

具有分形特性的背景序列及其概念格*

马 垣⁺, 刘占刚

辽宁科技大学 软件学院, 辽宁 鞍山 114051

Context Series and Its Concept Lattice with Fractal Properties*

MA Yuan⁺, LIU Zhangang

School of Software Engineering, University of Science and Technology Liaoning, Anshan, Liaoning 114051, China

+ Corresponding author: E-mail: mayuanas@sina.com

MA Yuan, LIU Zhangang. Context series and its concept lattice with fractal properties. Journal of Frontiers of Computer Science and Technology, 2009, 3(5):460-468.

Abstract: The association of fractal and concept lattice is firstly proposed in both domestic and international, a series of formal context is proposed which is corresponding to a graphic series of unit of square. The limitation of these graphic series is a fractal graph. The formalize method of representing fractal context is provided. Concept lattice series computing method of the fractal context series is researched too. It is proved that the number of concept in the concept lattice, which is the limitation of the fractal context series, has Hausdorff property, and then a new dimension is proposed for the graph of the limitation: Concept fraction dimension, which is used to show the more data catachrestic of the associations of the two big academic domains.

Key words: fractal; concept lattice; context series; fractal dimension; concept fractal dimension

摘 要: 在国内外首次研究分形与概念格这两大领域之间的联系,提出了一种形式背景序列,这个形式背景序列将对应一个单位正方形中的图形序列,而这个图形序列的极限是一个分形图形。给出了分形背景序列的形式化表示方法,研究了这种分形背景序列的概念格序列的计算方法。证明了这个分形背景序列的极限概念格中的概念个数具有 Hausdorff 特性;对相应的极限图形定义了一个全新的维数:概念分数维,展示了两大领域联系的更多数据表征。

* The National Natural Science Foundation of China under Grant No.06775036 (国家自然科学基金); the Natural Science Foundation of Liaoning Province of China under Grant No.20082189 (辽宁省自然科学基金).

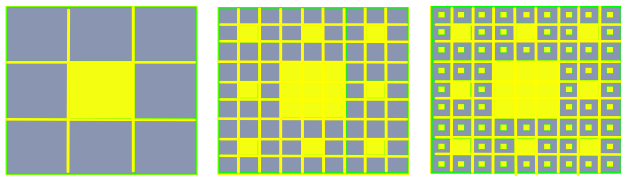
Received 2009-03, Accepted 2009-06.

关键词:分形;概念格;分形背景序列;分数维;概念分数维

文献标识码:A 中图分类号:TP18

1 引言

分形几何是 Mandelbrot 教授在 1982 年提出的^[1]。它的图形是具有以下 5 个特点的几何图形:(1)整体与局部自相似;(2)有无限细致的结构;(3)具有分数维;(4)可用简单的迭代方法生成;(5)度量与精度有关。例如 Sierpinski 地毯就是用以下方法迭代生成的分形图形:取一个单位正方形,把它分成 9 等份,去掉中间的一份(见图 1(a)),剩下的小正方形每个都再分成 9 等份,再去掉中间的一份(见图 1(b)),剩下的 64 个小正方形每个再分成 9 等份,再去掉中间的一份(见图 1(c)),如此下去,直到无穷,得到的就是 Sierpinski 地毯。



(a)The first graph (b)The second graph (c)The third graph
(a)第一个图形 (b)第二个图形 (c)第三个图形

Fig.1 The process generating Sierpinski carpet
图 1 Sierpinski 地毯的生成过程

分形几何提出的同一年,Wille 教授提出了形式概念的理论^[2],已应用到了很多领域^[3-14]。设 $U=\{u_1, \dots, u_p\}$ 是对象的集合, $M=\{m_1, \dots, m_q\}$ 是属性集合,而 $I \subseteq U \times M$ 是 U 与 M 之间的一个关系,则称三元组 $K=(U, M, I)$ 为一个形式背景(简称背景),给定背景 K ,对于每一个 U 的子集 α 及每一个 M 的子集 β ,都定义:

$$f(\alpha) = \{m \in M \mid \forall u \in \alpha, (u, m) \in I\}$$

$$g(\beta) = \{u \in U \mid \forall m \in \beta, (u, m) \in I\}$$

若 $\alpha = g(\beta), \beta = f(\alpha)$,则称二元组 (α, β) 为背景的一个概念,并称 α 为这个概念的外延, β 为这个概念的内涵。

文章提出了一种从任一个形式背景出发,利用迭

代方法生成一个形式背景序列,这个形式背景序列中的每一个背景都将对应一个单位正方形中的图形,从而形成一个图形序列,而这个图形序列的极限是个分形图形。给出了这种形式背景序列中各个背景概念格的求法。借助这些背景的概念格的序列,提出了一种分形图形的新的分数维:概念分数维。这种分数维是两大领域联系的一个重要数据表征。

本文的组织情况是:第 1 章是引言;第 2 章是形式背景序列;第 3 章是背景序列对应的图形;第 4 章是背景序列的概念格序列;第 5 章是分形图形的概念分数维;第 6 章是结论。

2 形式背景序列

由任何一个形式背景,都可用以下的迭代方法生成一个背景序列。先请看一个非形式化的实例。

例 1 设 $K=(U, M, I)$ 是一个形式背景。其中 $U=\{u_1, u_2\}, M=\{m_1, m_2, m_3\}, I$ 如表 1 所示。

Table 1 A formal context K
表 1 一个形式背景 K

I	m_1	m_2	m_3
u_1	×		×
u_2	×	×	

首先将 K 中 U 及 M 的元素都用从 0 开始的依次非负整数表示,即用 0 表示 $u_1, 1$ 表示 $u_2; 0$ 表示 $m_1, 1$ 表示 $m_2, 2$ 表示 m_3 ,则背景 K 被改写成背景 $K_1=(U_1, M_1, I_1)$,其中 $U_1=\{0, 1\}, M_1=\{0, 1, 2\}$,而 I_1 如表 2 所示。如果非形式的将 K_1 中的每个 \times 中都用 K 代替,则得到背景 $K_2=(U_2, M_2, I_2)$,其中的 U_2, M_2 仍用从 0 开始的依次非负整数表示, I_2 如表 3 所示。再把 K_2 中的每个 \times 都用 K 代替,则得到背景 $K_3=(U_3, M_3, I_3)$,其中 U_3, M_3 也仍用从 0 开始的依次非负整数表示。

