

# Ontology 综述

赵玲

(大连海事大学信息科学技术学院 大连 116026 )

**摘 要** 本体论是对存在或现实的本质以及它们的关系的基本类型的哲学研究。传统上被称为形而上学哲学主要分支的一部分, 本体常常涉及关于实体存在或可能被认为存在的问题, 以及这样的实体如何根据在层级内相关以及根据相似性和差异细分进行分组。本文作者对目前 Ontology 的研究与应用现状进行了综述性地介绍, 从 ontology 的定义, 起源, 应用领域等方面进行介绍。

**关键词** ontology; 本体语言; 本体多元论  
**中图法分类号** TP 301

## Ontology Summary

Zhao Ling

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

**Abstract** Ontology is the philosophical study of the nature of being, becoming, existence or reality as well as the basic categories of being and their relations. Traditionally listed as a part of the major branch of philosophy known as metaphysics, ontology often deals with questions concerning what entities exist or may be said to exist and how such entities may be grouped, related within a hierarchy, and subdivided according to similarities and differences. In this context, the author of the current research and application of the status of Ontology are reviewed in this paper, from the definition of ontology, the origin, application fields and so on.

**Key words** ontology; ontology and language; ontological pluralism

## 1 引言

本体论是对存在性、存在性或现实性的哲学研究, 是存在与存在的及其关系的基本范畴。传统上作为哲学来讲被称为形而上学的主要分支的一部分, “本体”是‘本’和‘体’两个独立的汉字组成的汉语复合词。‘本’字的基本含义是植物的根, ‘本’字的引申义是事物的本源或来源。

‘体’字的基本含义是人的身体, ‘体’字的引申义是事物的身体或形体。“本体”这个汉语复合词

在中国文化中的基本含义是事物的主体或自身, 事物的来源或根源。

二十世纪初, 有学者用“本体”这个汉语词翻译西方哲学所追求的“超越一切存在者或现象、具有创造各种事物或现象能力”的根本实体。从此, “本体”一词成为了一个哲学名词。随着计算机科学的发展, 计算机科学界用“本体”这个词表述和界定由 Gruber 提出的“本体是概念化的明确的规范说明”这个事物, 于是, “本体”一词又成为了一个计算机科学的名词。

本体 (Ontology) 是一个概念框架, 给出一套

词汇标识一套概念, 这些词汇就是术语。本体本身也需要标识, 比如我们说“物理学”, 物理学这三个字就标识了本体。对于本体这个形而上学的东西读者没必要纠结, 只需知道在数据交换平台中计算机是使用编码来标识本体的就可以了。如“JS(教师)”标识一个本体, “XS(学生)”标识一个本体, 在教师本体概念框架下“性别”指的是教师的性别, 而在学生本体概念框架下“性别”指的是学生的性别, 再比如: 在教师本体框架下有“所教学科”这样的概念但在学生本体下没有, 而在学生本体下会有“家长联系电话”这样的概念但在教师本体下没有。

虽然本体在哲学领域是高度理论化的, 但它在信息科学和技术, 如本体工程中有实际应用。

一些哲学家, 特别是柏拉图学派, 主张所有名词(包括抽象名词)是指存在的实体。其他的哲学家认为, 名词并不总是命名实体, 但一些提供了一种用于引用一个对象或事件的集合的一种速记。在这后一种观点, 思想, 而不是指一个实体, 指的是一系列的心理活动的个人经历; 社会指的是一系列有一些共同特征的人, 和几何指的是一系列的一种特定的智力活动。这两极唯实论与唯名论之间, 站各种各样的其他职位。本体可以给出一个帐户, 其中的一个词是指实体, 它不为什么, 和什么类别的结果。

本体的主要问题包括: “有什么可以说是存在的?”、“什么是一件事?”、“到什么类别, 如果有的话, 我们可以排序现有的东西吗?”、“存在的意义是什么?”、“实体的各种各样的模式是什么?”等。

