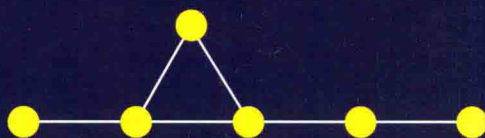


图对称性理论及其在数据管理中的应用

肖仰华 著



科学出版社

图对称性理论及其在数据 管理中的应用

肖仰华 著



科学出版社

北 京

内 容 简 介

本书系统地介绍了图对称性基础理论,包括基于图对称的复杂性度量模型与图距离度量模型、复杂对称网络生成模型以及基于图对称的网络约简理论等,介绍了一系列基于图对称性的应用方法,包括利用图对称实现社交网络隐私保护、利用图对称实现高效的最短路径索引与查询等。

本书理论和应用紧密结合,适合作为计算机等相关专业研究生和高年级本科生的参考书,为从事图数据管理与挖掘、复杂网络分析、复杂性理论等相关研究的工作者提供参考。

图书在版编目(CIP)数据

图对称性理论及其在数据管理中的应用/肖仰华著. —北京:科学出版社, 2018.11

ISBN 978-7-03-059137-1

I. ①图… II. ①肖… III. ①数据管理—研究 IV. ①TP274

中国版本图书馆 CIP 数据核字(2018)第 241922 号

责任编辑:赵艳春 / 责任校对:张凤琴
责任印制:师艳茹 / 封面设计:迷底书装

科学出版社 出版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

文林印务有限公司 印刷

科学出版社发行 各地新华书店经销

*

2018 年 11 月第 一 版 开本: 720 × 1000 1/16

2018 年 11 月第一次印刷 印张: 9 插页: 2

字数: 180 000

定价: 68.00 元

(如有印装质量问题,我社负责调换)

序

计算机科学的基本使命之一在于探索复杂性。一直以来，复杂性，更具体一点，数据或问题中的复杂性，长期困扰着计算机科学的研究者和实践者。什么是复杂性？是什么决定了某一数据(或问题)比另一数据(或问题)复杂？如何有效规避复杂性从而高效解决工程中的实际问题？征服复杂性更是成为了新旧世纪交替之时的时代最强音。复杂性的机理、测度与应用问题因而成为了计算机科学研究的核心使命之一。

复杂性的对立面是规则性。越为规则的数据与问题，越为简单。征服复杂性的关键因而也就归结为发现并应用数据与问题中的规则性。但是如何精准刻画规则性，则又成为了另一个难题。一个递增的数值序列是简单的，前提在于我们发现并能用数学公式描述这个数值序列。一堆看似杂乱的数据，从信息论的编码角度来看，则可能有着较短的编码长度，因而有着一定意义的规则性。虽然存在各种不同的规则性(或复杂性)刻画机制，但是具有普适的规则性刻画机制是对称，也即变换下的不变性。给定某数据或问题，如果存在越多能够保持不变性的变换，这个数据或问题则越为简单，也就越有可能发现高效的处理数据、解决问题的方法。

本书的主体内容是笔者博士阶段研究工作的总结。本书聚焦于一类特殊的对称性：图对称性，系统地介绍了图对称的概念、度量、模型与应用。内容从理论构建到工程落地，涉及代数图论、图数据库与隐私安全，相关成果发表在物理学、模式识别与数据管理等众多领域的国际重要期刊或会议上。以现在的眼光评判十年前的博士研究工作，不得不发出勇气可嘉的感叹。如此横贯理论与应用，不顾及学科藩篱，恣意驰骋于不同学科之间，这需要无知无畏的勇气，需要追根究底的决心，需要百折不挠的韧劲。现在想来，年轻的“我”是有足够资格睥睨现在的“我”。

当本书工作最早完成之时，21 世纪的第一个十年即将结束，大数据与人工智能的时代号角渐渐奏响。计算机学科的另一个重要问题——机器智能，渐渐重新进入很多计算机学者的视野之中。智能是什么，如何让机器具备智能成为了计算机学研究者的重要课题。近几年取得飞速发展的智能技术本质上是大数据智能，是大数据驱动下的机器智能。在 21 世纪的第二个十年即将结束之际重新盘点过去二十年的计算机技术进步，我发现复杂性问题与机器智能问题本质上是一奶同胞。

