

并发 Lambek 演算在时态查询中的语义转换*

刘冬宁^{1,2}, 汤庸¹

(1. 中山大学计算机科学系, 广东 广州 510275; 2. 中山大学数学系, 广东 广州 510275)

摘要: 为寻找时态逻辑的替代者, 以公理化的方式对时态查询语言做句法和语义的分析, 构建了并发的 Lambek 演算 (L_{CRQ}) 及其范畴语法用以作时态查询的句法分析。根据 L_{CRQ} 演算构造相应的 L_{CRQ} 标号自然演绎系统, 从而提供句法演算到语义计算的转换方式, 为时态查询语句的语义分析提供相关接口。

关键词: 时态查询; 并发的 Lambek 演算; 语义转换; 标号自然演绎系统

中图分类号: TP301 **文献标识码:** A **文章编号:** 0529-6579 (2009) 01-0137-04

Semantic Transform of Concurrency Lambek Calculus in Temporal Querying

LIU Dongning^{1,2}, TANG Yong²

(1. Department of Computer Science, Sun Yat-Sen University, Guangzhou 510275, China;

(2. Department of Maths, Sun Yat-Sen University, Guangzhou 510275, China)

Abstract: In order to find a replacement of temporal logic is necessary, which becomes the implements of syntactic and semantic analysis of temporal querying language by axiomatic methods, We built a axiomatic system used in syntactic analysis which is called Concurrency Lambek Calculus (L_{CRQ}). L_{CRQ} can be switched to its label natural deduction system, which is the interface from syntactic to semantic.

Key words: temporal querying; concurrency Lambek calculus; semantic transform; labeled natural deduction system

关于时间的表达、推理与查询一直是时态数据库中最为关键的研究点。目前许多的时态查询语言已经使用于时态数据库, 其理论基础也相对繁杂, 但其中最为直接关联并得到广泛应用的就是基于时态逻辑的查询语言^[1]。这主要是由于时态逻辑能充分体现时间的特性。然而, 1994年 Gabby 等论证了时态逻辑是不可递归公理化的, 由此它的公理化系统和证明论方法不适用于时态数据库查询语言的句法分析和语义分析^[2]。这使得相关形式化研究工作存在较大的遗憾, 时态逻辑无法利用公理化系统的良好性质及相关证明论方法对时态数据库的推

理和查询做更为严谨和细致的刻画。因此, 目前用于时态数据库查询语言的时态逻辑都是非公理化的, 无论其适用与否, 对于计算性语言这都是一个不足。

时态逻辑之所以与时态查询密切相关, 主要是因为其体现了时间的特性, 在句法分析和语义分析方面都对时间的处理给予极大支持。因此, 要取代时态逻辑在时态查询语言中的地位, 替代者也必须能恰如其分的体现时态属性、时态关系之间的运算, 并最终用于时态查询语言的句法分析和语义分析。针对此, 我们曾提出: 以 Lambek 演算^[3]为核

* 收稿日期: 2008-09-04

基金项目: 国家自然科学基金资助项目 (60673135、60373081); 国家自然科学基金重点资助项目 (60736020); 广东省自然科学基金资助项目 (7003721); 广东省科技攻关资助项目 (07B010200052); 广州市科技计划资助 (07Z3-D3191)

作者简介: 刘冬宁 (1979年生), 男, 博士后; 通讯作者: 汤庸; E-mail: issty@mail.sysu.edu.cn

心的范畴语法系统作为句法分析工具, 类型演算 (λ -演算^[4] 作为语义分析工具, 从而取代时态逻辑在时态查询中的地位。

据此, 我们构建了并发的 Lambek 演算 (L_{CTQ}) 及其范畴语法用以作时态查询的句法分析, 本文将据 L_{CTQ} 演算构造相应的 L_{CTQ} 标号自然演绎系统, 从而提供句法演算到语义计算的转换方式, 为时态查询的语义分析提供相关接口。

1 L_{CTQ} 的自然演绎系统

L_{CTQ} 系统系统的类型由原子类型 $p_1, p_2 \dots$ 等构成, 所有原子类型的集合表示为 T_p 。类型的有穷非空矢列集合写作 T_p^+ 。 L_{CTQ} 演算的矢列形如 $\Gamma \vdash A$, 并包含 4 个二元连接词 $\cdot, \rightarrow, \leftarrow$ (和 $|$)。

该演算系统的自然演绎系统如下所示:

