln 8)^{1/2}E_n) \leq x < E_x \\ f_2(x - (E_x + (\ln 8)^{1/2}E_n)), & E_x \leq x < (E_x + (\ln 8)^{1/2}E_n) \\ f_1((E_x + (\ln 8)^{1/2}E_n) - x), & x \geq (E_x + (\ln 8)^{1/2}E_n) \end{cases}$$

Step4 以 E_n 为均值, He' 为方差,随机产生 $\text{rand } E_n$,用于计算 x 点的隶属度;

Step5 计算 x 点的隶属度 $u \quad u = u(x) = \exp\left(-\frac{(x - E_x)^2}{2 \text{rand } E_n^2}\right)$.

Step6 重复上述步骤,直到产生 N 个点 (x, u) .

3.2 正向云修正算法的分析

正向云修正算法的核心就是根据云模型性质中认知歧义值的变化规律,实现对每个样本点处的认知歧

义值的计算. 主要体现在 Step2 和 Step3 上.

概念的认知歧义满足^[2]: 由最典型样本点(即 Ex) 向坐标轴两边, 认知歧义逐渐增加, 分别在 $Ex \pm (\ln 8)^{1/2} En$ 处达到最大后逐渐减小, 分别在 $Ex \pm 3En$ 处减小为 0. 其中, 增加和减小满足正态的变化规律. 对应到样本值区间上, 当 $x \in [Ex - 3En, Ex - (\ln 8)^{1/2} En]$ 时, 云层厚度满足 f_1 左半部分, 逐渐增加; $x \in [Ex - (\ln 8)^{1/2} En, Ex]$ 时满足 f_2 右半部分, 逐渐下降; $x \in [Ex, Ex + (\ln 8)^{1/2} En]$ 时满足 f_2 的左半部分, 逐渐增加; $x \in [Ex + (\ln 8)^{1/2} En, Ex + 3En]$ 时满足 f_1 的右半部分, 逐渐下降.

Step2 描述了变化规律中的 f_1 和 f_2 . 在 Step2 中, 两个正态分布乘以 He , 使得 f_1 和 f_2 的值域为 $[0, He]$. $Ex_1 = Ex_2 = 0$, 是因为这样取值简单, 通过坐标平移即可以实现满足要求的正态分布. En_1 为 $(3 - (\ln 8)^{1/2})En/4$ 是因为这段云层厚度增加的区间长度为 $|(Ex - 3En) - (Ex - (\ln 8)^{1/2} En)| = (3 - (\ln 8)^{1/2})En$. 对于任意正态分布, $x = Ex + 4En$ 时, y 值近似为 0; 对于 $f_1, x = Ex$ 时, $y = He$. 因此, $En_1 = (3 - (\ln 8)^{1/2})En/4$, 能够实现云层厚度由 0 到最大值 He 的变化. En_2 的取值原因同 En_1 .

Step3 根据上述的规律, 分别代入各段满足的正态分布 f_1 和 f_2 , 确定计算 He' 的分段函数, 实现对样本点认知歧义值的计算. Step3 通过限制 He' 的值, 使得随机生成的大部分 $\text{rand} En$ 都属于区间 $(En - 3He', En + 3He')$, 保证认知歧义点在 Ex 的邻域和远离 Ex 处, 认知歧义收敛.

3.3 正向云修正算法的仿真分析

在 Matlab 7.0 环境下对修正算法仿真, 同样实现 $Ex = 20, En = 3, He = 0.1$ 和 $Ex = 20, En = 3, He = 0.5$ 时的云模型如图 5 所示; 云层厚度曲线如图 6 所示.

