

过程模型中控制流反模式的定义和检测方法

韩兆刚¹, 巩 朋¹, 张 莉¹, 吕方兴²

(1. 北京航空航天大学 软件工程研究所, 北京 100191; 2. 菏泽学院 计算机与信息工程系, 山东 菏泽 274015)

摘要:提出一种过程模型中控制流反模式的定义和检测方法,抽取不同过程模型中的控制流结构并将其规约为语言无关的过程结构树,基于不同控制流反模式的 CAPDL 定义,在过程结构树上查找与之对应的控制流反模式。该方法既支持不同的过程建模语言,也允许用户自定义控制流反模式。基于 215 个实际 BPMN 过程模型的反模式检测试验表明,该方法可以有效地检测用户自定义的控制流反模式,与已有方法相比,该方法极大地提高了反模式检测的检测效率。

关键词: 过程建模; 控制流反模式; 反模式检测

中图分类号: TP 301 **文献标志码:** A

Definition and detection approach of control-flow anti-patterns in process models

HAN Zhao-gang¹, GONG Peng¹, ZHANG Li¹, LÜ Fang-xing²

(1. Institute of Software Engineering, Beihang University, Beijing 100191, China;

2. Department of Computer and Information Engineering, Heze University, Heze 274015, China)

Abstract: A control-flow anti-pattern definition and detection approach for process models was proposed. Firstly control-flow structures of different process models are extracted and reduced to process structure tree which is language-independent, then control-flow anti-patterns are detected according to their CAPDL (control-flow anti-pattern description language) definitions in the process structure tree. The proposed approach can support user-defined control-flow anti-pattern detection with different process modeling languages. The results of anti-pattern detection experiment based on 215 real-world BPMN process models show that the proposed approach can detect user-defined control-flow anti-patterns effectively. Compared with existing approaches, the efficiency of anti-pattern detection is greatly enhanced.

Key words: process modeling; control-flow anti-pattern; anti-pattern detection

过程模型^[1]中的反模式是指“导致过程模型错误的典型设计缺陷”^[2],它是评估过程模型正确性的有效手段。过程模型反模式以控制流反模式(控制结构方面的反模式)最为常见^[2]。文献[3]对过程模型进行化简并根据化简结果判断过程模型中是否存在死锁或乏同步,但该方法无法定位反模式。文献[4]使用编译技术中的控制流块分析技术^[5]分析过程模型并利用启发式方法检测并定位控制流反模式,但该方法只支持一些既定的控制流反模式,扩展性不足。文献[6]利用 BPMN-Q^[7]定义并查询

BPMN 模型中的常见控制流反模式,该方法允许用户自定义反模式,但它的效率较低,而且不支持 BPMN 以外的建模语言。为了提高过程模型的控制流反模式检测的实用性和灵活性,笔者提出一种过程模型中控制流反模式的定义和检测方法。

1 过程模型预处理

为了支持不同的过程建模语言,提取过程模型的控制流结构,并将其转化为过程建模语言无关的工作流图^[3]。为了方便,使用 BPMN^[8]的建模符号

来表示 workflow 图中的元素(图 1),并在下文中使用 ANDSplit、ANDJoin、XORSplit 和 XORJoin 来表示 workflow 图中的并行分支、同步合并、排他选择和简单合并。

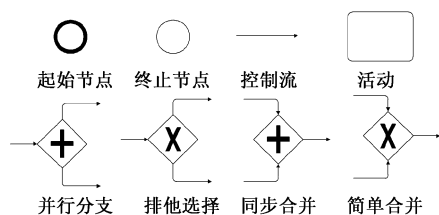


图 1 workflow 图中的基本元素

Fig. 1 Basic notations in workflow graph

使用文献[2]提出的过程结构树来分析过程模型的控制流结构。过程结构树由 workflow 图中的规则单入口单出口过程块和节点构成。通过环路等价算法^[4]对 workflow 图进行处理,可以在线性时间得到 workflow 图中的规则单入口单出口过程块(除非特殊说明,过程块均为规则单入口单出口过程块)。称该处理过程为 workflow 图归约。

图 2 给出了一个示例 workflow 图,过程从起始节点 start 出发到达 g1,然后并行地执行 g1 的两个后继分支。其中一个分支串行地执行 a1、a2 和 a3 这几个活动,另外一个分支则执行由 g2、a4 和 g3 构成的循环。接下来这两个分支在 g4 进行汇合,最后过程到达终止节点 end,该过程结束。值得注意的是,由于使用并行分支作为循环出口,过程块 C 中的 g2 和 g3 构成了无穷循环。图中的虚线框表示对 workflow 图实施规约后得到的不同过程块。

