

О синтаксическом анализе юридических текстов

Е. М. Перпер

В работе рассматривается задача создания компьютерной программы, осуществляющей синтаксический анализ юридических текстов на русском языке. На вход эта программа должна принимать текст одного предложения, а на выход должна выдавать построенное по этому предложению синтаксическое дерево, в вершинах которого находятся слова предложения, а рёбрам приписаны синтаксические отношения между словами предложения. В работе приведены правила, позволяющие проводить синтаксический анализ при наличии морфологической информации о словах предложения.

Ключевые слова: синтаксический анализ, синтаксическое дерево, морфологические характеристики, правила.

Введение

Синтаксический анализ является важным этапом обработки естественного языка. Автоматизация синтаксического анализа необходима для создания программ, работающих с естественным языком — вопросно-ответных систем, систем автоматического аннотирования текстов и т.д.

Написание осуществляющей синтаксический анализ программы, о которой идёт речь в данной статье, является частью проекта по автоматической обработке юридических документов, касающихся заполнения форм бухгалтерской отчётности [1].

Целью проекта является создание системы, которая самостоятельно заполняла бы формы отчётности в соответствии с поданными на вход системе текстами законов, касающихся заполнения этих форм. От пользователя такая система должна требовать лишь введения данных, необходимых для вычисления значений, которыми заполняется форма отчётности. Правила вычисления этих значений система должна определять самостоятельно, пользуясь лишь текстами юридических документов.

Данная статья посвящена исключительно синтаксическому анализу, а потому предполагается, что этапы, предшествующие ему — лексический, морфологический анализ — уже проведены, и программа получает на вход результаты морфологического анализа.

Результатом работы программы является синтаксическое дерево — таким образом, каждое слово должно присутствовать в каком-либо синтаксическом отношении в качестве зависимого. Программа проходит по всем словам предложения, и для каждого слова определяет, в каких синтаксических отношениях оно состоит. Это определение происходит с помощью специального набора правил.

Стоит отметить, что благодаря относительной бедности языка, используемого в юридических текстах, синтаксический анализ таких текстов проводить проще, чем семантический анализ произвольных текстов.

Автор выражает благодарность профессору Эльяру Эльдаровичу Гасанову за научное руководство и помощь в научной работе.

Основные понятия и формулировка результатов.

Синтаксическое отношение — это отношение между парой слов предложения, причём одно из слов является главным в этом отношении, а другое — зависимым. Каждое синтаксическое отношение в зависимости от частей речи участвующих в отношении слов, их морфологических характеристик и т.д.,

может быть отнесено к определённом классу синтаксических отношений.

На вход программе, осуществляющей синтаксический анализ, поступает последовательность *лексем*. Лексема, в свою очередь, представляет собой последовательность символов. Это может быть слово, число, знак препинания. Текст предложения разбивается на последовательность лексем в процессе *лексического анализа*.

Помимо последовательности лексем, на вход синтаксическому анализу поступает последовательность *токенов*. Токен создаётся для каждого слова предложения в процессе морфологического анализа и представляет собой тройку, в которую входят: само слово; *лемма* — каноническая форма слова (например, для существительного это будет то же слово, но в именительном падеже и единственном числе); набор *морфологических характеристик* (для существительного это род, падеж, число и т.д., для глагола это вид, время и т.д.).

Выходом синтаксического анализа является *синтаксическое дерево*, также называемое *деревом зависимостей*. Это ориентированное дерево. Каждой его вершине сопоставлен токен. Дуга в синтаксическом дереве ведёт из вершины А в вершину В тогда и только тогда, когда сопоставленные этим вершинам слова связаны в русском языке синтаксическим отношением, причём главным в этом отношении является слово, соответствующее вершине А. Дуге при этом сопоставлено название отношения.

В данной работе для синтаксического анализа предлагается использовать (возможно, в несколько упрощённом виде) подход, применяемый А.С. Подколзиным для автоматического решения различных математических задач [2]. В применении к синтаксическому анализу этот подход состоит в следующем. В программе имеется список *правил*, которые позволяют находить синтаксические связи между словами. Для каждого токена и каждого правила проверяется, применимо ли это правило к данному токену; если да, то создаётся продиктованное этим правилом синтаксическое отношение.