如此下去就得到了一个由背景 K 生成的背景序列。
现在严格定义这种生成方法:

Table 2 A rewritten context K_1

表 2 一个改写后的背景 K_1

I_1	0	1	2
0	×		×
1	×	×	

Table 3 The replaced context K_2 for the first time

表 3 一次代替后的背景 K_2

I_2	0	1	2	3	4	5	6	7	8
0	×		×				×		×
1	×	×					×	×	
2	×		×	×		×			
3	×	×		×	×				

定义 1 (背景序列) 设 $K=(U, M, I)$ 是一个具有 p 个对象, q 个属性的形式背景。则由 K 生成的形式背景序列 $\{K_n\}_{n=1,2,\dots}$ 可递归定义如下:

(1) 将 K 中 U 的元素 u_1, u_2, \dots, u_p , 分别用 $0, 1, \dots, (p-1)$ 表示, M 的元素 m_1, m_2, \dots, m_q , 分别用 $0, 1, \dots, (q-1)$ 表示, 并记 $\{0, 1, \dots, (p-1)\}$ 为 U_1 , 记 $\{0, 1, \dots, (q-1)\}$ 为 M_1 , 令 $I_1 = \{(i, j) | i \in U_1, j \in M_1, (u_{i+1}, m_{j+1}) \in I\}$, 则定义 $K_1 = (U_1, M_1, I_1)$ 。

(2) 若 $K_n = (U_n, M_n, I_n)$ 已定义 ($n=1, 2, \dots$), 则定义 $K_{n+1} = (U_{n+1}, M_{n+1}, I_{n+1})$ 如下:

$$U_{n+1} = \{pu_n + u_1 | u_n \in U_n, u_1 \in U_1\}$$

$$M_{n+1} = \{qm_n + m_1 | m_n \in M_n, m_1 \in M_1\}$$

$$I_{n+1} = \{(pu_n + u_1, qm_n + m_1) \in U_{n+1} \times M_{n+1} |$$

$$(u_n, m_n) \in I_n \wedge (n_1, m_1) \in I_1\}$$

用 $\lfloor ab \rfloor$ 表示 a 除以 b 略去小数取整, $\langle ab \rangle$ 表示 a 除以 b 取 $0 \sim (b-1)$ 之间的余数。注意到 $u_1 \in \{0, 1, \dots, (p-1)\}$, $m_1 \in \{0, 1, \dots, (q-1)\}$, 所以 $\lfloor qu_n + u_1/p \rfloor = u_n$, $\langle qu_n + u_1/p \rangle = u_1$, $\lfloor (qm_n + m_1)/q \rfloor = m_n$, $\langle (qm_n + m_1)/q \rangle = m_1$, 所以 I_{n+1} 的定义等价于:

$$I_{n+1} = \{(i, j) \in U_{n+1} \times M_{n+1} | (\lfloor i/p \rfloor, \lfloor j/q \rfloor) \in$$

$$I_n \wedge (\langle i/p \rangle, \langle j/q \rangle) \in I_1\}$$

定理 1 设 $K_n = (U_n, M_n, I_n)$ 是定义 1 中生成背景序列的第 n 个, 则有以下 3 个性质:

(1) $U_n = \{0, 1, \dots, (p^n - 1)\}$;

(2) $M_n = \{0, 1, \dots, (q^n - 1)\}$;

(3) $(i, j) \in I_n$ 是当且仅当: i 写成 p 进制数 $i_{n-1}i_{n-2} \dots i_2i_1i_0$, j 写成 q 进制数 $j_{n-1}j_{n-2} \dots j_2j_1j_0$ 时, $(i_{n-1}, j_{n-1}) \in I_1$, $(i_{n-2}, j_{n-2}) \in I_1, \dots, (i_1, j_1) \in I_1, (i_0, j_0) \in I_1$ 。例如在上例中表 3 所示的 $K_3 = (U_3, M_3, I_3)$ 。 U_3 中 6 的二进制形式为 110, M_3 中 14 的三进制形式为 112, 而 $(1, 1) \in I_1, (1, 1) \in I_1, (2, 0) \in I_1$, 则 $(6, 14) \in I_3$ 。 U_3 中的 5 的二进制形式为 101, M_3 中的 4 的三进制形式为 011, 而 $(1, 0) \in I_1, (0, 1) \notin I_1, (1, 1) \in I_1$, 则 $(5, 4) \notin I_3$ 。

证明 (1) 首先 $U_n = \{0, 1, \dots, (p^n - 1)\}$ 对 $n-1$ 正确, 因为由定义 $U_1 = \{0, 1, \dots, (p-1)\}$ 。 现设对 $n=k$ 正确, 即 $U_k = \{0, 1, \dots, (p^k - 1)\}$, 则对于 $n=k+1$ 有:

$$U_{k+1} = \{pu_k + u_1 | u_k \in U_k, u_1 \in U_1\} =$$

$$\{p \times 0 + 0, p \times 0 + 1, \dots, p \times 0 + (p-1), p \times 2 + 0, \dots, p \times 2 + (p-1), \dots, p \times (p^k - 1) + 0, \dots, p \times (p^k - 1) + (p-1)\} =$$

$$\{0, 1, \dots, p-1, p, p+1, \dots, 2p-1, 2p, \dots, 3p-1, \dots, p^{k+1} - p, \dots, p^{k+1} - 1\} = \{0, 1, 2, \dots, p^{k+1} - 1\}$$