当我们欣喜于大数据智能给我们带来的全新机会时，不要忘记了是复杂性的攻克给我们征服大数据的手段。征服复杂性的有效手段使得我们在新的数据与问题规模尺度上能够重新思考机器智能的实现问题。

因此，在这个时间点给读者呈上这本以复杂性为核心命题的专著，是我本人的偶然行为，但抑或也是在大数据智能技术飞速发展的今天对于复杂性的再次思考的一种必然趋势。我相信，复杂性问题与机器智能的思想碰撞会激荡出最为绚烂的火花。

肖仰华

2018年9月

前 言

随着真实网络数据的大量涌现,通过分析网络的统计性质,研究真实网络系统的结构和功能成为了众多学科的研究热点。虽然已经有大量的工作致力于探索网络的各类性质以及相应的网络形成机制,但这些研究(除了最近的少数工作之外)都忽视了网络结构的一个基本性质——对称性。真实网络研究的盛行也使得图数据管理问题成了数据管理领域的研究热点。然而现有的图数据管理领域的研究工作多是面向关系或 XML 数据的数据管理技术进行扩展,用于解决面向图的数据管理问题,而很少考虑图数据的基本性质——对称性。在很多图数据管理问题的方案中对于图对称的基本问题,多是采取一种回避的策略。

然而,网络对称性研究具有重要的理论和实践价值,是不应忽视的。一方面,从真实网络性质的研究角度来看:网络对称性是存在于各类真实网络中的普遍现象;网络对称性与网络的很多重要功能或性质密切相关,比如网络异构性、健壮性、网络约简等。另一方面,从图数据管理的角度来看:图对称是图数据管理的基本问题,很多数据管理问题包括图检索、图挖掘、社会网络隐私保护问题最终都归结为图对称相关问题,比如(子)图同构判定、设计图的完全不变量、构造自映射等价等问题;图对称刻画了图数据的规则性和简单性,因而寻找和利用图数据中的对称性是高效解决图数据管理问题的基本思路之一。

因此,直接以真实图或网络数据中的对称性为研究对象,发展真实网络对称性基本理论,研究对称网络生成模型,并将对称性应用到一系列具体应用,包括利用对称构造合理的网络度量,利用对称寻找网络骨架,利用对称压缩最短路径索引等,就成为了本书的主要研究内容。具体而言,本书包含以下研究工作并做出了相应的贡献。

(1) 针对普遍存在于真实网络中的对称性,阐述了真实网络对称产生的微观机制,并提出了相应的网络生成模型,成功地再生了真实网络中的对称性。具体而言,通过对网络中局部对称子团的统计分析,证实了相似链接模式,也就是网络中度相同或相近的节点倾向于有着相同或相似的邻居,是网络对称产生的微观机制。

(2) 针对现有的基于度的结构熵度量不能准确刻画网络异构性这一问题,提出了基于自映射分区结构熵度量;通过理论分析和大量真实网络数据上的统计分析证实了基于自映射分区结构熵能够较为准确地刻画网络异构性。针对现有

基于结构的图距离度量精度不高的问题,提出了基于图的子结构信息的图距离度量,并将其成功应用于人群结构分析之中。

(3) 提出了通过约简网络结构冗余信息,提取网络结构骨架的方法;通过理论分析以及各类真实网络关键指标的统计分析证实了使用新方法得到的骨架——网络商,能够在显著约简网络规模的同时,保持原网络重要性质包括复杂性以及通信性质。将网络商成功地应用于社会网络隐私保护问题中,提出了网络商的一个改进版本——B-骨架,并利用 B-骨架有效地保持了 k -对称匿名模型的可用性。

(4) 将网络对称性理论应用到最短路径索引问题;发展了局部对称理论、轨道邻接性理论,探索了子结构在自映射作用下的性质;针对面向大图的最短路径索引存储代价较大问题,提出了基于压缩宽度优先搜索树的索引结构,以及相应的查询方案;通过真实网络和模拟网络上的大量实验以及系统的理论分析论证了基于压缩的宽度优先搜索树的索引结构,可以在保证查询性能的前提下显著压缩索引空间。

