

**В. М. Левыкин**, доктор технических наук,  
заведующий кафедрой информационных  
управляющих систем Харьковского национального  
университета радиоэлектроники  
**О. В. Чалая**, кандидат экономических наук,  
доцент кафедры информационных  
управляющих систем Харьковского национального  
университета радиоэлектроники

## **МОДЕЛЬ МНОГОУРОВНЕВОГО ПРЕДСТАВЛЕНИЯ ТЕМПОРАЛЬНЫХ ЗНАНИЙ В ЗАДАЧАХ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ПРОЦЕССОВ**

*Предложена модель многоуровневого представления темпоральных знаний, которая разрешает отразить в отдельных временных потоках изменение объектов моделируемой предметной области на различных уровнях иерархии. В задачах интеллектуального анализа процессов модель позволяет единым образом представить лог процесса и его модель.*

*Ключевые слова: представление знаний; темпоральные знания; свойства времени; темпоральные логики; интеллектуальный анализ процессов.*

*A model of multi-level representation of temporal knowledge is proposed. Model allows to reflect the temporal changes in the modeled domain at different levels of the hierarchy. The model proposes a unified way to present and a log and a process model.*

*Key words: knowledge representation; temporal knowledge; properties of time; temporal logics; intelligent analysis process.*

**Постановка проблемы.** Представление знаний является областью искусственного интеллекта, связанной с документированием знаний в форме, пригодной для хранения и организации логического вывода в интеллектуальных информационных системах [1]. В настоящее время в данной области используют различные подходы к представлению пространственных и темпоральных знаний. Пространственные знания отражают взаимосвязи, свойства, структуру объектов предметной области; темпоральные – изменение пространственных знаний с течением времени.

**Анализ последних исследований и публикаций.** Для моделирования темпоральных знаний используются дискретное и интервальное представления [2; 3].

В качестве перспективного направления рассматривается построение общего темпорально-пространственного представления знаний на основе единых принципов.

Практическая необходимость построения единой темпорально-пространственной модели знаний возникает в частности при решении задач интеллектуального анализа процессов (processmining).

В рамках данного направления выполняется построение модели процесса на основе анализа его логов. Модель содержит либо процедурное, либо декларативное представление

© В. М. Левыкин, О. В. Чалая, 2015

---

знаний о последовательности действий по решению прикладной задачи в конкретной предметной области в различных условиях внешней среды. В качестве исходных данных для построения модели процесса выступает его лог-файл.

Лог состоит из набора следов (трасс) процесса. Каждый след содержит последовательность привязанных к моментам времени событий, отражающих однократное выполнение процесса. Многократная реализация процесса может выполняться как последовательно в одном месте предметной области (на одном объекте) в различные моменты времени, так и параллельно на различных объектах. Это означает, что в общем случае для каждого следа процесса задается своя шкала времени.

Традиционно при решении задач интеллектуального анализа процессов выполняется слияние временных шкал на основе сопоставления атрибутов событий. В результате получается модель процесса, отражающая его возможные реализации и использующая модель ветвящегося времени. При таком подходе в одной модели объединяются действия для различных уровней иерархии исходного объекта, на котором выполняется рассматриваемый процесс. В результате на практике при использовании логов действующих предприятий получаем так называемую “спагетти-модель”, понимание которой обычно затруднено [4].

Для анализа полученной “спагетти-модели” изменяют степень ее детализации, используя пороговые уровни для количества повторений пар событий из лога процесса. Это дает возможность изменять глубину рассмотрения модели путем отображения максимального количества связей между событиями либо удаления редко используемых связей. Однако при таком подходе не рассматривается семантика удаляемых связей.

Семантика связей основана на внутренней структуре предметной области, в которой выполняется процесс. Иными словами, семантика темпоральных зависимостей базируется на пространственной составляющей процесса и выражается через связи между объектами предметной области и их свойствами. При этом шкалы времени для различных объектов могут отличаться. Так, для иерархии объектов в предметной области события, возникающие для объектов верхнего уровня иерархии, обычно являются производными от последовательности событий для объектов нижних уровней той же иерархии, поскольку события отражают выполняемые с объектами действия. Например, события на уровне руководства предприятия происходят в результате цепочки событий на уровне исполнителей. Соответственно, оценка времени для этих уровней иерархии будет отличаться.

В связи с изложенным актуальной является проблема построения многоуровневого представления темпоральных знаний для задач processmining, позволяющего учесть статические взаимосвязи между объектами предметной области.