(2) 用与 (1) 类似的方法可证。

(3) 设 $(i, j) \in I_n$, i 写成 p 进制数为 $i_{n-1}i_{n-2} \dots i_2i_1i_0$, j 写成 q 进制数为 $j_{n-1}j_{n-2} \dots j_2j_1j_0$, 由定义 1(2) 知, $(i, j) \in I_n \Leftrightarrow (\lfloor i/p \rfloor, \lfloor j/q \rfloor) \in I_{n-1} \wedge (\langle i/p \rangle, \langle j/q \rangle) \in I_1$ 因 $\lfloor i/p \rfloor = i_{n-1}i_{n-2} \dots i_1$, $\lfloor j/q \rfloor = j_{n-1}j_{n-2} \dots j_2j_1$, $\langle i/p \rangle = i_0$, $\langle j/q \rangle = j_0$, 所以 $(i_0, j_0) \in I_1$ 及 $(i_{n-1}i_{n-2} \dots i_1, j_{n-1}j_{n-2} \dots j_2j_1) \in I_{n-1}$, 同理, 由 $(i_{n-1}i_{n-2} \dots i_1, j_{n-1}j_{n-2} \dots j_2j_1) \in I_{n-1}$, 知 $(i_{n-1}i_{n-2} \dots i_2, j_{n-1}j_{n-2} \dots j_2) \in I_{n-2}$, 及 $(i_1, j_1) \in I_1$, 如此下去知 $(i_{n-1}, j_{n-1}) \in I_1, (i_{n-2}, j_{n-2}) \in I_1, \dots, (i_1, j_1) \in I_1, (i_0, j_0) \in I_1$ 。 \square

为了实现上述背景序列与分形的联系,把该背景序列中的每一个背景,都与单位正方形中的一个图形对应起来。

3 背景序列对应的图形

定义 2 (图形序列) 设 $\{K_n\}_{n=1,2,\dots}$ 是由具有 p 个对象及 q 个属性的形式背景 $K=(U, M, I)$ 产生的背景序列, 则令背景 $K_n=(U_n, M_n, I_n)$ 对应的图形是:

$$G_n = \cup \{S(i, j) | i \in U_n \wedge j \in M_n \wedge (i, j) \in I_n\}$$

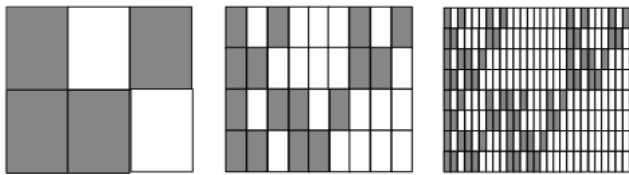
其中 $S(i, j)$ 是矩形

$$[j/q^n, (j+1)/q^n] \times [p^n - i - 1/p^n, p^n - i/p^n]$$

这里 $[a, b]$ 代表以 a 及 b 为端点的闭区间。

这样任何一个上述的背景序列都对应一个单位正方形中的图形序列。

例 2 对于例 1 中的背景序列对应的图形序列如图 2 所示, 其中 K_1 对应(a), K_2 对应(b), K_3 对应(c)。



(a)The graph relative to context K_1 (b)The graph relative to context K_2 (c)The graph relative to context K_3
 (a)背景 K_1 对应的图形 (b)背景 K_2 对应的图形 (c)背景 K_3 对应的图形

Fig.2 The graphs relative to context series

图 2 对应于背景序列的图形

由定义 2 还知道这个图形序列中每个 K_n 对应的图形都是由一些边长为 $1/p^n$ 及 $1/q^n$ 的彼此相等的矩形组成, 而且还知这些相等的矩形个数为 N^n , 其中 N 是关系 I 中的元素个数。

定理 2 设 $\{K_n\}_{n=1,2,\dots}$ 是由形式背景 $K=(U, M, I)$ 按以上方法生成的背景序列, 而 K 有 p 个对象, q 个属性。 $\{G_n\}_{n=1,2,\dots}$ 是 $\{K_n\}_{n=1,2,\dots}$ 对应的图形序列, 则 $\lim_{n \rightarrow \infty} G_n = G$ 是个分形图形(由此称这个背景序列为一个分形背景序列)。

证明 由 G 的生成过程(例如如图 2)知 G 具有分形图形的 5 个特性:

- (1)整体与局部自相似(由定义直接看出)。
- (2)有无限细致的结构(由定义直接看出)。
- (3)具有分数维: 设 I 中有 N 个元素, 令 $r = \min(p, q)$, 由前述知 G_n 中有 N^n 个边长为 $1/p^n$ 及 $1/q^n$ 的矩形, 这些矩形是整个图形的一个覆盖, 而这些矩形的最大直径是:

$$\delta_n = \sqrt{(1/p^n)^2 + (1/q^n)^2}$$

于是若 $p \leq q$, 则

$$\begin{aligned} \delta_n^m \times N^n &= (\sqrt{(1/p^n)^2 + (1/q^n)^2})^m \times N^n = \\ &= (1/p^n)^m \times N^n (\sqrt{1+(p^n+q^n)^{-2}})^m = \\ &= (N/p^m)^n (\sqrt{1+(p^n+q^n)^{-2}})^m \end{aligned}$$

所以当 $n \rightarrow \infty$ 时上式只有 $m = \log_p N$ 时才能有有限且不为 0 的极限, 同理 $q < p$ 时只有 $m = \log_q N$ 才能有有限且不为 0 的极限, 由 Hausdorff 定理^[15]知各种覆盖的这个极限都是相同的, 而这个值 $m = \log N / \log r$ 就是该图形 G 的 Hausdorff 分数维, 这里 $r = \min(p, q)$ 。

- (4)可用简单的迭代方法生成(由定义直接看出)。
- (5)度量与精度有关, 例如测量例 2 中产生的分形图形 G 的面积, 如果垂直精度只能到 $1/2^3$, 水平精度能只到 $1/3^3$, 则只能分辨图 2(c)中的那些小矩形, 于是面积为 $64 \times 1/2^3 \times 1/3^3 = 8/27$ 。如果垂直精度只到 $1/2^2$, 水平精度只到 $1/3^2$, 则只能分辨图 2(b)中的那些小矩形, 于是面积为 $16 \times 1/2^2 \times 1/3^2 = 4/9$ 。如果垂直精度只到 $1/2$, 水平精度只到, 则只能分辨图 2(a)中的那些小矩形, 于是面积为 $4 \times 1/2 \times 1/3 = 2/3$ 。可见测量与精度有关。