本书所涉及的大部分工作是我在攻读博士学位期间完成的。所以我首先要感谢培养了我的复旦大学。我在这里学习,在这里工作。复旦的一草一木对于我都有特别的感情。复旦一定程度上塑造了我的思想,而思想是一个人的灵魂。因此,复旦是我的再造者。其次,我要感谢我的博士生导师,复旦大学计算机学院汪卫教授。千里马常有,而伯乐不常有。汪卫教授的理解与肯定是取得本书中的一系列成果的前提。同时还要感谢与本书成果相关的几位合作者,包括 Momiao Xiong 教授(美国德州大学公共卫生学院)、Jian Pei 教授(Simon Fraser 大学、京东)、Ben D Macarthur 教授(英国南安普顿大学)、吴文涛博士(美国微软研究院)。没有与这些杰出学者的讨论,是不可能形成这些成果的。本书得以付梓,要感谢上海市科委大数据专项(No.13511505302)的资助。同时还要感谢我的博士生刘井平以及其他几位同学的文字校对工作。最后感谢我的家人,家人的理解与支持是我的最强后盾,与他们分享工作中的收获、生活中的喜悦是我人生中最美好的记忆片段。

目 录

序	
前言	
第 1 章 绪论	1
1.1 概述	1
1.2 一般对称性	5
1.2.1 对称性基本内涵	5
1.2.2 对称性的类别	6
1.3 研究背景	8
1.3.1 作为真实网络的重要性质之一的对称性	8
1.3.2 作为图数据管理重要问题之一的对称性	10
1.3.3 作为一种技术手段的图对称	13
第 2 章 网络对称性	14
2.1 图论基础	14
2.1.1 图的基本概念	14
2.1.2 图之间的关系	18
2.2 代数基础	24
2.2.1 集合和群	24
2.2.2 置换以及置换群	25
2.2.3 自映射	27
2.2.4 轨道与自映射等价性	28
2.2.5 置换群及其子群	29
2.2.6 群的生成集	30
2.3 图结构对称	31
2.3.1 图对称的概念	31
2.3.2 基本的变换操作	32
2.3.3 对称性度量	33
2.4 本章小结	35
第 3 章 对称网络模型	36
3.1 概述	36

3.2 相似链接模式	37
3.2.1 真实网络数据集介绍	37
3.2.2 对称二分子团	37
3.2.3 精确相似链接模式	39
3.2.4 非精确相似链接模式	43
3.3 对称网络生成模型	45
3.3.1 基于相似链接模式的择优链接	45
3.3.2 服从特定分布的初始度	46
3.3.3 基于相似链接模式的网络模型	47
3.4 实证分析	47
3.4.1 相似链接模式与网络对称性	47
3.4.2 没有相似链接模式时的网络对称性	50
3.4.3 SLP 模型的进一步讨论	53
3.5 本章小结	55
第 4 章 基于对称的网络度量	56
4.1 基于对称的网络结构熵	56
4.1.1 结构异构性	56
4.1.2 基于自映射分区结构熵	58
4.1.3 结构熵分析	61
4.1.4 基于对称的结构熵小结	65
4.2 基于对称的图距离度量	65
4.2.1 基于结构的图距离度量	66
4.2.2 子结构丰富性向量	68
4.2.3 基于 SAV 的图距离度量	69
4.2.4 基于子结构丰富性的图距离度量的变种	71
4.2.5 在人群结构分析中的应用	72
4.2.6 基于对称的图距离度量小结	74
4.3 本章小结	75
第 5 章 基于对称的网络结构约简	76
5.1 概述	76
5.2 基本概念	77
5.2.1 网络商	77
5.2.2 s-商	78
5.3 网络商的性质	79

5.3.1	网络商的规模	80
5.3.2	异构性	81
5.3.3	度分布	82
5.3.4	通信性质	84
5.4	网络商在社会网络隐私保护中的应用	86
5.4.1	k -对称	86
5.4.2	基于 B-骨架的可用性	88
5.5	网络商的其他应用	91
5.6	本章小结	92
第 6 章	利用图的对称性有效索引最短路径	93
6.1	概述	93
6.2	背景知识	95
6.3	算法框架	97
6.4	基于轨道的压缩	98
6.4.1	自映射作用下的子图	98
6.4.2	为每个轨道生成 BFS 树	101
6.5	压缩的 BFS 树	105
6.5.1	轨道邻接性和可达性	106
6.5.2	压缩的 BFS 树	110
6.5.3	基于压缩的 BFS 树的最短路径查询回答	112
6.6	实验结果	114
6.6.1	真实网络中的实验结果	114
6.6.2	模拟数据集上的实验结果	116
6.7	相关工作	119
6.8	本章小结	120
第 7 章	总结与展望	121
7.1	总结	121
7.2	对称技术应用局限性评述	121
7.3	进一步的研究工作	122
7.3.1	非精确对称理论及其应用研究	123
7.3.2	局部对称理论及实践研究	123
7.3.3	图稳定化过程研究	123
7.3.4	对称性在图查询中应用研究	124
参考文献		125
彩图		