Каждое правило состоит из трёх частей. Первая часть проверяет, подходит ли слово для этого правила: обладает ли оно нужным набором морфологических характеристик. В большинстве случаев значение леммы не проверяется, однако есть и правила, которые работают с конкретными леммами. В тех случаях, когда целью правила является построения синтаксического отношения, в котором рассматриваемое слово было бы зависимым, проверяется также, что слово ещё не является зависимым ни в каком построенном синтаксическом отношении. Объясняется эта проверка просто: каждое слово может входить в какое угодно число синтаксических отношений в качестве главного, но лишь в одно — в качестве зависимого.

Вторая часть заключается в поиске слова, которое может входить в синтаксическое отношение с рассматриваемым словом. В некоторых правилах ищется не одно слово, а несколько, притом таких, что каждое из них могло бы образовывать синтаксические отношения либо с другим найденным словом, либо с рассматриваемым словом.

Наконец, третья часть строит синтаксические отношения между рассматриваемым словом и найденными словами. В том случае, если слово прошло проверку в первой части правила, и для него были найдены подходящие слова во второй части правила, будем говорить, что правило *применимо* к слову. Таким образом, если правило применимо к слову, то в результате применения правила к этому слову строится одно или несколько новых синтаксических отношений.

Надо заметить, что описанный в работе алгоритм включает в себя элементы семантического анализа — это видно по названию некоторых отношений, по использованию списков списков существительных, формируемых исходя из смыслового значения этих существительных, и по нескольким другим признакам. Дело в том, что с учётом конечной цели проекта, частью которого является описываемая программа, определённую работу по семантическому анализу удобно выполнять одновременно с синтаксическим анализом.

Для большей части создаваемых алгоритмом синтаксических отношений обозначения классов, к которым относятся эти отношения, взяты из [3], но для некоторых отношений используются специальные названия, принятые только в данной работе: это позволяет облегчить работу с синтаксически разобранным предложением на следующих этапах.

Правила, с помощью которых программа создаёт синтаксические отношения, подробно описаны в следующем разделе.

Правила построения синтаксических отношений

Прежде чем перейти к описанию правил, заметим, что на слова, которые могли бы образовывать синтаксические отношения друг с другом, накладываются определённые ограничения. Если некоторая часть предложения заключена в круглые скобки, то никакое слово из этой части не может быть главным в синтаксическом отношении, в котором зависимое слово находится вне этих скобок. Применению основного алгоритма построения синтаксического дерева предшествует определение глубины вложенности каждого слова предложения в круглые скобки (в дальнейшем будем коротко именовать её глубиной). Эта глубина вычисляется как разность числа открывающих и числа закрывающих круглых скобок перед словом.

Подробнее остановимся на том, как осуществляется вторая часть правила. Происходит последовательный перебор всех слов предложения, осуществляемый либо в сторону начала предложения, либо в сторону его конца. Перебор начинается со слова, следующего после рассматриваемого слова в том направлении, в котором осуществляется перебор. Если рассматриваемое слово должно быть главным в синтаксическом отношении и имеет большую глубину, чем слово, до которого дошёл перебор, то перебор завершается неудачей. Если рассматриваемое слово должно быть зависимым в синтаксическом отношении и имеет меньшую глубину, чем слово, до которого дошёл перебор, то перебор также завершается неудачей. Если перебор дошёл до начала либо конца предложения, а в некоторых случаях — до начала либо

конца *клаузы* (простого предложения в составе сложного), в которой содержалось рассматриваемое слово, и нужное слово не найдено, перебор снова заканчивается неудачей. Если же нужное слово найдено, перебор успешно завершается, и найденное слово используется в следующей части правила для построения синтаксического отношения. В тех случаях, когда ищется несколько слов, может быть осуществлено несколько переборов.

Будем называть слово *допустимым*, если его глубина не является препятствием для включения этого слова в синтаксическое отношение.

Перед основным алгоритмом построения синтаксических отношений, совершается несколько подготовительных действий.