由于 G 满足以上 5 点, 所以 G 是分形图形。 □

4 背景序列的概念格序列

下面研究上述分形背景序列产生的概念序列的情况。

定义 3 (含空值的概念) 设 $K=(U, M, I)$ 是一个背景, $\mathfrak{B}(K)$ 是背景 K 中所有概念的集合, (α, β) 是背景 K 的一个概念, 若 $\alpha \neq \emptyset$ 且 $\beta \neq \emptyset$, 则称 $(\alpha, \beta) \in \mathfrak{B}(K)$ 是普通概念, 普通概念的集合记作 $\mathfrak{B}_0(K)$, 若 $\alpha = \emptyset$ 或 $\beta = \emptyset$, 则称 $(\alpha, \beta) \in \mathfrak{B}(K)$ 是含空值的概念, 含空值的概念的集合记作 $\mathfrak{B}_1(K)$ 。于是所有概念的集合为 $\mathfrak{B}(K) = \mathfrak{B}_0(K) \cup \mathfrak{B}_1(K)$ 。

定理 3 设 $\{K_n\}_{n=1,2,\dots}$ 是由 K 产生的分形背景序列, K 有 p 个对象及 q 个属性。则 $\mathfrak{B}_1(K_{n+1}) = \mathfrak{B}_0(K_n) \oplus \mathfrak{B}_1(K_1)$ 。

证明 证明若 $(\alpha_{n+1}, \beta_{n+1})$ 是 $\mathfrak{B}_0(K_{n+1})$ 中的任意概念, 则当且仅当存在 $(\alpha_n, \beta_n) \in \mathfrak{B}_0(K_n)$ 及 $(\alpha_1, \beta_1) \in \mathfrak{B}_0(K_1)$ 使得 $(\alpha_{n+1}, \beta_{n+1}) = (\alpha_n, \beta_n) \oplus (\alpha_1, \beta_1)$ 。

用 f_i 与 g_i 分别表示 K_i 中的 f 函数与 g 函数 ($i=1, 2, \dots$), 于是

$$(\alpha_{n+1}, \beta_{n+1}) \in \mathfrak{B}_0(K_{n+1})$$

等价于

$$\beta_{n+1} = f_{n+1}(\alpha_{n+1}) \text{ 且 } \alpha_{n+1} = g_{n+1}(\beta_{n+1})$$

即

$$\alpha_{n+1} = \{u_{n+1} \in U_{n+1} \mid \forall m_{n+1} \in \beta_{n+1}, (u_{n+1}, m_{n+1}) \in I_{n+1}\} \quad (1)$$

$$\beta_{n+1} = \{m_{n+1} \in M_{n+1} \mid \forall u_{n+1} \in \alpha_{n+1}, (u_{n+1}, m_{n+1}) \in I_{n+1}\} \quad (2)$$

由定义 1 中(2)知, 每一个 m_{n+1} 及 u_{n+1} 都存在 $m_n \in M_n, u_n \in U_n, m_1 \in M_1, u_1 \in U_1$ 使 $m_{n+1} = qm_n + m_1$ 及 $u_{n+1} = pu_n + u_1$, 且因 $(u_{n+1}, m_{n+1}) \in I_{n+1} \Leftrightarrow (u_n, m_n) \in I_n \wedge (u_1, m_1) \in I_1$, 于是式(1)、(2)等价于:

$$\alpha_{n+1} = \{pu_n + u_1 \in U_{n+1} \mid \forall (qm_n + m_1) \in \beta_{n+1}, (u_n, m_n) \in I_n \wedge (u_1, m_1) \in I_1\} \quad (3)$$

$$\beta_{n+1} = \{qm_n + m_1 \in M_{n+1} \mid \forall (pu_n + u_1) \in \alpha_{n+1}, (u_n, m_n) \in I_n \wedge (u_1, m_1) \in I_1\} \quad (4)$$

注意到, $m_1 \in \{0, 1, \dots, q-1\}, u_1 \in \{0, 1, \dots, p-1\}$, 所以 $\lfloor (qm_n + m_1)/q \rfloor = m_n, \langle (qm_n + m_1)/q \rangle = m_1, \lfloor (pu_n + u_1)/p \rfloor = u_n, \langle (pu_n + u_1)/p \rangle = u_1$ 。所以 $\forall (pu_n + u_1) \in \alpha_{n+1}$ 等价于: $\forall u_n \in \lfloor \alpha_{n+1}/p \rfloor$ 及 $\forall u_1 \in \langle \alpha_{n+1}/p \rangle \forall (qm_n + m_1) \in \beta_{n+1}$ 等

价于:

$$\forall m_n \in \lfloor \beta_{n+1}/q \rfloor \text{ 及 } \forall m_1 \in \langle \beta_{n+1}/q \rangle$$

所以式(3)、(4)等价于

$$\alpha_{n+1} = \{pu_n + u_1 \in U_{n+1} \mid \forall m_n \in \lfloor \beta_{n+1}/q \rfloor, (u_n, m_n) \in I_n \wedge \forall m_1 \in \langle \beta_{n+1}/q \rangle, (u_1, m_1) \in I_1\}$$

$$\beta_{n+1} = \{qm_n + m_1 \in M_{n+1} \mid \forall u_n \in \lfloor \alpha_{n+1}/p \rfloor, (u_n, m_n) \in I_n \wedge \forall u_1 \in \langle \alpha_{n+1}/p \rangle, (u_1, m_1) \in I_1\}$$

再注意到, 根据引言中关于 f 函数及 g 函数的定义:

$$\forall m_n \in \lfloor \beta_{n+1}/q \rfloor, (u_n, m_n) \in I_n$$

$$\text{即是 } u_n \in g_n(\lfloor \beta_{n+1}/q \rfloor), \forall m_1 \in \langle \beta_{n+1}/q \rangle, (u_1, m_1) \in I_1$$

$$\text{即是 } u_1 \in g_1(\langle \beta_{n+1}/q \rangle), \forall u_n \in \langle \alpha_{n+1}/p \rangle, (u_n, m_n) \in I_n$$

$$\text{即是 } m_n \in f_n(\lfloor \alpha_{n+1}/p \rfloor), \forall u_1 \in \langle \alpha_{n+1}/p \rangle, (u_1, m_1) \in I_1$$

$$\text{即是 } m_1 \in f_1(\langle \alpha_{n+1}/p \rangle), \text{ 所以}$$

$$\alpha_{n+1} = \{pu_n + u_1 \in U_{n+1} \mid u_n \in g_n(\lfloor \beta_{n+1}/q \rfloor) \wedge u_1 \in g_1(\langle \beta_{n+1}/q \rangle)\} \quad (5)$$

$$\beta_{n+1} = \{qm_n + m_1 \in M_{n+1} \mid m_n \in f_n(\lfloor \alpha_{n+1}/p \rfloor) \wedge m_1 \in f_1(\langle \alpha_{n+1}/p \rangle)\} \quad (6)$$