第1章 绪 论

现实世界中的很多复杂系统都可以描述为实体及实体之间的关系，这样一种认识世界的方式使图或网络成为广泛应用的一种数据建模方法。图或网络的普适性使利用网络来研究现实系统的功能与性质成为可能。基于这一事实，一个新兴交叉研究领域在过去十年取得了飞速发展：复杂网络。与传统图论侧重于理想图的性质不同，复杂网络以真实世界网络的统计特性和动力学特征为主要研究对象。然而一直以来，网络的一个重要特性——对称性被忽视了。网络对称性不仅提供了新的考察复杂网络的视角，也提供了相应的方法论(置换群论)支持，从这一新的角度考察真实网络将为人们展示真实网络系统新的景象。

图或网络的普适性也使图数据管理成为近年来数据管理领域最为热门的研究主题之一。虽然在图数据管理领域，已有大量工作致力于解决各类图数据管理问题。但是，现有工作多是将面向关系或树结构(如可扩展标记语言(extensible markup language, XML))的数据库技术向图数据管理问题的简单移植。而事实上，当尝试把这些方法移植到面向图数据的真实应用中时，我们发现图数据管理的基本问题——图对称问题被刻意回避了。

本书将针对真实网络对称性的若干具体理论问题以及应用问题展开研究。

1.1 概 述

图或网络被广泛应用于描述现实世界中实体及实体之间的关系，这样一种认识世界的方式使图或网络成为广泛应用的一种数据建模方法。图或网络的普适性使利用网络来研究现实系统的功能与性质成为近年来的研究热点，也使图数据管理成为近年来数据管理领域最为热门的研究主题之一。虽然真实网络性质的研究以及图数据管理领域的研究已经取得很大进展，但是图数据的一个重要性质——图对称，一直没有得到充分研究。本书以图对称理论及其在数据管理问题中的应用为研究内容。

本节以社会网络分析与计算为例，介绍本书的主要研究思路和主要的研究内容。

一个社会网络通常表达的是实体(人或组织)以及实体之间的社会关系。给定一个社会网络，给定网络中的两个实体，查询它们之间的最短路径是进行社会网络分析的一项基本的查询问题。为了实时回答面向较大社会网络的最短路径查询，

可以物化以每个顶点为根的宽度优先搜索(breadth-first search, BFS)树, 树中的每一条从根出发的路径都表达了图中相应的最短路径^[1]。然而, 直接实现这一物化方法需要消耗 $O(N^2)$ 的存储空间。

为了方便描述, 本节以如图 1.1 所示的一个简单的社会网络为例。通过对图 1.1 的观察, 我们发现顶点 v_1 和顶点 v_2 具有如下性质: 对于任意顶点 $v \in (V(G) - \{v_1, v_2\})$, v_1 到 v 的最短路径和 v_2 到 v 的最短路径几乎完全相同, 除了边 (v_1, v_3) 和边 (v_2, v_3) 不同。事实上, 在后面我们会介绍 v_1 和 v_2 之间的这种等价性的实质是自映射等价。当顶点自映射等价时, 它们的很多结构特征是相同的。这个例子说明的是它们的 BFS 树的等价性。事实上, 在后面的章节中我们会说明在图 1.1 中, 自映射等价关系还存在于 v_5, v_6 之间以及 v_7, v_8, v_9, v_{10} 之间。它们的 BFS 树之间也存在着类似的等价性。