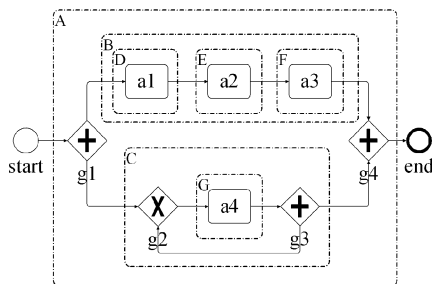


图 2 示例 workflow 图及其归约结果

Fig. 2 Sample workflow graph and its reduction result

对于给定过程块 F,称 F 的母过程块为包含 F 的最小过程块 F',并称 F 为 F'的子过程块。把过程块 F 的子过程块替换为一个活动节点不会影响该过程块的正确性^[6]。以图 2 为例,过程块 A 是过程块 B、C 的母过程块,而过程块 B、C 均为 A 的子过程块。处理过程块 A 时,可以把过程块 B、C 作为一个活动节点来处理,而无需考虑其内部结构,这可以显著地

提升处理速度。

文献[4]证明了规则单入口单出口过程块之间只存在嵌套和不相交这两种关系,因此 workflow 图中的过程块可以构成一个树形结构,这个树形结构就是过程结构树^[9]。图 3 给出了图 2 所对应的过程结构树。

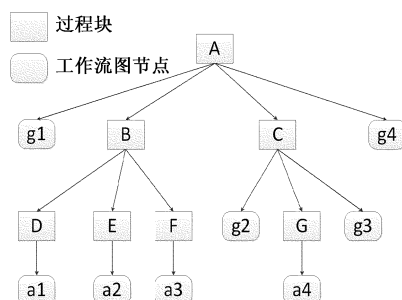


图 3 示例 workflow 图对应的过程结构树

Fig. 3 Corresponding process structure tree of sample workflow graph

过程结构树中的叶子节点对应于 workflow 图的节点,非叶子节点对应于与 workflow 图中的过程块。以图 2 的 workflow 图和图 3 的过程结构树之间的关系为例,图 3 中的根节点 A 对应图 2 中的过程块 A,图 3 中的叶子节点 g1 和 g4 对应于图 2 中的节点 g1 和 g4。workflow 图中的控制流信息则保存在过程结构树的叶子节点中。

由于过程结构树中的节点和 workflow 图中的元素之间存在这种映射,并且 workflow 图总可以转换为与之对应的过程结构树^[10],因此可以通过分析过程结构树来分析过程模型的控制流结构。

2 反模式描述语言 CAPDL

为了描述控制流反模式,设计了控制流反模式描述语言 CAPDL。

2.1 CAPDL 的设计思路

控制流反模式是一类导致过程模型异常的特殊控制流结构,描述控制流反模式也就是描述其对应的控制流结构。由于控制流结构由活动、分支、合并等模型元素构成,因此可以利用模型元素的属性来描述过程模型中的控制流结构。过程模型中的控制流结构可以分为两类:简单控制流结构和复合控制流结构。简单控制流结构对应于节点;复合控制流结构对应于过程块。对于简单控制流结构,利用模型节点的属性就可以对其进行描述;对于复合控制流结构,则需要利用过程块内部的节点信息对其进行描述。基于这个思路,本文设计了控制流反模式

描述语言 CAPDL。

2.2 CAPDL 的核心概念和元模型

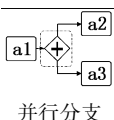
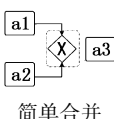
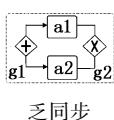
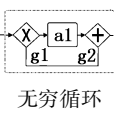
CAPDL 的基本概念是谓词,根据谓词的目标类型,谓词可以分为属性谓词和节点谓词;属性谓词应用于节点上的属性;节点谓词应用于过程块中的节点。表 1 给出了一些谓词的使用示例。

表 1 CAPDL 的谓词示例

Table 1 Examples of predicates in CAPDL		
谓词类型	谓词含义	CAPDL 定义
属性谓词	节点包含一条以上的出边	outputedgescount> 1
属性谓词	节点类型为 AndGateway	type = AndGateway
节点谓词	起始节点满足 XORJoin 规则的过程块	start with XORJoin
节点谓词	终止节点满足 ANDSplit 规则的过程块	end with ANDSplit