为了方便定理的证明, 定义两个记号: 设 S 是一个非负整数的集合, a 是一个正整数, 则令:

$$\lfloor S/a \rfloor = \{\lfloor s/a \rfloor \mid s \in S\}$$

$$\langle S/a \rangle = \{\langle s/a \rangle \mid s \in S\}$$

例如: $S = \{0, 2, 8, 9, 11\}, a = 3$, 则

$$\lfloor S/a \rfloor = \{0, 0, 2, 3\} = \{0, 2, 3\}$$

$$\langle S/a \rangle = \{0, 2, 2, 0, 2\} = \{0, 2\}$$

按这种记号式(5)、(6)等价于:

$$\alpha_{n+1} = pg_n(\lfloor \beta_{n+1}/q \rfloor) + g_1(\langle \beta_{n+1}/q \rangle) \quad (7)$$

$$\beta_{n+1} = qf_n(\lfloor \alpha_{n+1}/p \rfloor) + f_1(\langle \alpha_{n+1}/p \rangle) \quad (8)$$

由定义 1 中(2)知:

$$f_1(\langle \alpha_{n+1}/p \rangle) \in \{0, 1, \dots, (p-1)\}$$

$$g_1(\langle \beta_{n+1}/q \rangle) \in \{0, 1, \dots, (q-1)\}$$

因此, 得到

$$\lfloor \beta_{n+1}/q \rfloor = f_n(\lfloor \alpha_{n+1}/p \rfloor) \quad (9)$$

$$\langle \beta_{n+1}/q \rangle = f_1(\langle \alpha_{n+1}/p \rangle) \quad (10)$$

$$\lfloor \alpha_{n+1}/p \rfloor = g_n(\lfloor \beta_{n+1}/q \rfloor) \quad (11)$$

$$\langle \alpha_{n+1}/p \rangle = g_1(\langle \beta_{n+1}/q \rangle) \quad (12)$$

由式(9)、(11)知

$$(\lfloor \alpha_{n+1}/p \rfloor, \lfloor \beta_{n+1}/q \rfloor) \in \mathfrak{B}_0(K_n) \quad (13)$$

由式(10)、(12)知

$$(\langle \alpha_{n+1}/p \rangle, \langle \beta_{n+1}/q \rangle) \in \mathfrak{B}_0(K_1) \quad (14)$$

于是从式(7)、(8)知存在 $(\alpha_n, \beta_n) \in \mathfrak{B}_0(K_n)$ 及 $(\alpha_1, \beta_1) \in \mathfrak{B}_0(K_1)$ 使 $\alpha_{n+1} = p\alpha_n + \alpha_1$ 及 $\beta_{n+1} = q\beta_n + \beta_1$, 这里 $\alpha_n = \lfloor \alpha_{n+1}/p \rfloor, \alpha_1 = \langle \alpha_{n+1}/p \rangle, \beta_n = \lfloor \beta_{n+1}/q \rfloor, \beta_1 = \langle \beta_{n+1}/q \rangle$ 。所以 $(\alpha_{n+1}, \beta_{n+1}) \in \mathfrak{B}_0(K_{n+1})$, 当且仅当

$$(\alpha_{n+1}, \beta_{n+1}) = (\alpha_n, \beta_n) \oplus (\alpha_1, \beta_1)$$

由此还进一步得到:

$$\mathfrak{B}_0(K_{n+1}) = \mathfrak{B}_0(K_n) \oplus \mathfrak{B}_0(K_1)$$

下面再研究 $\mathfrak{B}_1(K_n)$ 的迭代情况:

定理 4 设 $\{K_n\}_{n=1,2,\dots}$ 是由 K 产生的分形背景序列, K 有 p 个对象及 q 个属性。对于含空值概念, 有:

- (1) 若 $\mathfrak{B}_1(K_1) = \emptyset$, 则 $\mathfrak{B}_1(K_n) = \emptyset, n=2, 3, \dots$;
- (2) 若 $\mathfrak{B}_1(K_1) = \{(\emptyset, M_1)\}$, 则 $\mathfrak{B}_1(K_n) = \{(\emptyset, M_n)\}, n=2, 3, \dots$;
- (3) 若 $\mathfrak{B}_1(K_1) = \{(U_1, \emptyset)\}$, 则 $\mathfrak{B}_1(K_n) = \{(U_n, \emptyset)\}, n=2, 3, \dots$;
- (4) 若 $\mathfrak{B}_1(K_1) = \{(U_1, \emptyset), (\emptyset, M_1)\}$, 则 $\mathfrak{B}_1(K_n) = \{(U_n, \emptyset), (\emptyset, M_n)\}, n=2, 3, \dots$ 。

证明 先证明对任何 $n=2, 3, \dots$ 都有:

$$(\emptyset, M_n) \in \mathfrak{B}_1(K_n) \Leftrightarrow (\emptyset, M_{n-1}) \in \mathfrak{B}_1(K_{n-1}) \vee (\emptyset, M_1) \in \mathfrak{B}_1(K_1)$$

