

基于本体的分层抽象模型

王楠^{1,2} 欧阳彤¹ 孙善武³

(吉林大学计算机科学与技术学院符号计算与知识工程教育部重点实验室 长春 130012)¹
(吉林财经大学信息学院 长春 130117)² (吉林财经大学网络实验中心 长春 130117)³

摘要 KRA 表示模型给出了对物理世界 W 进行形式化建模的一般框架。在扩展后的广义 KRA 模型中引入本体概念,将抽象对象库扩展为本体类,在本体类中实现 KRA 模型框架中的感知层、语言层和理论层的知识共享和重用,简化了 KRA 模型表示。分别定义了作用在物理世界 W 和本体类上的 3 种本体抽象算子(集):基本本体抽象算子、实体本体抽象算子集和连接本体抽象算子集,并给出了本体类之间的映射关系,通过这种映射关系能够实现模型抽象和模型还原。对本体类的抽象度进行形式化的定义,并给出了相关定理,指出在这样的框架中利用抽象映射可以自动构建物理世界 W 的抽象度不同的模型。

关键词 KRA 抽象模型,本体类,本体抽象算子,抽象度,分层抽象模型

中图分类号 TP391 文献标识码 A

Ontology-based Hierarchical Abstraction Model

WANG Nan^{1,2} OUYANG Dan-tong¹ SUN Shan-wu³

(Key Laboratory of Symbolic Computation and Knowledge Engineering of MOE, College of Computer Science and Technology, Jilin University, Changchun 130012, China)¹

(Department of Information, Jilin University of Finance and Economics, Changchun 130117, China)²

(Network and Laboratorial Center, Jilin University of Finance and Economics, Changchun 130117, China)³

Abstract The KRA Model offers a generic framework to construct the model of the world W. In this paper, the ontology of concept was introduced in the General KRA model and the Abstraction Object Database was extended to Ontology Class which makes the knowledge sharing and reuse possible in the Perception level, the Language level and the Theory level of the KRA model. The three kinds of ontology abstraction operator(set) were also defined: fundamental operator, entity operator set and connection operator set, working on W and ontology class. Furthermore, the mapping between two different ontology classes was given to realize model abstraction and model reverse-abstraction. The abstraction degree of ontology class was proposed to represent the degree of ontology class and it is as well pointed out that in such a framework, the models of W with different abstraction degrees can be automatically constructed by using abstraction mapping.

Keywords KRA abstraction model, Ontology class, Ontology abstraction operator, Abstraction degree, Hierarchical abstraction model

1 引言

基于模型的推理(Model-Based Reasoning)是以物理世界的模型为基础的一种推理方法,其中很重要的一步是对物理世界 W 进行抽象建模。很多学者研究了抽象理论^[1-6], KRA (Knowledge Reformulation and Abstraction)模型^[7,8]将抽象的表示一般化,给出了对 W 进行形式化建模的一般性框架。我们将 KRA 模型扩展为广义 KRA 模型(G-KRA)^[9],将感知分成基本感知和抽象感知,利用人工建立抽象对象库和抽象

映射,实现 W 的基本感知中基本部件到抽象部件的映射,从而生成抽象感知。本文在 G-KRA 抽象模型框架中引入本体的概念,将抽象对象库扩展为本体类,并实现本体类的自动创建、更新、共享和重用。

本文第 2 节引入本体概念简化 KRA 抽象模型表示框架,并给出了本体类的定义,实现建模知识的共享和重用;第 3 节定义了 3 种本体抽象算子以实现不同抽象程度的本体类的创建和扩展,并给出了本体类间的映射关系;第 4 节给出了本体类的抽象度的形式化定义以及相关的定理,最后给出了总结与

到稿日期:2010-03-03 返修日期:2010-06-09 本文受国家自然科学基金重大项目基金(60496320, 60496321),国家自然科学基金(60973089, 60773097, 60873148),吉林省科技发展计划项目基金(20060532, 20080107),欧盟合作项目(155776-EM-1-2009-1-IT-ERAMUNDUS-ECW-L12)资助。

王楠(1980-),女,博士生,主要研究方向为基于模型的诊断等,E-mail:ctuwangnan@126.com;欧阳彤(1968-),女,博士生导师,主要研究方向为基于模型的诊断等,E-mail:ouyangdantong@163.com(通信作者);孙善武(1969-),男,副教授,主要研究方向为计算机网络。