- 1) Для каждого тире в предолжении создаётся токен, который в качестве леммы содержит само тире, а его набор морфологических характеристик состоит из одного элемента — указания на то, что тире следует рассматривать как глагол. Действительно, в большинстве случаев тире заменяет глагол «есть» в какой-либо форме.
- 2) Предложение разбивается на сегменты по некоторым знакам препинания — запятым, точкам с запятой, двоеточиям, притом рассматриваются только те из них, что имеют нулевую глубину. Во многих случаях сегмент совпадает с клаузой, но не всегда, так как запятые не только разделяют простые предложения в составе сложного, но и выполняют другие роли (например, отделяют друг от друга однородные члены предложения).

В основном алгоритме синтаксического анализа, помимо попыток применения к слову правил построения синтаксических отношений, для этого слова производятся некоторые вычисления, которые могут быть в дальнейшем использованы при рассмотрении не только этого слова, но и последующих слов. Например, таким образом находится и запоминается последний встретившийся в рассмотренной части предложения глагол, а также последний встретившийся глагол перед последним встретившимся двоеточием.

Приступим, наконец, к описанию правил, из которых состоит основной алгоритм синтаксического анализа. Заметим, что некоторые части речи в процессе применения правил приравниваются к другим частям речи. В большинстве случаев это явно оговаривается. Исключение — местоимения: местоимения-существительные во всех правилах приравниваются к существительным, а местоимения-прилагательные — к прилагательным, если находятся непосредственно перед прилагательным или существительным, и к существительным в противном случае. В правилах это явно не оговаривается, но, например, когда речь в правиле идёт о существительном, предполагается, что это может быть также и местоимение-существительное, и местоимение-прилагательное, если оно приравнено к существительному.

- 1) Правило, создающее для тире синтаксическое отношение «ВАРИАНТ» между этим тире и последним глаголом, находящимся перед последним двоеточием перед тире. Если такой глагол существует, то он был найден ранее. В создаваемом отношении этот глагол является главным элементом, а тире — зависимым.
- 2) Правило, которое для тире проверяет, является ли союз первым элементом сегмента, в котором это тире находится. Если является, то происходит перебор сегментов от предыдущего к первому в предложении, пока первым элементом сегмента является тот же союз и не достигнуто начало предложения. Если у сегмента, на котором перебор прекратился, вторым словом является тот же союз, а первым — глагол или причастие, то между этим глаголом (причастием) и тире создаётся синтаксическое отношение «ВАРИАНТ», в котором глагол (причастие) — главный элемент, а тире — зависимый.
- 3) Правило, проверяющее, не образует ли рассматриваемое слово вместе с несколькими идущими за ним предлог, состоящий из нескольких слов (к таким предлогам относятся, например, «исходя из», «в течение», «в отношении» и

др; в программе имеется список таких предлогов). Если образует, то каждая пара идущих подряд слов из этого предлога соединяется отношением «НЕДЕЛИМ», где главным объявляется то из двух слов, которое находится ближе к началу предложения. Кроме того, после применения этого правила все слова, входящие в предлог, считаются предлогами.

- 4) Правило, которое для слова «не» ищет ближайшее находящееся после него допустимое слово, не являющееся предлогом или союзом, и создаёт синтаксическое отношение «ОТР» («отрицание»), в котором найденное слово является главным элементом, а «не» — зависимым.
- 5) Правило, которое для существительного ищет ближайшее находящееся после него допустимое слово, не являющееся прилагательным в родительном падеже. Если такое слово найдено, и это — существительное в родительном падеже, то создаётся синтаксическое отношение «ГЕНИТ-ИГ» («генитивная именная группа») с рассматриваемым словом в качестве главного элемента и найденным существительным в родительном падеже в качестве зависимого элемента.
- 6) Правило, ищущее для существительного, не находящегося в именительном падеже, ближайшее к нему среди находящихся перед ним слов сегмента допустимое слово, являющееся глаголом, существительным, причастием (не согласованным с существительным), кратким причастием, предлогом либо союзом (но не союзом «или», «и», «а»). Если такого слова нет и сегмент отделён от предыдущего точкой с запятой, то в качестве искомого слова рассматривается последний глагол в первом сегменте предложения, а при наличии предлога непосредственно после этого глагола — этот предлог. Создаётся синтаксическое отношение «ДОП» («дополнение»), в котором главным является найденное слово, а зависимым — рассматриваемое существительное.