公理: $A \vdash A$

规则:

$$\frac{X \vdash A \quad Y, A, Z \vdash B}{Y, X, Z \vdash B} \text{ (Cut)}$$

$$\frac{X \vdash A \cdot B \quad Y, A, B, Z \vdash C}{Y, X, Z \vdash C} \text{ (}\cdot\text{E)}$$

$$\frac{X \vdash A \quad X \vdash B}{X \vdash A \cdot B} \text{ (}\cdot\text{I)}$$

$$\frac{X \vdash A \quad Y \vdash A \rightarrow B}{X, Y \vdash B} \text{ (}\rightarrow\text{E)}$$

$$\frac{A, X \vdash B}{X \vdash A \rightarrow B} \text{ (}\rightarrow\text{I)}$$

$$\frac{X \vdash A \leftarrow B \quad Y \vdash B}{X \vdash B \leftarrow A} \text{ (}\leftarrow\text{E)}$$

$$\frac{X \vdash A \leftarrow B \quad Y \vdash B | C}{X, Y \vdash A \cdot C} \text{ (|E}^l\text{)}$$

$$\frac{X \vdash A | B \quad Y \vdash B \rightarrow C}{X, Y \vdash A \cdot C} \text{ (|E}^r\text{)}$$

$$\frac{X \vdash A \quad X \vdash B}{X \vdash A | B} \text{ (|I)}$$

除 (|E^l)、(|E^r)、(|I) 三条规则外, 其余规则是经典 Lambek 演算 L 系统的全部规则, 它们组成原 L 系统^[3]。(|E^l)、(|E^r)、(|I) 三条规则是为新添加的连接符“|”加入的消去和引入规则。在消去规则中, (|E^l) 为左消去规则, (|E^r) 为右消去规则。(|I) 为引入规则。

定义 1 如果矢列 $\Gamma \vdash A$ 在 L_{CTQ} 演算中是可推导的, 则写作 $\Gamma \vdash_{L_{CTQ}} A$ 。

2 L_{CTQ} 的句法分析

至此, 根据 L_{CTQ} , 我们可以时态查询语句做相

应句法分析。举例如下:

表 1 人员工作履历表

Table 1 Personnel employment Record

Works Valid	Name	Company
[1997-2001]	John	DEC
[2003-forever]	John	HP
[1999, 2005]	Marry	HP
[2005-forever]	Marry	Microsoft
.....

我们希望查询表 1 中没有失业过的人员名单, 该查询语句如下所示:

```
Select r1. Name
from works r1, works r2
where r1. Name = r2. Name
and Valid (r1) meets Valid (r2)
```

在这里 Valid (r1) 和 Valid (r2) 表示有效时间, 这是基于 ATSQL2 和 TempDB 语法习惯的^[5]。

对上例, 我们可建立词汇表 (LEX), 如表 2 所示。

表 2 词汇表

Table 2 LEX

词项	类型	简要描述
Select	(n→s) ←a	动作类型
r1. Name	a	属性类型
r2. Name	a	属性类型
from	(s→s) ←f	介词类型
works	f 或 f←f	表示“表”的时候类型为 f, 修饰表的时候类型为“f←f”
r1, r2	f	表类型
Where, and	(s→s) ←s	连词类型
Valid (r1)	a←t	时态属性类型
Valid (r2)	t a	时态属性类型
Meets	t (a→ (s←a)) t	时态运算类型
,	f→ (f←f)	连词类型
=	a→ (s←a)	运算类型
System	n	系统公设类型

在这里, 相关词汇类型的设置是根据范畴语法对自然语言的处理相应转换至半结构 SQL 语言进行处理的, 我们额外添加的类型主要体现在时态类型、时态关系运算以及系统公设 (即 System) 上。

各部分类型演算步骤分列如下:

- 1) “System”: 系统公设, 类型为 n;

- 2) “Select r1. Name” 类型演算为：((n→s) ←a) · a ⊢ n→s;
- 3) “from works r1, works r2 ” 类型演算为：((s→s) ←f) · ((f←f) · (f) · (f→(f←f)) · ((f←f) · f) ⊢ s→s);
- 4) “where r1. Name = r2. Name” 类型演算为：((s→s) ←s) · a · (a→(s←a)) · a ⊢ s→s;
- 5) “and Valid (r1) meets Valid (r2)” 类型演算为：((s→s) ←s) · (a←t) · (t | (a→(s←a)) | t) · (t→a) ⊢ s→s;
- 6) 全句最终类型演算为：n · (n→s) · (s→s) · (s→s) · (s→s) ⊢ s。