展望。

2 简化 KRA 抽象模型表示框架

本体(Ontology)作为一种说明机制通过不同的应用增强了知识共享和重用^[10],近年来,本体在人工智能、知识工程等领域获得了广泛的关注,很多学者对本体给出了不同的定义^[11-15]。

从这些定义可知,本体通过对概念、术语及其相互关系的规范化描述,勾画出某一领域的基本知识体系和描述语言。我们将本体概念引入到 G-KRA 抽象模型框架中,通过对本体类的创建、扩展、抽象和重用,实现对物理世界自动抽象建模。

KRA 抽象模型用 4 个层次来表示物理世界:感知层、结构层、语言层和理论层,并且在 G-KRA 抽象模型中,通过人工建立抽象对象库—实现由基本模型到抽象模型的映射。这里将抽象对象库的概念扩展为“本体类”(Ontology Class),用来描述构成物理世界的一组实体(包括实体属性、实体行为以及一般推理机制)。本体类中包括两类本体:实体本体和推理本体。其中实体本体包含构成物理世界的实体类型、实体属性、实体行为以及实体间的关联关系;推理本体包括一组预定义的推理语言和推理公理(类似于 KRA 抽象模型框架内的 L 层和 T 层)。

定义 1 本体类 OC 是一个三元组,即 $OC=(E, L, T)$,其中 $E=\{OBJ, ATT, FUNC, REL\}$, E 中各部分的定义引用 KRA 抽象模型中感知 P 的构成要素 $OBJ, ATT, FUNC, REL$ 的定义, L 和 T 对应于 KRA 抽象模型中的语言层 L 和理论层 T 的定义,详见文献[7,8]。

本体类重用了 KRA 模型中的部分构成要素,包括感知层中不同类型实体的属性、行为模式以及语言层和理论层,因此简化的 KRA 模型框架描述如下。

定义 2 基于本体的 KRA 模型框架 R 是一个四元组,即 $R=(P^*, S, L^*, T^*)$,其中 P^*, L^* 和 T^* 表示实际物理世界抽象模型中感知层,语言层和理论层到本体类中对应元素 E, L 和 T 上的映射关系,需要指出的是感知层中的元素 OBS 和结构层 S 与实际系统有关,需要实际表示和存储。

经过这样简化后的 KRA 抽象模型框架只需要表示待建模的物理世界的具体知识,对于多种可以抽象成同类模型的物理世界,可以实现模型知识的共享和重用。

根据本体类的构建过程,将其广义上分为基本本体类(Fundamental Ontology Class,简称 FOC)和抽象本体类(Abstraction Ontology Class,简称 AOC)。FOC 直接根据物理世界 W 构建,在构建过程中,感知 W 中的所有实体类型以及相关实体类型构成的模型片段,生成 FOC 中的基本实体本体,同时定义 FOC 中的基本推理本体。AOC 是指根据已有本体类通过本体抽象算子(在第 3 节给出相应定义)作用学习生成的本体类。

第 4 节将定义抽象度的概念来表示本体类的抽象程度,通过本体抽象算子的作用可以实现本体类的分层构造,本体类的生成和扩展过程可以独立于建模过程,通过输入不同的表示世界 W 和本体抽象算子的运算使多层本体类得到不断更新和扩展;而创建 W 的某个抽象度的本体类对应的抽象模型时,可以直接利用 G-KRA 模型中的抽象建模过程分层实现,在建模过程中,也可同时利用本体抽象算子实现本体类的

动态扩展。

3 本体抽象算子

本体抽象算子是以创建或扩展本体类为目标的独立于建模过程的运算,这里定义 3 种本体抽象算子(集):基本本体抽象算子、实体本体抽象算子集和连接本体抽象算子集。下面分别介绍 3 种算子的定义以及运算对象、运算功能和运算过程。

3.1 基本本体抽象算子

基本本体抽象算子 Ψ 作用于待表示的物理世界 W 和基本本体类 FOC ,运算结果是创建或扩展基本本体类,即 $\Psi(W, FOC_1)=FOC_2$ 。 Ψ 对 W 进行学习,将其中包括的新的实体类型、实体类型间的新型连接关系以及相应的推理机制存储于基本本体类中。