这是因为:

$$\begin{aligned} (U_n, \emptyset) \in \mathfrak{B}_1(K_n) &\Leftrightarrow \forall u_n \in U_n \exists m_n \in M_n (u_n, m_n) \notin I_n \Leftrightarrow \forall (pu_{n-1} + u_1) \in U_n \exists (qm_{n-1} + m_1) \in M_n (u_{n-1}, m_{n-1}) \notin I_{n-1} \vee (u_1, m_1) \notin I_1 \Leftrightarrow (\forall u_{n-1} \in U_{n-1}, \exists m_{n-1} \in M_{n-1} (u_{n-1}, m_{n-1}) \notin I_{n-1}) \vee (\forall u_1 \in U_1, \exists m_1 \in M_1, (u_1, m_1) \notin I_1) \Leftrightarrow (\emptyset, M_{n-1}) \in \mathfrak{B}_1(K_{n-1}) \vee (\emptyset, M_1) \in \mathfrak{B}_1(K_1) \end{aligned}$$

显然这也就是

$$f_{n+1}(U_{n+1}) = \emptyset \Leftrightarrow f_n(U_n) = \emptyset \vee f_1(U_1) = \emptyset \quad (15)$$

$$f_{n+1}(U_{n+1}) \neq \emptyset \Leftrightarrow f_n(U_n) \neq \emptyset \vee f_1(U_1) \neq \emptyset \quad (16)$$

同理对 $n=2, 3, \dots$ 都有: $(U_n, \emptyset) \in \mathfrak{B}_1(K_n) \Leftrightarrow (U_{n-1}, \emptyset) \in \mathfrak{B}_1(K_{n-1})$ 或写作

$$g_{n+1}(M_{n+1}) = \emptyset \Leftrightarrow g_n(M_n) = \emptyset \vee g_1(M_1) = \emptyset \quad (17)$$

$$g_{n+1}(M_{n+1}) \neq \emptyset \Leftrightarrow g_n(M_n) \neq \emptyset \vee g_1(M_1) \neq \emptyset \quad (18)$$

下面证明这个定理:

(1) 若 $\mathfrak{B}_1(K_1) = \emptyset$, 则 (U_1, \emptyset) 及 (\emptyset, M_1) 都不存在, 即 $f_1(U_1) \neq \emptyset$ 且 $g_1(M_1) \neq \emptyset$ 。在式(16)、(18)中令 $n=1$, 得

$$f_2(U_2) \neq \emptyset \Leftrightarrow f_1(U_1) \neq \emptyset \wedge f_1(U_1) \neq \emptyset$$

$$g_2(M_2) \neq \emptyset \Leftrightarrow g_1(M_1) \neq \emptyset \wedge g_1(M_1) \neq \emptyset$$

所以 $f_2(U_2) \neq \emptyset$ 及 $g_2(M_2) \neq \emptyset$ 成立。从而 $\mathfrak{B}_1(K_2) = \emptyset$ 现设 $f_n(U_n) = \emptyset$ 及 $g_n(M_n) \neq \emptyset$ 成立。注意到 $f_1(U_1) \neq \emptyset$ 且 $g_1(M_1) \neq \emptyset$, 所以由式(16)、(18)知 $f_{n+1}(U_{n+1}) \neq \emptyset$ 及 $g_{n+1}(M_{n+1}) \neq \emptyset$ 成立。从而 $\mathfrak{B}_1(K_{n+1}) = \emptyset$ 。于是由归纳法知对所有 $n=1, 2, 3, \dots$ 都有 $\mathfrak{B}_1(K_n) = \emptyset$ 。

(2) 若 $\mathfrak{B}_1(K_1) = \{(\emptyset, M_1)\}$, (U_1, \emptyset) 不存在, 而 (\emptyset, M_1) 存在, 即 $f_1(U_1) \neq \emptyset$ 且 $g_1(M_1) = \emptyset$ 。根据 $g_1(M_1) = \emptyset$, 由图 2(c) 立即知对所有 $n=1, 2, 3, \dots$ 都有 $g_n(M_n) = \emptyset$ 。再根据 $f_1(U_1) \neq \emptyset$, 用(1)证明的类似方法知, 所有 $n=1, 2, 3, \dots$ 都有 $f_n(U_n) \neq \emptyset$, 因此所有 $n=1, 2, 3, \dots$ 都有 $\mathfrak{B}_1(K_n) = \{(\emptyset, M_n)\}$ 。

(3) 若 $\mathfrak{B}_1(K_1) = \{(U_1, \emptyset)\}$, 即 (\emptyset, M_1) 不存在, 而 (U_1, \emptyset) 存在, 即 $g_1(M_1) \neq \emptyset$ 且 $f_1(U_1) = \emptyset$ 。根据 $f_1(U_1) = \emptyset$, 由图 2(a) 立即知对所有 $n=1, 2, 3, \dots$ 都有 $f_n(U_n) = \emptyset$ 。再根据 $g_1(M_1) \neq \emptyset$, 用(1)证明的类似方法知, 所有 $n=1, 2, 3, \dots$ 都有 $g_n(M_n) \neq \emptyset$, 因此所有 $n=1, 2, 3, \dots$ 都有 $\mathfrak{B}_1(K_n) = \{(M_n, \emptyset)\}$ 。

(4) 若 $\mathfrak{B}_1(K_1) = \{(U_1, \emptyset), (\emptyset, M_1)\}$, 即 (U_1, \emptyset) 与 (\emptyset, M_1) 都存在, 即 $f_1(U_1) = \emptyset$ 且 $g_1(M_1) = \emptyset$ 。根据 $f_1(U_1) =$

\emptyset , 由图 2(a) 立即知对所有 $n=1, 2, 3, \dots$ 都有 $f_n(U_n)=\emptyset$, 根据 $g_1(M_1)=\emptyset$, 由图 2(c) 立即知对所有 $n=1, 2, 3, \dots$ 都有 $g_n(M_n)=\emptyset$ 。这样对所有 $n=1, 2, 3, \dots$ 都有 $\mathfrak{B}_1(K_{n+1})=\{(U_{n+1}, \emptyset), (\emptyset, M_{n+1})\}$ 。 □