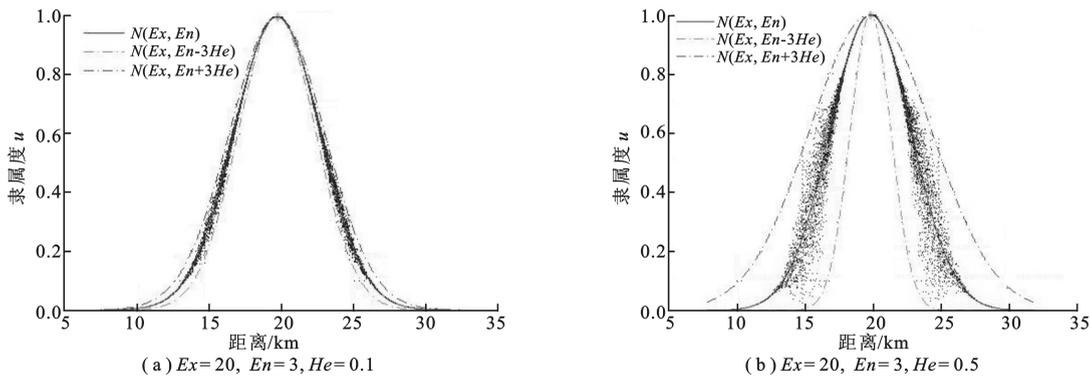


图 5 修正的正向云算法形成的云模型

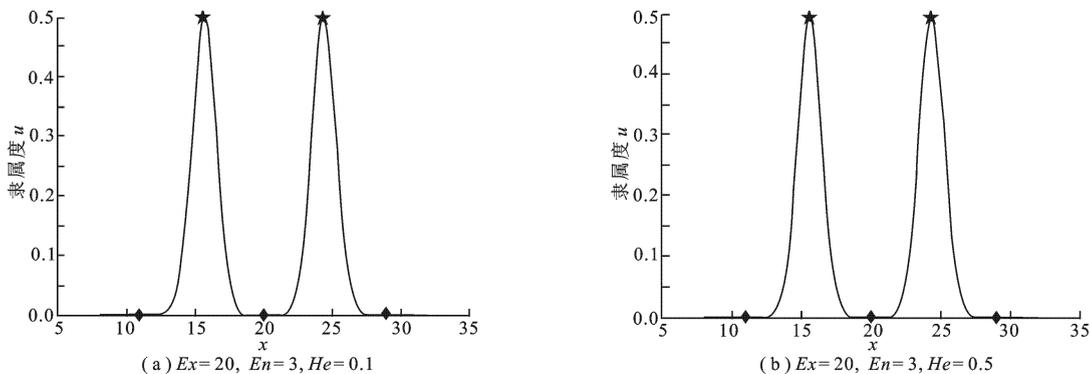


图 6 修正的正向云算法形成的云模型的云层厚度曲线

可以看出, 正向云修正算法修正了现有算法存在的 3 个问题:

(1) 云图在期望较小邻域内和远离期望处都能够实现收敛, 不再离散. 这表示在期望较小邻域内和远离期望处认知没有歧义. 云层厚度曲线中, $x = Ex = 20$ 附近曲线变化趋势和菱形点的值证明了这两处没有认知歧义的结论;

(2) 比较五角星形点与云层厚度曲线的峰值可得: 修正算法中, 云图的 He 在位置和大小上都满足理论

要求.

(3) 对比图 6 中两图,可以看出云层厚度曲线满足正态分布,并不随着 He 的变化而变化. 云层厚度曲线上菱形点和五角星形点的位置也没有随着 He 变化. 这符合人类社会中的语言概念.

此外,在图 6(b)中, He 偏大时,云图中部分点严重偏离期望曲线,云图出现了异常. 这表现了该概念在模棱两可的论域区中认知歧义较大,但并没有影响其他点,不与人类认知矛盾.

为了更好地说明修改算法的效果,分别使用现有正向云算法和笔者提出的正向云修正算法产生文献[5]中描述的“青年”这一概念在样本点 $x = 17$ 和 $x = 25$ 处的云滴,如表 1(各产生 10 个点)所示,其中, $Ex = 26$, $En = 3$, $He = 0.03$.

$x = 25$ 时,现有正向云算法生成的 10 个点的隶属度变化范围为 $[0.9471, 0.9441]$,正向云修正算法生成的 10 个点的隶属度变化范围为 $[0.9459, 0.9460]$. $x = 17$ 时,现有算法的变化范围为 $[0.0095, 0.0119]$,而修正算法保持为 0.0111.

表 1 正向云算法中的云滴

算法类型	样本点 x	隶属度 u									
现有正向云算法	17	0.0107	0.0113	0.0097	0.0112	0.0118	0.0114	0.0119	0.0103	0.0095	0.0113
	25	0.9446	0.9441	0.9471	0.9447	0.9458	0.9453	0.9451	0.9462	0.9452	0.9455
正向云修正算法	17	0.0111	0.0111	0.0111	0.0111	0.0111	0.0111	0.0111	0.0111	0.0111	0.0111
	25	0.9460	0.9459	0.9460	0.9460	0.9460	0.9460	0.9460	0.9460	0.9460	0.9460

由此可以看出,在期望附近和远离期望处,修正算法的认知歧义能够收敛. 综上,可以得到:修正后的正向云算法更加符合人们对定性概念的认知.

4 结束语

正向云算法发展至今主要使用的只有一种,如 2.1 节所示. 现有正向云算法在远离期望值处和期望较小邻域内无法收敛,特殊点的位置发生变化,特殊点的值和云层厚度曲线随着 He 变化而变化. 在 He 较大时,这些问题尤为明显.

笔者以云模型满足的 3 个正态分布为依据,从 He 的物理意义出发,结合对现有算法的分析,指出了 He 为认知歧义的最大值而非各点认知歧义通用值这一问题,对现有算法做出修正. 将现有算法中忽略的 He 满足的两个半正态分布加入,得到了正向云修正算法. 对修正算法仿真,并与现有算法的仿真结果对比可得,修正后的算法避免了现有算法中无法收敛的问题,更符合人们的认知,能够更准确地刻画定性概念.

参考文献:

- [1] Li D, Cheung D, Shi X M, et al. Uncertainty Reasoning Based on Cloud Models in Controllers [J]. Computers and Mathematics with Applications, 1998, 35(3): 99-123.
- [2] 李德毅, 孟海军, 史雪梅. 隶属云和隶属云发生器 [J]. 计算机研究和发展, 1995, 32(6): 15-20.
Li Deyi, Meng Haijun, Shi Xuemei. Membership Cloud and Membership Cloud Generators [J]. Journal of Computer Research and Development, 1995, 32(6): 15-20.
- [3] 李德毅, 刘常昱. 论正态云模型的普适性 [J]. 中国工程科学, 2004, 6(8): 28-34.
Li Deyi, Liu Changyu. Study on the Universality of the Normal Cloud Model [J]. Engineering Science, 2004, 6(8): 28-34.
- [4] 杨朝晖, 李德毅. 二维云模型及其在预测中的应用 [J]. 计算机学报, 1998, 21(11): 961-969.
Yang Zhaohui, Li Deyi. Planar Model and Its Application in Prediction [J]. Chinese Journal of Computer, 1998, 21(11): 961-969.
- [5] 田玉敏, 马勇. P2P 网络下基于联合投票的成员资格协议 [J]. 西安电子科技大学学报, 2007, 34(1): 110-114.
Tian Yumin, Ma Yong. A Collaborative Membership Protocol Based on Joint Votes for P2P network [J]. Journal of Xidian University, 2007, 34(1): 110-114.

(下转第 180 页)