3.2 实体本体抽象算子集

定义实体本体抽象算子集合 $\Omega=\{\omega_E, \omega_L, \omega_T\}$,实体本体抽象算子根据实体类型的属性和行为特征进行合并运算,其作用对象为任意抽象度的本体类 OC_1 中的构成要素,运算结果是生成与 OC_1 抽象度不同的抽象本体类 OC_2 ,即 $\Omega(OC_1)=OC_2$ 。 ω_E 可以利用并扩展 G-KRA 抽象模型建模过程中的抽象映射 SMP 和 BMP^[16] 来实现具有相同属性特征和相同行为模式的实体类型的抽象,而另外两个算子则作用于本体类中的构成要素 L 和 T 上,生成对应的抽象。

3.3 连接本体抽象算子集

定义连接本体抽象算子集合 $\Phi=\{\lambda_E, \lambda_L, \lambda_T\}$,连接本体抽象算子的作用对象是任意抽象度的本体类,根据本体类中定义的实体类型间的连接关系,聚合相互关联的实体类型,生成超实体类型,以超实体类型的基本属性和行为模式替代相互关联的实体类型,并更新对应的 L 层和 T 层内容,约简被作用的本体类,从而得到抽象度不同的本体类,即 $\Phi(OC_1)=OC_2$ 。

3.4 本体类间的抽象映射关系

经过实体抽象算子集和连接抽象算子集运算后的两个本体类之间存在映射关系,这种映射可以用于实际物理世界建模时的模型抽象与模型还原。

(1) 实体映射

将本体类 OC_1 到由 Ω 算子作用得到的本体类 OC_2 之间的映射称为实体映射 EM(Entity Mapping), OC_1 中的每一个实体类型都对应着 OC_2 中的一个抽象实体类型,而 OC_2 中的每个抽象实体类型可以对应 OC_1 中的多个不同的实体类型,故 EM 映射是一个一对多的映射。

设本体类 $OC_1=(E_1, L_1, T_1)$, $OC_2=(E_2, L_2, T_2)$,且 $OC_2=\Omega(OC_1)$,则

EM:任意实体类型 $e_1 \in E_1$,存在一个并且仅存在一个实体类型 $e_2 \in E_2$ 与 e_1 对应;

任意实体类型 $e_2 \in E_2$,存在一个或多个实体类型 $e_1 \in E_1$ 与 e_2 对应。

图 1 为实体映射表示图。

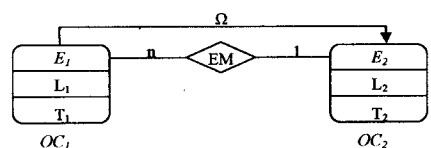


图 1 实体映射表示图

可以看出,通过实体映射能够得到模型 M_1 的唯一的抽象模型 M_2 ;反之,由抽象模型 M_2 根据实体映射则可能得到由多个功能相同的基本模型构成的模型集合。

(2)连接映射

定义本体类 OC_1 到由 Φ 算子作用得到的本体类 OC_2 之间的映射为连接映射 CM(Connection Mapping),具体描述为:设本体类 $OC_1 = (E_1, L_1, T_1)$ 和 $OC_2 = (E_2, L_2, T_2)$,且 $\Phi(OC_1) = OC_2$,则

1)对于任意实体类型 $e_1 \in E_1$ 并且 e_1 不属于 REL_1 中的任何关系,其中 $REL_1 \in E_1$,则存在一个且仅存在一个实体类型 $e_2 \in E_2$ 与 e_1 一一对应。

2)对于任意的实体类型 e_1, \dots, e_n ,若存在关系 $r \in REL_1$,使得 $e_1, \dots, e_n \in r$,则存在一个且仅存在一个实体类型 $e \in E_2$ 与 $\{e_1, \dots, e_n, r\}$ 对应。

图 2 为连接映射的表示图。

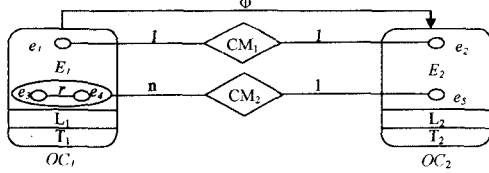


图 2 连接映射表示图

与实体映射类似,由于映射中存在一对多的关系,在模型还原时,可能会得到多个组成不同但是功能相同的抽象模型。

4 本体类的抽象度

定义 3 本体抽象算子的作用改变了本体类的抽象程度,即抽象度(Abstraction Degree),用符号 γ 表示。

设本体类 OC 中的抽象实体数为 N_{α} ,基本本体类中的基本实体数为 N_{RX} ,则我们定义本体类 OC 的抽象度 $\gamma_{\alpha} = N_{\alpha} / N_{RX}$ 。基本本体类 FOC 的抽象度 $\gamma_{RX} = 1$ 。

