

МОДЕЛИРОВАНИЕ ПРИЧИННО-СЛЕДСТВЕННЫХ ОТНОШЕНИЙ В ТЕКСТЕ С ПОМОЩЬЮ СЕТЕЙ ПЕТРИ

Желтов П.В.¹

¹ФГБОУ ВПО «Чувашский государственный университет им. И.Н. Ульянова», Чебоксары, Россия (428015, г. Чебоксары, Московский проспект, 15), e-mail: chnk@mail.ru

В данной работе предпринята попытка смоделировать с помощью сетей синтаксические структуры, выражающие причинно-следственные отношения в тексте. Причинно-следственные отношения вводятся в текст следующими двумя основными способами: явно – с помощью служебных слов; неявно – в контексте соотносительности предложений текста. В первом случае речь идет о двучленных предложениях, состоящих из двух частей, связанных с помощью семантических союзов (предлогов) – потому что, так как, вследствие (того что), из-за (того что), в связи с (тем что), после (того как), по мере (того как), благодаря (тому что), ввиду (того что), оттого что, в связи с (тем что), по причине (того что) и др. (причинные), так что (следствия); предлогов – от, ради, благодаря, изъяснительных союзов – поэтому. Во втором случае речь идет о группе предложений (синтаксически независимых, разделенных знаками пунктуации), связанных тематически – содержащих ситуационно связанные между собой слова.

Ключевые слова: моделирование, причинно-следственные отношения, текст, сеть Петри.

SIMULATION OF A CAUSAL RELATIONSHIP TO TEXT USING PETRI NETS

Zhelto P.V.²

²I.N. Ulyanov's Chuvash State University Cheboksary, Russia (428015, Cheboksary, Moscow prospect, 15), e-mail: chnk@mail.ru

In this paper an attempt to simulate the network using syntactic structures, expressing causal relationships in the text. Causal relationships are entered in the text of the following two basic ways: 1. explicitly - with the help of auxiliary words; 2. implicitly - in the context of relatedness of the proposed text. In the first case it is a dimeric sentence composed of two parts linked by semantic associations (preposition) - because, since, due to (that of) due to (the fact that), in connection with (so that) after (as), as (As) due to (the fact that) due to (the fact that), because that, in connection with (so that), because (of which) et al. (causal) so that (the investigation); prepositions - by, for, thanks izjasnitelnyh unions - so. In the second case it is about a group of sentences (syntax independent, separated by punctuation marks), related thematically - situationally containing unrelated words.

Keywords: modeling, causal relationships, text, Petri net.

Моделирование лингвистических структур с помощью математических методов и средств имеет давнюю историю [4]. В области моделирования, синтаксиса, как правило, используется либо анализ по непосредственным составляющим (деревья) и формальные грамматики (линейные правила вывода с ветвлениями/вариантами, которые также могут быть сведены к деревьям), либо сети и сетевые грамматики [2].

В общем случае сеть является более удобным средством моделирования, т. к., помимо ветвлений, может содержать циклы, а динамические сети позволяют ещё и оперировать временем, т.е. моделировать синтаксические структуры с категориями времени, выражающие не только последовательные действия (которые можно моделировать и с помощью деревьев, правил вывода и обычных статических сетей), но и параллельные (таксисные) [1].

В то же время сети, вследствие наличия циклов, требуют более сложных структур для компьютерного представления, что усложняет временные характеристики в поиска в таких структурах.

В данной работе предпринята попытка смоделировать с помощью сетей синтаксические структуры, выражающие причинно-следственные отношения в тексте.

Причинно-следственные отношения вводятся в тексте следующими двумя основными способами:

- 1) явно – с помощью служебных слов;
- 2) неявно – в контексте соотнесенности предложений текста.

В первом случае речь идет о двучленных предложениях, состоящих из двух частей, связанных с помощью семантических союзов (предлогов) – *потому что, так как, вследствие (того что), из-за (того что), в связи с (тем что), после (того как), по мере (того как), благодаря (тому что), ввиду (того что), оттого что, в связи с (тем что), по причине (того что)* и др. (причинные), *так что* (следствия); предлогов – *от, ради, благодаря*, изъяснительных союзов – *поэтому* [7].