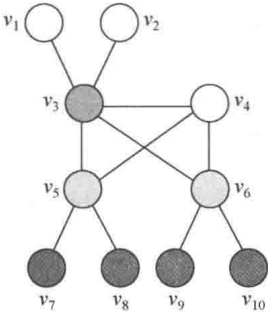


图 1.1 一个简单的社会网络

直观地来讲, 在一个图中如果能找到自映射等价关系, 那么这个图就是对称的; 并且如果找到的自映射等价关系越多, 图越对称。那么能否利用图对称性, 具体来说自映射等价的顶点在 BFS 树上的等价性, 降低最短路径查询所需的 BFS 树索引空间消耗, 这成为本书第 6 章的主要研究内容。

在社会网络中, 结构等价的顶点通常扮演着相同的角色, 因而这些等价的顶点之间通常是可以互相替换的, 因此结构等价性可以理解为一种结

构上的冗余。那么一个很自然的想法就是, 我们能够从原网络中约简掉这些冗余信息而不损失网络必要的结构信息。寻找这样一种能够保持原网络重要性质的骨架在很多问题中有着实际需求。社会网络隐私保护的可用性問題就是其中之一。本书第 5 章将对基于对称的网络约简及其在社会网络隐私保护中的应用展开研究。

当我们发布一个社会网络时, 通常需要对网络中的顶点进行匿名化处理。然而, 攻击者仍然可以根据顶点的结构特征识别出那些结构特征独特的个人(这种攻击称作结构再识别(structural re-identification)^[2])。例如, 如果知道 Bob 的邻居数目为 5, 那么在图 1.1 中, v_3 必定就是 Bob, 因为图中仅有 v_3 拥有 5 个邻居。因此, 发布社会网络数据时, 人们自然地想知道, 在结构再识别这种攻击下, 网络实体隐私泄露的风险有多大。显然, 网络实体隐私泄露的风险与网络的结构异构性有着直接关系。因此, 网络隐私泄露风险的度量就可以转换为网络结构异构性度量问题。现有的基于顶点度的网络异构性度量不能很好地刻画网络隐私泄露风险。

例如,在图 1.1 中顶点 v_1 和 v_7 的度都为 1,但是它们的邻居的度不一样,攻击者仍然可以利用这一信息进行攻击。而利用自映射等价关系可以很自然地将网络顶点划分为结构特征不同的等价类,因而成为度量网络的结构异构性的理想方法。这构成了本书第 4 章的研究内容之一。

在对社会网络中的顶点进行角色分析时,通常需要评价两个顶点在网络中是否扮演着相同或者相似的角色,其中角色相似性判定的一个重要的依据是顶点的邻居图(neighbor graph)^[3](由某个顶点的邻居集导出的子图)在结构上的相近程度。这其中的关键问题是如何度量两个顶点邻居图的相似程度或者它们之间的距离。相同规模的图,由于对称性的不同,其子结构的模式也就是不同构的子结构数量也不相同。通过对比图中蕴含的子结构模式来计算图的相似程度,不仅可以提高图距离度量精度,同时在生物网络分析中能够较好地体现生物学含义(如某个子结构模式通常可以理解为某种生物功能模块)。基于这些事实,本书第 4 章提出了一个子结构信息的图距离度量。

本书研究的另一个基本问题是对称网络生成模型,这是研究网络对称的首要问题。给定一个对称的社会网络,如何生成一个对称性接近真实社会网络的模拟网络,是将网络对称性理论进一步应用的基础问题。而为了生成对称网络,必须探索真实网络的对称产生机制,并提出相应的网络生成算法。这些研究构成了本书第 3 章的研究内容。

在深入介绍这些内容之前,必须在数学层面理解网络对称性的精确含义,了解网络对称性的基本性质。这构成了本书第 2 章的主要内容。

需要指出的是,上述网络对称性相关问题的研究不局限于社会网络,在一般网络上也存在着同样的问题。虽然上述很多问题是从社会网络分析的角度引入的,但事实上,这些问题有其独立的研究价值。例如,图距离度量就是一个在多个学科如模式识别、图数据管理、化学信息学中都有着重要研究价值的基础问题。

概括来讲,本书的研究内容为:以真实网络数据,如生物网络、社会网络数据为研究对象;从网络结构对称性角度研究网络的基本性质,提出对称网络生成模型;并将网络对称性理论应用于具体问题,包括网络度量(网络异构性度量和图距离度量)、网络结构约简以及降低最短路径索引空间等。本书的研究框架如图 1.2 所示。