- 7) Правило, ищущее для не находящегося в именительном падеже существительного, чей сегмент отделён от предыдущего запятой, первое в предыдущем сегменте (считая от его начала) существительное в том же падеже. Если такое существительное найдено, и оно является зависимым в каком-либо синтаксическом отношении, то добавляется такое же отношение, в котором зависимым словом является рассматриваемое правилом существительное.
- 8) Правило, ищущее для предлога «на» ближайшее к нему среди находящихся перед ним слов сегмента допустимое слово из списка существительных, обозначающих величину или значение («стоимость», «цена», «размер» и т.д.). Если между этим существительным и рассматриваемым предлогом «на» не расположены глагол или тире, создаётся синтаксическое отношение «ОПР» («определятельное») с найденным существительным в качестве главного слова и предлогом «на» в качестве зависимого слова.
- 9) Правило, которое ищет для предлога «о» ближайшее к нему среди находящихся перед ним слов сегмента допустимое слово из списка существительных, обозначающих информацию («законодательство», «положение» «приказ» и т.д.). Если между этим существительным и рассматриваемым предлогом «о» не расположены глагол или тире, создаётся синтаксическое отношение «ОПР», где главным словом является найденное существительное, а зависимым — предлог «о».
- 10) Правило, проверяющее для существительного в именительном падеже, есть ли глагол в сегменте, где это существительное находится. Пусть глагола там нет, и непосредственно перед последним двоеточием, встретившимся до этого сегмента, находится то же самое существительное (возможно, в другом падеже). Тогда создаётся отношение «А_ИМЕННО», где главным словом является найденное перед двоеточием существительное, а зависимым — рассматриваемое слово.

- 11) Правило, которое для предлога, а также для слов «исходя» (если следующее за ним слово – «из») и «как», ищет допустимый глагол, чтобы связать их синтаксическим отношением «ГЛ_ДОП» («глагол – дополнение»), в котором глагол будет главным словом, а рассматриваемое правилом слово – зависимым. Глагол ищется в том же сегменте, где находится рассматриваемое слово, сначала в направлении от этого слова к концу сегмента, а в случае неудачи – в направлении от этого слова к началу сегмента. В данном правиле к глаголу приравниваются отглагольные существительные и краткие причастия. Если нужного слова найти не удалось, а рассматриваемое правилом слово – «по» или «при», ищется в направлении от него к концу предложения ближайший допустимый глагол, не находящийся в клаузе, начинающейся со слова «который». В случае, когда нужное слово всё ещё не найдено, в качестве него берётся последний глагол, находящийся перед последним (до рассматриваемого правилом слова) двоеточием, если такой существует.
- 12) Правило, ищущее для существительного в именительном падеже допустимый глагол, чтобы связать их синтаксическим отношением «ПОДЛ» («подлежащее»), в котором глагол будет главным словом, а существительное – зависимым. Глагол ищется так же, как и в предыдущем правиле, за тем исключением, что отглагольное существительное в данном случае к глаголу не приравнивается.
- 13) Правило, которое ищет для наречия допустимый глагол, чтобы связать их синтаксическим отношением «ГЛ_ОБСТ» («глагол – обстоятельство»), где глагол будет главным словом, а наречие – зависимым. Глагол ищется так же, как и в правиле 11.
- 14) Правило, ищущее для слова «если» допустимый глагол, чтобы связать их синтаксическим отношением «ЕСЛИ», в котором глагол будет главным словом, а существительное – зависимым. Глагол ищется так же, как и в правиле 11.