Причем, как правило, причинно-следственные предложения этого типа бывают двух подтипов:

1) прямые – <причина> ',' <следствие>, например: «Потому что в лесу уже стемнело, мы решили зажечь фонарь», структура которых выглядит так: <служебное слово> <причинное предл.>',' <следств. предл.> или, например: «Жар такой, что плавится песок на пляже» – <причинное предл.> <служебное слово> <следств. предл.>.

2) инвертированные – <следствие>&<причина>, например: «Машины засветили фары, потому что в лесу уже стемнело», «Он играет от безделья», структура которых выглядит как <следств. предл.>.

Во втором случае речь идет о группе предложений (синтаксически независимых, разделенных знаками пунктуации), связанных тематически – содержащих ситуационно связанные между собой слова, например: «Шел дождь (1). Было холодно (2). Люди кутались в шарфы (3) и раскрывали зонтики (4)» - здесь связаны между собой синтагмы 1, 4: «Шел дождь» → «раскрывали зонтики» и 2, 5 «Было холодно» → «Люди кутались в шарфы».

Если предложения первого типа выделяются из текста, как правило, на этапе синтаксического анализа и для их выделения достаточно лишь иметь классификацию их типов в виде регулярной грамматики, то предложения второго типа требуют семантического анализа текста с помощью тезауруса на предмет соотносимости предложений или их частей по ключевым словам [3].

В обоих случаях конечному пользователю необходимо предоставить результаты анализа в удобном для зрительного восприятия виде.

Очевидно, что в общем случае схема причинно-следственных связей в тексте представляет собой либо древовидную структуру, либо сеть. Последняя имеет место, когда речь идет о циклических причинно-следственных процессах, что часто наблюдается в научных текстах, а также в художественной и специальной литературе при описании правил различных игр, технологических процессов и т.п.

В качестве формализма для достаточно полного интерфейса, способного описывать результаты анализа динамически, возможно использовать сети Петри.

Сеть Петри – это набор из пяти элементов, $\langle P, T, I, O, M_0 \rangle$, где P – множество позиций, T – множество переходов, I – множество входных функций, O – множество выходных функций, M_0 – начальная маркировка сети [5; 6].

Графически сеть Петри выражается в виде двудольного ориентированного графа, в котором существуют вершины двух типов – позиции (P) и переходы (T), связанные между собой дугами со стрелками, которые могут связывать только вершины разных типов, а также при помощи маркеров – фишек, которые могут переходить из одной позиции в другую через переход. Начальное расположение маркеров называется начальной маркировкой (M_0).

При этом для моделирования с помощью сети Петри какой-либо предметной области удобно соотносить позиции с состояниями системы или объекта, а переходы с действиями. При этом предшествующие переходу позиции, в зависимости от наличия в них маркеров, моделируют условия для действия, которое моделируется с этим переходом, т.е. в нашем случае причину, а последующие моделируют пост-состояния (состояния после действия), т.е. в нашем случае следствие.

Так, предложение «Машины засветили фары, потому что в лесу уже стемнело», можно представить следующей сетью Петри (рис. 1).

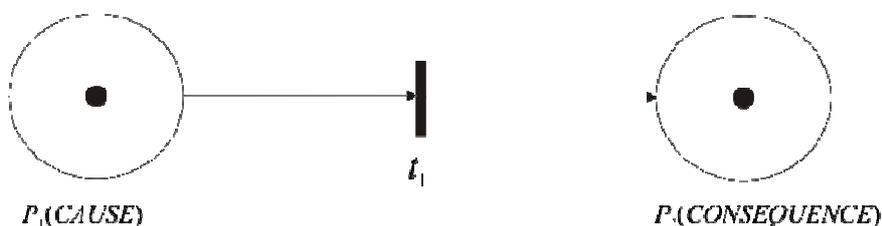


Рис. 1 :

$P_1(CAUSE)$ – причинная позиция;

$P_2(CONSEQUENCE)$ – результирующая позиция, содержащая следствие.

