电　子　学　报
ACTA ELECTRONICA SINICA

基于时间 Petri 网的工作流系统模型的线性推理

刘　婷，林　闯，刘卫东
(清华大学计算机科学与技术系，北京 100084)

摘　要：目前工作流理论的研究主要集中在工作流管理模型的结构及正确性分析，很少有人研究与时间有关的工作流模型的性质，特别是模型中的时序关系推理及性能计算问题. 本文重点研究了这方面的问题，用时间 Petri 网表示工作流模型并对基本工作流模型进行时序分析，给出线性时间推理的规则，运用这些规则，可对复杂的工作流模型进行逐步化简，并在线性时间复杂度内解决时间推理问题.

关键词：　工作流管理系统；时间 Petri 网；线性推理
中图分类号：　TP391.7　　**文献标识码**：　A　　**文章编号**：　0372-2112 (2002) 02-0245-04

Linear Temporal Inference of Workflow Management System Based on Timed Petri Net Models

LIU Ting，LIN Chuang，LIU Wei-dong
(Department of Computer Science and Technology, Tsinghua University, Beijing 100084*, China)*

Abstract：　The research on workflow has been focused on the verification of workflow task structure，but few have paid attention to the properties of the models，which are closely related to the time，especially the problems of temporal inference and performance calculation. In this paper，we research on this aspect. Timed Petri Nets are used to model workflows as well as analyze the temporal relations of the model. A set of linear inference rules is put forward. Using these rules，we can simplify the complicated model step by step，and solve the inference problem within linear time complexity.

Key words：　workflow management systems；timed Petri nets；linear inference

1　引言

工作流是业务流程的全部或部分自动化，在此过程中，文档，信息或任务按照一定的过程规则流转，实现组织成员间的协调工作以期达到业务的整体目标. 从理论上概括来说，工作流管理是一个包括定义，监测，控制，优化，以及相关逻辑支持的复杂过程. 它是一个三维的概念，即过程，事件，资源的有机结合[1].

近年来，随着用工作流的方法管理业务流程的需求的不断增加，人们提出了各种各样的模型和分析方法[2,3]. 在工作流理论方面，W. M. P. van der Aalst 做出了重要贡献. 他对复杂工作流模型结构性质进行了分析[4]，提出了工作流模型的顺序，并行，选择，循环等基本结构及复合模型的有效性(Soundness)，提出了对模型正确性的判定规则及方法. 此外，Petri 网(TPN)模型方法[5,6]已被用来描述工作流模型. 主要是因为 Petri 网可以很好的描述模型的静态和动态特性——顺序，并行和冲突. 同时也是因为 Petri 网有有效的数学分析技术，可以给出问题求解的算法[7,8].

对工作流模型结构的研究已经发展到相当水平，然而工作流作为一个过程流，运用于实际系统时多数系统与时间有密切的关系，但是系统模型中的这种时序关系及性能计算问题却常常被人们忽视. 本文重点考虑工作流模型在这方面的性质. 本文首先基于时间 Petri 网对工作流的基本模型进行分析，提出线性推理的规则. 然后推广到由基本模型复合构成的复杂工作流模型中. 同时，利用线性推理规则的组合对此种模型进行压缩推理，可在线性时间复杂度内对模型进行时序关系的计算，进而实现对工作流系统的时间性能计算.

2　工作流基本的时间 Petri 网模型与线性推理规则

在时间 Petri 网(TPN)模型中，每个变迁被赋予一个时间区间 $X(t) = [t_l, t_u]$，$(0 \le t_l \le t_u)$. t_l，t_u 是相对时间，它们相对于变迁可实施的时刻，这个起始时间由系统决定. 假如变迁在 s 时刻可实施，实际实施的时间记为 t^*，则 $s + t_l \le t^* \le s + t_u$. TPN 模型的这种性质可对应于工作流过程模型中的一个由时间触发的基本事件的表示. 相应 TPN 模型的变迁可实施时刻 s 为事件可实施时刻，事件在时段 $[s + t_l, s + t_u]$ 内实施并完成. 基本事件 X 对应的 TPN 模型如图 1 所示.