比谓词高一级的概念是规则,规则由一系列谓词组成。规则根据其目标类型可以分为节点规则和块规则;节点规则由若干属性谓词组成,用于描述过程模型中的模型节点;块规则由若干节点谓词组成,用于描述过程模型中的过程块。规则的使用示例见表 2。

表 2 CAPDL 的规则示例

Table 2 Examples of rules in CAPDL		
规则类型	目标控制流结构	CAPDL 定义
节点规则		rule AND Split on node; type = AND inputedgescount = 1 outputedgescount>1
节点规则		rule XORJoin on node; type = XOR inputedgescount = 1 outputedgescount>1
块规则		rule LackOfSync on block; start with AND split end with XORJoin
块规则		rule InfiniteLoop on block; start with XORJoin end with ANDSplit

比规则高一级的概念是反模式。由于过程模型中的控制流反模式可能具有多种实现形式^[11],因此一个反模式由一个到多个规则组成。对于指定的反模式,只要过程块或模型节点满足该反模式规则集合中的一条规则,就称该过程块或模型节点满足该反模式。

为了复用已有的反模式定义和规则定义,CAPDL 中引入了模块的概念。模块可以包含一个到多个规则或反模式;模块可以通过导入其他模块来引用已定义的规则或反模式。以“错误循环”反模式为例,它由 InfiniteLoop 规则和 DeadLoop 规则组成,并引用了 BasicPatterns 模块中的节点规则来定义自己的谓词。具体如下:

```
import BasicPatterns \\引用 BasicPatterns 模块
antipattern Badloop:\\定义 Badloop 反模式
rule InfiniteLoop on block:\\定义 InfiniteLoop 规则
start with BasicPatterns. XORJoin \\ 引用 BasicPatterns 模块中的节点规则
end with BasicPatterns. ANDSplit
rule DeadLoop on block:\\定义 DeadLoop 规则
start with BasicPatterns. ANDJoin
end with BasicPatterns. XORSplit
```

通过谓词、规则、反模式、模块等概念,CAPDL 提供了一种简洁的控制流反模式定义方法。图 4 给出了 CAPDL 的元模型。

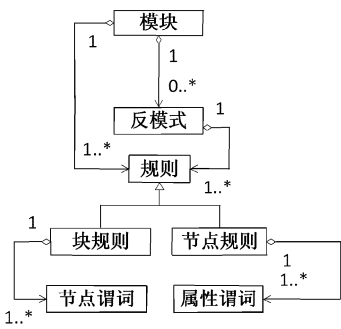


图 4 CAPDL 元模型

3 控制流反模式查询算法

利用预处理过程中得到的过程结构树,以及控制流反模式的 CAPDL 定义,可以在过程结构树上查询满足目标反模式的过程模型节点和过程块。

3.1 查询算法

基于 CAPDL 的控制流反模式查询算法步骤如下:

- (1) 读取工作流图对应的过程结构树 PST,并初始化 PST 上每个叶子节点的谓词集合。
- (2) 分析待查询反模式的 CAPDL 定义,对其中节点规则 NodeRule 中的属性谓词和块规则 BlockRule 中的块谓词进行等价谓词合并,并将去重后的

谓词保存在属性谓词集合和块谓词集合中。

(3) 遍历 PST 上的叶子节点 LeafNode, 记录 LeafNode 所满足的属性谓词并将其添加至节点的谓词集合, 并根据该节点满足的谓词判断其所属的 NodeRule。

(4) 遍历 PST 的非叶子节点 Non-LeafNode, 并根据该节点满足的块谓词判断其所属的 BlockRule。

(5) 根据 PST 节点所属的 NodeRule 或 BlockRule 判断其所属的控制流反模式, 并记录其信息。

3.2 反模式查找过程

当使用二叉链表作为存储结构时, 过程结构树的后序遍历可借用二叉树的先序遍历和中序遍历的算法实现, 由于二叉树的遍历算法的时间复杂度与节点数存在线性关系, 因此过程结构树上的反模式检测算法的时间复杂度也与节点数存在线性关系。

以检测图 2 的示例 workflows 中的错误循环为例, 介绍基于 CAPDL 的控制流反模式查询过程。图 2 的工作流图可以使用图 3 中的过程结构树表示。

