

基于知识的可重用构件查询系统 KCQ *

余江 全炳哲 金淳兆

(吉林大学计算机科学系, 长春 130023)

摘要 本文提出了一个基于知识的、支持面向对象程序设计方法的软件重用构件查询系统, 本系统采用了语义网络知识表示模型, 提供了构件知识表示语言 CKRL, 来描述可重用构件的特征和用于检索的知识, 提供了语义网络联想式查询、条件查询和构件评价。

关键词 软件重用, 构件查询, 基于知识。

基于构件合成的软件重用技术要解决的主要问题之一, 是对于可重用构件的检索, 目前国内外在这方面都开展了工作, 取得了一些成果, 构件检索方法可分为三种^[1]: 第一种是图书与信息科学索引方法, 如以关键词为主的检索系统^[2,3], 第二种是基于知识的方法, 常用的有语义网和框架表示形式, 如日本学者 Hiroyuki Tarnmi 等^[4]提出的支持面向对象软件重用的环境中使用了关系描述, Murray Wood 等^[5]应用概念依赖模型, 采用框架形式, 描述和检索构件, 需要相当丰富的框架, 第三种是超级文本方法, 它是前两种方法的结合, 表示形式直观, 易于使用。

我们吸收了以上方法的优点, 提出了支持面向对象程序设计方法的, 基于知识的可重用构件检索系统 (Knowledge-based reusable Component Query system—KCQ), 与国内外其他工作相比, 具有以下特点:

1. 可描述更复杂的构件, 可表达更丰富的知识, 增加了一些新关系形式, 对于构件进行了更深入的刻划。
2. 查询使用的词更丰富, 有同义词, 包含关系词, 领域关系词。
3. 提供的语义网联想式查询、条件查询两种交互式及专家评价式检索, 适合多种用途的使用。

1 系统概况

软件重用构件的查询是与软件开发方法密切相关的, 即支持特定的软件重用模型。

1.1 基于面向对象程序设计的软件重用

* 本文 1991-10-17 收到, 1992-06-23 定稿

本文是国家 863 计划资助项目。作者余江, 34 岁, 副教授, 主要研究领域为软件工程环境, 软件工具, 人工智能, 全炳哲, 33 岁, 讲师, 主要研究领域为软件重用技术, 面向对象语言, 软件生产自动化。金淳兆, 57 岁, 教授, 主要研究领域为软件工程。

本文通讯联系人: 余江, 长春 130023, 吉林大学计算机科学系

面向对象的程序设计方法采用数据抽象和继承等机制较好地反映出现实世界中存在的对象及对象之间的关系,我们采用的软件重用模型是以面向对象的程序设计方法为基础的,可重用构件用 RECOS^[6]语言描述.语言定义的可重用构件有两种:类(class)和模块(module),模块与类不同的是模块中无状态变量表,且不产生对象,为了增加构件的抽象度,这两种构件均可带有类属参数.

1.2 检索模型

检索模型与重用模型、软件开发过程密切相关.为了增加检索的灵活性和有效性,可检索的构件除了类和模块之外,还可以对操作进行检索.

系统提供了两个层次上的检索,即“粗”检索和“细”检索.“粗”检索的目的在于为软件设计方案的确定作准备.“细”检索的目的往往是在设计方案基本确定之后,为方案的实施搜集构件,当然由于设计和实现阶段不能截然划分,这两种检索目的也不总是十分明确的.因此,我们提供了三种查询方式,其中前两种方式用于“粗”检索.(1)交互式语义网络联想查询,可以以某个结点(话题)为基础,查询与它有关系的其它结点,可在语义网络中自由地搜索(自由谈话);(2)条件查询,根据用户提供的描述词及它们之间的与或关系,选出用这些描述词描述其特征的构件;(3)专家评价查询,由用户提供候选条件及对象内容的详细描述,系统将求出候选对象集合,对集合中的每一对象(类,操作,或模块)与用户描述对象的近似度进行评价,根据近似程度排序输出.

1.3 KCQ 系统的组成

组成 KCQ 系统的几个主要部分是:

1. 可重用构件知识库 CKL 为了便于构件检索和知识的管理,构件的实体与构件的知识分别存于构件库和构件知识库中. CKL 中保存关于构件特征,构件间关系,各对象活动规律等知识,还有用于检索的知识.

2. 检索控制规则库 RL 对于各种检索单位进行检索的控制规则存于此库.

3. 词典工具库 为了扩大检索范围,增加有效性,检索过程中使用了一些词典作为辅助工具,如同义词词典,包含关系词词典,领域关系词词典等.