**Цель статьи.** Поставленная в работе задача заключается в построении модели представления темпоральных знаний, представляющего протекание полученного в результате анализа логов процесса в форме набора временных потоков, соответствующих различным уровням иерархии предметной области.

Для решения поставленной задачи необходимо выделить свойства времени, применимые к многоуровневой темпоральной модели, и представить реализацию этих свойств как для исходного лога, так и для результирующей модели процесса.

**Изложение основного материала.** Представление темпоральных знаний базируется на двух возможных шкалах событий [5]: “прошлое – настоящее – будущее” и “раньше – позже”.

---

Для первой шкалы существует точка отсчета – текущий момент времени, которая перемещается по шкале. Событие, заданное в будущем, может реализоваться в настоящем. В таком случае данное событие фиксируется в прошлом. Если же событие не реализовалось, то оно исчезает со шкалы.

Данная шкала может содержать либо одну, либо несколько версий будущего. Свойства времени при этом отличаются. В первом случае время линейно. Последовательность событий в будущем может быть описана в виде линейной темпоральной логики LTL [6]. Во втором случае из всех возможных вариантов последовательностей событий в процессе перехода через текущий момент времени остаются только реализовавшиеся события. Поэтому второй случай соответствует логике ветвящегося времени (CTL) [7; 8].

Шкала “раньше – позже” фиксирует последовательность уже произошедших событий. События на данной шкале уже упорядочены и более не изменяются. Данная шкала обладает свойствами линейности и упорядоченности.

Для каждой шкалы используется свое подмножество набора свойств времени. При моделировании времени в интеллектуальных системах учитывают его свойства направленности, линейности, непрерывности, неоканчиваемости, гомогенности (однородности) [5].

Направленность отражает течение процессов (последовательности событий) из прошлого в будущее. Фиксация направления определяет семантику описания знаний, например в виде “обязательно произойдет в будущем”, “возможно случится в будущем”. Поскольку прошлое и будущее обладают разными свойствами, то при формализации направленности времени их целесообразно разделять. Формализация может быть выполнена на базе темпоральных операторов  $U^+$  и  $U^-$  [9].  $e'U^+e''$  означает, что событие  $e'$  будет повторяться до тех пор, пока не произойдет событие  $e''$ .  $e'U^-e''$  означает, что после события  $e''$ , произошедшего в прошлом, до настоящего времени многократно происходили события  $e'$ .

Отметим, что через оператор  $U^+$  выражаются темпоральные операторы  $X$  (событие произойдет в следующий момент времени),  $F^+$  (событие произойдет в будущем) и другие. Аналогично через  $U^-$  можно выразить  $F^-$ .

Тогда при представлении направленности имеем 3 интервала времени: только будущее; только прошлое; только настоящее.

Соответственно, в модели процесса момент настоящего определяется в момент моделирования, лог же содержит только события из прошлого.

Линейность времени обычно рассматривается в виде множества, на котором задано отношение порядка, включающее отношение перехода, достижимости и равенства соответственно:  $R = \{<, <=, =\}$ .

Модель времени для следов процесса обладает свойством линейности, поскольку след содержит упорядоченную по времени последовательность событий, соответствующую однократному выполнению процесса.

В то же время модель процесса содержит набор возможных вариантов реализации. Иными словами, будущее в модели процесса ветвится и каждая ветвь соответствует альтернативной последовательности событий. Поэтому свойство линейности для модели нарушается.

---

Свойство непрерывности времени свидетельствует о возможности бесконечного увеличения масштаба времени. То есть между двумя текущими событиями всегда можно найти еще одно. Применительно к рассматриваемой задаче свойство непрерывности нарушается в силу ограниченного количества событий в логге.

Свойство неоканчиваемости означает, что время можно бесконечно продолжать как в направлении прошлого, так и в направлении будущего. В рамках рассматриваемой задачи представляется разумным сузить данное свойство и свести его к свойству цикличности.

Свойство гомогенности означает, что любую последовательность событий можно перенести по шкале времени из прошлого в будущее и наоборот, при этом последовательность событий не изменится.

При реализации интеллектуального анализа процессов происходит перенос последовательности событий из прошлого (лог) в будущее (модель процесса). Поэтому данное свойство истинно для модели и не поддерживается логом.

Обобщение изложенных выше возможностей формализации свойств времени при представлении темпоральных знаний в задачах processmining представлено в табл. 1.

Таблица 1

**Отражение свойств времени при представлении темпоральных знаний  
в задачах processmining**

Свойства времени	Модель процесса	След процесса
1. Направленность	+	+
2. Линейность	-	+
3. Непрерывность	-	-
4. Неоканчиваемость	-- (+ цикличность)	-
5. Однородность	+	-

В результате анализа представленных в таблице данных можно сделать следующие выводы.