定理 5 设 $\{K_n\}_{n=1,2,\dots}$ 是由 K 产生的分形背景序列, 如果 $(\alpha_n, \beta_n)(\gamma_n, \delta_n) \in \mathfrak{B}_0(K_n)$, 而

$$\alpha_n = p\alpha_{n-1} + \alpha_1, \beta_n = q\beta_{n-1} + \beta_1$$

$$\gamma_n = p\gamma_{n-1} + \gamma_1, \delta_n = q\delta_{n-1} + \delta_1$$

这里

$$(\alpha_1, \beta_1)(\gamma_1, \delta_1) \in \mathfrak{B}_0(K_1)$$

$$(\alpha_{n-1}, \beta_{n-1})(\gamma_{n-1}, \delta_{n-1}) \in \mathfrak{B}_0(K_{n-1})$$

则当且仅当 $(\alpha_{n-1}, \beta_{n-1})$ 是 $(\gamma_{n-1}, \delta_{n-1})$ 的直接父概念, 且 (α_1, β_1) 是 (γ_1, δ_1) 的直接父概念时 (α_n, β_n) 是 (γ_n, δ_n) 的直接父概念。

证明 因为 (α_n, β_n) 是 (γ_n, δ_n) 的直接父概念也就是 $(p\alpha_{n-1} + \alpha_1, q\beta_{n-1} + \beta_1)$ 是 $(p\gamma_{n-1} + \gamma_1, q\delta_{n-1} + \delta_1)$ 的直接父概念。于是当且仅当 $p\alpha_{n-1} + \alpha_1 \supset p\gamma_{n-1} + \gamma_1$, 且不存在 $(\sigma_1, \tau_1) \in \mathfrak{B}_0(K_1)$ 及 $(\sigma_{n-1}, \tau_{n-1}) \in \mathfrak{B}_0(K_{n-1})$ 使 $p\alpha_{n-1} + \alpha_1 \supset p\sigma_{n-1} + \tau_1 \supset p\gamma_{n-1} + \gamma_1$ 。注意到

$$\lfloor (p\alpha_{n-1} + \alpha_1) / p \rfloor = \alpha_{n-1}$$

$$\lfloor (p\gamma_{n-1} + \gamma_1) / p \rfloor = \gamma_{n-1}$$

$$\lfloor (p\sigma_{n-1} + \sigma_1) / p \rfloor = \sigma_{n-1}$$

且

$$\langle (p\alpha_{n-1} + \alpha_1) / p \rangle = \alpha_1$$

$$\langle (p\gamma_{n-1} + \gamma_1) / p \rangle = \gamma_1$$

$$\langle (p\sigma_{n-1} + \sigma_1) / p \rangle = \sigma_1$$

所以这是当且仅当不存在 $(\sigma_1, \tau_1) \in \mathfrak{B}_0(K_1)$ 使 $\alpha_1 \supset \sigma_1 \supset \gamma_1$ 且不存在 $(\sigma_{n-1}, \tau_{n-1}) \in \mathfrak{B}_0(K_{n-1})$ 使 $\alpha_{n-1} \supset \sigma_{n-1} \supset \gamma_{n-1}$ 。

这样 (α_n, β_n) 是 (γ_n, δ_n) 的直接父概念是当且仅当 $(\alpha_{n-1}, \beta_{n-1})$ 是 $(\gamma_{n-1}, \delta_{n-1})$ 的直接父概念且 (α_1, β_1) 是

(γ_1, δ_1) 的直接父概念。 □

例 3 根据定义 3、定理 3 和定理 4, 对于例 1 中的背景序列, 可得背景 K_1, K_2 的概念格的 Hasse 图如图 3、图 4 所示。

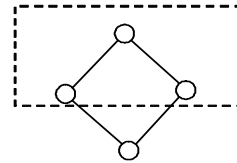


Fig.3 Concept lattice of K_1

图 3 K_1 所对应的概念格

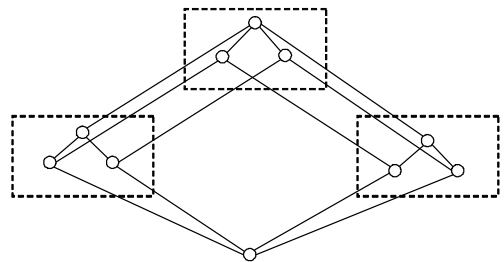


Fig.4 Concept lattice of K_2

图 4 K_2 所对应的概念格

5 分形图形的概念分数维

根据定义 3、定理 3 和定理 4, 下面证明任一个背景 K 生成的背景序列的概念格序列的概念个数具有 Hausdorff 特性。首先有以下定理:

定理 6 设 $\{K_n\}_{n=1,2,\dots}$ 是一个由 K 生成的背景序列, 而 $\mathfrak{B}(K)$ 中有 N 个概念, 其中含空值概念有 N_0 个(显然 $N_0=0$ 或 1 或 2), 则 $\mathfrak{B}(K)$ 中的概念个数是 $(N-N_0)^n + N_0$ 。

证明 首先定理对 $n=1$ 正确, 因为 $(N-N_0)^1 + N_0 = N$ 。现设定理对 n 正确, 注意到定理 3 知, 不论 n 是多少, $\mathfrak{B}(K_n)$ 中都仍有 N_0 个含空值的概念。于是 $\mathfrak{B}(K_n)$ 有 $(N-N_0)^n$ 个普通概念。

再由 \oplus 运算的定义知, 每个 $\mathfrak{B}(K_n)$ 中的普通概念都可与 $\mathfrak{B}(K_1)=\mathfrak{B}(K)$ 中的普通概念进行 \oplus 运算, 而且显然运算的结果各不相同。于是产生

$$(N-N_0)^n \times (N-N_0) = (N-N_0)^{n+1}$$

个 $\mathfrak{B}(K_{n+1})$ 中的普通概念。这样 $\mathfrak{B}(K_{n+1})$ 中的概念总数就是 $(N-N_0)^n + N_0$ 。□

进而有:

定理 7 设 $\{K_n\}_{n=1,2,\dots}$ 是一个由 p 个对象, q 个属性的背景 K 生成的背景序列, 则这个背景序列生成的概念格序列的概念个数具有 Hausdorff 特性。

