

《智能信息处理》课程作业

## 基于多模态图像检索的形式概念分析

李怀清

作业	分数[20]
得分	

2021 年 11 月 28 日

# 基于多模态学习方法的形式概念分析

李怀清

(大连海事大学 信息科学技术学院, 大连 116026)

**摘要:** 形式概念分析是从形式背景进行数据分析和规则提取的强有力的工具。形式概念分析以数学为基础而建立, 对本体中的概念、属性以及相互关系等用形式化的语言表述, 然后根据语境构造出概念格, 清晰的表达出本体的结构。本文首先介绍了几种常见的多模态学习方法, 并对形式概念分析进行详细描述, 最后对几种常见的多模态学习方法进行形式概念分析, 给出了从概念得到形式概念, 从背景转化为形式背景, 从形式背景得到约简形式背景, 之后构造出概念格的全过程。

**关键词:** 形式概念分析; 概念格; 形式背景; 多模态学习

## A formal conceptual analysis based on multimodal learning methods

Li Huaqing

( School of Information Science and Technology, Dalian Maritime University, Dalian 116026)

**Abstract:** Formal conceptual analysis is a powerful tool for data analysis and rule extraction from formal background. Formal concept analysis is based on mathematics, formal language expression of concepts, attributes and interrelations in the noumenon, and then the conceptual grid is constructed according to context, and the structure of the noumenon is clearly expressed. This paper first introduces several common multimodal learning method, and the formal concept analysis are described in detail, finally, several common methods of multimodal learning formal concept analysis, gives a formal concept, from concept from background into form background, from the form background get reduction form background, and then constructs the concept lattice.

**Key words** Formal concept analysis; Formal background; Concept lattice; Multimodal learning.

## 0. 引言

早在公元前 4 世纪,多模态的相关理念和理论即被哲学家和艺术家所提出,用以定义融合不同内容的表达形式与修辞方法。20 世纪以来,这一概念被语言学家更为广泛地应用于教育学和认知科学领域[3]。近年来,描述相同、相关对象的多源数据在互联网场景中呈指数级增长,多模态已成为新时期信息资源的主要形式。人类的认知过程是多模态的。个体对场景进行感知时往往能快速接受视觉、听觉乃至嗅觉、触觉的信号,进而对其进行融合处理和语义理解。多模态机器学习方法更贴近人类认识世界的形式。

形式概念分析(Formal Concept Analysis, FCA)这一概念是在 1982 年由 R.Wille 提出的。形式概念分析是应用数学的一个分支,是信息处理、知识处理的一种理论。本文首先介绍了几种常见的多模态学习方法,并对形式概念分析进行详细描述最后对几种常见的多模态学习方法进行形式概念分析。

## 1. 多模态学习方法

在这一部分,我们介绍了四种具有代表性的多模态学习方法:

### 1.1 Relationship<sup>[1]</sup>

一个关系推理模块。是在 VQA 任务中捕获关系推理的方法。它采用从 CNN 的最后一层提取的特征图和从 RNN 提取的文本特征,然后创建一组关系特征,每个关系特征是二维特征图中文本特征  $\phi_t$  和 2 个局部特征的拼接,成对的关系简单地求和;然后通过一个 MLP 来学习跨模态关系。应用另一个 MLP 来获得  $\phi_{xt}$  输出。

### 1.2 FiLM<sup>[2]</sup>

特征式线性调制组件。它包含了一个堆叠的三个电影层级联后,一个 CNN。从 RNN

中提取文本特征来表示文本信息,通过仿射变换对每个特征映射进行调制。FiLM 只是一个简单的仿射变换,但是有相当大的容量,可以广泛应用在 NLP、CV 任务中。实际上,其大容量是借助 FiLM 层后面的深度神经网络实现的。后续的神经网络层可以将简单的线性变化转化为复杂的非线性变化。例如,在 CNN 早期使用一层 FiLM 可以获得与在整个网络中使用 4 层 FiLM 大致相同的性能。如果 FiLM 层后面没有任何神经网络层,FiLM 的容量将非常有限,正好是简单仿射变换的容量。