(1) 首先读取图 3 中的过程结构树, 并初始化过程结构树中每个叶子节点的谓词集合。以叶子节点 a1、a2、a3、a4 为例, 它们的谓词集合均被初始化为: inputedgescount = 1; outputedgescount = 1。

(2) 为了提高查询效率, 本文对 CAPDL 代码中重复的谓词/规则进行了合并处理, 从而保证同样的谓词/规则不会查询多次。以“错误循环”反模式的 CAPDL 定义为例, 谓词 inputedgescount = 1 在规则 Sequential 和规则 ANDSplit 均有出现(这些规则定义在模块 BasicPatterns 中)。

(3) 通过遍历过程结构树上的叶子节点, 判断其满足的属性谓词, 可以得到它们所满足的节点规则。以图 3 中的叶子节点 g2 为例, 由于谓词 type = XOR、inputedgescount > 1 和 outputedgescount = 1 在 g2 上均为真, 因此 g2 满足规则 XORJoin。同理 a1、a2、a3、a4 满足规则 Sequential, g3 满足规则 ANDSplit。节点规则的结果见图 5, 图中所有叶子节点满足的节点规则均由节点规则查询得到。

(4) 通过遍历过程结构树上的非叶子节点, 并利用其子节点满足的节点规则来判断其满足的节点谓词, 可以得到它们满足的块规则。以过程块 C 为例, 由于 C 的起始节点 g2 满足规则 XORJoin, 终止节点 g3 满足规则 ANDSplit, 因此 C 满足规则 InfiniteLoop。块查询的结果见图 5, 图中过程块 C 满足规则 InfiniteLoop 这一结果由块规则查询得到。

(5) 根据过程结构树中节点的节点规则和块规则, 得到过程模型的控制流反模式信息: 由于满足“InfiniteLoop”规则, workflows 中的过程块 C 构成了一个错误循环。

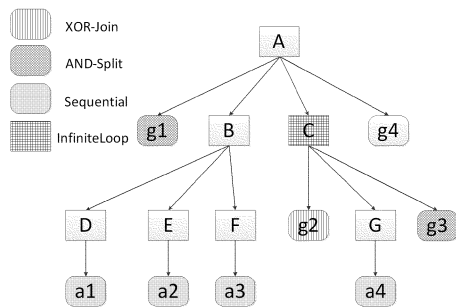


图 5 节点规则和块规则查询结果

Fig. 5 Querying results of node rules & block rules

4 试验结果分析

为了验证本文方法的有效性, 对 215 个实际过程模型进行控制流反模式检测试验, 试验以文献[2]提出的 4 种常见控制流反模式为检测目标。这 4 种反模式分别是: 异或分支 (XORSplit) 和同步合并 (ANDJoin) 的组合 (会产生死锁错误)、并行分支 (ANDSplit) 和简单合并 (XORJoin) 的组合 (会产生同步错误)、同步合并 (ANDJoin) 作为循环入口 (会产生死锁错误) 和并行分支 (ANDSplit) 作为循环出口 (会产生无穷循环错误)。这些控制流反模式的详细信息参见文献[2]。使用 CAPDL 定义了这些反模式, 并基于这些反模式的 CAPDL 定义, 对 ILog 建模工具 (<http://www-01.ibm.com/software/integration/visualization/java/>) 绘制的 215 个 BPMN 过程模型进行了控制流反模式检测。图 6 给出了这些模型的规模信息。

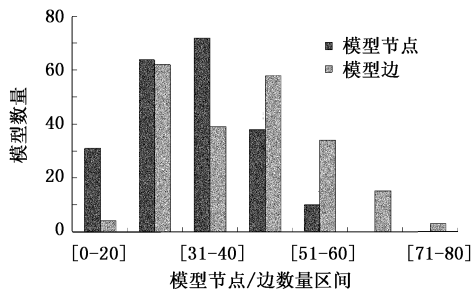


图 6 模型的规模信息

Fig. 6 Model scale information

检测程序运行在一台装有 4G 内存和 2.93 GHz 的双核 Intel 处理器的台式机上, 检测环境为 Windows 7 SP1。检测过程分为预处理过程和反模式检测过程。图 7 给出了检测过程所耗时间和模型节点