- 15) Правило, которое для слова «если», являющегося зависимым в каком-либо синтаксическом отношении, ищет ближайший к нему допустимый глагол, находящийся в предложении до сегмента, в котором находится рассматриваемое слово. Если такое слово найдено, то создаётся отношение «УСЛ» («условие»), где найденный глагол будет главным словом, а зависимым будет слово, являющееся главным в отношении, где «если» является зависимым.
- 16) Правило, которое для полного причастия, находящегося непосредственно после запятой, ищет ближайшее к нему в сторону начала предложения существительное, имеющее тот же падеж и число, причём если это число — единственное, то и тот же род. Если такое существительное нашлось, создаётся синтаксическое отношение «ПРИЧ_СУЩ» («причастие – существительное»), в котором главным словом является существительное, а зависимым — причастие.
- 17) Правило, которое для прилагательного или полного причастия ищет ближайшее к нему в сторону конца предложения существительное, имеющее тот же падеж и число, причём если это число — единственное, то и тот же род. Если такое существительное нашлось, создаётся синтаксическое отношение «ПРИЛ_СУЩ» («прилагательное – существительное») (если рассматриваемое слово — прилагательное) или «ПРИЧ_СУЩ» (если рассматриваемое слово — причастие), где главным словом является существительное, а зависимым — рассматриваемое слово.
- 18) Правило, обрабатывающее союзы «и», «или», «а», не находящиеся в начале сегмента. Сначала ищется слово w_1 — допустимое слово, ближайшее к рассматриваемому союзу по направлению к началу предложения. Кроме того, ищется слово w_2 — допустимое слово, ближайшее к рассматриваемому союзу по направлению к концу предложения. Если w_2 — полное причастие, краткое причастие, глагол либо отглагольное существительное, то от рассматриваемого со-

юза в сторону начала предложения ищется слово w_3 — ближайшее допустимое слово, удовлетворяющее следующим условиям: если w_2 — полное причастие, то w_3 должно быть полным причастием либо прилагательным, имеющим тот же падеж, что w_2 ; если w_2 — глагол либо краткое причастие, то найденное слово также должно быть глаголом либо кратким причастием и иметь то же число, что и w_2 ; если w_2 — отглагольное существительное, то w_3 должно быть существительным в том же падеже.

Если w_2 не является полным причастием, кратким причастием, глаголом либо отглагольным существительным, ищется ближайшее к союзу по направлению к концу предложения допустимое слово w_4 , удовлетворяющее следующим условиям: если w_1 — прилагательное либо полное причастие, то найденное слово также должно быть прилагательным либо полным причастием и иметь тот же падеж, что и w_1 ; если w_1 — существительное, то найденное слово должно быть существительным и иметь тот же падеж, что и w_1 ; если w_1 — глагол либо краткое причастие, то найденное слово также должно быть глаголом либо кратким причастием и иметь то же число, что и w_1 . Заметим, что w_1 может совпадать с w_3 , а w_2 — с w_4 .

Если подходящее слово найдено, то создаётся два синтаксических отношения «ОДНОР_ИГ» («однородные именные группы»), которые содержит рассматриваемый союз в качестве главного слова и найденные слова w_1 и w_4 (либо w_2 и w_3 , если w_2 — полное причастие) в качестве зависимых. Если слово w_1 (w_3) до применения данного правила участвовало в каком-либо синтаксическом отношении в качестве зависимого, то оно заменяется в этом отношении на рассматриваемый союз.

- 19) Правило, обрабатывающее союзы «и», «или», «а», находящиеся в начале сегмента. Правило в основном аналогично предыдущему, отличие заключается в том, что в качестве w_2 выбирается не любое допустимое слово, ближайшее к

рассматриваемому союзу по направлению к концу предложения, а только глагол или краткое причастие. Затем так же, как это делается в предыдущем правиле, для w_2 ищется слово w_3 , и с участием этих слов и рассматриваемого союза создаются синтаксические отношения «ОДНОР_ИГ».

