



119270, Москва, Лужнецкая наб., д. 6,
стр.1, офис 214, ООО «ЭР СИ О»
Тел./факс: (495) 287-98-87
E-mail: info@rco.ru
<http://www.rco.ru>

Руководство администратора

RCO Fact Tuner – программа для настройки семантических шаблонов

Версия 1.5

(Microsoft Windows)

Москва, 2007

В содержание данного документа могут быть внесены изменения без предварительного уведомления. Названия организаций, имена и даты, используемые в качестве примеров, являются вымышленными, если не оговорено обратное.

© ООО «ЭР СИ О», 2007. Все права защищены.

ЭР СИ О, Russian Context Optimizer, RCO являются охраняемыми товарными знаками.

ООО «ЭР СИ О» может являться правообладателем патентов и заявок, поданных на получение патента, товарных знаков и объектов авторского права, которые имеют отношение к содержанию данного документа.

Предоставление вам данного документа не означает передачи какой-либо лицензии на использование данных патентов, товарных знаков и объектов авторского права, за исключением использования, явно оговоренного в лицензионном соглашении ООО «ЭР СИ О».

Все другие названия юридических лиц и изделий являются охраняемыми товарными знаками или товарными знаками, принадлежащими их владельцам.

Содержание

Обзор	4
Настройка описаний фреймов – работа с RCO Fact Tuner	5
Интерфейс пользователя.....	5
Подключение описаний фреймов	6
Работа с перечнем описаний фреймов	6
Управление перечнем описаний фреймов	6
Свойства описания фрейма.....	8
Создание/Удаление описания фрейма.....	8
Настройка описания фрейма	9
Просмотр описания	9
Свойства элементов описания фрейма	9
Добавление/Удаление элементов описания фрейма	12
Анализ текста и тестирование описаний фреймов	13
Преобразование предложения в описание фрейма	13
Сценарий создания описаний фреймов	14
Приложение 1. Описание атрибутов	15
Приложение 2. Синтаксис описания ограничений	18

