

不确定数据上的 Reverse Top- k 查询

王晓伟 贾 焰

(国防科技大学计算机学院 长沙 410073)

摘 要 传统的 top- k 查询为顾客返回符合其偏好的产品集合, reverse top- k 查询则返回将给定产品作为 top- k 结果的偏好集合。reverse top- k 查询由于能帮助生产者评估产品对顾客的影响, 因此在商业分析中具有重要价值。现有的 reverse top- k 查询假设数据是精确的, 许多现实应用中, 数据的不确定性广泛存在。将 reverse top- k 查询扩展到不确定数据上, 并给出了基于物化视图的高效查询算法 GMV。实验结果表明, GMV 算法能够减少需要计算的偏好数量, 具有较高的计算效率。

关键词 不确定数据, 偏好, reverse top- k 查询, 物化视图

中图法分类号 TP311.13 **文献标识码** A

Reverse Top- k Query on Uncertain Data

WANG Xiao-wei JIA Yan

(School of Computer Science, National University of Defense Technology, Changsha 410073, China)

Abstract Traditional top- k query returns the products to customers according to their preferences, whereas reverse top- k query returns the preferences for which a given product is in the top- k result. Reverse top- k query is valuable in business analysis because it can help the manufacturers evaluate the impact of a product on customers. Existing reverse top- k query assumes the underlying data is certain, however, uncertainty arises in many real applications. In this paper, we extended reverse top- k query to uncertain data, and provided an efficient query algorithm named GMV based on materialized views. Our experimental evaluation demonstrates that, GMV can reduce the preferences which need to be computed, and achieves relatively high computational efficiency.

Keywords Uncertain data, Preference, Reverse top- k query, Materialized view

1 引言

随着网络应用的日益个性化, 根据用户行为研究用户偏好成为商业分析的迫切需求。top- k 查询^[1]在数据集 D 中部分属性 $\{A_1, \dots, A_n\}$ 上定义一个聚集函数 $f(A_1, \dots, A_n)$ 来表示用户偏好, 为顾客返回 $f(\cdot)$ 值最高的 k 个对象。reverse top- k 查询^[2]则返回以给定对象为 top- k 结果的聚集函数集合, 用于研究用户偏好。top- k 查询是面向顾客的, 查询结果是符合某个顾客偏好的产品。reverse top- k 查询则是面向生产者的, 查询结果是偏爱某个产品的顾客偏好集合。

现有的 reverse top- k 查询假定数据是精确的。由于仪器精度、数据冲突、环境干扰、隐私保护等因素, 因此数据的不确定性普遍存在。例如, 在房产交易网站中, 每套房产的信誉度或推荐度往往是顾客是否购买该房产所需要考虑的重要因素, 可将其视为该房产是否应该存在于数据库中的概率值。为了将最受欢迎的房产推荐给用户, 经纪人需要了解每套房产的可能用户群体。这类应用可以建模为不确定数据上的 reverse top- k 查询, 即给定不确定数据集 U 、用户偏好集 W , 返回以 U 中某个对象为 top- k 结果的概率大于某个阈值的偏

好。

本文首先提出了不确定数据上的 reverse top- k 查询的定义, 其次研究了查询优化算法。朴素的查询算法需要扫描每个偏好, 并对每个偏好计算给定对象的 top- k 概率, 计算代价较大。提出了基于物化视图的改进算法, 预先计算网格顶点的查询结果, 后续查询可以利用物化结果避免对偏好集合的全盘扫描。在真实数据集和模拟数据集上进行的实验验证了本算法的有效性。

本文第 2 节介绍背景知识和相关工作; 第 3 节引入不确定数据上 reverse top- k 查询的定义; 第 4 节提出基于物化视图的查询算法; 第 5 节分析实验结果; 最后总结全文。

2 背景知识与相关工作

2.1 背景知识

数据的不确定性广泛存在, 如何有效地管理不确定数据成为近年来数据库领域的研究热点。不确定数据上的查询普遍采用可能世界语义^[3], 即每个不确定数据库中的关系包含了其中不确定对象的所有合法组合, 每个组合称为一个可能世界, 每个可能世界都是一个传统关系。如图 1 所示, 假设有