本书的主要贡献可以概括如下。

(1) 系统地综述了(图)对称性研究的理论及实践意义;结合图数据管理问题,对于图论领域、代数领域的重要理论结果进行了系统的评述。

(2) 首次针对存在于真实网络中的对称性,阐述了网络对称性的产生机制,并提出了相应的对称网络生成模型。该模型可以广泛用于改善现有的真实网络模拟

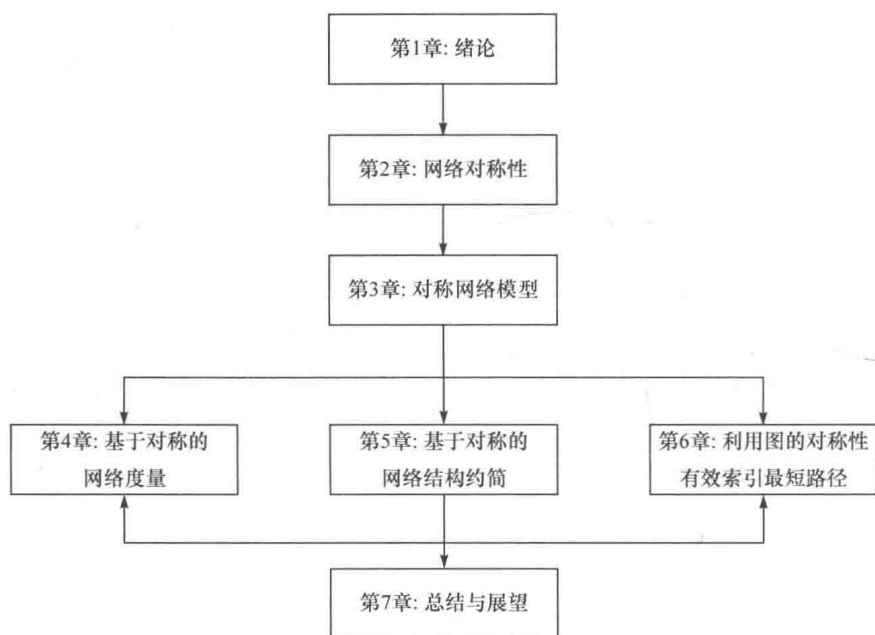


图 1.2 研究框架

以及基于此的一系列实际应用问题。

(3) 首次提出了结构异构性概念,并基于网络对称性,提出了能够较为合理地度量网络异构性的新的结构熵度量。这一新的度量,可以应用于刻画社会网络隐私泄露风险,刻画网络复杂性等一系列问题。基于子结构丰富性,提出了泛化的基于结构的图距离度量,提高了图距离度量的精度。

(4) 首次提出了通过约简对称性所刻画的网络结构冗余提取网络骨架的方法;通过实证证实,所得到的网络骨架在规模上显著小于原网络,且能够保持原网络的通信性质以及复杂性等重要性质。将网络商应用于社会网络隐私保护问题,提出了网络商的一个改进版本——B-骨架,并成功利用B-骨架保持了 k -对称匿名模型的可用性。网络商的研究还可以广泛应用于图数据管理各类问题,如网络的精简存储与表达等,也可以应用于各类物理系统、生物系统结构的优化设计等。

(5) 首次将网络对称性理论应用到图数据管理问题。发展了局部对称理论、轨道邻接性理论,探索了子结构在自映射作用下的性质。针对最短路径索引开销较大问题,基于网络对称性理论,提出了基于压缩BFS树的索引结构,大量实验和理论分析说明基于对称的方法在实现实时最短路径查询的同时,有效地压缩了最短路径索引所需要的空间开销。此项成果为图对称技术在其他图数据管理问题中的应用奠定了基础,具有示范意义。

本章剩余内容的安排如下。图结构对称是对称性的一种,1.2节将详细阐述对

称性基本内涵、对称变换以及对称性基本原则。1.3 节将主要从复杂网络研究和图数据管理研究两个角度阐述图对称研究的背景, 论述网络对称性研究的价值。

1.2 一般对称性

1.2.1 对称性基本内涵

对称(symmetry)这个词源自希腊语 *sun*(相同的)和 *metron*(尺寸), 最初用来表示相同的尺寸。德国数学家 Weyl^[4]曾对对称性作了如下定义: 如果一个操作能使某系统从一个状态变换到另一个与之等价的状态, 即系统的状态在此操作下保持不变, 则该系统对这一操作对称, 这一操作称为该系统的一个对称操作。简而言之: 对称刻画了系统在一组变换下的不变性。现实生活中普遍存在着对称性, 如人和一些动物的形体, 建筑物的结构, 各种花瓶、花边等是对称的典型例子。对称性自从被提出, 就成为很多自然学科包括生物学、物理学、化学等的重要研究内容。