定理 1 本体抽象算子作用生成的本体类的抽象度小于或等于基本本体类的抽象度。

基本本体抽象算子 Ψ 作用于物理世界 W ,感知 W 中的构成实体类型,创建或扩展基本本体类 FOC,其运算结果为新的基本本体类 FOC*,故 $\gamma_{RX^*} = \gamma_{RX} = 1$ 。

若本体类 OC_1 中存在属性和行为模式相同的实体类型 e_1 和 e_2 ,则实体本体抽象算子集 Ω 对其进行同类归并,生成本体类 OC_2 中的一个抽象实体类型 e ,此时 $\gamma_{\alpha_1} < \gamma_{\alpha_2}$;若本体类 OC_1 中不存在任何属性和行为模式相同的两个实体类型,则 Ω 的作用过程中未发生任何实体类型的归并,此时 $\gamma_{\alpha_1} = \gamma_{\alpha_2}$ 。

若本体类 OC_1 中存在关系 $r_1, \dots, r_m \in REL_1, e_1, \dots, e_n \in E_1$,其中 $OC_1 = (E_1, L_1, T_1), REL_1 \in E_1$,使得 $Relation(e_1, \dots, e_n) = (r_1, \dots, r_m)$,则连接本体抽象算子集 Φ 作用 OC_1 将实体 e_1, \dots, e_n 以及关系 r_1, \dots, r_m 聚合生成具有相同行为的超实体类型 e ,得到新的本体类 OC_2 ,此时 $\gamma_{\alpha_1} < \gamma_{\alpha_2}$;反之,若本体类 OC_2 中不存在任何两个相互关联的实体,即 REL_1 为空,则 $\gamma_{\alpha_1} = \gamma_{\alpha_2}$ 。

定理 2 实体本体抽象算子集的运算保持实体间的连接关系不变。

设本体类 $OC_1 = (E_1, L_1, T_1), OC_2 = (E_2, L_2, T_2)$,且

$OC_2 = \Omega(OC_1)$ 。存在两个实体类型 $e_1, e_2 \in E_1$,且 $Relation(e_1, e_2) = r_1, r_1 \in REL_1, REL_1 \in E_1$ 。

若 e_1 和 e_2 的属性与行为模式相同,则实体本体算子集将 e_1 和 e_2 合并为实体类型 $e, e \in E_2$,同时修改关系 r_1 为 $r_2 = Relation(e, e), r \in REL_2, REL_2 \in E_2, r_1$ 和 r_2 的连接实体类型相同,故 r_1 和 r_2 相同。

若 e_1 或 e_2 与 OC_1 中的其它实体进行了类型合并,则关系 $r_2 = Relation(e_1, e)$ 或者 $r_2 = Relation(e, e_2)$,仍然满足 r_1 和 r_2 的连接实体类型相同,故 r_1 和 r_2 相同。

定理 3 实体本体抽象算子集和连接本体抽象算子集按照不同顺序进行有限次的交替作用,有可能获得抽象度不同的本体类。

显然地,由于连接本体抽象算子除了可能生成新的本体类型外,还会丢弃参与运算的本体类型,因此会导致下一步的实体本体抽象过程的不可用。

例如本体类 OC_1 中包含 3 个属性和行为模式相同的实体类型 e_1, e_2 和 e_3 ,而 e_1 和 e_2 之间存在连接关系 r ,若先应用连接算子,则 e_1, e_2 和关系 r 聚合成结果本体类 OC_2 中的超实体类型 e ,并且 e_1 和 e_2 不在 OC_2 中出现,导致对 OC_2 的实体抽象运算中实体类型 e_3 得不到约简,仍存在于 OC_2 中;若先进行实体抽象运算,则 e_1, e_2 和 e_3 首先得到约简,生成结果本体类中的实体类型 e^* ,并且由定理 2,关系 r 得以保留,因此下一步的连接抽象运算由实体类型 e^* 以及关系 r 共同生成结果本体类 OC_2^* 中的超实体 e ,同时 OC_2^* 中不包含 e_1, e_2 和 e_3 。