据 L_{CTQ} 演算，上述句型的类型演算最终能推出终止符类型 s ，使得句法分析演算停止并得到正确结果。而对于不符合句法的查询，则不能得到终止符类型 s ，这是由系统的可靠性和完全性保证的，限于文章篇幅，在此不做介绍。

3 L_{CTQ} 标号自然演绎系统与语义转换

Lambek 演算一个最重要的优点，就是可以利用“Curry-Howard”对应理论^[6-7]，将处理句法分析的演算与应用于语义计算（ λ -演算对应衔接起来，其对应的公理系统为标号自然演绎系统。

根据“Curry-Howard”对应理论，我们给出 L_{CTQ} 演算对应的语义演算系统的 L_{CTQ} 标号自然演绎系统如下所示：

公理：

$$\frac{}{x: A \vdash x: A} \text{ (id)}$$

规则：

$$\frac{X \vdash M: A \quad Y, x: A, Z \vdash N: B}{Y, X, Z \vdash N [M/x]: B} \text{ (Cut)}$$

$$\frac{X \vdash M: A \cdot B \quad Y, x: A, y: B, Z \vdash C}{Y, X, Z \vdash N [(M)_0/x] [(M)_1/y]: C} \text{ (}\cdot\text{E)}$$

$$\frac{X \vdash M: A \quad X \vdash N: B}{X \vdash \langle M, N \rangle: A \cdot B} \text{ (}\cdot\text{I)}$$

$$\frac{X \vdash M: A \quad Y \vdash N: A \rightarrow B}{X, Y \vdash NM: B} \text{ (}\rightarrow\text{E)}$$

$$\frac{x: A, X \vdash B}{X \vdash \lambda x M: A \rightarrow B} \text{ (}\rightarrow\text{I)}$$

$$\frac{X \vdash M: A \leftarrow B \quad Y \vdash N: B}{X, Y \vdash MN: A} \text{ (}\leftarrow\text{E)}$$

$$\frac{X, x: A \vdash B}{X \vdash \lambda x M: B \leftarrow A} \text{ (}\leftarrow\text{I)}$$

$$\frac{X \vdash M: A \leftarrow B \quad Y \vdash N: B | C}{X, Y \vdash MN: A \cdot C} \text{ (}\mid\text{E}^1\text{)}$$

$$\frac{X \vdash M: A | B \quad Y \vdash N: B \rightarrow C}{X, Y \vdash NM: A \cdot C} \text{ (}\mid\text{E}^r\text{)}$$

$$\frac{X \vdash M: A \quad Y \vdash N: B}{X, Y \vdash \langle M, N \rangle: B} \text{ (}\mid\text{I)}$$

特殊地，在这里我们只对时态属性和时态关系运算做阐述和说明，其余仍和经典 Lambek 演算的标号系统一致^[4]。

针对 $r1. Year$ 和 $r2. Year$ 的类型，我们可用标号系统的 (\leftarrow I) 规则和 (\rightarrow I) 将其 λ 抽象为 $\lambda x M: a \leftarrow t$ 和 $\lambda y N: t \rightarrow a$ 。而 before 的标号系统类型为 $P: t | (a(\rightarrow s \leftarrow a))t$ ，再据 (\mid E¹) 和 (\mid E^r) 规则， $r1. Year$ before $r2. Year$ 的句法分析类型为： $\langle \lambda x M, \langle P, \lambda y N \rangle \rangle : a \cdot ((a \rightarrow (s \leftarrow a)) \cdot a)$ 或 $\langle \langle x M, P \rangle, \lambda y N \rangle : (a \cdot (a \rightarrow (s \leftarrow a))) \cdot a$ (由于 \cdot 是可结合的，实际上两者等价)，从语义上看，这是一个时态关系代数绑定了两个时间属性的 λ 函数，并由此转入相应 λ 语义计算，将对时间值的处理转入相关 λ 抽象处理。

λ 语义的抽象处理实际上是函数化和参数化的^[4]，其与 SQL 的 bag 语义处理方式相似程度较高，由此可以作为句法分析和语义分析的良好接口。在国际上 (λ -演算和 Lambek 演算及线性逻辑的许多交叉研究成果已融入到语义网 (Semantic Web)^[8]、数据流程序^[9] (Dataflow Programming) 和模糊分析信息系统^[10] 等人工计算语言的语义分析工作上，证明了其是有效的分析方法，值得借鉴。