证明 由以上定理知 $\mathfrak{B}(K_{n+1})$ 中的概念总数是 $(N-N_0)^n + N_0$ 。在 $\mathfrak{B}(K_{n+1})$ 中用边长为 $(1/p)^n$ 及 $(1/q)^n$ 的矩形来覆盖, 则矩形的最大直径为其对角线, 长度为:

$$\delta = \sqrt{(1/p^n)^2 + (1/q^n)^2} = \sqrt{(1/p)^{2n} + (1/q)^{2n}}$$

考虑: $\delta^m((N-N_0)^n + N_0)$, 当 $p \leq q$ 时:

$$\delta_n^m((N-N_0)^n + N_0) = (\sqrt{(1/p)^{2n} + (1/q)^{2n}})^m ((N-N_0)^n + N_0) = \frac{(\sqrt{1+(p/q)^{2n}})^m ((N-N_0)^n + N_0)}{p^{mn}}$$

只有当 $m = \log_p(N-N_0)$ 时 $t \rightarrow \infty$ 有有限的不为 0 的极限。

当 $p \geq q$ 时:

$$\delta_n^m((N-N_0)^n + N_0) = \frac{(\sqrt{(p/q)^{2n} + 1})^m ((N-N_0)^n + N_0)}{q^{mn}}$$

只有当 $m = \log_q(N-N_0)$ 时 $t \rightarrow \infty$ 有有限的不为 0 的极限。

这样看到只有 m 为某一个值时 $\delta_n^m((N-N_0)^n + N_0)$ 才有有限的不为 0 的极限, 所以概念格序列的概念个数具有 Hausdorff 特性。□

根据定理 7 可以给极限背景定义一种新的维数: 概念分数维。

定义 4 设 $\{K_n\}_{n=1,2,\dots}$ 是一个由背景生成的自相似背景序列, K 有 p 个对象及 q 个属性。而 $\mathfrak{B}(K)$ 中有 N 个概念, 其中含空值概念有 N_0 个, 而 $r = \min\{p, q\}$,

则这个背景序列对应的图形序列的极限分形图形的概念分数维是: $\log(N-N_0) \log r$ 。

6 结论

文章在国内外首次探讨了概念格与分形之间的联系, 提出了一种从任意一个原始形式背出发形成分形形式背景序列的方法, 研究并给出了这种自相似形式背景序列的每个背景的概念格的求法。建立了这种形式背景序列与单位正方形中图形序列的对应, 证明了这个图形序列的极限图形是分形图形。还研究了这种形式背景序列中每个背景的概念格中的概念个数的求法, 并以此定义了表征分形背景形成概念能力的一种分数维: 概念分数维。

References:

- [1] Mandelbrot B B. The fractal geometry of nature[M]. San Francisco: Freeman, 1982.
- [2] Wille R. Restructuring lattice theory: An approach based on hierarchies of concepts[M]//Rival I. Ordered Sets. Dordrecht-Boston: Reidel, 1982:445-470.
- [3] Oosthuizen G D. The application of concept lattice to machine learning[R]. South Africa: University of Pretoria, 1996.
- [4] Ho T B. Incremental conceptual clustering in the framework of Galois lattice [M]//Lu H, Motoda H, Liu H. KDD: Techniques and Applications. Singapore: World Scientific, 1997:49-64.
- [5] Kent R E, Bowman C M. Digital libraries, conceptual knowledge systems and the nebula interface[R]. University of Arkansas, 1995.
- [6] Corbett D, Burrow A L. Knowledge reuse in SEED exploiting conceptual graphs[C]//International Conference on Conceptual Graphs (ICCS'96). Sydney: University of New South Wales, 1996:56-60.
- [7] Schmitt I, Saake G. Merging inheritance hierarchies for scheme integration based on concept lattices[EB/OL]. <http://www.mathematic.tu-darmstadt.de/ags/ag1>.

- [8] Siff M, Reps T. Identifying modules via concept analysis[C]// Harrold M J, Visaggio G. International Conference on Software Maintenance, Bari, Italy. Washington, DC: IEEE Computer Society, 1997:170-179.
- [9] Hu Keyuen, Lu Yuchang, Shi Chuenyi. Advances in concept lattice and its application[J]. J Tsinghua Univ: Sci and Tech, 2000,40(9):77-81.
- [10] Ho T B. An approach to concept formation based on formal concept analysis[J]. IEICE Trans Information and Systems, 1995, E782D(5):553-559.
- [11] Carpineto C, Romano G. Galois: An order-theoretic approach to conceptual clustering[C]//Utgo P. Proceedings of ICML 2003. Amherst: Elsevier, 1993:33-40.
- [12] Godin R. Incremental concept formation algorithm based on Galois (concept) lattices[J]. Computational Intelligence, 1995,11(2):246-267.
- [13] Yao Y Y. Concept lattices in rough set theory[C]//Dick S, Kurgan L, Pedrycz W. Proceedings of 2004 Annual Meeting of the North American Fuzzy Information Processing Society (NAFIPS 2004), IEEE Catalog Number: 04TH8736, June 27-30, 2004:796-801.
- [14] Oosthuizen G D. Rough sets and concept lattices[C]//Ziarko W P. Rough Sets, and Fuzzy Sets and Knowledge Discovery (RSKD'93). London:Springer-Verlag, 1994:24-31.
- [15] Zhang Jizhong. Fractal[M]. Beijing: Beijing Tsinghua University Press, 1995.

附中文参考文献:

- [9] 胡可云, 陆玉昌, 石纯一.概念格及其应用进展[J].清华大学学报:自然科学版,2000,40(9):77-81.
- [15] 张济忠.分形[M].北京:清华大学出版社,1995.



MA Yuan was born in 1941. He is a professor and doctoral supervisor at University of Science and Technology Liaoning. His research interests include database theory, data mining and knowledge discovery, rough set theory, formal concept analysis.

马垣(1941-),男,北京人,教授,辽宁科技大学研究生导师,主要研究领域:数据库理论,数据采掘知识发现,粗糙集理论,形式概念分析。



LIU Zhangang was born in 1970. He is a M.S. candidate at School of Software Engineering, University of Science and Technology Liaoning. His research interests include formal concept analysis, knowledge engineering, and knowledge-based software engineering.

刘占刚(1970-),男,辽宁科技大学软件学院硕士研究生,主要研究领域:形式概念分析,知识工程,基于知识的软件工程。