近年来, 对称性的研究日益成为自然科学领域的研究热点。在自然科学哲学领域, 先后出版了好几部专著讨论广泛存在于生活以及自然学科中的对称性现象以及基本原则。其中, 德国哲学家 Mainzer 于 2005 年出版了 *Symmetry and Complexity: The Spirit and Beauty of Nonlinear Science*^[5]一书, 系统综述了存在于文化、经济、社会、物理、化学、数学以及信息科学领域中的对称性现象。2007 年, Springer-Verlag 出版的自然科学前沿系列丛书, 其中有两本都是关于对称的, 包括 Muller 编著的 *Asymmetry: The Foundation of Information*^[6]和 Rosen 编著的 *Symmetry Rules: How Science and Nature are Founded on Symmetry*^[7]。Muller 的书从对称角度向人们揭示了信息的本质是非对称。Rosen 的书详细阐述了自然科学与对称之间的紧密关系, 向人们展示了自然科学是如何建立在对称基础之上的。无独有偶, 2008 年诺贝尔物理学奖, 授予了发现亚原子物理学中自发对称性破缺机制的南部阳一郎和发现有关对称性破缺的起源的小林诚、益川敏英。此外, 2003 年来自全世界的一些知名科学家包括数学家、物理学家、化学家、生物学家以及人文社会科学领域的专家共同发起成立了国际对称协会(International Symmetry Association), 并发行了专刊 *Symmetry*, 每年组织召开学术年会(称为 *Symmetry Festival*), 旨在推动在自然科学、工程、人文、社会等领域的对称性研究。

对称性受到如此广泛的重视不是偶然的。对称性研究的重要意义可以归纳为以下几点。

(1) 对称性是普遍存在的现象。对称性是复杂系统包括生物系统、社会系统、经济系统、技术系统中普遍存在的现象^[4, 5, 8, 9]。对称现象在日常生活中也是随处

可见的。

(2) 事物的演化过程可以理解为对称不断破缺的过程，从对称走向非对称的过程^[5, 10, 11]。各种不同的系统的演化动因从对称角度都可以归结为对称破缺，是对称破缺产生了世界的多样性和复杂性。

(3) 复杂性可以刻画为对称破缺^[12]。对称破缺的程度决定了事物的复杂程度；事物的非对称性越强，事物越复杂。对称性与复杂性之间的对立关系，为理解和征服复杂性提供了一种新的途径。影响系统复杂性的因素往往是纷繁多变的，通常难以准确刻画各因素之间的复杂关系。因而直接理解和征服复杂性面临巨大挑战。而刻画系统对称性，只需寻找各类变换操作与相应的不变量。因而，寻找隐藏于系统中的对称性则从复杂性的对立面为征服复杂性提供了全新途径。

1.2.2 对称性的类别

Weyl 对于对称性的定义只是一个一般性的框架，随着研究对象的不同，对称性的具体含义也不完全相同。但是不论对于哪种研究对象，理解对称性含义的关键都是两点：变换和不变量。目前根据对称性的研究对象，可以将对称分为三大类：空间对称、时间对称和尺度对称。

首先介绍时间对称和尺度对称。如果一个系统的某一变量满足 $F(t) = F(-t)$ ，则称该系统是时间对称的。时间对称本质上体现了系统的一种周期性。其变换操作是时间的推移。很多真实系统，如商品价格都呈现出这种周期性。尺度对称是指系统在不同尺度下拥有相同或相似的性质。显然这里的变换是考察系统的尺度变换。尺度对称的本质是分形。例如，图 1.3(a)所示的分形图案，在不同尺度下考察都有着相同的基本组成单元。再如图 1.3(b)，不同粒度下网络的基本组织方式都是四个点的完全图。