- 20) Правило, проверяющее, является ли последнее допустимое слово предложения, идущее в предложении до количественного числительного в именительном падеже, существительным. Если да, то создаётся отношение «НОМЕР», в котором найденное существительное — главное слово, а числительное — зависимое.
- 21) Правило, проверяющее, является ли последнее допустимое слово предложения, идущее в предложении до числительного, словом, которое может обозначать количественное отношение («больше», «ниже», «равное» и т.д.). Если это так, то строится синтаксическое отношение «КОЛИЧ» («количественное»), где найденное слово является главным, а числительное — зависимым словом.
- 22) Правило, проверяющее, является ли существительным последнее отличное от частицы «не» допустимое слово предложения, идущее в предложении до слова, обозначающего количественное отношение. Если это так, то строится синтаксическое отношение «ОПР» («определяющее»), где найденное существительное — главный элемент, а слово, обозначающее количественное отношение — зависимый.
- 23) Правило, обрабатывающее слово, чья лемма — «который». Сначала в предыдущем сегменте ищется ближайшее к рассматриваемому слову допустимое существительное, имеющее то же число, что и рассматриваемое слово. Затем ищется ближайшее к рассматриваемому слову по направлению к концу предложения допустимое слово, являющееся глаголом, кратким причастием или тире. Если нужные слова найдены, создаётся синтаксическое отношение «ПРИДАТ_ОПР» («придаточное определяющее»), в ко-

тором главным словом является найденное существительное, а зависимым — найденный глагол, краткое причастие или тире.

Описанные выше правила расположены в том же порядке, в каком их рассматривает алгоритм. Если несколько правил независимо друг от друга пытаются создать разные отношения, в которых одно и то же слово было бы зависимым, сработает только то из правил, которое было рассмотрено раньше. Таким образом, при изменении порядка применения правил результат может измениться.

Помимо правил, касающихся построения синтаксических отношений, имеется правило, изменяющее уже созданные синтаксические отношения. Это правило для каждого союза «и», «или», «а» просматривает все отношения «ОДНОР_ИГ», где этот союз является главным словом. Если среди зависимых слов есть слова $w_1, w_2, \dots, w_k, k \in N$ из списка существительных, обозначающих величину или значение, и ровно одно слово не из этого списка, то отношения, в которых рассматриваемый союз — главный, разрываются, а в том отношении, где этот союз является зависимым, он заменяется на единственное найденное слово не из списка существительных, обозначающих величину или значение. Далее происходит перебор вершин ориентированного дерева от вершины, которой соответствует рассматриваемый союз, по направлению к корню дерева. Пусть вершине, получаемых при этом переборе, соответствует слово w из списка существительных, обозначающих величину или значение. Тогда перебор прекращается; найденное слово в том отношении, где оно является зависимым, заменяется на рассматриваемый союз; создаются отношения «ОДНОР_ИГ», в которых главным словом является рассматриваемый союз, а зависимыми — слова w, w_1, w_2, \dots, w_k .

Данное правило введено по следующей причине: оказывается полезным считать, что существительные, обозначающие значение, могут быть однородными только с существительными, обозначающими значение.

Приведём пример предложения, для которого описанный алгоритм построил дерево зависимостей. Рассмотрим следующее предложение, взятое из [4]: «Стоимость объектов основных средств погашается посредством начисления амортизации, если иное не установлено настоящим Положением».

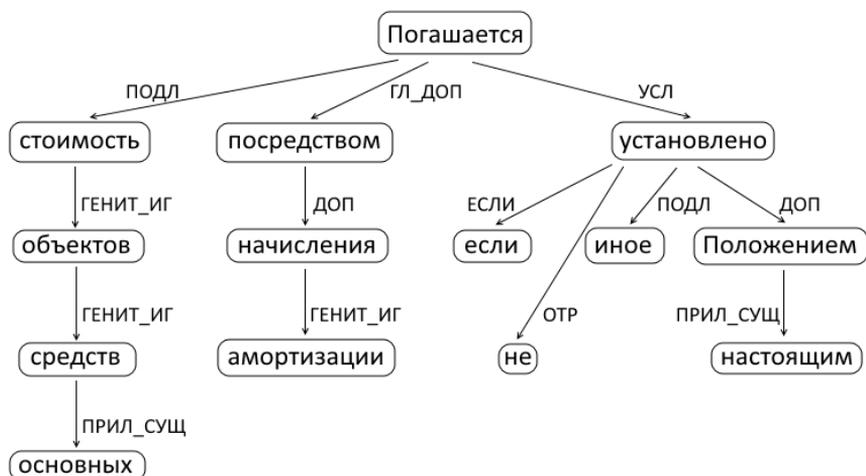


Рис. 1. Пример дерева зависимостей предложения

Будем считать, что на вход программы, осуществляющей синтаксический анализ, поступила последовательность лексем, на которые было разбито предложение, а также следующие токены (они получены благодаря морфологическому разбору, произведённому программой, созданной на проекте АОТ [3]):