Обзор

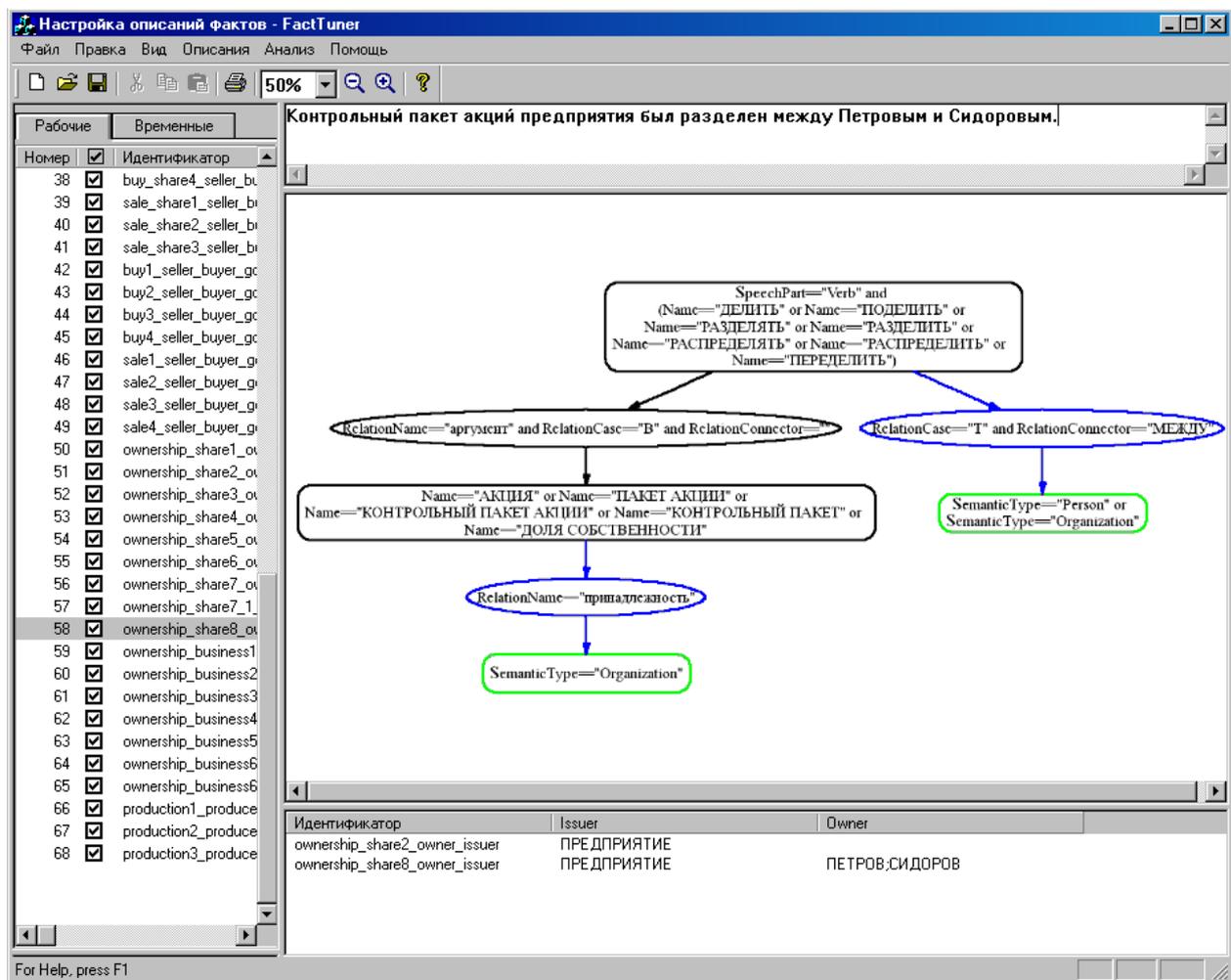
В этом руководстве содержатся основные сведения о способах создания и настройки семантических шаблонов (описаний фреймов, используемых для выделения в тексте ситуаций и их участников) при помощи графического приложения **RCO Fact Tuner**.

Для установки приложения просто скопируйте все файлы дистрибутива в папку **RCO Fact Tuner** или любую иную папку на вашем компьютере.

Настройка описаний фреймов – работа с RCO Fact Tuner

Интерфейс пользователя

Общий вид интерфейса пользователя **RCO Fact Tuner** приведен на рисунке ниже.



Основу интерфейса приложения составляют четыре окна:

- *Окно 1* – в левой части экрана. Содержит две закладки, позволяющие просмотреть перечень описаний фреймов;
- *Окно 2* – вверху экрана. Позволяет вводить текст для анализа;
- *Окно 3* – в центре экрана. Отображает граф выбранного описания фрейма или граф семантической сети текста;
- *Окно 4* – внизу экрана. Отображает результат выделения фреймов в тексте – идентификаторы сработавших описаний и участников ситуаций сгруппированы по ролям участников (именам слотов фреймов).

Подключение описаний фреймов

Семантические шаблоны (описания фреймов) создаются в ходе работы с программой и автоматически сохраняются в файлах ***.dot** подкаталога **\ld** папки установки **RCO Fact Tuner**. Перечень описаний, загружаемых вместе с приложением, указывается в секции `semantic-expounder\se-templates` файла конфигурации **config.xml** из папки установки **RCO Fact Tuner**. При создании/удалении/переименовании описаний содержимое секции соответствующим образом модифицируется.

Работа с перечнем описаний фреймов

Перечень описаний фреймов представлен в *окне 1*. Описания разделены на группы – *рабочие* и *временные*, доступ к которым осуществляется через одноименные закладки.

Список рабочих описаний автоматически загружается в начале работы программы и может пополняться новыми описаниями в процессе работы.