### 1.3 MRN<sup>[3]</sup>

一个多模态残差学习组件。它通过融合 CNN 和 RNN 的视觉和文本特征来学习多模态表示。跨模态特征通过元素相乘和残差学习三种方法得到。

### 1.4 TIRG<sup>[4]</sup>

一种用于图像检索的图像-文本合成方法。它通过连接来组成视觉和文本特征,然后学习一个门控连接和一个用于跨模态融合的残余连接。其背后的关键思想是,文本应该修改查询图像的特征,但我们希望得到的特征向量仍然“存在”于目标图像的相同空间中。我们通过一个门控剩余连接使文本修改图像特征来实现这个目标

## 2. 形式概念

### 2.1 形式概念的定义

形式概念分析(Formal Concept Analysis, FCA)这一概念是在 1982 年由 R.Wille 提出的。形式概念分析是应用数学的一个分支,是信息处理、知识处理的一种理论。形式概念即形式对象集以及形式对象集所具有的形式属性集的总和。其定义为:

定义 1<sup>[5]</sup> 设  $X$  为形式对象集  $U$  的一个子集,  $B$  为形式属性集  $A$  的一个子集,有二元关系  $R \subseteq U \times A$ , 若  $X = \{x | x \in U, \forall a \in B, xRa\}$ ,  $B = \{a | a \in A, \forall x \in X, xRa\}$ , 则二元组  $(X, B)$  被称为形式概念。

举例如下:  $(\{\text{猪}, \text{牛}\}, \{\text{哺乳}, \text{嗅觉}\})$ 。

在该形式概念中，猪和牛都具有嗅觉，且都是哺乳动物；反过来说，具有嗅觉的是猪和牛，是哺乳动物的也是猪和牛。

2.2 形式背景

单个的形式对象集以及形式对象集所具有的形式属性集构成了形式概念，形式概念的作用是构建自然概念的层次连通结构。而自然概念则是把人们认为很自然的事物用概念的形式表现出来<sup>[6]</sup>。人脑对客观世界的认知可以构成许许多多的形式概念机，这就决定着单一的形式概念不足以描述自然界万事万物之间的关系，当不同的形势概念集相互联系时，人们就需要引入新的概念来描述它，将人脑中的形势概念集置于一张二维表中，就构成了形式背景（formal context），如图 1 所示。

	形式属性1	形式属性2	.....	形式属性n
形式对象1	√			√
形式对象2		√		
.....			.....	
形式对象n	√			√

图 1 形式背景

在实际应用过程中，一些复杂的形式背景可能会出现问题，比如对象或者属性的冗余。例如冗余信息可能表现在两个或多个不同的对象拥有相同的属性，人们在实际应用中把具有相同属性的不同对象合并来进行约简，称为“约简行”；同理，有时候不同的对象之间虽然具有不同的属性，但是属性与属性之间却存在着异构的现象，人们把这样的属性进行合并来进行约简，称为“约简列”；有时候多个属性之间能相互关联，某一个或多个决策属性可以由一个或者多个其他条件属性推出，人们通过“关联规则抽取”来进行约简。通过行或列的约简人们就可以进行信息的合并，从而消除形式背景之间的冗余。

2.3 概念格的概念

格（lattice）的意义是任两个元素的上确界和下确界都存在的偏序集。完备格为任一子集的上确界和下确界存在的偏序集，其特点是只有一个最高点，且只有一个最低点，且图中任何两点连通。概念格是元素为概念的完备格。概念格，也称为 Cralois 格，它提供了一种支持数据分析的有效工具。概念格的每个节点是一个形式概念，每一个形式概念都是由外延和内涵两部分组成<sup>[7]</sup>。概念格通过 Hasse 图生动和简洁地体现了这些概念之间的泛化和特化关系。从形式背景中生成概念格的过程实质上是一种概念聚类过程。目前，概念格已被广泛应用于机器学习、模式识别、专家系统、数据挖掘、信息检索等领域<sup>[2]</sup>。

2.4 概念格的构建

概念格的表示形式是 Hasse 图，概念格的构建的基础是形式背景，形式背景描述了多个形式概念之间的关系，单个形式概念描述了形式对象以及形式对象所具有的形式属性之间的关系<sup>[8]</sup>。所以，概念格的构建必须明确不同形式对象以及不同形式对象所具有的形式属性。