数的关系,其中反模式检测时间和模型节点数呈预期的线性关系;预处理时间与模型节点数也呈现线性关系,这印证了文献[4]中的论断;通过环路等价算法对 workflow 图进行处理,可以在线性时间得到与之对应的过程结构树。

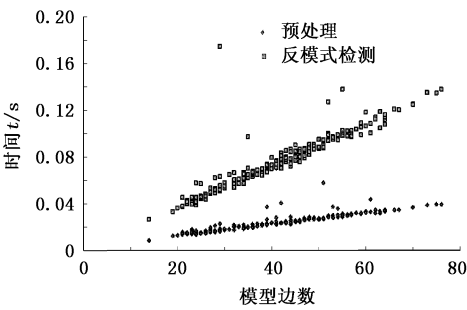


图 7 检测时间和模型节点数的对应关系

Fig. 7 Relationship between detection time and model edge number

与文献[6]提出的利用 BPMN-Q 检测控制流反模式的方法相比,本文的检测方法具有较快的速度。对相近规模的过程模型检测同样的控制流反模式,本文的平均检测时间为 0.33 s,文献[6]的平均检测时间为 1.9~2.1 s。表 3 给出了具体的比较数据(本文的方法会对四种控制流反模式进行同时检测,这里计入的是检测总时间,所以检测时间相同)。

表 3 两种反模式检测方法的平均查询时间比较
Table 3 Average querying time comparison of two anti-pattern detection approaches

反模式名称	文献[6]方法	本文方法
XORSplit 和 ANDJoin 的组合	1.90	0.33
ANDSplit 和 XORJoin 的组合	1.94	0.33
ANDJoin 作为循环入口	2.02	0.33
ANDSplit 作为循环出口	2.11	0.33

文献[6]中的模式查询方法受系统架构限制,每查询一个控制流模式就需要遍历一次过程模型,因此在查询多个控制流模式时需要遍历过程模型多次,从而造成相当的性能损耗。本文方法通过对过程模型进行预处理,将过程模型转化为过程结构树,并对控制流反模式的 CPDL 描述进行等价谓词合并以减少不必要的查询判断。从而在查询多个控制流反模式时只需一次预处理和四次过程结构树遍历,从而显著提升了控制流模式查找的效率。

文献[6]采用图匹配的方式进行控制流模式查询。每一次查询均需要遍历整个过程模型;本文利用文献[9]的过程结构树,把对指定模式的查询限制在过程模型的子过程块中,从而减少了查询时间。

笔者选取了 215 个过程模型中的 100 个模型进行人工检测,然后与检测结果进行比较。结果见表 4。可以看出,出现最多的控制流反模式是 XORSplit 和 ANDJoin 的组合(15 次),其次是 XORJoin 和 ANDSplit 的组合和 ANDJoin 作为循环入口(各出现 6 次),同文献[6]中的过程模型中常见控制流反模式统计结果相一致。

表 4 反模式检测结果
Table 4 Anti-pattern detection results

控制流反模式	反模式出现的次数	检测方法误判次数	检测方法漏判次数
XORSplit 和 ANDJoin 的组合	15	2	0
XORJoin 和 ANDSplit 的组合	6	1	0
ANDJoin 作为循环入口	6	0	0
ANDSplit 作为循环出口	2	0	0

从表 4 可以看出,检测过程中存在少数误判现象。一般来讲,只有基于状态空间搜索的正确性验证方法才能完全避免错判和漏判。

5 结束语

为了提高过程模型中控制流反模式检测的实用性和灵活性,本文提出一种过程模型中控制流反模式的定义和检测方法。与已有方法相比,该方法既支持不同的过程建模语言,又可以有效地检测过程模型中用户自定义的控制流反模式。同时极大地提高了反模式检测的检测效率。

参考文献:

[1] SMITH H, FINGAR P. Business Process Management: The Third Wave [M/OL]. Meghan-Kiffer Press, 2003. <http://www.bpmi.org/downloads/LIB-2002-08-2.pdf>.
[2] KOEHLER J, VANHATALO J. Process anti-patterns: how to avoid the common traps of business process modeling[J]. IBM WebSphere Developer Technical Journal, 2007,10(2):4.
[3] SADIQ W, ORLOWSKA M E. Analyzing process models using graph reduction techniques [J]. Information Systems, 2000,25(2):117-134.
[4] VANHATALO J, VOLZER H, LEYMANN F. Faster and more focused control-flow analysis for business process models through sese decomposition [M/OL]. Service-Oriented computing-ICSOC 2007, Vienna, Austria. Springer Berlin Heidelberg, 2007:43-55. http://link.springer.com/chapter/10.1007/978-3-540-74974-5_4.
[5] AHO A V, LAM M S, SETHI R, et al. Compilers: principles, techniques, and tools [M/OL]. Pearson/Addison Wesley, 2007. <http://www.lavoisier.fr/livre/notice>.