不同的哲学家对这些问题提供了不同的答案。一种常见的方法包括将现存的主题和谓词划分成被称为类别的组。当然, 这样的类别列表彼此不同, 它是通过不同的分类方案的协调, 本体涉及到图书馆学和人工智能等领域。这样的理解本体论范畴, 但是, 仅仅是分类, 分类。亚里士多德的类别是一个存在的方式, 可以简单地作为一个存在, 如: 它是什么、它是多少(定量)、它在哪里、它与其他生物的关联。

存在本体论问题的进一步的例子包括: 什么是存在, 也就是说, 它意味着什么呢? 是否存在一个属性? 是否存在一个属或一般类, 只是简单地划分了具体的差异? 哪一实体, 如果有的话, 是基本的? 都是实体对象吗? 一个对象的属性如何与对

象本身有关? 物理属性实际上是存在的吗? 什么样的功能是必不可少的, 而不是仅仅是一个给定的对象的偶然属性?

基本本体论的二分法, 包括: 共性的详情、物质和事故、抽象和具体的对象、本质与存在、决定论与非决定论、一元论和二元论、唯心主义与唯物主义。

哲学家可以分类本体的各种方式, 使用的标准, 如抽象程度和应用领域;

上层本体论: 支持本体论发展的概念、元本体论;

领域本体: 一个特定的主题或感兴趣的领域的概念, 例如, 信息技术或计算机语言, 或特定的科学分支;

接口本体论: 与两个学科衔接的概念;

过程: 输入, 输出, 约束本体。

## 2 Ontology

### 2.1 ontology 定义

Ontology 最早是一个哲学上的概念, 从哲学的范畴来说, Ontology 是客观存在的一个系统的解释或说明, 关心的是客观现实的抽象本质。在人工智能界, 最早给出 Ontology 定义的是 Neches 等人, 他们将 Ontology 定义为给出构成相关领域词汇的基本术语和关系, 以及利用这些术语和关系构成的规定这些词汇外延的规则的定义。

1993 年, Gruber 给出了 Ontology 的一个最为流行的定义, 即 Ontology 是概念模型的明确的规范说明。后来, Borst 在此基础上, 给出了 Ontology 的另外一种定义: Ontology 是共享概念模型的形式化规范说明。Studer 等对上述两个定义进行了深入的研究, 认为 Ontology 是共享概念模型的明确的形式化规范说明。这包含 4 层含义: 概念模型(conceptualization)、明确(explicit)、形式化(formal)和共享(share)。概念模型是指通过抽象出客观世界中一些现象(Phe-nomenon)的相关概念而得到的模型。概念模型所表现的含义独立于具体的环境状态。明确指所使用的概念及使用这些概念的约束都有明确的定义。形式化指 Ontology 是计算机可读的(即能被计算机处理)。共享指 Ontology 中体现的是共同认可的知识, 反映的是相关领域中公认的概念集, 即 Ontology 针对的是团体而非个体的共识。Ontology 的目标是捕获相关领域的知

识, 提供对该领域知识的共同理解, 确定该领域内共同认可的词汇, 并从不同层次的形式化模式上给出这些词汇(术语)和词汇间相互关系的明确定义。

## 2.2 起源

“本体”这个哲学概念是人类思想认识活动产生的必然结果。求知是人类的天性, 每个人经过思想认识活动都会从个别事物的产生都有一定原因的经验知识出发, 经过演绎推理获得一个全部事物的产生必然有一个本源或本体的结论, 在大脑中产生出“本体(本体事物)”的概念, 都会产生出寻求天地万物的本体或本源, 获得本体事物全面知识的意愿或意识, 都会提出天地万物的本体或本源究竟是什么? 这样的哲学问题。

早在两千多年以前, 人类就已经从天地万物中发现、界定、彰显、区分和产生出了本体(本体事物), 创立了关于事物“本体(本体事物)”的哲学概念。在中国文化中“本体”这个词经常被用来指事物的来源或根源, 本体就是本源。中国古代哲学家老子在《道德经》中指出: “道”具有产生天地万物的能力, 是天地万物的母亲, 是天地万物的根源或来源。“有名”和“无名”是“道”的两个具体存在和表现形式。