收稿日期：2000-03-19；修回日期：2001-08-08
基金项目：国家重点基础研究发展规划项目(No. G1999032707)；国家自然科学基金(No. 60173012)

一般变迁称为时间变迁,在 TPN 模型中用矩形表示. 当 $X(t) = [0,0]$时,变迁在可实施时刻即被触发,这种变迁被称为瞬时变迁,在 TPN 模型中用竖线表示.

复杂的工作流模型都可以由一些基本的结构复合构成[1],这几种基本结构为顺序,选择,并行,和循环等. 从工作流过程的角度,用几个基本流控制结构 AND-split, AND-join, OR-split 和 OR-join 等组成顺序、选择、并行和循环等基本模型.

在下面,将定义这四种基本模型,并给出线性推理的规则,并对模型进行时间压缩等效变换,变换原则如下:

初始模型为工作流基本模型,统一表示为从唯一的源位置 A 出发,经过一子网到达汇位置 C. 经压缩变换后达到的目标模型为从 A 出发经等效时间变迁 T 到达 C,变换过程如下图:

® 顺序

如果一个变迁的实施直接导致另一个变迁可以被实施,则称此二变迁的发生关系是顺序的. 在图 3 中变迁 $T2$ 的可实施时刻即为 $T1$ 实际实施的结束时刻,即 $T1$, $T2$ 是顺序实施的.

推理分析:

图 3 中 $T1$, $T2$ 构成子网. 设等效变迁为 $T[t_l, t_u]$,需求出 t_l, t_u.

记 s_1、s_2、s 分别为 T_1, T_2, T 的可实施时刻, $t1^*$、$t2^*$、t^* 分别为 $T1$、$T2$、T 的实际实施时刻. 由 TPN 定义:

$s_1 + t_{1l} \le t1^* \le s1 + t_{1u}$;

$s_2 + t_{2l} \le t2^* \le s2 + t_{2u}$;

$s + t_l \le t^* \le s + t_u$.

由顺序实施的定义: $t1^* = s2$.

由 T 是等价变迁, $s = s1$, $t^* = t2^*$.

由不等式的性质得:

$s1 + s2 + t_{1l} + t_{2l} \le t1^* + t2^* \le s1 + s2 + t_{1u} + t_{2u}$,进一步,由 $t1^* = s2$, $t^* = t2^*$, $s = s1$ 可得:

$$s + t_{1l} + t_{2l} \le t^* = t2^* \le s + t_{1u} + t_{2u},$$

即可得等效变迁 $T[t_l, t_u]$,其中 $t_l = t_{1l} + t_{2l}$; $t_u = t_{1u} + t_{2u}$.

顺序等效压缩过程由下面规则 1 表示:

规则 1 $(T1, T2) \xrightarrow{1} T[t_l, t_u] = T[t_{1l} + t_{2l}, t_{1u} + t_{2u}]$.

® 并行

在图 4 中,我们称变迁 $T1$ 和 $T2$ 是并行的关系,即 $T1$ 和 $T2$ 之间没有直接的时间约束关系. 此模型由两个基本流控制结构构成: (1) AND-split, (2) AND-join. 如图 4 所示,两个结构都用瞬时变迁表示,起到对工作流控制的作用. AND-split 的执行使 $T1$ 和 $T2$ 同时可实施. $T1$ 和 $T2$ 都执行完后, AND-join 可实施,它的作用是同步两个子工作流,并在 C 中产生新标识. $T1$ 和 $T2$ 是并行执行的关系.