4. 构件知识表示语言 CKRL 用于描述可重用构件的知识.它以 Recos 语言的注解形式表示,通过构件知识抽取工具获得构件本身的特征的知识 and 用于检索的知识,并以语义网表示形式存入知识库中.

5. 构件检索语言 CQL 用户使用 CQL 描述要评价的对象的细节.

6. 推理机 推理机是 KCQ 的核心,它根据查询要求,利用构件知识库内容、查询控制规则、各种词典工具等进行推理,找出满足要求的目标.

KCQ 的主要功能包括对于构件知识库、查询控制规则库及词典工具库的管理和维护,以及对于构件进行多种形式的查询.

2 知识表示与构件知识表示语言 CKRL

2.1 知识表示

KCQ 系统中的知识表示可分为三种:

1. 词关系知识

在各词典中使用的与词相关的,词与词之间关系的知识,这种知识采用关系式形式 $\text{relation}(\text{word1}, \text{word2})$ 来表示.

2. 规则知识

规则表示形式为:

IF 〈条件〉 THEN 〈结果〉

以上两种知识为基本知识,与构件无关.

3. 构件知识

与构件相关的知识用 CKRL 语言描述.

2.2 CKRL 语言

对于类构件知识的描述分为五个部分,即类的概括知识,类属参数知识,状态变量知识,操作知识,对象活动规律知识,模块构件无状态变量知识.

用 CKRL 表示的各部分内容和形式表示如下:

1. 类的概括知识

“INHERIT” 〈类名〉“;” /* 继承说明 */
 “ALIAS” 〈名字表〉“;” /* 别名表 */
 “DOMAIN” 〈名字表〉“;” /* 应用领域表 */
 “KEYWORD” 〈名字表〉“;” /* 关键词表 */

2. 类属参数知识

若构件含有类属参数,要指出参数的属性(类型,常量或操作),每个类属参数表示为:

“GPARA”〈类属参数名〉“:”〈属性值〉“;”

3. 状态变量知识

每个状态变量的知识表示为:

“SVAR”〈状态变量名〉“;”
 “STRUCT”“<”〈结构特征〉“,”〈内部元素类〉“>”;
 “KEYWORD”〈名字表〉“;”
 “RESTRICT”〈限制词表〉“;”

当状态变量是结构型时,〈结构特征〉指它的结构形式如:数组,文件,树等,〈内部元素类〉指结构内部的元素类.“限制词”指对于它使用的限制条件,如文件是有序的等.

4. 操作知识

每个操作的知识表示为:

“OPR” 〈操作名〉“;”
 “FUNCTION” 〈功能描述〉“;”
 “KEYWORD” 〈名字表〉“;”
 操作中的参数的知识表示为:
 “PARA” 〈参数名〉“;”
 “STRUCT” “<”〈结构特征〉“,”〈内部元素类〉“>”“;”
 “KEYWORD” 〈名字表〉“;”
 “RESTRICT” 〈限制词表〉“;”

“RELATION” (关系名表)“;”

关系描述了该操作对于该参数的影响,各关系描述词及它们的意义如表 1 所示:

表 1 操作与其参数的关系

关系名称	意义	例子
ACTIVATE	使对象可以使用	打开文件
CHANGE	对象发生了变化了	赋值
CREAT	创建一个新对象	拷贝产生的新文件
DECREASE	元素对象个数的减少	弹出元素后的栈
DISPOSE	消除一个对象	文件删除
GET	取值	表达式中出现的状态名
INCREASE	元素对象个数的增加	添加记录后的文件
TEST	测试	出现在条件表达式中的对象
UNACTIVATE	使对象不能使用	文件关闭

例如:有操作 $\max(x, y, z)$, 功能是把 x, y 的最大值放入 z 中。

各参数的关系描述部分为:

X: RELATION TEST, GET;

Y: RELATION TEST, GET;

Z: RELATION CHANGE;

为了处理方便,接收消息的对象及作为结果返回的对象都采用了与参数描述类似的结构。

5. 对象活动规律知识

对象活动规律知识采用与 BNF 类似的表示形式,例如,一个直接访问的文件的活动规律描述为:

ACTION

```
creat {open {read | write | seek} close}
delete;
```

表示文件创造之后可多次打开使用,每次打开后可多次进行读或写或移动文件指针的操作等。

3 构件查询

构件查询的基本原理是:根据用户提出的各种要求,生成系统内部的提问形式,启动推理机求出结果,以用户易读的形式显示。

3.1 联想式语义网查询

系统提供了如图 1 所示的一个由语义网中的主要关系名称组成的查询控制面板。查询过程是:先由用户提出一个讨论的话题,这个话题可以是应用领域名称、类名称、示例变量、操作等,它是语义网络中的一个结点;然后按动控制面板上的关系按钮(button)(这种关系按钮共有二十多个),若有满足关系的结点存在,就会在查询结果窗口中显示出查询结果,然

后重复上述过程就可以在语义网中遍历. 设语义网中的结点 N_1 和 N_2 , 从 N_1 到 N_2 用关系 R 连接, 表示为 $R(N_1, N_2)$.

