

关于图同构复杂性的分析^{*})

戴琼¹ 邹潇湘² 谭建龙³

(中国科学院软件研究所 北京 100080)¹ (国家计算机网络与信息安全管理中心 北京 100029)²
(中国科学院计算技术研究所 北京 100080)³

摘要 图同构问题是指对两个图寻找顶点之间的一个一一映射,使得两图的边在该映射下也保持对应关系,该问题得到许多研究者的关注。在一些论文中对图同构问题的复杂性给出了错误的描述,有的给出了多项式时间算法。本文对此进行了讨论,并给出了一些反例来证明其算法的错误。根据图同构国内外目前的研究进展,图同构既未被归入 P 问题,也未被归入 NPC 问题,是一个尚未解决的问题,有待进一步研究。

关键词 图同构, NP 问题, P 问题, NPC 问题, 图同构完备

Analysis on Complexity of Graph Isomorphism Problem

DAI Qiong¹ ZOU Xiao-Xiang² TAN Jian-Long³

(Institute of Software, Chinese Academy of Sciences, Beijing 100080)¹
(National Computer Network & Information Security Administration Center, Beijing 100029)²
(Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080)³

Abstract The graph isomorphism is to find a bijection between the vertexes of two graphs that preserve the edges. This problem has been paid much attention by many researchers. In some papers the complexity of the graph problem has been wrong described, and polynomial time algorithm is given. In this paper, we discuss that algorithm. Some examples are proposed to show the error of that algorithm. Graph isomorphism problem is one of the problems that have not been classified as NP complete, or within P. So it is an open problem until now and need further studies.

Keywords Graph isomorphism, NP problem, P problem, NPC problem, Graph isomorphism complete

1 背景

图同构问题可以分为 4 类:精确图同构、精确子图同构、不精确图同构、不精确子图同构。其中后面 3 个问题已被证明属于 NP 完备问题,是困难的问题。而对于精确图同构问题的复杂性,目前还没有明确的结果。本文所讨论的图同构,都是指精确的图同构。文[1]中指出,图同构问题既未被证明是多项式时间可解的,也未被证明属于 NP 完备的,正是由于这一点,得到许多学者的关注。

国内的一些学者也对图同构问题进行了研究,但其中一些论文有失严谨。如文[2]中指出:“图的同构问题,判断两个图是否同构是一个 NPC-问题”,疑为笔误。在文[3]和文[4]中,基于这样的性质:两个图同构,当且仅当这两个图的任意点连通子图的关联度序列相等,提出了一种图同构判断算法,复杂度为 $O(2^n)$,其中 n 为图的顶点数,但是对于作为算法基础的前面那条性质,文中只证明了充分性,而对必要性没有给出严格证明,因此算法的正确性有待证明。而且,即使该性质成立,对于较大的 n , $O(2^n)$ 的复杂度仍然是难以接受的。

另外,基于文[5]中提出的函数的概念,文[6]中提出了一种多项式时间复杂度的图同构判别算法,从而认为图同构问题是 P 问题。下面围绕这一算法展开讨论,首先介绍基本概念以及该算法,然后给出反例证明其错误,最后对图同构的复杂性进行讨论。

2 基本概念及方法

图 G 定义为 $G(V, E)$, 其中 $V = \{v_0, v_1, \dots, v_{n-1}\}$ 表示顶点集合, $E = \{(u, v) | u, v \in V\}$ 表示连接顶点的边的集合,这里考虑的是无向图,即连接顶点的边是没有方向的。顶点所邻接的边数称为顶点的度。仅由孤立点(即没有边)构成的图称为零图。对于两个图 G_1 和 G_2 , 如果存在两图的顶点之间的一个一一映射,使得两图的边在该映射下也保持对应关系,则称这两个图是同构的。