到稿日期: 2011-03-30 返修日期: 2011-06-04 本文受“863”国家高技术研究发展计划(2007AA010502, 2007AA01Z474)资助。

王晓伟(1980-), 男, 博士生, 主要研究方向为海量数据管理与挖掘, E-mail: gfkdw@yaho.com.cn; 贾 焰(1961-), 女, 博士, 教授, 主要研究方向为网络与信息安全、数据库与数据挖掘等。

包含 2 个不确定对象的关系 U , 则 U 的可能世界包含 4 个关系: $PW(U) = \{pw_1, pw_2, pw_3, pw_4\}$ 。通常假设对象之间的存在相互独立, 则 pw 的概率 $\Pr(pw) = \prod_{o \in pw} p(o) \times \prod_{o' \notin pw} (1 - p(o'))$ 。首先介绍确定数据上 top- k 查询^[1]和 reverse top- k 查询^[2]的定义。

定义 1(确定数据上的 top- k 查询) 给定偏好 w 、正整数 k , 确定数据集 D 上的 top- k 查询 $Top_k^D(w)$ 定义为

$$Top_k^D(w) = \{o \in D \mid Rank_w^D(o) < k\}$$

$$Rank_w^D(o) = |\{o' \in D \mid f_w(o') > f_w(o)\}|$$

$Rank_w^D(o)$ 表示 o 根据偏好 w 在 D 中的序号, 即聚集函数值大于 o 的对象个数, 可知 $Rank_w^D(o) \in [0, |D| - 1]$ 。

$Top_k^D(w)$ 查询返回序号小于 k 的所有对象。

定义 2(确定数据上的 reverse top- k 查询) 给定偏好集合 W 、正整数 k 、确定数据集 D 中的对象 o , reverse top- k 查询 $RTOP_k(o)$ 定义为

$$RTOP_k(o) = \{w \in W \mid o \in Top_k^D(w)\}$$

$RTOP_k(o)$ 返回以给定对象 o 为 top- k 查询结果的偏好集合。若忽略图 1 所示的不确定数据集 U 中对象的概率, 并将其视为确定数据集, 则 $RTOP_1(o_1) = \{w_1\}$, $RTOP_1(o_2) = \{w_2\}$ 。

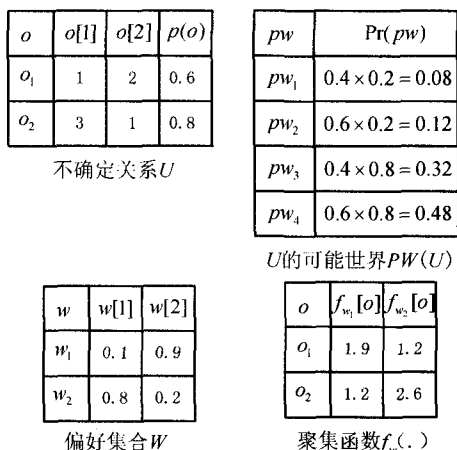


图 1 不确定数据集 U 的可能世界 $PW(U)$

2.2 相关工作

如何从海量数据中提取最有价值的信息, 一直是数据库领域的重要挑战, 关系数据库中 top- k 查询因此得到了深入研究^[1]。最近 Vlachou^[2] 研究了 reverse top- k 查询问题, 但是他们假设数据是确定的, 无法将其应用于不确定数据上。Skyline^[4] 是另外一种偏好查询。Lian^[5] 首先提出不确定数据上的 reverse skyline 查询, Wu^[6] 研究了确定数据上的 reverse skyline 优化方法。但是由于 skyline 和 top- k 查询语义的差别, 两者的技术无法直接借鉴。