Временных описаний всего два, с идентификаторами *source* и *result*. Они автоматически формируются программой в результате анализа текста, заданного в *окне 2*. Затем их можно модифицировать и скопировать в рабочие описания. Первое временное описание изначально содержит граф с результатом разбора введенного текста – семантическую сеть. Второе временное описание содержит граф, формирующийся вследствие применения рабочих описаний к графу разбора введенного текста. Таким образом, граф *result* либо в точности совпадает с графом *source*, либо содержит дополнительные узлы и связи, образующиеся в результате срабатывания описаний, содержащих порождаемые элементы графа.

Основное назначение временных описаний – получение графа-прототипа для построения целевого описания путем анализа типового предложения, представляющего один из способов описания ситуации в тексте.

Управление перечнем описаний фреймов

Вид перечня из *окна 1* приведен на рисунке ниже и содержит поля:

- *Номер* – определяет порядок применения описания;
- *Пометка* – для анализа текста только помеченных описаний;
- *Идентификатор* – строка латинских символов для идентификации описания;
- *Краткое описание и полное описание* – произвольные строки, содержащие пояснения к описанию фрейма;
- *Число вершин* – число узлов и связей в описании;
- *Опц. вершин* – число необязательных узлов и связей в описании;
- *Порожд. вершин* – число порождаемых узлов и связей описания;
- *Совпавш. вершин* – число узлов и связей в графе описания фрейма, успешно отождествленных с вершинами семантической сети при анализе последнего текста;
- *Изоморфизмов* – число подграфов семантической сети, изоморфных графу описания при анализе последнего текста.

Порядок следования полей в списке можно менять, перетаскивая мышью названия столбцов.

Номер	Идентификатор	Краткое описание	Полное описание	Число...	Опц...	Порожд
0	serv_expl_verbstr_verb1	Служебное: субъект - модальный или т...	Субъект хочет или спешит или начинает ...	6	0	1
1	serv_expl_verbstr_noun1	Служебное: субъект - экспликатор или ...	Субъект совершает или начинает действие	6	0	1
2	serv_expl_verbstr_noun2	Служебное: субъект - темповой глагол ...	Субъект торопится с действием	6	0	1
3	serv_expl_verbstr_noun3	Служебное: субъект - связочный глагол...	Субъект является уточнением	6	0	1
4	serv_expl_verbspec_n...	Служебное: субъект - особый глагол - ...	Субъект объявляет или мечтает о действ...	6	0	1
5	serv_expl_verbspec_n...	Служебное: субъект - особый глагол - ...	Субъект решается на действие	6	0	1
6	serv_expl_verbspec_n...	Служебное: субъект - особый глагол - ...	Субъект готовится или стремится к дейс...	6	0	1
7	serv_expl_verbspec_n...	Служебное: субъект - особый глагол - ...	Субъект сосредотачивается на действии	6	0	1
8	serv_expl_verbspec_n...	Служебное: субъект - особый глагол - ...	Субъект занимается действием	6	0	1
9	contract1_signer1_sign...	договор: партнер - Signer1, партнер - Si...	Партнер1 заключает контракт о предмет...	10	5	0
10	contract2_signer1_sign...	договор: партнер - Signer1, партнер - Si...	Заключение контракта о предмете партн...	10	5	0
11	contract2_1_signer1_si...	договор: партнер - Signer1, партнер - Si...	Заключение контракта о предмете межд...	7	4	0
12	contract2_2_signer1_si...	договор: партнер - Signer1, партнер - Si...	Контракт партнера1 и партнера2 о пред...	7	4	0
13	contract3_signer1_sign...	договор: партнер - Signer1, партнер - Si...	Партнер1 и партнер2 договариваются о ...	7	4	0
14	contract4_signer1_sign...	договор: партнер - Signer1, партнер - Si...	Партнер1 и партнер2 приходят к соглаш...	10	5	0
15	contract5_signer1_sign...	договор: партнер - Signer1, партнер - Signer2, договор - Contract, содержание - Subject		10	5	0
16	contract6_signer1_sign...	договор: партнер - Signer1, партнер - Si...	Партнер1 и партнер2 ставят подпись под...	12	5	0
17	contract7_signer1_sign...	договор: партнер - Signer1, партнер - Si...	Партнер1 и партнер2 работают над конт...	10	5	0
18	contract8_signer1_sign...	договор: партнер - Signer1, партнер - Si...	Работа партнера1 и партнера2 над контр...	10	5	0
19	contract9_signer1_sign...	договор: партнер - Signer1, партнер - Si...	У партнера1 есть контракт с партнером...	10	5	0
20	meeting1_participant1_...	встреча: участник1 - Participant1, участн...	Участник1 встречается с участником2	5	0	0
21	meeting2_participant1_...	встреча: участник1 - Participant1, участн...	Встреча участника1 с участником2	5	0	0
22	meeting3_participant1_...	встреча: участник1 - Participant1, участн...	Участник1 посещает участника2	5	0	0
23	meeting4_participant1_...	встреча: участник1 - Participant1, участн...	Участник1 проводит встречу с участнико...	7	0	0
24	meeting5_participant1_...	встреча: участник1 - Participant1, участн...	У участника1 происходит встреча с участ...	8	3	0
25	meeting6_participant1_...	встреча: участник1 - Participant1, участн...	Визит участника1 к участнику2	5	0	0
26	meeting7_participant1_...	встреча: участник1 - Participant1, участн...	Участник1 наносит визит участнику2	7	0	0
27	meeting8_participant1_...	встреча: участник1 участник2	Участник1 и участник2 на встрече	7	0	0
28	traveling1_traveler_des...	поездка: путешественник - Traveler, ме...	Путешественник едет в место	5	0	0
29	traveling2_traveler_des...	поездка: путешественник - Traveler, ме...	Поездка путешественника в место	5	0	0
30	traveling3_traveler_des...	поездка: путешественник - Traveler, ме...	Путешественник едет что-то делать в ме...	7	0	0
31	traveling4_traveler_des...	поездка: путешественник - Traveler, ме...	Путешественник посещает место	5	0	0
32	traveling5_traveler_des...	поездка: путешественник - Traveler, ме...	Путешественник оказывается в месте	5	0	0
33	traveling6_traveler_des...	поездка: путешественник - Traveler, ме...	Пребывание путешественника в месте	5	0	0