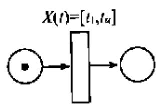

图 1 基本事件 X 的 TPN 模型

图 2 等效变换图形

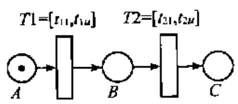

图 3 顺序关系模型

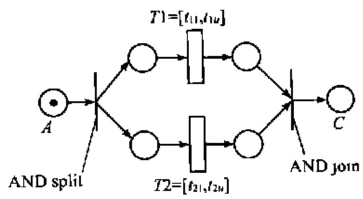

图 4 并行关系模型

推理分析:

由图 4 可知两个瞬时变迁和 $T1$, $T2$ 共同构成子网. 记 $s1$、$s2$、s 分别为 $T1$, $T2$, T 的可实施时刻, , $t1^*$、$t2^*$、t^* 分别为 $T1$、$T2$、T 的实际实施时刻.

从 A 经瞬时变迁使 $T1$, $T2$ 同时可触发,即 $s1 = s2 = s$. 由 TPN 定义:

$s1 + t_{1l} \le t1^* \le s1 + t_{1u}$;

$s2 + t_{2l} \le t2^* \le s2 + t_{2u}$;

即 $T1$, $T2$ 的实际触发时间的时间范围分别为 $[s + t_{1l}, s + t_{1u}]$和 $[s + t_{2l}, s + t_{2u}]$. $T1$, $T2$ 全部结束的最早时刻为 $s + \max(s + t_{1l}, s + t_{2l})$,最晚时刻为 $s + \max(s + t_{1u}, s + t_{2u})$,即 T 的等效最早触发时间为 $s + \max(s + t_{1l}, s + t_{2l})$,最晚时间为 $s + \max(s + t_{1u}, s + t_{2u})$.

得到: $s + \max(t_{1l}, t_{2l}) \le t^* \le s + \max(t_{1u}, t_{2u})$;所以 $t_l = \max(t_{1l}, t_{2l})$, $t_u = \max(t_{1u}, t_{2u})$.

并行等效压缩过程由下面规则 2 表示:

规则 2

$$(T1, T2) \xrightarrow{2} T[t_l, t_u] = T[\max(t_{1l}, t_{2l}), \max(t_{1u}, t_{2u})].$$

® 自由选择

在自由选择关系模型中, $T1$, $T2$ 两个变迁处于时间选择冲突的关系,不能同时实施,见图 5. 每次系统选择 $T1$, $T2$ 之一执行,究竟选择哪个变迁由系统随机决定. 采用两个基本流控制结构: (1) OR-split 和 (2) OR-join. 一个 OR-split 结构用一个含有多个输出弧的位置表示;一个 OR-join 结构用一个含有多个输入弧的位置表示. 标识从 A 出发,经瞬时变迁后,执行 $T1$ 或 $T2$, $T1$ 或 $T2$ 执行结束后,经瞬时变迁,到达 C.

推理分析:

$T1$, $T2$ 是冲突的关系. 等效变迁 T 或者按 $T1$ 实施,或者按 $T2$ 实施. 根据时间选择的定义, T 的最早实施时间为 $\min[t_{1l}, t_{2l}]$,最晚实施时间为 $\min[t_{1u}, t_{2u}]$. 即 $t_l = \min[t_{1l}, t_{2l}]$, $t_u = \min[t_{1u}, t_{2u}]$. 特别当 $t_{1u} < t_{2l}$时, $T1$ 总被实施, $T2$ 不被实施.

自由选择等效压缩过程写成如下规则 3:

规则 3

$(T1, T2) \xrightarrow{3} T[t_l, t_u] = T[\min(t_{1l}, t_{2l}), \min(t_{1u}, t_{2u})]$.

® 条件选择

此模型与条件选择模型不同的一点是:自由选择模型中,当标识到达 OR-split 处时,选择哪一条路是不确定的;而在条件选择模型中,选择是由工作流的性质决定的,是一种确定的选择.条件选择 Petri 网模型见图 6.

推理分析:

等效变迁 T 或者按 $T1$ 实施,或者按 $T2$ 实施. T 的可实施时段应为 $[t_{1l}, t_{1u}]$ 与 $[t_{2l}, t_{2u}]$ 的并.于是可分为两种情况. (1) 当 $[t_{1l}, t_{1u}]$ 与 $[t_{2l}, t_{2u}]$ 有交集的时候,二者的并仍构成一个连续时段 T,以最早的最小实施时间 $\min[t_{1l}, t_{2l}]$ 实施,以最大的最晚实施时间 $\max[t_{1u}, t_{2u}]$ 实施. (2) 当 $[t_{1l}, t_{1u}]$ 与 $[t_{2l}, t_{2u}]$ 没有交集时,等效变迁的可实施时段变为两个时段的或.