近年来, 对不确定数据的管理和挖掘^[7,8] 已成为数据管理领域的研究热点之一。由于 top- k 查询的重要性, 使得不确定数据上的 top- k 查询得到了广泛关注。不确定数据的排序既要考虑聚集值, 又要考虑数据的概率值。结合这两个维度进行数据排序的不同方式就产生了不同的 top- k 查询语义, 比较重要的有 U-Topk^[9]、UkRanks^[9]、PT- k ^[10]、Global-Topk^[11]、Expected Rank^[12] 等。U-Topk 返回概率最大的 top- k 结果。UkRanks 则返回在前 k 个序号上概率最大的对象。PT- k 定义为阈值形式, 返回属于 top- k 结果的概率超过阈值的所有对象。Global-Topk 返回属于 top- k 结果的概率最高

的 k 个对象。每种查询语义都有一定的应用背景。什么样的查询语义才最符合 top- k 查询的本质需求? Cormode^[12] 提出不确定数据上的 top- k 查询应该满足 5 个特性, 并提出了满足这些特性的 Expected Rank 语义。Li^[13] 从多维优化的角度出发, 提出了一种统一的排序函数来近似上述排序语义。该函数能够处理有复杂关联的数据集, 并且可以通过对用户数据的学习来得到函数的参数。综上所述, 不确定数据上 top- k 查询的研究都从用户角度出发, 对不确定数据上的 reverse top- k 查询尚属空白。

3 问题定义

假设有不确定关系 U , U 中任意对象 o 表示为 $\langle o[1], \dots, o[d] \rangle$, o 的存在是不确定的, 用概率 $p(o)$ 表示。沿用文献^[9-11] 的合理假定, 假设对象之间的存在是相互独立的。另有用户偏好集合 W , W 中任意偏好 w 表示为权值向量 $\langle w[1], \dots, w[d] \rangle$, $w[i]$ 表示 w 在第 i 维上的权重值。本文考虑最常见的线性聚集函数 $f_w(\cdot)$, 对象 o 的聚集函数值 $f_w(o)$ 为 $\sum_{i=1}^d w[i] \times o[i]$ 。

给定对象 o , 任意偏好 w 以一定的概率令 $o \in Top_k(w)$ 。考虑到现实应用中通常忽略概率值较低的结果, 定义概率 reverse top- k 查询如下。

定义 3(不确定数据上的 reverse top- k 查询 $RTOP_k^U(o)$)

给定参考对象 o 、参数 k 和阈值 $\sigma (\sigma \in [0, 1])$, $RTOP_k^U(o)$ 返回以 o 为 top- k 查询结果的概率大于 σ 的偏好, 即 $RTOP_k^U(o) = \{w \mid \Pr(o \in Top_k^U(w)) > \sigma\}$, 其中 $\Pr(o \in Top_k^U(w)) = \sum_{pw \in PW(U) \wedge o \in Top_k^{pw}(w)} \Pr(pw)$ 。

$Top_k^{pw}(w)$ 表示在可能世界 pw 上的使用定义 1 的确定数据 top- k 查询结果。简便起见, 下文中概率 $\Pr(o \in Top_k^U(w))$ 均简写为 $\Pr_w^U(o)$ 。根据定义 3, 在图 1 中所示的数据集中, $\Pr_{w_1}^{U1}(o_1) = \Pr(pw_2) + \Pr(pw_4) = 0.6$, $\Pr_{w_1}^{U1}(o_2) = \Pr(pw_3) = 0.32$ 。同理, $\Pr_{w_2}^{U2}(o_1) = 0.12$, $\Pr_{w_2}^{U2}(o_2) = 0.8$ 。

对偏好集合 $W = \{w_1, w_2\}$, 假设需要计算 $RTOP_1^{U1}(o_1)$, 由于 $\Pr(o_1 \in Top_1(w_1)) = 0.6 > 0.5$, 而 $\Pr(o_1 \in Top_1(w_2)) = 0.12 < 0.5$, 故 $RTOP_1^{U1}(o_1) = \{w_1\}$ 。同理可得 $RTOP_1^{U2}(o_2) = \{w_2\}$ 。

4 概率 reverse top- k 查询算法

根据定义 3, 可以得到一个基本的查询算法 Naive。如算法 1 所示, 对 W 中的每个偏好 w 计算 $\Pr_w^U(o)$, 并将其与 σ 比较, 来判断 w 是否属于 $RTOP_k^U(o)$ 。