В $P_1(CAUSE)$ находится синтагма «В лесу уже стемнело», а в $P_2(CONSEQUENCE)$ синтагма «Машины засветили фары», а переход t_1 со стрелками

моделирует их логическую связь, реализованную синтаксически с помощью союза *потому что*.

Для того чтобы подчеркнуть инверсный тип причинно-следственного отношения, выраженного в данном предложении, возможно инвертировать графически сеть на рис. 1, получив при этом следующую сеть Петри (рис. 2).

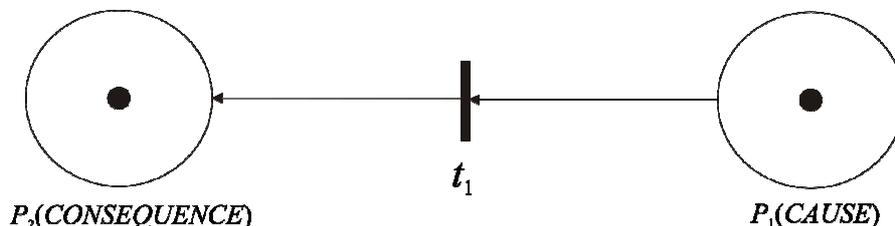


Рис. 2.

Эта сеть совершенно идентична первой, отличается только графическое представление. При компьютерной реализации этих двух сетей будет отличаться только последовательность элементов при выводе этих сетей на экран или форму.

В связи с динамичностью сети Петри, т.е. с наличием в сети времени, возможна группировка синтагм по времени выраженных в них действий.

Например, приведенный выше текст «Шел дождь (1). Было холодно (2). Люди кутались в шарфы (3) и раскрывали зонтики (4)» можно представить следующей сетью Петри:

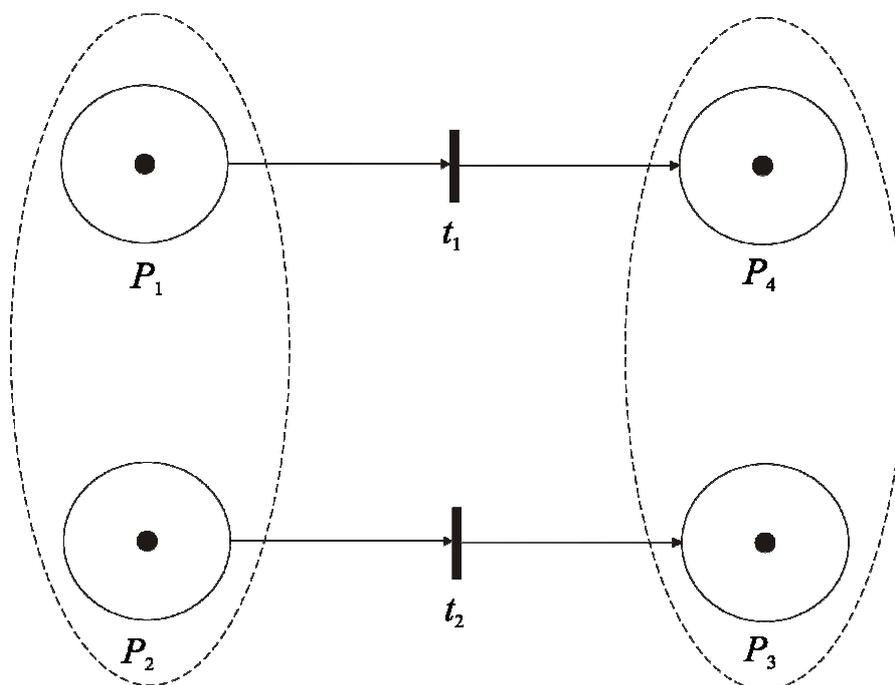


Рис. 3.

где номера позиций соответствуют номерам выражаемых ими синтагм (рис. 3). Прерывистой линией (овалом) объединены действия параллельные по времени, выражаемые синтагмами 1 и 2, 3 и 4.

Для представления причинно-следственных отношений в тексте был разработан интерфейс в среде MS Visual Studio (на языке программирования C++). Скриншоты форм с результатами анализа примеров приводятся ниже (рис. 4).