Чтобы скрыть отобранные описания, снимите с них флажок в поле «Пометка», затем в меню «Описания» выберите «Фильтр→Только помеченные» (для возвращения описаний выберите «Фильтр→Все описания»). При анализе текста работают только помеченные описания.

Порядок следования описаний в перечне и порядок их применения при выделении фреймов совпадают. Изменение порядка полезно в случаях:

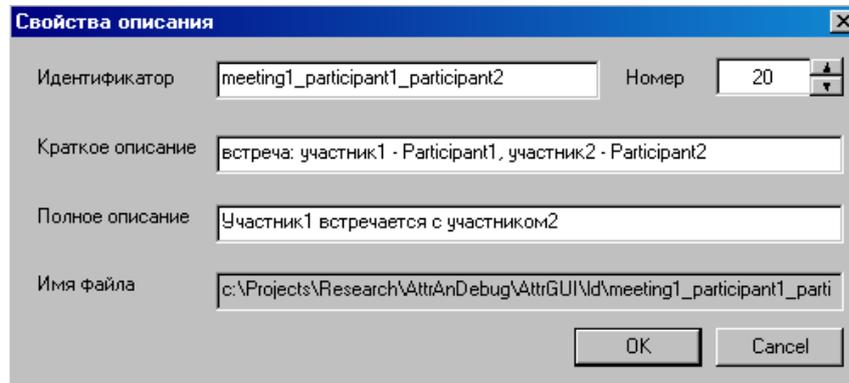
- группировки описаний для облегчения визуального восприятия,
- обеспечения нужного порядка при поиске фреймов в случае зависимости одних описаний от других – некоторые описания порождают в семантической сети текста дополнительные узлы и связи, приводящие к выделению фреймов другими описаниями.

Изменить порядок следования шаблонов можно двумя способами:

- изменить значение поля **Номер** в **свойствах описания**,
- выбрать описание в *окне 1* и нажать правую кнопку мыши. В открывшемся контекстном меню выбрать пункт «переместить», и далее по желанию – «выше», «ниже», «в начало», «в конец».