结束语 本文对表示物理世界的 KRA 抽象模型进行了深入探讨,引入了本体概念,探讨了抽象建模时的知识共享和知识重用,简化了 KRA 表示框架中的感知层、语言层和理论层表示。提出了 3 种本体抽象算子来创建和扩展本体类,并建立了本体类之间的映射关系,为模型抽象和模型还原提供了基础。用抽象度表示本体类的抽象程度,并给出了相关定理,说明了本体类在分层构建上的特点。

文中对静态的物理世界建模给出了基于本体的具有知识共享和知识重用特点的抽象表示,下一步工作将继续研究动态物理世界的抽象理论,并生成相应抽象表示。

参考文献

- [1] Sacerdoti E. Planning in a hierarchy of abstraction spaces[J]. Artificial Intelligence, 1974, 5: 115-135
- [2] Lowry M. The abstraction/implementation model of problem reformulation[C]//Int. Joint Conf. on Artificial Intelligence, Milano, Italy, 1987; 1004-1010
- [3] Ibrededeche N, Shi Z, Zucker J-D. Perceptual learning and abstraction in machine learning: an application to autonomous robotics [J]. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, 2006, 36(2): 172-181
- [4] Mozetic I. Hierarchical model-based diagnosis[J]. Int. Journal of Man-Machine Studies, 1991, 35(3): 329-362
- [5] Chittaro L, Ranon R. Hierarchical model-based diagnosis based on structural abstraction[J]. Art. Intell., 2004, 155(1/2): 147-182
- [6] 欧阳丹彤, 欧阳继红, 程晓春, 等. 分层的基于模型诊断方法[J]. 计算机科学, 2004, 31(10A): 254-277

(下转第 213 页)

构成的合成服务来评价发现的服务关系的准确度。

从数据集中,根据经验人工生成 10 个合成服务作为测试集,每个合成服务包含的组件服务个数最少 5 个,最多 10 个。实验中,在生成的 RT 中根据发现的语义关系,分别搜索由几个组件服务组合而成的合成服务从而评价准确度。准确度被定义如下:

$$acc_2 = \frac{|R_s|}{|Q_s|} \quad (8)$$

式中,|Q_s|是预定义的合成服务 Q 包含的组件服务的个数,|R_s|是搜索 Q 时,返回的组件服务和预定义的 Q 中组件服务匹配的个数,实验的结果如图 5 所示。实验结果显示,通过设计的实验方法获取的准确度值较高,并且相对稳定,表明了提出的服务语义关系发现方法是有效的。

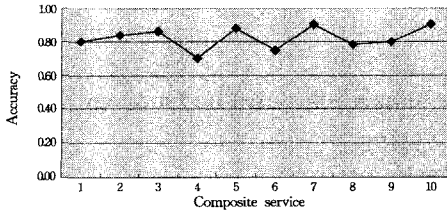


图 5 服务搜索的精度

因此,分级结构是一种能有效地表示不同粒度服务间语义关系的模型。这些实验结果验证了本文提出的发现不同粒度的服务语义关系的方法是有效的、可行的和实用的。

结束语 本文探讨了一个基于分级的结构发现服务语义关系的方法,使得具有相关和合成关系的服务被聚集在相应的位置。主要目的是方便服务发现,增强服务搜索引擎的能力,能自动搜索出语义功能上相关的服务以便执行协作的任务,最大限度地提高服务网络的效用。方法紧紧围绕分级结构所能表示的自然语义,即自然地表示了不同粒度的服务关系。简单介绍了 Web 服务相关的语义模型的简单定义,基于分级结构讨论了度量不同粒度服务间相关度的方法,并进一步根据服务相关度,将语义功能上相关的服务汇集在分级结构的层次,进而逐步发现不同粒度服务间的语义关系。

由于 Web 服务之间复杂的相互关系,服务之间的关系会不时地动态变化。如何捕获动态的服务关系是值得探讨的另一个重要问题,也是将来我们需进一步研究的工作。

参考文献

[1] Liu F, et al. Discovery of Web services based on Collaborated Se-

mantic Link Network [C]//Proc. IEEE International Workshop on Semantic Computing and Systems, 2008