概念格的构建包含以下几个步骤：生成形式背景，约简形式背景，生成单值形式背景，确定父子关系，绘制 Hasse 图，补充各形式概念的上确界和下确界，最后获得概念格。

3. 常见多模态学习方法的形式

概念分析

对常见的多模态学习算法进行归纳总结得出，算法在是否基于特征连结，是否使用残差网络，是否适用小数据集，模型复杂度和线性上均表现为不同差异。其中表 1 为几种常见的多模态学习方法的形式背景。

表 1 几种常见的多模态学习方法形式背景

方法	特征 连结	残差 网络	适用 小数 数据集	模型 复杂 度	线性
Relationship	是	未使用	不适用	低	非线性
FILM	否	未使用	适用	中	线性
MRN	否	使用	适用	中	线性
TIRG	是	使用	适用	高	线性

因为该形式背景中不存在具有相同属性的不同对象，所以无需进行“行约简”。同理，因为不同的对象之间适用小数据集与线性两个属性之间存在异构的现象，因此需进行“列约简”。列约简后形成的形式背景如表 2。

表 2 列约简后的形式背景

方法	特征连 结	残差网 络	模型复 杂度	线性
Relationship	是	未使用	低	非线性
FILM	否	未使用	中	线性
MRN	否	使用	中	线性
TIRG	是	使用	高	线性

由表 2 形式背景导出概念格，如图 2 所示。

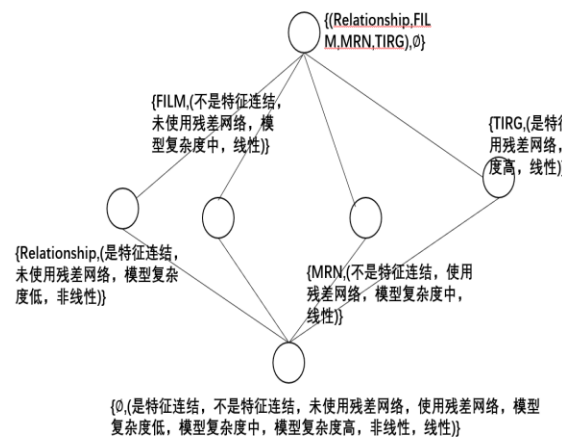


图 2 多模态学习方法概念格

## 4. 总结

在使用形式概念分析过程中，我们可以利用已知的知识或者常识，经过形式化以后得到新的知识和常识，从而帮助我们更好的理解研究的系统发现潜在的知识。

多模态学习方法的核心思想就是多个模态特征的提取和融合。随着科研工作者的进一步深入研究，未来会有越来越多的多模态学习方法。

## 参考文献

- [1] Adam Santoro, David Raposo, David G Barrett, Mateusz Malinowski, Razvan Pascanu, Peter Battaglia, and Timothy Lillicrap. A simple neural network module for relational reasoning. 2017.4,5,6
- [2] Ethan Perez, Florian Strub, Harm De Vries, Vincent Dumoulin, and Aaron Courville. Film: Visual reasoning with a general conditioning layer. In AAAI Conference on Artificial Intelligence, 2018.5,6
- [3] Jin-Hwa Kim, Sang-Woo Lee, Donghyun Kwak, Min-Oh Heo, Jeonghee Kim, Jung-Woo Ha, and Byoung-Tak Zhang. Multimodal residual learning for visual qa. In Advances in Neural Information Processing Systems, 2016.5,6
- [4] Nam Vo, Lu Jiang, Chen Sun, Kevin Murphy, Li-Jia Li, Li Fei-Fei, and James Hays. Composing text and image for image retrieval - an empirical odyssey. In IEEE Conference on Computer Vision and Pattern Recognition, 2019.2,5,6
- [5] 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.
- [6] 王娜. 基于概念格的知识获取 [J]. 科技创业, 2010, 6(4): 118-120.
- [7] 曲开社, 翟岩慧. 偏序集、包含度与形式概念分析[J]. 计算机学报, 2006, 29(2):32- 33.
- [8] 张云中. 基于形式概念分析的领域本体构建方法研究[D]. 吉林大学, 吉林, 2009.