同“本体”、“道”、“有名”、“无名”这些中国哲学名词含义相同的希腊词语主要有三个: 第一个是“onta(音译: 安特)”, 第二个是“logos(音译: 逻各斯、逻辑)”, 第三个是“nuos(音译: 奴斯)”。希腊语词“onta(意译: 本体、道、存在、有、是)”, “logos(意译: 言说、思想活动、道、本体)”, “nuos(意译: 心智、思维、认识、本体、道)”相当于英语的“being(音译: 被)”一词。

古希腊的哲学家赫拉克利特用 logos(逻各斯、逻辑)这个词表述和界定能够形成或产生现象的根本实体。赫拉克利特指出: logos(言说、道、思想活动)是 onta(万物生灭的尺度和根据、本体、存在), 一切存在者均在 onta(存在、万物生灭的尺度或根据、本体)之中”。巴门尼德提出了“onta(存在)之外并无非存在”、“onta(存在、本体、思想)是一”、“onta(存在、本体、道、有、是)与思维同一”的思想观点。从柏拉图开始, 希腊哲学对“onta(本体、道、存在)”的思考降格为对“存在者(万物)”的思考, 对“存在者(万物)”的讨论代替了对“存在(本体、道)”本身的追问,

这种忽略存在者和存在、万物和道、本体事物和非本体事物差别的思想行为, 是自柏拉图以来哲学探讨本体或存在问题的主流思想行为方式。

中国的语言学家把希腊文单词 Onta 和英文单词 being 翻译为“存在”、“本体”、“有”、“是”。日本学者把 17 世纪的德国经院学者郭克兰纽(Goclenius, 1547-1628)首先使用“Ontology”一词翻译为本体论。我认为应该把希腊文的 Ont 和英文的 being 翻译为中文的“本体”、“道”、“良知”, 不应该把希腊文单词 Ont 和英文单词 being 翻译为“存在”、“有”或“是”。

西方哲学对“本体(本体事物)”事物的研究始于古希腊时期, 从古希腊的米利都学派开始, 哲学家们就努力寻找万物的始基, 后来发展为对 onta(本体、道)和 logos(逻各斯、言说、道)的探讨。巴门尼德通过对 onta(本体、道)的探讨, 明确了找到 onta(本体、道)的正确方向和道路。他指出: 对于被万物所分有的本体(被“是者”所分有的“是”、被“存在者”所分有的“存在”)只有通过人类的思想认识活动向超验领域中寻找, 不能通过感觉从经验领域中获得。

## 2.3 本体多元论

与电子一元论相反的是存在的多元概念。在公元前 5 世纪, Anaxagoras 和 Leucippus 取代了存在的现实(独特和不变)与成为的现实, 因此更基本和基本的多元。本论文起源于希腊世界, 由 Anaxagoras 和 Leucippus 以两种不同的方式表述。第一种理论处理各种物质的“种子”(亚里士多德称之为“homeomerics”)。第二个是原子论, 它处理现实是基于真空, 原子及其内在运动。

Leucippus 提出的唯物主义原子主义是不确定的, 但后来由 Democritus 以确定性的方式发展起来的。后来(公元前 4 世纪), 原始的原子主义再次被作为不确定的由 Epicurus。他证实了现实是由无限的不可分割的, 不可改变的小体或原子组成的(atomon, lit.'uncuttable'), 但他重视原子的特征, 而对于 Leucippus, 他们的特点是“数字”, “秩序”在宇宙中的“位置”。它们除了在真空中产生整体运动, 产生多样化的流动。他们的运动受到薄壁组织的影响(Lucretius 命名它是门诊), 这是由机会决定的。这些想法预示着我们对传统物理学的理解, 直到原子的性质在 20 世纪被发现。