Свойство направленности времени используется при представлении как модели процесса, так и для исходных логов, позволяя построить модель на основе объединения логов. Свойство линейности характерно только для логов процесса. Модель процесса содержит альтернативные варианты будущих последовательностей событий, поэтому в данном случае используется не линейное, а ветвящееся время. Свойство непрерывности не формализуется в данных задачах в силу дискретного представления событий в исходных логах. Аналогично свойство неоканчиваемости времени неприменимо при решении задач processmining. Процесс реализует решение некоторой задачи в предметной области и, следовательно, обязательно содержит начальное и одно или несколько (в случае множественности решения) конечных событий. Применение свойства однородности времени приводит к появлению “спагетти-моделей”, объединяющих события процесса вне зависимости от степени их важности для пользователя.

Представленные в таблице возможности отображения свойств времени при представлении темпоральных знаний позволяют обосновать возможность построения модели представления знаний на основе отображения параллельно протекающих ли-

нейных потоков времени, используя свойства 1 и 2. Тогда в модели альтернативные варианты развития событий формализуются через переход на иной поток времени.

Как показано на рис. 1, при представлении темпоральных знаний в форме трех потоков времени  $T_0, T_1, T_2$  каждый из потоков обладает свойством линейности, т. е. не содержит альтернативных версий будущего.

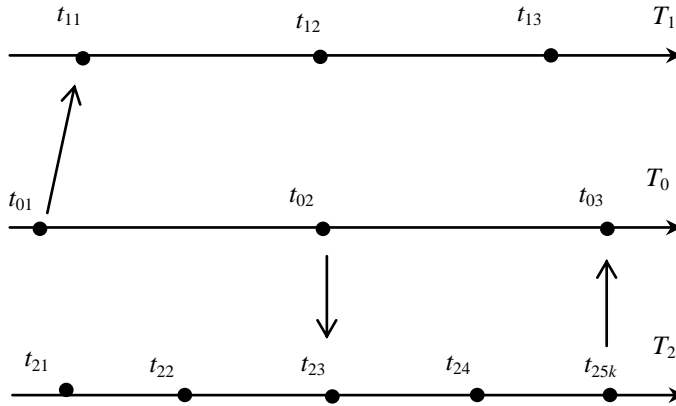


Рис. 1. Пример представления темпоральных знаний в виде трех параллельных потоков времени

Будем считать поток  $T_0$  главным, представляющим знания об основном ходе течения процесса, а  $T_1, T_2$  – потоками, содержащими альтернативные варианты развития событий. Тогда при реализации альтернативных версий осуществляется переход на другие потоки времени. На рис. 1 альтернативной базовой последовательности  $t_{01} \rightarrow t_{02} \rightarrow \dots$  является последовательность  $t_{01} \rightarrow t_{11} \rightarrow \dots$ , связанная с переходом к потоку  $T_1$ .

Альтернативной последовательности  $t_{02} \rightarrow t_{03}$  является цепочка событий  $t_{02} \rightarrow t_{23} \rightarrow t_{24} \rightarrow t_{25} \rightarrow t_{03}$ , связанная с переходом в иной поток времени и возвратом в базовый поток  $L_2$ .

Из сравнения последовательностей  $t_{02} \rightarrow t_{03}$  и  $t_{02} \rightarrow t_{23} \rightarrow t_{24} \rightarrow t_{25} \rightarrow t_{03}$  видно, что масштабы времени в различных потоках отличаются. Действительно, метки времени для событий не используются, поскольку в таком случае получим описание конкретных свершившихся событий, а не модель. Следовательно, необходимо использовать относительный масштаб времени, который можно выражать через количество произошедших событий. Тогда для любого потока времени можно найти минимальный интервал времени  $(t', t'')$ , который задается через пару событий, обладающих свойством (1).

$$\exists(t', t'') : t' < t'' \wedge \forall t((t < t'' \rightarrow t < t') \wedge (t > t' \rightarrow t > t'')). \quad (1)$$

Согласно (1) между  $t''$  и  $t'$  не существует промежуточного события.

При использовании интервала (1) для оценки масштаба времени абсолютное значение разности времени  $t''$  и  $t'$  не учитывается, поскольку пара  $(t', t'')$  в данном

случае указывает просто на наличие связи между событиями в моменты  $t''$  и  $t'$ . Иными словами, темпоральный аспект позволяет лишь найти связи между объектами предметной области на основе связей между изменяющими эти объекты событиями.