- asp? ouvrage=1184227.
- [6] GRUHN V, LAUE R. A heuristic method for detecting problems in business process models [J]. Business Process Management Journal, 2010, 16(5):806-821.
- [7] AWAD A. BPMN-Q: a language to query business processes[C/OL]//Proceedings of EMISA 2007, Nanjing China. 2007:115-128. <http://dbis.eprints.uni-ulm.de/205/1/EMISA-Proceedings-Komplett.pdf#page=117>.
- [8] Business Process Modeling Notation (BPMN) Version 2.0 [S]. Object Management Group (OMG), Jun 2010. <http://www.omg.org/spec/BPMN/2.0/PDF/>.
- [9] VANHATALO J, VOLZER H, KOEHLER J. The refined process structure tree[J]. Data & Knowledge Engineering, 2009, 68(9):793-818.
- [10] POLYVYANY A, VANHATALO J, VOLZER H. Simplified computation and generalization of the refined process structure tree [M]//Web Services and Formal Methods. 2010, Hoboken, NJ, USA: Springer Berlin Heidelberg, 2011:25-41.
- [11] WHITE S A. Process modeling notations and workflow patterns[J]. Workflow Handbook, 2004:265-294.
- [12] VAN DER AALST W M P, TER HOFSTEDE A H M. YAWL: yet another workflow language[J]. Information Systems, 2005, 30(4):245-275.
- (编辑 修荣荣)

(上接第 171 页)

- [6] 陈洪, 韩利娟, 徐鹏, 等. 疏水改性聚丙烯酰胺的增黏机制研究[J]. 物理化学学报, 2003, 19(11):1020-1024.
- CHEN Hong, HAN Li-juan, XU Peng, et al. The thickening mechanism study of hydrophobically modified polyacrylamide [J]. Acta Phys-Chim Sin, 2003, 19(11):1020-1024.
- [7] 欧阳坚, 朱卓岩, 王贵江. TS 系列新型耐温抗盐聚合物的研究[J]. 石油勘探与开发, 2003, 30(2):79-81.
- OUYANG Jian, ZHU Zhuo-yan, WANG Gui-jiang, et al. Research on the high temperature and salt resistance (TS series) polymers [J]. Petroleum Exploration and Development, 2003, 30(2):79-81.
- [8] 钟传蓉, 黄荣华, 马俊涛. 疏水缔合水溶性聚合物的分子结构对疏水缔合的影响[J]. 化学世界, 2003(12):660-664.
- ZHONG Chuan-rong, HUANG Rong-hua, MA Jun-tao. Influences of molecular structure of hydrophobically associating water soluble copolymers on hydrophobic association[J]. Chemical World, 2003(12):660-664.
- [9] 任鲲, 姜桂元, 林梅钦, 等. NaCl 对疏水缔合聚合物溶液性质的影响研究[J]. 功能高分子学报, 2005, 18(2):321-324.
- REN Kun, JIANG Gui-yuan, LIN Mei-qin, et al. Influence of salt concentration on the properties of the hydrophobically associating polymer solution [J]. Journal of Functional Polymers, 2005, 18(2):321-324.
- [10] 张瑞, 叶仲斌, 罗平亚. 原子力显微镜在聚合物溶液结构研究中的应用[J]. 电子显微学报, 2010, 29(5):475-481.
- ZHANG Rui, YE Zhong-bin, LUO Ping-ya. The atomic force microscopy study on the microstructure of the polymer solution [J]. Journal of Chinese Electron Microscopy Society, 2010, 29(5):475-481.
- [11] 任鲲, 姜桂元, 徐春明, 等. 疏水缔合聚合物溶液形态的研究[J]. 石油大学学报: 自然科学版, 2005, 29(2):117-120.
- REN Kun, JIANG Gui-yuan, XU Chun-ming, et al. Solution conformation of hydrophobically associating polymers [J]. Journal of the University of Petroleum, China (Edition of Natural Science), 2005, 29(2):117-120.
- (编辑 刘为清)