### 3 Ontology 应用

#### 3.1 本体语言

一些哲学家认为,“什么是什么?是(至少部分是)使用问题,而不是关于事实的问题。这个观点由唐纳德·戴维森(Donald Davidson)所做的比喻来传达:假设一个人将“杯子”称为“椅子”,并对杯子做出一些评论,但是使用“椅子”一词而不是“杯子”。人们可能很容易地发现,这个人简单地称为“杯子”和“椅子”,奇怪的是解释。类似地,如果我们发现人们断言“有”这样的人,而我们不是自己认为“这样的”存在,我们可以得出结论,这些人不是坚果(戴维森称这种假设为“慈善”),他们只是使用“有”与我们不同。什么是的问题?至少部分是语言哲学中的一个主题,并不完全是关于本体本身。这个观点已经由 Eli Hirsch 表达。

过去的几年里已经开发了五种基于 XML 的 Ontology 语言。有些是直接基于 XML 语法,例如基于 XML 的本体交换语言(XOL, XML-Based Ontology Exchange Language)、简单的 HTML 本体扩展(SHOE, Simple HTML Ontology Extension)和本体标记语言(OML, Ontology Markup Language),另外的两种语言是建立于 RDF(S)RDF 和 RDFS 的并集之上,以便改善 RDF(S)的特征:本体交互语言(OIL, Ontology Inter-change Language)和 DAML(DARPA Agent Markup Language)+OIL。

基于 XML 的本体交换语言(XOL)为了在异构的软件系统中进行本体定义的交换,美国的生物信息学领域设计了 XOL。在研究了生物信息学专家的代表性需求后,研究人员创造了 XOL。他们基于 XML,选取 Ontolingua 和 OML 作为生成了 XOL 的基础,并结合开放知识库的连通协议(Open KnowledgeBase Connectivity protocol)的子集 OKBC-Lite 的高级表现和 OML 的语法。没有工具支持使用 XOL 开发本体论。然而,由于 XOL 文件使用 XML 语法,因此可以使用 XML 编辑器来生成 XOL 文件。

美国 Maryland 大学开发了 SHOE,并用它来开发 OML。SHOE 是 HTML 的扩展,在 HTML 文档或其它 Web 文档中结合了机器可读的语义知识。最近, Maryland 大学已经把 SHOE 语法适应到 XML。SHOE 能够使得 Agent 收集 Web 页面和文档

的有意义的信息,改善搜索机制和知识聚集。这个过程由三个阶段构成:定义本体;用本体论的信息注释 HTML 页面,以便描述自身和其它的面; Agent 通过搜索现有的全部网页来语义地检索信息,并一直保持信息的更新。

#### 3.2 本体和人类地理

在人类地理中,存在两种类型的本体:小的“o”,其说明了实际取向,描述了作为组的一部分的功能,认为过度简化和忽略关键活动。另一个“O”或大的“O”,系统地,逻辑地和合理地描述了本质特征和普遍特征。这个概念与柏拉图的观点密切相关:人类思想只有在他们继续生活在他们的“洞穴”的范围内时才能感觉到一个更大的世界。然而,尽管存在差异,本体论依赖于成员之间的象征性协议。也就是说,本体论对于公理语言框架至关重要。

根据 A.N.Whitehead,对于本体论,区分术语“现实”和“现实”是有用的。

在这种观点中,“实体”具有基本本体论优先性的哲学地位,而“真实实体”是可能是实际的实体,或者可以从与某些实际实体的逻辑关系得出其现实。例如,苏格拉底生活中的一个场合是一个实体。但是苏格拉底作为一个人不会使“人”成为一个实体,因为它不确定地指向许多实体,例如在苏格拉底生活中的几个场合,以及在阿尔喀比亚德和其他人的生活中的几个场合。但人的概念是真实的;它从它提到的那些许多实际场合得出其现实,每一场合都是一个实际的实体。实际情况是具体实体,而诸如“人”的术语是来自许多具体相关实体的抽象。

根据 Whitehead,实体必须通过满足几个哲学标准获得其基本本体论优先的哲学地位,没有一个实体的背后,找到事实或功效更基本的东西。这个标准被认为是表达一个公理或假定的尊敬原则。