- 1) Стоимость; стоимость; С,но,жр,им,ед;
- 2) объектов; объект; С,но,мр,рд,мн;
- 3) основных; основной; П,но,од,рд,мн;

- 4) средств; средство; С,но,ср,рд,мн;
- 5) погашается; погашаться; Г,дст,нп,нс,Зл,нст,ед;
- 6) посредством; посредством; ПРЕДЛ;
- 7) начисления; начисление; С,но,ср,рд,ед;
- 8) амортизации; амортизация; С,но,жр,рд,ед;
- 9) если; если; СОЮЗ;
- 10) иное; иной; МС-П,но,од,ср,вн,им,ед;
- 11) не; не; ЧАСТ;
- 12) установлено; установить; КР_ПРИЧАСТИЕ,
стр,пе,св,но,од,прш,ср,ед;
- 13) настоящим; настоящий; П,но,од,ср,тв,ед;
- 14) Положением; положение; С,но,ср,тв,ед.

В результате применения описанного алгоритма создаются следующие синтаксические отношения (у отношения сначала указывается главное слово, а потом зависимое):

- 1) ГЕНИТ_ИГ(Стоимость, объектов) по правилу 5;
- 2) ПОДЛ(погашается, Стоимость) по правилу 12;
- 3) ГЕНИТ_ИГ(объектов, средств) по правилу 5;
- 4) ПРИЛ_СУЩ(средств, основных) по правилу 17;
- 5) ГЛ_ДОП(погашается, посредством) по правилу 11;
- 6) ГЕНИТ_ИГ(начисления, амортизации) по правилу 5;
- 7) ДОП(посредством, начисления) по правилу 6;
- 8) ЕСЛИ(установлено, если) по правилу 14;
- 9) УСЛ(погашается, установлено) по правилу 15;

- 10) ПОДЛ(установлено, иное) по правилу 12;
- 11) ОТР(установлено, не) по правилу 4;
- 12) ПРИЛ_СУЩ(Положением, настоящим) по правилу 17;
- 13) ДОП(установлено, Положением) по правилу 6.

Получившееся дерево зависимостей можно увидеть на рис. 1.

Описанный набор правил ещё не является полным и продолжает пополняться, однако в настоящий момент этих правил достаточно, чтобы осуществить синтаксический анализ части Положения по бухгалтерскому учёту ПБУ 6/01 [4], касающейся амортизации основных средств. При подаче на вход алгоритма произвольного предложения из какого-либо юридического текста в подавляющем большинстве случаев программа также осуществляет синтаксический разбор без ошибок.

Список литературы

- [1] Кудрявцев В. Б., Гасанов Э. Э., Перпер Е. М. Автоматическая генерация компьютерной программы, моделирующей нормативно-правовой акт // Интеллектуальные системы. Теория и приложения — 2014. —Т. 18, вып. 2. — С. 133–156.
- [2] Подколзин А.С. Самообучение интеллектуальной системы // Интеллектуальные системы. Теория и приложения — 2014. —Т. 18, вып. 2. — С. 197–266.
- [3] А. Сокирко. Семантические словари в автоматической обработке текста (по материалам системы ДИАЛИНГ). <http://www.aot.ru/docs/sokirko>
- [4] ПРИКАЗ МИНИСТЕРСТВА ФИНАНСОВ РОССИЙСКОЙ ФЕДЕРАЦИИ от 30 марта 2001 г. N 26н ОБ

Е. М. Перпер

УТВЕРЖДЕНИИ ПОЛОЖЕНИЯ ПО БУХГАЛТЕРСКО-
МУ УЧЕТУ "УЧЕТ ОСНОВНЫХ СРЕДСТВ" ПБУ 6/01.

<http://base.consultant.ru/cons/cgi/online.cgi?req=doc;base=LAW;n=111>