文[5]中定义了一种 VC 算法,首先选择图 G 中度最大的顶点 x_1 , 从图中去掉该顶点及所关联的边。在余下的子图中选择度最大的顶点 x_2 , 从子图中去掉该顶点及所关联的边。如此进行下去,直到余下的子图为零图为止。设所删除的这些顶点依次为 x_1, x_2, \dots, x_i , 设顶点 x_i 在对应的子图中所邻接的顶点为 $x_{i1}, x_{i2}, \dots, x_{ip_i}$, 用上标的形式表示顶点的度,即如果顶点 x_i 在图 G 中的度为 n_i , 则表示为 $x_i^{(n_i)}$, 定义

$$VC(G) = \{V_{x_1^{(n_1)}}^{(n_{11}), \dots, x_{1p_1}^{(n_{1p_1})}}, V_{x_2^{(n_2)}}^{(n_{21}), x_{22}^{(n_{22})}, \dots, x_{2p_2}^{(n_{2p_2})}}, \dots, V_{x_s^{(n_s)}}^{(n_{s1}), \dots, x_{sp_s}^{(n_{sp_s})}}\}$$

也就是说, VC 函数记录了在不断删除度为最大的顶点的过程中当前的顶点以及所邻接的顶点的一些信息。

基于上述定义,文[6]中提出了如下定理:

^{*}) 本课题研究得到中科院计算所青年基金课题 20056600-4 支持。戴琼 硕士,主要研究领域为网络与信息安全技术;邹潇湘 博士,主要研究领域为网络与信息安全技术;谭建龙 博士,主要研究领域为网络与信息安全技术。

定理 1 简单连通图 $G(V, E)$ 和 $F(U, W)$ 的同构的充要条件是有 VC 算法, 使得在

$$VC(G) = \{V_{x_1}^{(n1)}, x_{12}^{(n12)}, \dots, x_{1P_1}^{(n1P_1)}, V_{x_2}^{(n21)}, x_{22}^{(n22)}, \dots, x_{2P_2}^{(n2P_2)}, \dots, V_{x_s}^{(ns1)}, x_{s2}^{(ns2)}, \dots, x_{sP_s}^{(nsP_s)}\}$$

和

$$VC(F) = \{V_{y_1}^{(m11)}, y_{12}^{(m12)}, \dots, y_{1P_1}^{(m1P_1)}, V_{y_2}^{(m21)}, y_{22}^{(m22)}, \dots, y_{2P_2}^{(m2P_2)}, \dots, V_{y_s}^{(ms1)}, y_{s2}^{(ms2)}, \dots, y_{sP_s}^{(msP_s)}\}$$

中, 有

$$|\{x_{i1}^{(n1)}, x_{i2}^{(n12)}, \dots, x_{iP_i}^{(n1P_i)}\}| = |\{y_{i1}^{(m1)}, y_{i2}^{(m12)}, \dots, y_{iP_i}^{(m1P_i)}\}|$$

且

$$\{n_{i1}, n_{i2}, \dots, n_{iP_i}\} = \{m_{i1}, m_{i2}, \dots, m_{iP_i}\}, n_i = m_i$$

其中 $0 \leq i \leq s$.

由于 VC 函数表示在不断删除度数最大的顶点及关联的边的过程中当前的顶点以及所邻接的顶点的一些信息, 因此该定理的含义是: 如果两个图是同构的, 那么一定存在一个对应相同的删除顶点的过程; 反之, 如果删除顶点过程是对应相同的, 那么反过来也能够重建两个同构的图。

称在 $VC(G)$ 中的 P_1, P_2, \dots, P_s 为图的度序列, 定义 $\{n_{i1}, n_{i2}, \dots, n_{iP_i}\}$ 为顶点 x_i 关联的度集合。文章证明了图的度序列数量为 $O(n)$, 根据这一性质及定理 1, 得到一个图同构算法, 我们称为 FGI。基本思路为:

- (1) 生成两个图 G 和 F 的所有度序列, 分别记为 $g(1), g(2), \dots, g(q)$ 和 $f(1), f(2), \dots, f(q)$;
- (2) 如果两个度序列中没有相同的, 则输出“两图不同构”, 结束, 否则转 3;
- (3) 如果存在满足定理 1 条件的 $VC(G)$ 和 $VC(F)$, 则输出“两图同构”, 否则输出“两图不同构”, 结束。

最后证明该算法的复杂度为 $O(n^4)$ 。

3 算法的分析及举例

在 FGI 算法中, 尝试将顶点的度进行某种排序, 通过计算、比较两个图的度序列、被删除的顶点的度数以及它们关联的度集合, 来判读两个图是否同构。然而, 在选择度数最大的顶点时, 由于一个图中往往同时存在多个顶点具有最大的度, 特别是对于正则图更是如此, 因此使得当前选择的顶点不是唯一的, 正如文[6]中所指出的那样。在文[5]发表之后, 发现 $VC(G)$ 并非唯一。下面是一个例子:

例 1 考虑图 G

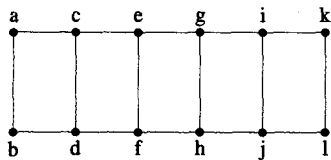


图 1 具有多个度序列的图

从图中依次删除顶点 c, f, g, j, b, k , 得到 $VC(G)$

$$\{V_{a,d,e}^{(a,d,e)}, V_{f,g,h}^{(f,g,h)}, V_{i,j,k}^{(i,j,k)}\}$$

或者依次删除顶点 c, h, i, b, f, l, e , 得到 $VC(F)$

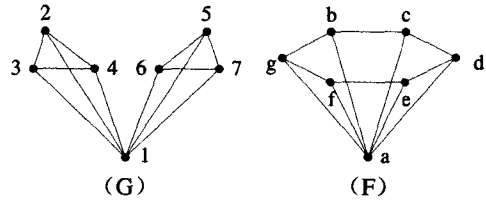
$$\{V_{a,d,e}^{(a,d,e)}, V_{f,g,j}^{(f,g,j)}, V_{i,k}^{(i,k)}\}$$

为简单起见, 这里没有标出顶点的度。这两个 $VC(G)$ 是不同的。

当然, 这并没有证伪定理 1 的必要性, 因为可能存在多个 $VC(G)$ 与 $VC(F)$, 定理 1 只需要其中存在一个相同的即可, 而对于同构的图是不难证明这一点的。进一步, 定理 1 的必要性可以加强为: 两个图 G 和 F 同构, 则所有 $VC(G)$ 和 $VC(F)$ 对应相等。

定理 1 必要性成立, 说明 FGI 算法对于同构的图, 是能够正确识别为同构的。那么定理 1 的充分性是否满足呢? 答案是否定的, 下面是一个反例。

例 2 (充分性反例) 考虑图 2 的两个图 G 和 F ,



(G) (F)

图 2 两个不同构的图

我们有 $VC(G)$:

$$\{V_1^{(2,3,4,5,6,7)}, V_2^{(3,4)}, V_3^{(6,7)}, V_4^{(3)}, V_5^{(7)}\}$$

$$\{V_a^{(b,c,d,e,f,g)}, V_b^{(c,g)}, V_d^{(d,f)}, V_f^{(g)}, V_e^{(d)}\}$$

注意到顶点 1 和 a 的度均为 6, 其它顶点的度均为 3, 因此满足定理 1 给出的条件。但是图 G 和 F 是不同构的, 因为图 G 中去掉顶点 1 就不连通了, 但图 F 中去掉顶点 a 仍然是连通的。