实际实体必须是完全确定的,在某种意义上,可能没有关于其身份的混淆,这将允许它与另一个实际实体混淆。在这个意义上,一个实体是完全具体的,没有潜力成为别的东西。就是这样。它当然是创造其他实体的潜力的来源,它可以说是一个部分原因。同样,它是其它实际实体的潜力的固化或实现,这是其部分原因。

实体之间的因果关系对其实际性至关重要。因此,对于 Whitehead,每个实际实体在物理闵可夫斯基空间中具有其明确和确定的扩展,因此是唯一

可识别的。Minkowski 空间中的描述支持特定观察者在时间和空间上的描述。

这是 Whitehead 这样一个本体的哲学的目标的一部分, 实际实体应该都是一样的, 是实体; 他们应该都满足一个明确的一套明确的现实本体论标准。

Whitehead 建议他的经验场合的概念满足其作为实际实体的哲学上首选定义的地位的标准。从纯逻辑的角度来看, 每种经验都充分地体现了客观和主观现实的特征。主观性和客观性是指一个经验场合的不同方面, 而不是它们彼此排斥。

在这个观点中, 其他哲学建议或候选人作为实体的例子是亚里士多德的“物质”, 莱布尼兹的原子和笛卡尔的“现代”, 以及更现代的“事物状态”。亚里士多德的物质, 如苏格拉底, 已经作为更基本的“主要物质”, 在这个意义上, 不符合怀特海的标准。Whitehead 对 Leibniz 的 Monad 不满意, 因为他们是“无窗”的, 不会导致对方。“事务状态”往往没有被严格定义, 往往没有具体提及在实体间可夫斯基空间的延伸; 他们因此不一定是成为的过程, 但可能是如他们的名字所暗示的, 在某种意义上简单的静态。事务状况取决于具体情况, 因此有其背后的事情。Whiteheadian 实体的一个总结是, 它是一个成熟的过程。另一个总结, 指的是其与其他实体的因果联系, 是与“莱布尼茨”的无窗原子相比, 是“所有窗口”的含义。

## 4 结束语

目前, ontology 工程是一门艺术, 而非一门科学, 其原因是现在还没有统一的生命周期的定义和标准, 也没有统一的驱动 ontology 开发的方法学和技术。ontology 的核心思想主要包括两个方面: 一是实现共同认识; 另一个就是实现可重用性。

Plato.stanford.edu. Retrieved 2013

- [4]. Yan, Z., Chakraborty, D. Parent, C, Spaccapietra, S. and Aberer, K. Semantic trajectories: mobility data computation and annotation[J]. ACM Trans. Intell. Syst. Technol. 2013, Article 49, 38 pages.
- [5]. Lawson, C. Latsis, J. S., & Martins, N. (Eds.). Contributions to social ontology. Routledge, 2013
- [6]. Johannes Wollbold, Ren Huber, Raimund Kinne, and Karl Erich Wolff. Conceptual Representation of Gene Expression Processes[C]// Knowledge Processing and Data Analysis, 2011:79-100

**Zhao ling** born in 1993

E-mail: 2495662643@qq.com, The main research direction is intelligent information processing

## Background

During the last three decades, formal concept analysis (FCA) became a well-known formalism in data analysis and knowledge discovery because of its usefulness in important domains of knowledge discovery in databases (KDD) such as ontology engineering, association rule mining, machine learning, as well as relation to other established theories for representing knowledge processing, like description logics, conceptual graphs, and rough sets. In early days, FCA was sometimes misconceived as a static crisp hardly scalable formalism for binary data tables. In this paper, we will try to show that FCA actually provides support for processing

## 参考文献

- [1]. Barry Smith: Objects and Their Environments: From Aristotle to Ecological Ontology The Life and Motion of SocioEconomic Units (GISDATA 8), London: Taylor and Francis, 2013
- [2]. Michal Devaux and Marco Lamanna, "The Rise and Early History of the Term Ontology (1606-1730)", Quaestio. Yearbook of the History of the Metaphysics, 9, 2009
- [3]. Ancient Atomism (Stanford Encyclopedia of Philosophy)".