Для учета же чисто временной составляющей процесса в качестве общей метрики времени для каждого из потоков целесообразно использовать продолжительность минимального для всех потоков временного интервала между событиями, что позволит в дальнейшем соотносить масштабы времени в каждом из потоков.

$$\Delta t = \min_{i,j} \Delta t_{ij} = (t_{ij}'' - t_{ij}') \quad (2)$$

$$\forall t_{ij}'', t_{ij}', t_{ij}' \in T_j \quad t_{ij}' < t_{ij}'' \wedge (t_{ij}' < t_{ij}'' \rightarrow t < t_{ij}') \wedge (t_{ij}' > t_{ij}'' \rightarrow t_{ij}' > t_{ij}'')$$

Таким образом, появляется возможность сочетать интервальное и дискретное представление времени, поставив в соответствие интервал времени в одном временном потоке дискретному событию в другом. Различные временные потоки можно связать с разными уровнями иерархии в предметной области. Тогда многоуровневое представление времени позволит построить многоуровневую модель процесса. В такой модели мы можем представлять темпоральные знания с различной степенью детализации, скрывая подробности взаимосвязи между событиями на текущем потоке времени. Указанные связи раскрываются через взаимосвязи между событиями на другом временном потоке.

Пример взаимодействия между потоками времени приведен на рис. 2.

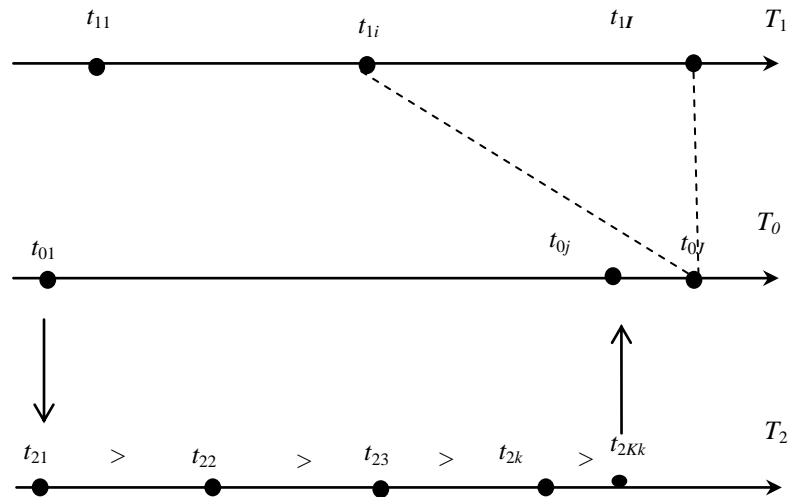


Рис. 2. Пример взаимосвязей между потоками времени

В данном примере в потоке  $T_1$  используется интервальное представление времени, а в потоках  $T_0$  и  $T_2$  – дискретное. Событие  $t_{0j}$  отражает в потоке  $T_0$  интервал  $(t_{1j}, t_{1l})$  продолжительностью  $\Delta t_1 = t_{1l} - t_{1j}$ . Следовательно, при дискретном представлении времени в потоке  $T_0$  и интервальном в потоке  $T_1$  событие в первом потоке произойдет мгновенно, а во втором – будет выполняться в течение времени  $\Delta t_1$ .

Аналогично связь между событиями  $t_{01}$  и  $t_{0j}$  в потоке  $T_0$  раскрывается через последовательность событий  $t_{21} \rightarrow t_{22} \rightarrow \dots \rightarrow t_{2K}$ .

Представленные метрики и примеры многоуровневого представления времени показывают, что для многопоточного представления темпоральных знаний необходимо использовать свойства направленности и линейности, а также гранулярности [10] времени. Последнее заменяет свойство непрерывности, задавая пределы дробления времени парой событий  $(t', t'')$  согласно (1) или (2).

Представленные свойства времени позволяют определить модель многоуровневого представления времени в задачах processmining, которая дает возможность единым образом представить и исходные логи, и модель процесса.

При представлении логов (фиксированных последовательностей событий) рассматриваются только конечные линейные последовательности, для которых задано отношение следования. Изначально данное отношение задается для каждого следа (трассы) реализации процесса. Лог состоит из набора трасс.