Свойства описания фрейма

Описание фрейма обладает набором общих свойств, задать которые автоматически предлагается при создании нового рабочего описания. Свойства любого существующего описания можно изменить. Для доступа к свойствам описания в *окне 1* наведите на него мышью и нажмите правую кнопку, после чего в открывшемся контекстном меню выберите пункт «**свойства**». Вид окна свойств приведен на рисунке ниже.



Описание содержит следующие свойства:

- Идентификатор (строка из латинских букв, цифр, символов подчеркивания), уникальный в рамках всего перечня описаний, позволяет определить класс фрейма при его выделении в тексте;
- Номер – порядковый номер описания в перечне. При анализе текста описания применяются в порядке возрастания номеров;
- Краткое описание – произвольная строка текста. Обычно содержит название типа ситуации и перечень ролей ее участников с указанием идентификаторов ролей;
- Полное описание – произвольная строка текста. Обычно содержит типовое предложение, описывающее фрейм;
- Имя файла – устанавливается автоматически.

Создание/Удаление описания фрейма

Создавать или удалять можно только рабочие описания.

Наиболее часто новое описание создают, копируя рабочее или временное описание, близкое по структуре к создаваемому. Для получения копии выберите описание в *окне 1* и нажмите правую кнопку мыши, затем в контекстном меню выберите «**новое описание**→**копия текущего**». Через пункты «**новое описание**→**пустое**» или «**новое описание**→**из файла**», создаются пустое описание или копия описания из выбранного файла.

Для удаления описания выберите описание в *окне 1* и нажмите правую кнопку мыши, после чего в открывшемся контекстном меню выберите пункт «**удалить**» – описание будет удалено из перечня и переименовано на диске – к имени файла с описанием будет добавлен символ “~” слева.

Настройка описания фрейма

Просмотр описания

Структура описания фрейма автоматически отображается в виде графа в *окне 3* при щелчке мышью на описании в *окне 1*.

Для изменения масштаба изображения графа используйте элементы на панели инструментов приложения – кнопки “+”, “-”, или выпадающий список с указанием масштаба изображения в процентах.

Узлы в графе представлены в виде прямоугольников со скругленными краями, связи между ними – в виде направленных дуг с эллипсами посередине. Текст в узлах и связях формируется на основе их **свойств**.

Зеленым цветом выделяются целевые узлы, синим – необязательные связи, красным цветом – порождаемые узлы и связи, а черным – все остальные (вспомогательные) узлы и связи описания фрейма.