4 结 语

在本文中，我们据 Curry-Howard 对应理论为 L_{CTQ} 演算构造了相应的 L_{CTQ} 标号自然演绎系统，从而提供了句法演算到语义计算的转换方式，为时态查询语句的语义分析提供相关接口。在下一步工作中，将主要研究类型演算 λ -演算在时态查询中的语义演算和推理，并将其作为时态查询的语义机理论基础，建立相应的原型系统或直接在 TempDB^[5] 上进行验证。

参考文献：

- [1] CHOMICKI J, SAAKE G. Logics for databases and information systems: Temporal logic information system [M]. Boston/Dordrecht/London: Kluwer Academic Publishers, 1998.
- [2] GABBAY D M, HODKINSON I, REYNOLDS M. Tem-

- poral logic: Mathematical foundations and computational aspects [M]. Oxford: Oxford University Press, 1994.
- [3] LAMBEK J. The mathematics of sentence structure [J]. Amer Math Monthly, 1958, 65: 154 - 170.
- [4] BARENDECT H P. The lambda calculus its syntax and semantics [J]. Studies in Logic and The foundations of Mathematics, 2004, 103(6).
- [5] <http://www.cosoft.sysu.edu.cn/TempDB/index.htm>, [2008].
- [6] CURRY H B. Some logic aspects of grammatical structure [C]// JAKOBSON R. Structure of language and its mathematical aspects. AMS, Providence, RI.
- [7] HOWARD W. The formulae-as-types notion of construction [M]. Manuscript, Published in Seldin and Hindley, 1980.
- [8] RAO J H, KÜNGAS P. Logic-based web services composition from service description to process model [C]//In 2nd Intl. Conference on Web Services, 2004.
- [9] UUSTALU T, VENE V, The essence of dataflow programming [C]// In 3rd Asian Symp. on Programming Languages and Systems, APLAS' 05, Number 3780 in Lect. Notes Comp Sci, 2005.
- [10] HAJEK P. Metamathematics of fuzzy logic [C]//Trends in Logic 4. Kluwer, 1998.

(上接第 136 页)

- [6] FUGATE M, GATTIKER J R. Anomaly detection enhanced classification in computer intrusion detection [C] // LNCS 2388. Berlin, Heidelberg: Springer-Verlag, 2002: 186 - 197.
- [7] KIM D, PARK J. Network-based intrusion detection with support vector machines [C] // LNCS 2662. Berlin, Heidelberg: Springer-Verlag, 2003: 747 - 756.
- [8] PARK J, SHAZZAD K, KIM D. Toward modeling lightweight intrusion detection system through correlation-based hybrid featureselection [C] // FENG D, LIN D, YUNG M. Proceedings of the CISC. Heidelberg: Springer-Verlag, 2005: 279 - 289.
- [9] TAYLOR C, ALVES-FOSS J. NATE: Network analysis of anomalous traffic events, a low-cost approach [C] //Proceedings of the 2001 Workshop on New Security Paradigms. New Mexico: ACM, 2001: 89 - 96.
- [10] HORNG S, FAN P, CHOU Y, et al. A feasible intrusion detector for recognizing IIS attacks based on neural networks [J]. Computers & Security, 2008, 27 (3 - 4): 84 - 100.
- [11] ZHANG Changyou, CAO Yuanda, Yang Minghua, et al. The immune recognition method based on analogy reasoning in IDS [J]. Wuhai University Journal of Natural Sciences, 2006, 11 (6): 1839 - 1843.
- [12] 焦李成, 杜海峰. 人工免疫系统进展与展望 [J]. 电子学报, 2003, 31 (10): 1540 - 1548.
- JIAO Licheng, DU Haifeng. Development and prospect of the artificial immune system [J]. Acta Electronica Sinica, 2003, 31 (10): 1540 - 1548.
- [13] 肖人彬, 王磊. 人工免疫系统: 原理、模型、分析及展望 [J]. 计算机学报, 2002, 25 (12): 1281 - 1293.
- XIAO Renbin, WANG Lei. Artificial immune system: principle, models, analysis and perspectives [J]. Chinese Journal of Computers, 2002, 25 (12): 1281 - 1293.