算法 1 基本的查询算法 Naive

Input: W ; the preference set; o ; the input object;

σ : threshold of probability; k : number of objects returned

Output: the preference set W'

1. $W' = \emptyset$;
2. for each $w \in W$
3. if $(\Pr_w^U(o) > \sigma)$
4. $W' = W' \cup \{w\}$;
5. end if
6. end for
7. return W' ;

由于 $\Pr_w^U(o)$ 表示对象 o 属于 top- k 结果的概率, 因此可以直接使用文献^[10] 给出的算法进行计算, 即首先将 U 中对

象按照 $f_w(\cdot)$ 的值从大到小排序,再依此顺序进行增量式计算,故 $\text{Pr}_k^w(o)$ 的计算复杂性为 $O(k \times N_o)$,其中 N_o 表示 U 中 $f_w(\cdot)$ 值大于 $f_w(o)$ 的对象数量。由于需要对 W 中的每个偏好进行概率计算,因此基本算法的计算复杂性为 $O(k \times N_o \times |W|)$,对偏好集合规模的可扩展性差。

4.1 基于物化视图的查询算法

为了提高计算效率,需要避免对每个偏好 w 计算 $\text{Pr}_k^w(o)$ 。受文献[2]启发,提出基于物化视图的方法,通过物化视图来提高查询的 I/O 效率。基本思想是将数据空间划分为网格,预先计算网格顶点 c 的 $\text{RTOP}_k^w(c)$ 结果。对某个对象 o 执行 $\text{RTOP}_k^w(o)$ 查询时,只需要考虑 o 落入的网格 c 。顶点的查询结果所限定范围内的偏好。为了阐述方便,首先给出对象间支配关系的概念。

定义 4(支配关系) $\forall o, o' \in D, o > o'$ 当且仅当 $\forall i \in [1, d], o[i] \leq o'[i]$, 且 $\exists j \in [1, d]$, 使得 $o[j] \neq o'[j]$ 。

定理 1 $\forall o_1, o_2 \in U$, 若 $o_1 > o_2$, 且 $p(o_1) \leq p(o_2)$, 则 $\forall w, \text{Pr}_k^w(o_1) \leq \text{Pr}_k^w(o_2)$ 。

证明:记 DS_o 为 U 中聚集值大于 $f_w(o)$ 的对象集合,即 $DS_o = \{o' \in D | f_w(o') > f_w(o)\}$ 。若 $o_1 > o_2$, 则由定义 4 知, $\forall w, f_w(o_1) < f_w(o_2)$, 故 $DS_{o_1} \supseteq DS_{o_2}$ 。由定义 3, $\forall o \in U, \forall w, \text{Pr}(o \in \text{Top}_k^w(w)) = p(o) \prod_{o' \in DS_o} (1 - p(o'))$ 。又 $p(o_1) < p(o_2)$, 故 $\text{Pr}(o_1 \in \text{Top}_k^{DS_{o_1}}(w)) \leq \text{Pr}(o_2 \in \text{Top}_k^{DS_{o_2}}(w))$ 。故 $\text{Pr}(o_1 \in \text{Top}_k^w(w)) \leq \text{Pr}(o_2 \in \text{Top}_k^w(w))$, 从而有 $\text{Pr}_k^w(o_1) \leq \text{Pr}_k^w(o_2)$ 。

对象与所在网格的两个顶点之间存在支配关系。如图 2 所示,对象 o 落入的网格记为 c_o , c_o^+ 表示 c_o 的右上角顶点, c_o^- 表示 c_o 的左下角顶点,由定义 4 得 $c_o^- > o > c_o^+$, 由此可以得到定理 2。

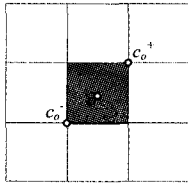


图 2 对象与网格顶点之间的支配关系

定理 2 若 $p(c_o^-) \leq p(o) \leq p(c_o^+)$, 则以下公式成立:

$$\text{RTOP}_k^w(c_o^-) \subseteq \text{RTOP}_k^w(o) \subseteq \text{RTOP}_k^w(c_o^+)$$