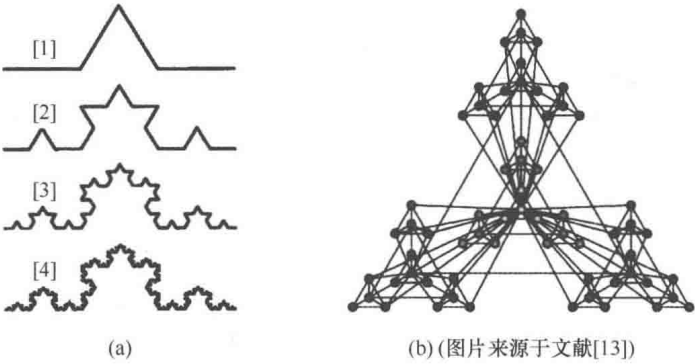


图 1.3 尺度变换下的对称

空间对称是在日常生活中经常可以见到的对称类型，例如，图 1.4 所示的古

印度宗教图案和中国的太极图案在直觉上都是相当对称的,蝴蝶的形状也是比较对称的。这种直觉上的对称可以理解为,当我们沿着某个中心点旋转图案,或者沿着某个轴线左右折叠交换时,可以得到与原图完全重叠的图案。因此,空间对称实质上是指在镜像变换或旋转变换等操作下,图案严格保持不变。

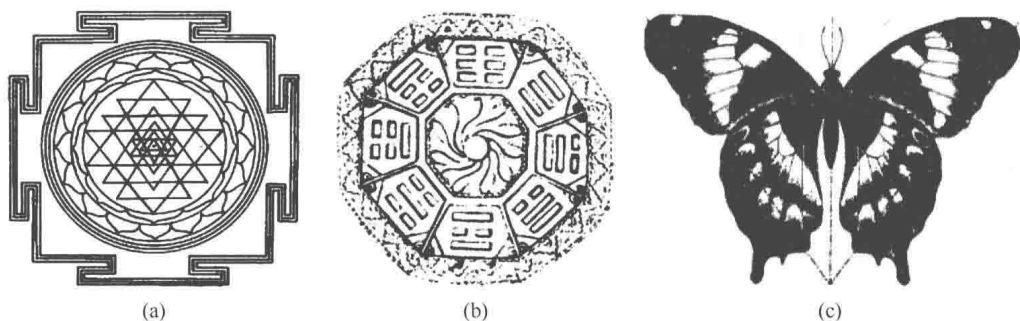


图 1.4 空间对称(图片来源于文献[8])

事实上,空间对称根据变换操作的不同可以分为三类,除了上面提及的镜像对称、旋转对称,还包括平移对称。需要指出的是,平移对称理论上只适用于存在于无限空间的图形。图 1.5 给出了镜像对称、旋转对称和平移对称的例子。图 1.5(a)旋转特定角度之后可以和原图重叠,沿着虚线左右交换也可以得到与原图重叠的图案。图 1.5(b)所示的图案是一个无限图的局部,在无限空间中,沿着特定方向平移可以与原图重叠。

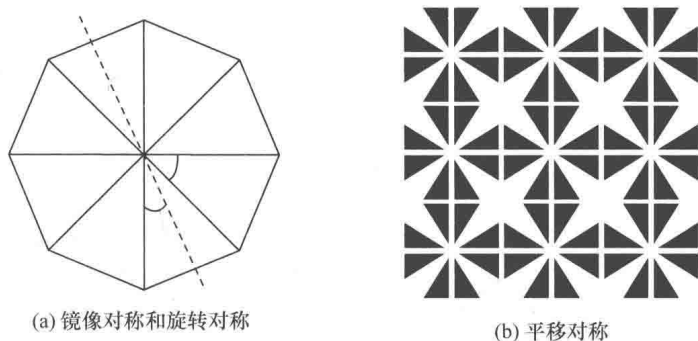


图 1.5 镜像对称、旋转对称与平移对称(图片来源于文献[8])

图结构对称可以理解为空间对称的一种。由于真实的图或网络都是有限的,所以图结构对称一般只包括镜像对称和旋转对称。与一般的空间对称不同的是,图结构对称的不变量是顶点集上的邻接关系,而不再是空间图形的形状。如图 1.6 所示的有着四个顶点的环,具有高度对称性。沿着顺时针或者逆时针方向旋转交换四个顶点不会改变顶点之间的邻接关系。沿着图中虚线所示的任意轴线交换顶点,顶点之间的邻接关系也可以得到保持。