LIST PART	
(ALL DOMAINS)	(ALL COMPONENTS) (RETURN)
RELATION PART	
(Doma. Super () (Doma. Comp. () (Comp. Super) (Class Svar ()	
(Comp. Opr () (Comp. Import () (Comp. Export () (Comp. Keyw. ()	
(Svar. Type) (Svar. Keyw. () (Svar. Restr. () (Comp. Others)	
(Opr Func.) (Opr Keyw. () (Opr Para.) (Comp. G. para.)	
(Para. Type) (Para Keyw. () (Para Restr. () (Para Rel.) \ = INFER	
DERECTION	
(known) → → → → (unknown)	
	系统窗口
	用户窗口

图 1 语义网查询控制面板

若已知 N_1 或 N_2 中的任何一个结点都可以在此查询中查到另一个结点, 即提供了正、逆两种方向上的查询.

3.2 条件查询

这是一种以构件描述词为主的查询, 可以对于类、模块和操作进行查询, 以交互方式输入查询词, 使用多个查询词时, 应指出这些词之间的“与”、“或”关系. 可以根据查询结果集合元素的多少, 选择描述词或调整这些描述词之间的关系.

3.3 构件评价

构件评价是指对于候选构件集合中的每个元素, 衡量它与用户提出要查找的构件之间的匹配程度. 用户应对要查询的构件给出候选条件和构件细节的规格说明. 这里采用了与人工智能中类比学习相似的方法.

设用户构件 U 的特征集合为 US , 库中构件 L 的特征集合为 LS . 定义 U 与 L 的匹配度为

$$MUL = \frac{\text{card}(US \cap LS)}{\text{card}(US)};$$

定义 L 与 U 的匹配度为

$$MLU = \frac{\text{card}(US \cap LS)}{\text{card}(LS)}$$

例如, 一个类构件的特征集合包括它的概括知识, 状态变量知识, 操作知识和活动规律知识. 最终的匹配度是加权后的用户类与库中类的对应部分的匹配度之和.

当候选构件较多时, 使用评价方法求出各构件的匹配度, 然后按匹配度从高到低排序后输出, 必要时可以查看构件内部各部分的匹配情况. 如类中的示例变量、操作及参数之间的匹配情况.

4 总 结

本文提出的 KCQ 系统是基于知识的、可重用构件的检索系统,是人工智能技术在软件重用研究中的一个尝试,采用专家系统结构使事实与规则分离,便于检索功能的丰富和完善,更适合于实验性系统的开发. KCQ 系统在 SUN 工作站上用 C 语言开发完成.

参考文献

- 1 Frakes W B, Gandel P B. Representing reusable software. *Information and Software Technology*, 1990, 32(10).
- 2 RUBEN PRIETO—DIAZ. Classification of reusable modules. *IEEE Software*, 1987(4).
- 3 董韫美,李开德. 一种面向可理解性的可复用软件开发方法. *软件学报*, 1990, 1(1): 31—38.
- 4 Hiroyuki Tarum *et al.* A programming environment supporting reuse of object—oriented software. *Proc. Tenth INT'l Software Engineering Conf.*, 1988.
- 5 Murray Wood *et al.* An information retrieval system for software components. *Software Engineering*. Sep. 1988.
- 6 全炳哲,余江,金淳兆. 可重用构件及其描述语言. *软件学报*, 1994, 5(1).

A KNOWLEDGE—BASED RETRIEVAL SYSTEM FOR SOFTWARE COMPONENTS

Yu Jiang, Quan Bingzhe and Jin Chunzhao

(*Department of Computer Science, Jilin University, Changchun 130023*)

Abstract A knowledge—based system for querying reusable software components is presented in this paper. It supports object—oriented programming method. In this system, the authors used a semantic network model for knowledge representation, and provided a knowledge representation language, CKRL, for describing the features of the reusable components and the knowledge for querying. The authors provided three kind of ways for querying, or semantic net associate querying, condition querying, and components evaluating.

Key words Software reuse, component querying, knowledge—based.