证明:由于 $o > c^+$ 且 $p(o) \leq p(c_o^+)$, 则由定理 1 可知, $\forall w \in W$, 有 $\text{Pr}_k^w(o) \leq \text{Pr}_k^w(c_o^+)$ 。再由定义 2 知, $\forall w \in \text{RTOP}_k^w(o)$, 有 $\text{Pr}_k^w(c_o^+) \geq \text{Pr}_k^w(o) > \sigma$, 故 $w \in \text{RTOP}_k^w(c_o^+)$, 即 $\text{RTOP}_k^w(o) \subseteq \text{RTOP}_k^w(c_o^+)$ 。同理可得 $\text{RTOP}_k^w(c_o^-) \subseteq \text{RTOP}_k^w(o)$ 。证毕。

记 $W_o = \text{RTOP}_k^w(c_o^+) - \text{RTOP}_k^w(c_o^-)$ 。通过定理 2, 对落入网格 c_o 的对象 o , 需要计算的 w 限定在 W_o 范围内。算法 2 给出了基于物化视图的查询算法 GM(Grid-based Materialized)。对每个对象 o , 首先计算 W 的子集 W_o 中的每个偏好 w 的概率 $\text{Top}_k^w(w)$, 将符合概率阈值的 w 加入结果集合 W' 。算法最后将必定属于 W' 的 $\text{RTOP}_k^w(c_o^-)$ 加入最终的查询结果集。

算法 2 基于物化视图的算法 GM

Input: W : the preference set; o : the input object;

σ : threshold of probability; k : number of objects returned

Output: the preference set W'

1. $W' = \emptyset$;
2. for each $w \in W_o$.

3. if $(\text{Pr}(o \in \text{Top}_k^w(w)) > \sigma)$
4. $W' = W' \cup \{w\}$;
5. end if
6. end for
7. $W' = W' \cup \text{RTOP}_k^w(c_o^-)$;
8. return W' ;

4.2 生成基于网格的物化视图

前一节给出了基于物化视图的查询算法,本节将介绍对数据进行网格划分和生成物化视图的方法。对给定的不确定数据集 U , 采用与文献[2]相同的方法进行网格划分,即从覆盖整个数据空间的单网格开始,递归地划分查询代价最高的网格。网格查询 c 的代价 $\text{Cost}(c)$ 定义为

$$\text{Cost}(c) = (|c^-| - |c^+|) \int_{o \in c} f(o)$$

所不同的是,对网格顶点执行物化查询时,需要将顶点作为不确定对象,赋予其概率值。注意到定理 2 成立的前提是 $p(c_o^-) \leq p(o) \leq p(c_o^+)$ 。为此,对网格顶点执行物化查询时,顶点 c_o^- 和 c_o^+ 的概率应做如下设置

$$p(c_o^-) = \min\{p(o') | o' \in c_o\}$$

$$p(c_o^+) = \max\{p(o') | o' \in c_o\}$$

这样就可以保证每个落入网格中的对象都能够利用顶点的查询结果。

5 实验

5.1 实验设置

实验采用 CPU 2.0G, 内存 2G 的 PC 机, 在 Windows XP 上用 C++ 语言编写试验程序, 编译运行环境为 VC 2005。不确定对象集 U 采用模拟的方法生成, 每个维度上的值服从均匀分布, 各维度上的值相互独立。为 U 中每条数据随机赋予服从正态分布 $N(\mu, \epsilon)$ 的概率。偏好集合 W 中的每个维度上的权值 w_i 服从均匀分布。每次实验均匀、随机地生成 $\text{RTOP}_k^w(o)$ 的输入对象 o , 实验结果取 100 次实验的平均值。除非特别说明, 参数的默认设置如表 1 所列。