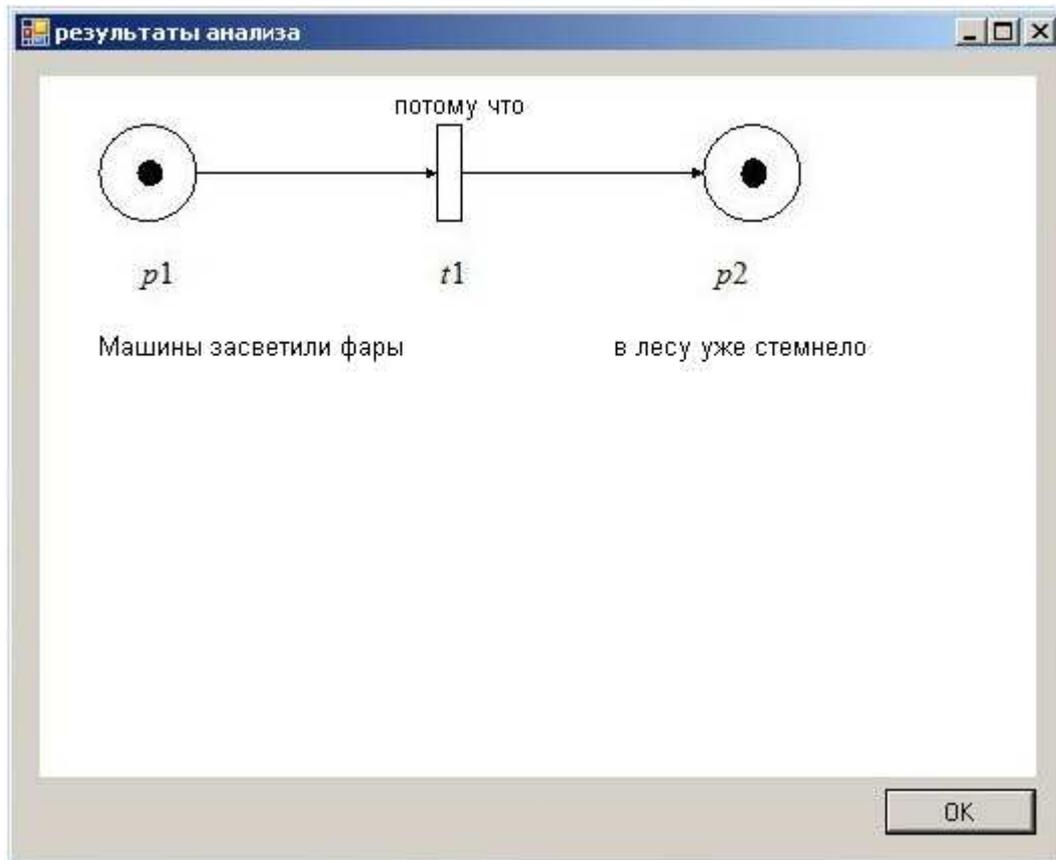


Рис. 4.

С точки зрения программной реализации отображение сети Петри в динамике реализовано как формы перерисовки холста формы (объект Canvas), т.е. при изменении состояния сети изображение на форме обновляется.

В дальнейшем целесообразно создание специальных компонент, т.к. это убыстряет процесс отображения изменяющейся сети на форму.