条件选择等效压缩过程得到如下规则 4:

规则 4

$(T1, T2) \xrightarrow{4} T[t_l, t_u] = T[\min(t_{1l}, t_{2l}), \max(t_{1u}, t_{2u})]$ $(T1 \cap T2 \neq \Phi)$

$(T1, T2) \xrightarrow{4} T = [t_{1l}, t_{1u}] \cup [t_{2l}, t_{2u}], (T1 \cap T2 = \Phi)$.

® 循环

循环执行的 Petri 网模型见图 7.

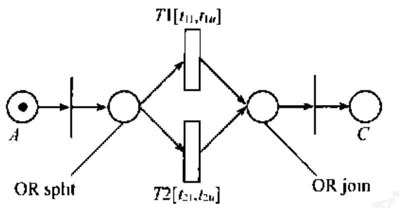

图 5　自由选择关系模型

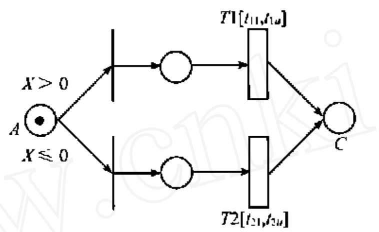

图 6　条件选择关系的 Petri 网模型

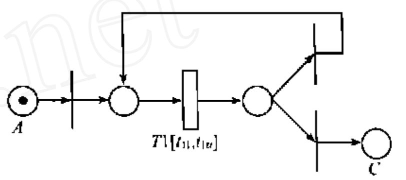

图 7　$T1$ 可能被执行多次的循环模型

推理分析:

$T1$ 可能被执行一次或多次,一般使用 $k \geq 1$ 表达执行的次数.显然等效执行的结果是 T 的最早实施时间和最晚实施时间都被扩大 k 倍.

循环等效压缩过程使用规则 5 计算:

规则 5　$(T1, k) \xrightarrow{5} T[t_l, t_u] = T[kt_{1l}, kt_{1u}]$.

3　复合工作流模型的线性推理

在对复合工作流模型进行推理计算之前,先对复合模型的构成加以限制如下:

在本文中所考虑的工作流过程应该满足有效性(Soundness),即以下几条基本要求[1]:

(1) 一个工作流网只有一个源位置 I (起始状态) 和一个汇位置 O (结束状态).

(2) 每一个变迁/位置在一条从 I 到 O 的路径上.

这两个要求可以静态考察,因为它们只与 Petri 网的结构有关,不仅如此,我们还有以下两条基本规定:

(3) 在任何情况下,过程将最终终结,在过程终结时,标识到达位置 O (结束状态),其它位置没有标识.

(4) 不存在死的变迁,也就是说,对于任意一个工作流网络中的变迁经过某种途径必可以实施.

以上附加的两个性质,也就是我们所说的有效性.

定理 1　对于用时间 Petri 网模型表示的满足有效性的工作流模型,必可运用线性推理规则在线性时间复杂度内定量的解决时间推理计算问题.

证明　在系统中进行时间推理求解一般是回答两类问题:一类是求标识是否可到达某位置(由有效性判定已可解决此问题),另一类是求标识到达某位置所用时间,即性能评价问题,也是线性规则所解决的问题.不妨设 X 是系统中的一个事件,第一类问题是求 X 是否可发生,而第二类问题是定量的求出 X 可实施的时间 s.