表 1 参数的默认值

参数	W	D	k	σ	d	μ	ϵ
默认值	10k	10k	10	0.3	5	0.3	0.3

5.2 实验结果

首先比较不同偏好规模 $|W|$ 下, GMV 与 Naive 需要计算 $\text{Pr}_k^w(o)$ 的次数。如图 3 所示, Naive 算法需要的计算次数等于 $|W|$, 而 GMV 需要的计算次数远小于 Naive 算法, 这得益于物化视图带来的剪枝效果。GMV 随着 $|W|$ 的增加其扩展性也更好, 随着 $|W|$ 从 2k 增加到 10k 时, GMV 所需要的计算次数分别为 Naive 的 5.1%, 4.6%, 4.5%, 4.3%, 4.1%, 说明偏好集规模越大, 其剪枝效果越明显。

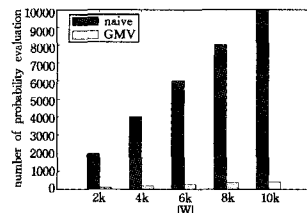


图 3 GMV 相对于 $|W|$ 的扩展性

图 4 比较两种算法在不同的不确定数据集规模 $|U|$ 下所需要的计算次数。Naive 算法需要的概率计算次数与 $|U|$ 无

关,始终为 $|W|=10000$ 。而 GMV 随着 $|U|$ 的增加,需要的计算次数单调减少,这是由于不确定数据集规模增大,网格顶点的 top- k 概率均会变小,其 reverse top- k 的结果集合也相应变小,从而 GMV 需要进行的概率计算次数减少。随着 $|U|$ 从 2k 增加到 10k, GMV 需要的概率计算次数分别为 Naive 的 17.4%, 9.25%, 7.83%, 4.72%, 4.1%, 说明 U 的规模越大, GMV 优势越明显。

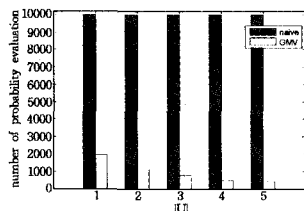


图 4 GMV 相对于 $|U|$ 的扩展性

结束语 本文研究了不确定数据上的 reverse top- k 查询,提出了基于物化视图的查询优化算法 GMV。实验结果表明,GMV 算法对数据集规模有较好的可扩展性。

参考文献

- [1] Ilyas I F, Beskales G, Soliman M A. A survey of top- k query processing techniques in relational database systems[J]. ACM Computer Survey, 2008, 40(4): 1-58
- [2] Vlachou A, Doukeridis C, Kotidis Y, et al. Reverse Top- k Queries[C]//Proc. of ICDE. 2010: 365-376
- [3] Abiteboul S, Kanellakis P, Grahne G. On the Representation and

Querying of Sets of Possible Worlds[J]. Theoretical Computer Science, 1991, 78(1): 34-48

- [4] Borzsony S, Kossmann D, Stocker K. The skyline operator[C]//Proc. of ICDE. 2001: 421-430
- [5] Lian X, Chen L. Monochromatic and bichromatic reverse skyline search over uncertain databases[C]//Proc. of SIGMOD. 2008: 213-226
- [6] Wu X, Tao Y, Wong R C, et al. Finding the Influence Set through Skylines[C]//Proc. of EDBT. 2009: 1030-1041
- [7] Aggarwal C C, Yu P S. A Survey of Uncertain Data Algorithms and Applications[J]. IEEE Transactions on Knowledge and Data Engineering, 2009, 21(5): 609-623
- [8] 周傲英, 金澈清, 王国仁, 等. 不确定性数据管理技术研究综述[J]. 计算机学报, 2009, 32(1): 1-16
- [9] Soliman M A, Ilyas I F, Chang K C-C. Top- k query processing in uncertain databases[C]//Proc. of ICDE. 2007: 896-905
- [10] Hua M, Pei J, Zhang W, et al. Ranking Queries on Uncertain Data: A Probabilistic Threshold Approach[C]//Proc. of SIGMOD. 2008: 673-686
- [11] Zhang X, Chomicki J. On the Semantics and Evaluation of Top- k Queries in Probabilistic Databases[C]//Proc. of DBRank. 2008: 556-563
- [12] Cormode G, Li F, Yi K. Semantics of Ranking Queries for Probabilistic Data and Expected Ranks[C]//Proc. of ICDE. 2009: 305-316
- [13] Li J, Saha B, Deshpande A. A Unified Approach to Ranking in Probabilistic Databases[C]//Proc. of VLDB. 2009: 502-513