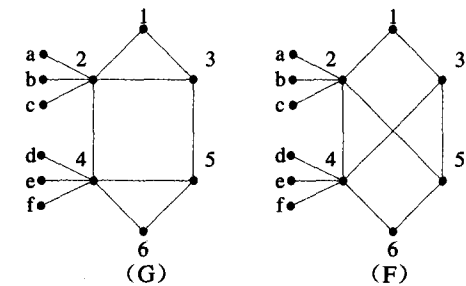
这样, 我们就找到了两个不同构的图, 满足定理 1 中所给出的 VC 函数对应相等的条件, 因此 FGI 算法会给出“两图同构”的错误判断。

类似于必要性的讨论, 能否对定理 1 的条件进行加强, 使充分性得以成立, 即是否有: 两个图 G 和 F 所有 $VC(G)$ 和 $VC(F)$ 对应相等, 则两个图同构。例 2 中两个不同构的图并非所有 $VC(G)$ 和 $VC(F)$ 都对应相等, 因为对于 $VC(F)$:

$$\{V_a^{(b,c,d,e,f,g)}, V_b^{(c,g)}, V_d^{(c,e)}, V_f^{(e,g)}\}$$

不存在与之对应的 $VC(G)$ 。然而, 考虑下面的例子:

例 3 (定理加强条件下的反例) 如图 3 的两个图 G 和 F :



(G) (F)

图 3 定理加强条件下的两个图

由于 $VC(G)$ 为:

$$\{V_1^{(1,3,4,a,b,c)}, V_4^{(5,6,d,e,f)}\} \cup VC(G \setminus \{2,4\})$$

或

$$\{V_4^{(2,5,6,d,e,f)}, V_1^{(1,3,a,b,c)}\} \cup VC(G \setminus \{2,4\})$$

$VC(F)$ 为:

$$\{V_2^{1,5,4,a,b,c}, V_4^{3,6,d,e,f}\} \cup VC(F \setminus \{2,4\})$$

或

$$\{V_4^{2,3,6,d,e,f}, V_2^{1,5,a,b,c}\} \cup VC(F \setminus \{2,4\})$$

其中 $G \setminus S$ 表示图 G 去掉顶点集合 S 后得到的子图。显然

$$VC(G \setminus \{2,4\}) = VC(F \setminus \{2,4\})$$

因此所有的 $VC(G)$ 和 $VC(F)$ 对应相等。但是,由于图 G 中存在三角形,而图 F 中不存在三角形,因此两图是不同构的。这说明,即使将定理 1 的条件进行加强,定理 1 的充分性仍然不能得到满足。

4 分析和总结

从前面的分析我们证明了,FGI 对于同构的图能给出正确的判断,但对于不同构的图可能会给出错误的判断。而识别出不同构的图,恰恰是图同构算法研究的关键和难点所在。这是因为,对于两个有 n 个顶点的图 G 和 F ,顶点间所有可能的映射数为 $n!$,要判断两图是不同构的,等同于判断所有这 $n!$ 个映射都不构成同构映射。对于较大的 n ,如 $n=1000$,用穷举法进行搜索显然是不可能的,因此必然需要利用一些限制条件来降低搜索空间。

先引入几个定义。一个图到自身的同构称为自同构。顶点集合 V 的划分(Partition),是将 V 分为一些不相交的集合,每个集合称为一个单元。顶点集合 V 的划分称为自同构划分(Automorphism Partitioning),如果顶点 u, v 属于该划分的同一个单元,当且仅当存在 G 的一个自同构映射,将 u 映射到 v 。

设 M 为顶点间的一个一一映射,设 $u \in G, M(u) = v \in F$,如果 u 和 v 的度不等,那么 M 一定不是同构映射。也就是说,同构映射一定使得 G 的一顶点映射到 F 中某个度相同的顶点。这样根据顶点的度,可以得到顶点的一个划分,使得属于同一单元的顶点具有相同的度。进一步,属于同一单元的两个顶点 u 和 v 所邻接顶点的度也应满足对应关系,这样可以将单元进一步细分。这个过程可以不断进行下去,直到划分不能再细分为止。设这样得到的划分由如下单元构成:

$$\{v_1, v_2, \dots, v_{c_1}\}, \{v_1, v_2, \dots, v_{c_2}\}, \dots, \{v_1, v_2, \dots, v_{c_n}\}$$