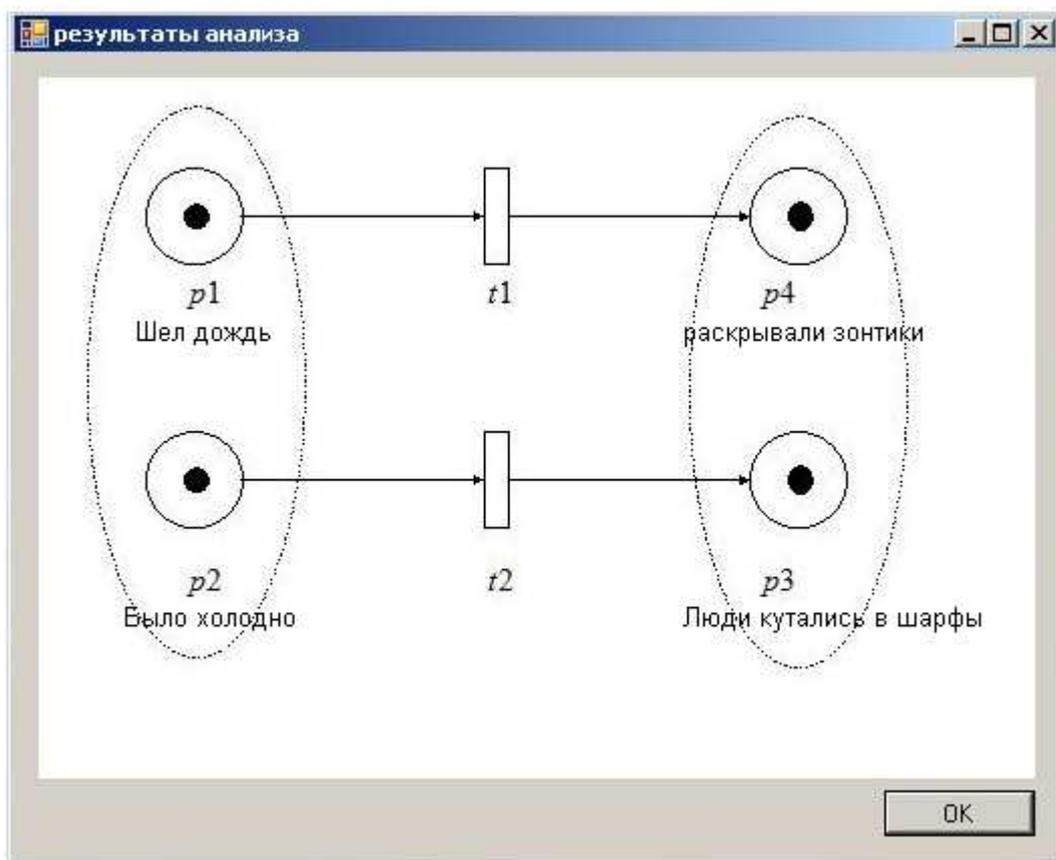


Рис. 5.

Показано применение простых сетей Петри для моделирования синтаксических структур, выражающих причинно-следственные отношения. Смоделированы два типа причинно-следственных отношений: прямые и инвентированные.

Создана программа с пользовательским интерфейсом, выводящая результаты анализа подобных синтаксических структур.

Публикация подготовлена в рамках поддержанного РГНФ научного проекта № 15-04-00532.

Список литературы

1. Бондарко А.В. Теория функциональной грамматики. Введение. Аспектуальность. Временная локализация. Таксис. – Л. : Наука, 1987. – 204 с.
2. Вудс В.А. Сетевые грамматики для анализа естественного языка // II Кибернетический сборник. – М., 1976. – Вып. 13. – С. 121-158.
3. Желтов П.В. Текстологический анализ на основе семантико-лексических соответствий // Компьютерные технологии и моделирование : сб. науч. тр. – Казань : Изд-во КГТУ, 2008. – Вып. 1. – С. 65-70.

4. Желтов П.В. Формальные методы и модели в сравнительно-составительном языкознании : монография. – Чебоксары : Изд-во Чувашского университета, 2006. – 252 с.
5. Котов В.Е. Сети Петри. – М. : Наука, 1984. – 180 с.
6. Питесон Дж. Теория сетей Петри и моделирование систем сетями Петри / пер. с англ. – М. : Мир, 1984. – 246 с.
7. Русская грамматика. Т. I. Фонетика. Фонология. Ударение. Интонация. Словообразование. Морфология. Т. II. Синтаксис. – М. : Наука, 1960. – 709 с.

Рецензенты:

Славутский Л.А., д.ф.-м.н., профессор кафедры автоматки и управления в технических системах ФГБОУ ВПО «ЧГУ им. И.Н. Ульянова», г. Чебоксары;

Охоткин Г.П., д.т.н., профессор, заведующий кафедрой автоматки и управления в технических системах ФГБОУ ВПО «Чувашский государственный университет имени И.Н. Ульянова», г. Чебоксары.