考虑如何求 s.由有效性要求 (1) 知,整个系统只有一个起始点 I.可把标识到达 I 的时间定为整个系统的起始时间.标识从 I 到达事件 X 的起始点间必存在一个相关事件集合 $Ux = \{Z1, Z2, \cdots, Zj\}$, ($j \leq$ 模型中总的变迁个数 n) Ux 中事件的实施或者是 X 可实施的前提条件,或者与 X 有冲突关系.由系统有效性,标识以某种方式从 I 出发,经 Ux 中的一些事件到达 X,除循环结构外标识不会重复流经同一位置,而循环结构也可经一步等效压缩,求得等效变迁,即无论哪种情况,最多经 j 步线性推理计算可求得 I 到达 X 的等效时间.而 $j \leq n$,于是证明了对于任何给出数量关系的时间推理问题,可以在线性时间复杂度内解决.　　证毕

下面给出一个例子,说明复合模型推理计算的方法.

例 1:一个旅行社帮助客户预定旅行,首先要对客户进行登记.在这步工作流程中将要记录客户的目的地,要求往返的时间以及其它一些需求并收取一定费用,总共需要花费 20～30 分钟.然后旅行社根据客户提供的信息进行查询,并将得到方案提供给客户进行选择,这个过程需要 60 分钟.结果将有三种可能:(1) 客户选择了一种方案. (2) 客户不满意,要求更多的方案. (3) 客户不满,取消预定.如果客户要求更多的方案,则旅行社将再次按照客户的要求进行查询,并再次提交给客户.如果客户取消约定,则需 5～10 分钟办理退款手续离开.一旦客户选中了一种方案,则进入下一步准备工作.旅行社将通知合作伙伴(机场,旅馆) (需 10～30 分钟).另一方面预约旅行(需 20～40 分钟).在准备工作过程中,客户将说明是否需要保险,并签署协议(需 15～30 分钟).注意:通知,预约和加入保险三步工作可以同时进行.当通知和预约两步工作做完后,相应的文件将被提交给客户,客户离开.

求:(1) 旅行社查询两次仍不能满足客户要求,客户取消约定离开的等效时段.

(2) 客户定好行程所需最短时间.

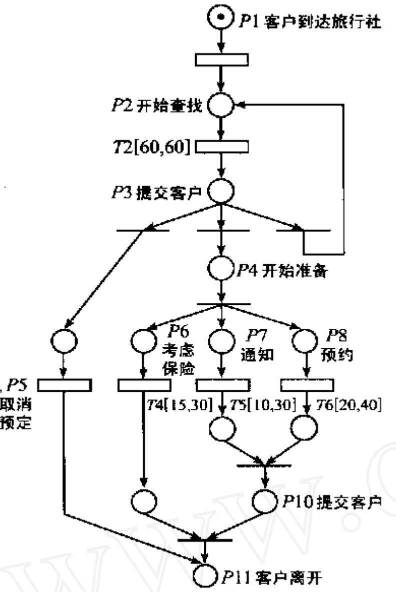

图 8 例 1 的 Petri 网模型表示

解:首先得出如图 8 所示的等价 TPN 模型. 此模型满足有效性. 其中, $P1$ 是起始点, $P11$ 是结束点.

从 $P1$ 到 $P3$ 是顺序和循环模型的组合. 运用规则 1 和 5 的组合可得从 $P1$ 到 $P3$ 的等效变迁 $T1'$:

$(T2, k) \xrightarrow{5} T[kt_{1l}, kt_{1u}] = T[60k, 60k]$

$(T1, T[60k, 60k]) \xrightarrow{1} T[20+60k, 30+60k] = T1'$

$(T1, T2) \xrightarrow{1} T[t_l, t_u] = T[t_{1l}+t_{2l}, t_{1u}+t_{2u}]$.

通知和预约是并行关系,运用规则 2 可得等效变迁 $T2'$:

$(T5, T6) \xrightarrow{2} T[\max(10,20), \max(30,40)] = T[20,40]$.