对两个图 G 和 F 如上得到划分,那么现在只要考虑将一个顶点映射到对应单元中某个顶点的映射即可,这样的映射数为:

$$c_1! \times c_2! \times \dots \times c_n!$$

其中 $c_1 + c_2 + \dots + c_n = n$ 。这样所需考虑的映射数可以大为降低。实际上,这一思想也是目前许多常见的图同构算法^[7~11]的基本思想。

然后,上述基于度的顶点划分并不能对顶点进行充分的细分,也就是得到的并不是自同构映射划分,因此这些算法不可避免地都需要进行回溯,也就是构造一棵搜索树,不断进行试探和剪枝,因此这些算法都不是多项式时间算法。

对于一些特殊类型的图,如平面图、区间图(Interval Graph)、树等,存在多项式算法。但对于一般的图,还没有得

到多项式时间算法。有学者^[12]提出了一种图同构完备的概念,来表示可以在多项式时间内归约为图同构的一类问题。正则图、二分图、弦图等类型的图的同构问题是图同构完备的^[13]。

图同构的复杂性如图 4 表示。

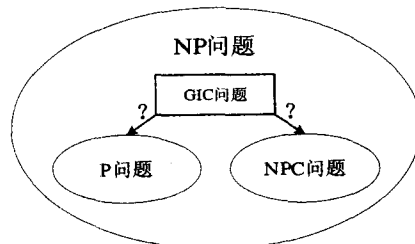


图 4 图同构完备问题

从目前的研究进展来看,图同构问题仍然未被归入 P 或者 NPC,是一个有待进一步研究的问题。

参考文献

- 1 Toran J. On the hardness of graph isomorphism. In: Proc. 41st Annual Symposium on Foundations of Computer Science (FOCS), Redondo Beach, CA Nov. 2000. 180~186
- 2 刘桂真,禹继国,谢力同. 两类图同构的充分必要条件. 山东大学学报(理学版), 2003,38(3): 1~3
- 3 李锋,商慧亮. 图同构判定算法:出入度序列法. 应用科学学报, 2002,20(3)
- 4 李锋,李晓艳. 图同构判定算法:关联度序列法及其应用. 复旦学报(自然科学版), 2001,40(3): 318~325
- 5 罗示丰. 两图同构的判别准则及其复杂性. 计算机科学, 1997(10): 148~153
- 6 罗示丰. 关于图同构复杂性的一点补充. 广西科学院学报, 2004, 20(3): 133~135
- 7 McKay B D. Practical graph isomorphism. Congressus Numerantium, 1981(30): 45~87
- 8 Ullmann J R. An algorithm for subgraph isomorphism. Journal of the Association for Computing Machinery, 1976, 23(1): 31~42
- 9 Schmidt D C. A fast backtracking algorithm to test directed graphs for isomorphism using distance matrices. Journal of the Association for Computing Machinery, 1976, 23(3): 433~445
- 10 Cordella L P, Foggia P, Sansone C, et al. Subgraph transformations for the inexact matching of attributed relational graphs. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(10): 1367~1372
- 11 Cordella L P, Foggia P, Sansone C, et al. An improved algorithm for matching large graphs. In: International Workshop on Graph-based Representation in Pattern Recognition, Ischia, Italy, May 2001. 149~159
- 12 Chen Wen Chi. Hierarchy of isomorphism testing; [Technical Report], [CaltechCSTR; 1984. 5140-tr-84]. California Institute of Technology. <http://caltechctr.library.caltech.edu/351/>
- 13 Uehara R, Toda S, Nagoya T. Graph isomorphism completeness for chordal bipartite graphs and strongly chordal graphs. Discrete Applied Mathematics, 2005, 145(3): 479~482