(上接第 182 页)

证明方法,即通过共归纳扩展射将两个不同的元素映射到终结共代数中的同一状态上。这不仅可以降低系统对行为等价关系的证明难度,也便于借助形式化验证工具进行证明。而强 Monads 可进一步提高共代数在面向对象方法的形式化描述中的抽象性。

在下一步的工作中,我们将继续探讨不同类规范下的对象行为等价关系及其行为精化语义。

参考文献

- [1] Papathomas M. Language Design Rationale and Semantic Framework for Concurrent Object-oriented Programming [D]. Dept. of Computer Science, University of Geneva, 1992
- [2] Hameurlain N, Sibertin-Blanc C. Behavioural Types in CoOperative Objects [C]//Moreira A M D, Demeyer S, eds. Object-Oriented Technology, ECOOP 99 Workshop Reader, ECOOP 99 Workshops, Panels, and Posters, Lisbon. Volume 1743 of Lecture Notes in Computer Science, Springer, 1999: 320-321
- [3] Peter C, Puntigam F. A Concurrent Object Calculus with Types that Express Sequences [C]//Proceedings of the ECOOP Workshop on Semantics of Objects as Processes (SOAP'99). Lisbon, Portugal, 1999: 321
- [4] Goel A, Roychoudhury A, Thiagarajan P S. Interacting Process Classes [J]. ACM Transactions on Software Engineering Methodology, 2009, 18(4): 1-47
- [5] Burstall R, Diaconescu R. Hiding and Behaviour: an Institutional Approach [M]//Roscoe A W, ed. A Classical Mind. Essays in honour of C. A. R. Hoare, 1994: 75-92
- [6] Berrima M, Rajeb N B. Linking Algebraic Observational Equiva-

lence and Bisimulation [C]//Lecture Notes in Computer Science, Volume 6224. 2010: 76-87

- [7] Rutten J. Universal Coalgebra: a Theory of Systems [J]. Theoretical Computer Science, 2000: 3-80
- [8] 周晓聪, 舒忠梅. 计算机科学中的共代数方法的研究综述[J]. 软件学报, 2003, 4(10): 1661-1671
- [9] Reichel H. An Approach to Object Semantics Based on Terminal Coalgebras [J]. Mathematical Structures in Computer Science, 1995, 5(2): 129-152
- [10] Jacobs B. Objects and Classes, Co-algebraically [M]//Freitag B, Jones C B, Lengauer C, eds. Object-orientation with Parallelism and Persistence Kluwer Acad. Publ., 1996: 83-103
- [11] Jacobs B, Rutten J. Inheritance and Cofree Construction [C]//Cointe P, ed. European Conference on Object-oriented Programming, Lecture Notes in Computer Science, Vol. 1098. Springer, Berlin, 1996: 210-231
- [12] Jacobs B. Invariants, Bisimulations and the Correctness of Coalgebraic Refinements [C]//Proceedings of the 6th International Conference on Algebraic Methodology and Software Technology. Lecture Notes In Computer Science, Vol. 1349. 1997: 276-291
- [13] Jacobs B. Introduction to Coalgebra: Towards Mathematics of States and Observations [M]. Book Draft. <http://www.cs.ru.nl/~bart>, 2005
- [14] Moggi E. Notions of Computation and Monads [J]. Inf. & Comp., 1991, 93(1): 55-92
- [15] Tews H. The Coalgebraic Class Specification Language CCSL-Syntax and Semantics [R]. TU Dresden technical report TUD-FI02-08-August, 2002
- [16] Jacobs B, Rutten J. A Tutorial on (Co)Algebras and (Co) Induction [J]. EATCS Bulletin, 1997, 62: 222-259