- [2] Zhuge H, et al. An automatic semantic relationships discovery approach [C]//Proc. WWW 2004, 2004; 278-279
- [3] Kun Y, et al. Discovering semantic associations among Web services based on the qualitative probabilistic network [J]. Expert Systems with Applications, 2009, 36(5): 9082-9094
- [4] Platzer C, et al. Web service clustering using multidimensional angles as proximity measures [J]. ACM Transactions on Internet Technology, 2009, 9(3): 11-37
- [5] Paliwal A V, et al. Web Service Discovery via Semantic Association Ranking and Hyperclique Pattern Discovery [C]//Proc. the 2006 IEEE/WIC/ACM International Conference on Web Intelligence, 2006: 649-654
- [6] Yin X, Han J, et al. LinkClus: Efficient Clustering via Heterogeneous Semantic Links [C]//Proc. VLDB '06, 2006
- [7] Zheng Q. Collaborative Semantic Association Discovery from Linked Data [C]//Proc. IEEE International Conference on Information Reuse & Integration, 2009
- [8] Luo X, et al. Automatic discovery of semantic relations Based on Association Rule [J]. Journal of software, 2008, 3(8): 11-18
- [9] Zhuge H. Communities and Emerging Semantics in Semantic Link Network: Discovery and Learning [J]. IEEE Transactions on Knowledge and Data Engineering, 2009, 21(6): 785-799
- [10] Rouached M, et al. Web service mining and verification of properties: An approach based on event calculus [C]//Proc. Coop-IS, LNCS, 2006, 4275: 408-425
- [11] Zhao A, et al. Semantic message link based service set mining for service composition [C]//Proc. 5th International Conference on Semantics, Knowledge, and Grid, 2009: 338-341
- [12] Salton G, Buckley C. Term-weighting approaches in automatic text retrieval [J]. Information processing & Management, 1988, 24(5): 513-523
- [13] Al-Masri E, Mahmoud Q H. The qws dataset. Web page [OL]. Available: <http://www.uoguelph.ca/~qmahmoud/qws/index.html>
- [14] Al-Masri E, Mahmoud Q H. Qos-based discovery and ranking of web services [C]//Proc. the IEEE International Conference on Computer Communications and Networks, 2007
- [15] Al-Masri E, Mahmoud Q H. Investigating web services on the world wide web [C]//Proc. the International World Wide Web Conference, 2008

(上接第 186 页)

- [7] Saitta L, Zucker J. Semantic Abstraction for Concept Representation and Learning [C]//Proc. SARA, 1998: 103-120
- [8] Saitta L, Zucker J-D. A Model of Abstraction in Visual Perception [J]. Applied Artificial Intelligence, 2001, 15(8): 761-776
- [9] 孙善武, 王楠, 欧阳丹彤. 广义 KRA 抽象模型 [J]. 吉林大学学报: 理学版, 2009, 47(3): 537-542
- [10] Neches R, Fikes R, Finin T, et al. Enabling technology for knowledge sharing [J]. AI Magazine, 1991, 12: 36-56
- [11] Uschold M. Knowledge level modelling: concepts and terminology [J]. The Knowledge Engineering Review, 1998, 13(1): 5-29
- [12] Guarino N, Giaretta P. Ontologies and Knowledge Bases: Towards a Terminological Clarification [M]//N. Mars, ed. Towards Very Large Knowledge Bases: Knowledge Building and

Knowledge Sharing. Amsterdam: IOS Press, 1995: 25-32

- [13] Gruber T. Towards principles for the design of ontologies used for knowledge sharing [J]. International Journal of Human-Computer Studies, 1995, 43(5/6): 907-928
- [14] William S, Austin T. Ontologies [J]. IEEE Intelligent Systems, Jan/Feb 1999: 18-19
- [15] Chandrasekaran J, Josephson R, Benjamins V R. What Are Ontologies, and Why Do We Need Them? [J]. IEEE Intelligent Systems, Jan/Feb 1999: 20-25
- [16] Wang Nan, Ouyang Dan-tong, Sun Shan-wu, et al. Formalizing the Modeling Process of Physical Systems in MBD [C]//The 2009 International Conference on Web Information Systems and Mining (WISM'09) and the 2009 International Conference on Artificial Intelligence and Computational Intelligence (AICI'09), 2009: 685-695