$T2$ 与 $T4$ 仍是并行的关系,运用规则 2 可得整个准备工作的等效变迁 $T3'$:

$(T4, T2') \xrightarrow{2} T[\max(15,20), \max(30,40)] = T[20,40]$

取消预定 $T3$ 和准备工作 $T3$ 是条件选择关系,运用规则 4:

$(T3, T3') \xrightarrow{4} T = [5,10] \cup [20,40]$

再运用规则 1,可得从客户到达旅行社到取消预定离开的等效变迁 $T4'$:

$(T1', T[5,10]) \xrightarrow{1} T[60k+5, 60k+10]$.

(1) 当 k 取 2 时,可得旅行社查询两次仍不能满足客户要求,客户取消约定离开的等效时段为 2 小时 5 分钟到 2 小时 10 分钟之间.

从客户到达旅行社到定好行程离开的等效变迁 $T5'$:

$(T1', T[20,40]) \xrightarrow{1} T[60k+20, 60k+40]$.

(2) 当 k 取 1 时,得到客户定好行程的最短时间为 80 (60 + 20 = 80) 分钟.

4 结论

本文考虑了有时间约束的工作流模型的结构性质,用时间 Petri 网模型描述了工作流模型,提出线性推理的规则. 然后推广到由基本模型复合构成的复杂工作流模型中,证明了对于满足有效性的复杂工作流模型,可以在线性时间复杂度内解决时间推理计算问题,进而可实现对工作流系统的性能评价. 本文只是将推理问题及性能评价的方法引入工作流模型的初步尝试,进一步还可以考虑时间段是随机不确定的情况,应用 Petri 网模型的数学分析的基础,可深入对工作流模型的时间推理分析和评价.

参考文献:

[1] W M P van der Aalst. The application of Petri nets to workflow management [J]. Journal of Circuits, Systems, and Computers, 1998, 8(1): 21 - 66.

[2] W M P van der Aalst. Chapter 10; Three Good reasons for Using a Petri-net-based Workflow Management System. In T. WaKaYaMa et al., editor, Information and Process Integration in Enterprises; Rethinking documents [M]. The Kluwer International Series in Engineering and Computer Science, Kluwer Academic Publishers, Norwell, 1998: 161 - 182.

[3] C A Ellis, K Keddara, G Rozenberg. Dynamic change within workflow systems [J]. In N. Comstock and C. Ellis, editors, Conf. on Organizational Computing Systems, ACM, SIGOIS, Milpitas, CA (1995): 10 - 21.

[4] W M P van der Aalst. Verification of workflow tast structures [A]. Information Systems, 2000, 25(1): 43 - 69.

[5] Zaidi A K. On temporal logic programming using Petri nets [J]. IEEE Transactions on Systems, Man and Cybernetics, May 1999, 29(3): 245 - 254.

[6] Yao Y. A Petri net model for temporal knowledge representation and reasoning [J]. IEEE Trans. Systems, Man, and Cybernetics, 1994, 24(9): 1374 - 1382.

[7] Lin C, Chanson S T. Logical inference of clauses based on Petri net models [J]. International Journal of Intelligent Systems, John Wiley & Sons, August 1998, 13: 821 - 840.

[8] Lin C, Chaudhury A, Whinston A B, Marinescu D C. Logical inference of Horn clauses in Petri net models [J]. IEEE Trans. on Knowledge and Data Engineering, June 1993, 5(4): 416 - 425.

作者简介:

刘 婷 女. 1978 年生于北京. 2001 年毕业于清华大学计算机科学与技术系,获学士学位. 感兴趣的研究方向是计算机网络、时序逻辑和工作流模型.

林 闯 男. 1948 年生于辽宁沈阳. 清华大学计算机系教授,博士导师,《计算机学报》编委,中科院网络中心和北京科技大学兼职教授. 主要研究领域为计算机网络、系统性能评价、随机 Petri 网、逻辑推演和推理系统. 已在 IEEE Transactions on Computers, IEEE Transactions on Knowledge and Data Engineering, ACM Journal of Wireless Networks, International Journal of Intelligent Systems, IEICE Transactions of Fundamentals, 计算机学报, 软件学报、电子学报、通信学报等国内外核心期刊上和 IEEE Computer Society 的学术年会上发表论文 70 多篇.