Тогда модель  $i$  следа логa представляет собой модель Крипке [11–12], для которой задано только одно отношение между элементами:

$$M_i = (U_i, I_i(R), t_0), R = \{<\}, \quad (3)$$

где  $U$  – непустое множество событий, произошедших в конкретные моменты времени;  
 $I$  – каждому отношению  $R$  ставит в соответствие подмножество  $I(R) \subseteq U \times U$ ,  $t_0$  – начальный момент времени, в который произошло первое событие следа процесса.  
 Отношение перехода  $I(<)$  – представлено в виде выражения (4).

$$I(<) = I(R) | (t' < t'') \Rightarrow \neg \exists t : t' < t \vee t < t'', t, t', t'' \in U. \quad (4)$$

Модель многоуровневого представления темпоральных знаний объединяет модели последовательностей событий  $M_i$ , для каждой из которых заданы отдельные временные соотношения  $I_i(R)$ , а также включает отношения упорядоченности  $I_{ij}(R)$  между событиями из различных моделей  $M_i$ .

Между событиями из различных моделей могут быть отношения перехода в случае дискретного представления времени в связываемых временных потоках, а также отношение  $d_i$  (contains) [3] в том случае, если на одном из потоков используется дискретное представление времени, а на другом – интервальное. Данное отношение позволяет представить интервал времени на более низком уровне иерархии в виде отдельного события на более высоком уровне.

$$M = (U, I(R), t_0), \quad (5)$$

$$U = \bigcup_i U_i, I(R) = (\bigwedge_i I_i(R) | R = \{<\}) \wedge (\bigwedge_{ij} I_{ij}(R)) | R = \{<, d_i\}.$$

---

**Выводы из данного исследования и перспективы дальнейших разведок в данном направлении.** В статье предложена модель многоуровневого представления темпоральных знаний, которая объединяет упорядоченные наборы событий, представляющие различные уровни реализации последовательности действий процесса в параллельно протекающих потоках времени. Предложенная модель позволяет отразить в отдельных временных потоках изменение объектов моделируемой предметной области на различных уровнях иерархии, формируя различные уровни детализации темпоральных и пространственных знаний.

В задачах интеллектуального анализа процессов модель позволяет единым образом представить и лог процесса, и его модель, обеспечивая различную степень детализации модели с учетом взаимосвязей объектов предметной области.

#### **Список использованных источников:**

1. Brachman R. J. Knowledge Representation and Reasoning / R. J. Brachman, H. J. Levesque // Amsterdam : Elsevier, 2004. – 381p.
2. Towards a general theory of action and time / J. Allen // Artif Intell, 1984. – № 23. – P. 123–154.
3. Allen J. F. A common-sense theory of time / J. F. Allen, P. J. Hayes // JoshiA(ed) Proceedings of the ninth international joint conference on artificial intelligence (IJCAI'85). International joint conferences on artificial intelligence, inc. Morgan Kaufmann San Francisco. – 1985. – P. 528–531.
4. W.M.P. van der Aalst. Process Mining: Discovery, Conformance and Enhancement of Business Processes // Springer-Verlag. – Berlin, 2011. – 352 p.
5. Кандрашина Е. Ю. Представление знаний о времени и пространстве в интеллектуальных системах / Е. Ю. Кандрашина, Л. В. Литвинцева, Д. А. Поспелов ; под ред. Д. А. Поспелова. – М. : Наука, 1989. – 328 с.
6. Pnueli A. The temporal logic of programs / A. Pnueli // Proceedings of the 18th Annual Symposium on Foundations of Computer Science (FOCS). – 1977. – P. 46–57.
7. Emerson E. A. Temporal and modal logic / E. A. Emerson // by ed. : J. van Leeuwen editor : Handbook of Theoretical Computer Science, Elsevier. – 1990. – P. 995–1072.
8. Emerson E. A. Decision procedures and expressiveness in the temporal logic of branching time / E. A. Emerson, J. Y. Halpern // Journal of Computer and System Sciences. – 1985. – № 30 (1). – P. 1–24.
9. Clarke E. M. Model checking / E. M. Clarke, O. Grumberg, D. A. Peled // MIT Press Cambridge, MA, USA. – 1999.
10. Euzenat J. Time granularity / J. Euzenat, A. Montanari // Handbook of temporal reasoning in artificial intelligence. – Elsevier, Foundations of artificial intelligence, 2005. – P. 59–118.
11. Kripke S. A. Semantical Considerations on Modal Logic / S. A. Kripke // Acta Philosophica Fennica. – № 16. – 1963. – P. 83–94.
12. Kripke S. A. Outline of a Theory of Truth / S. A. Kripke // Journal of Philosophy. Sets his theory of truth (against Alfred Tarski), where an object language can contain its own truth predicate. – 1975. – P. 690–716.