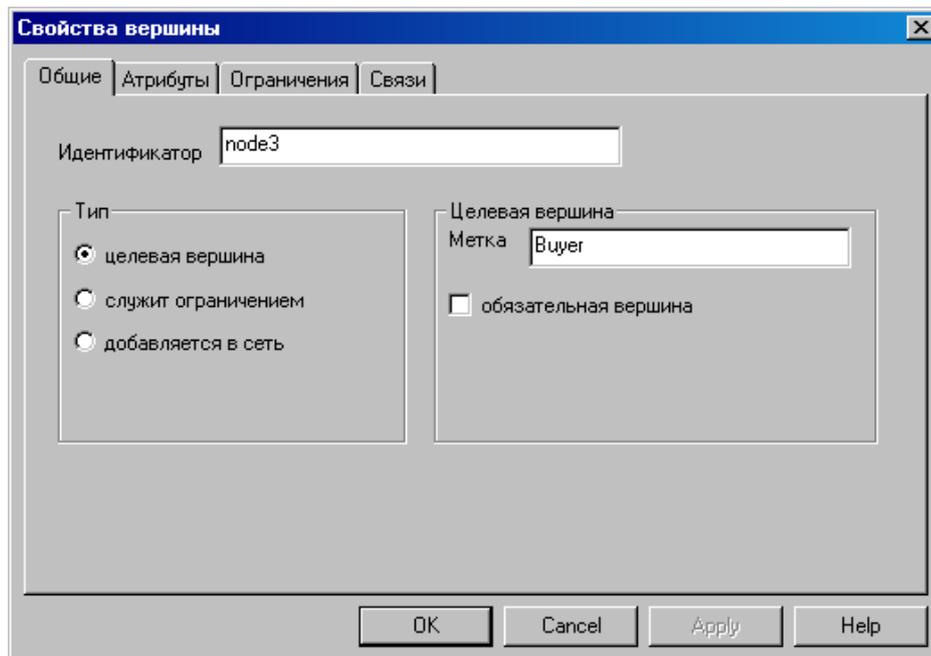
Свойства элементов описания фрейма

Каждый элемент семантического шаблона – описания фрейма (узел или связь графа) имеет набор свойств.

Для просмотра или изменения свойств щелкните правой кнопкой мыши на выбранном элементе в *окне 3* на изображении графа. Открывшееся диалоговое окно содержит четыре закладки.

Общие

Свойства, содержащиеся в закладке «**Общие**», отражены на рисунке ниже. Набор свойств этой группы одинаков для вершин и для связей.



Идентификатор – произвольный идентификатор узла или связи, написанный латинскими символами, уникальный в рамках множества всех элементов данного графа.

Тип:

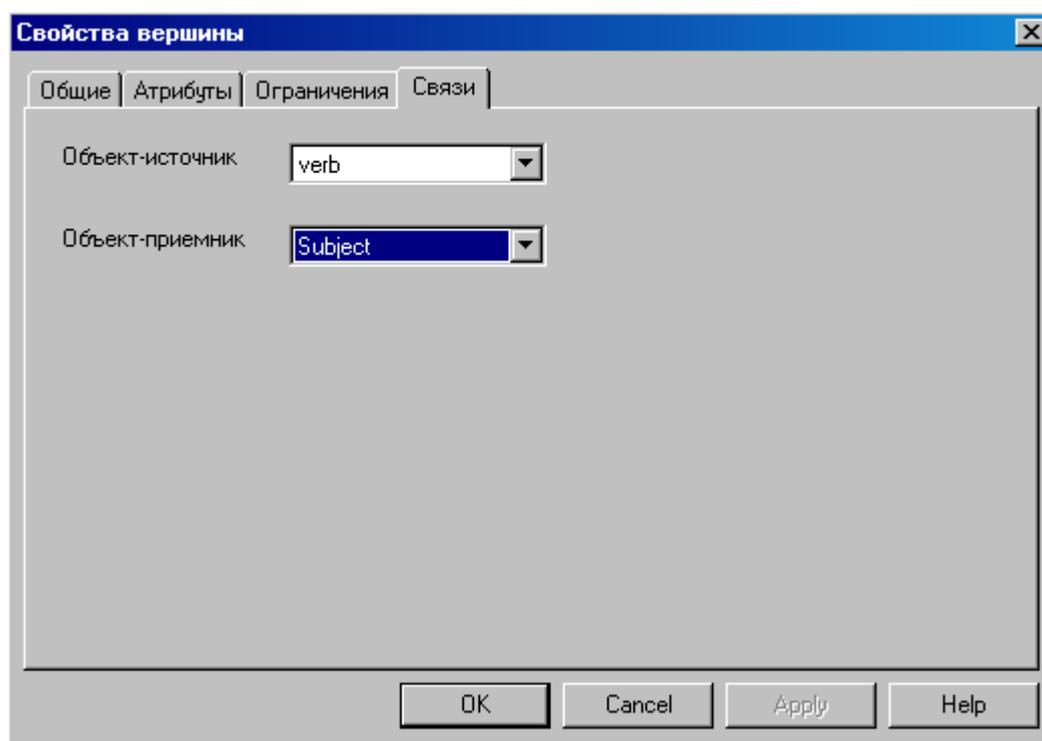
- *целевая вершина* – узел такого типа соответствует участнику ситуации, чье имя извлекается при нахождении фрейма в тексте;
- *служит ограничением* – вспомогательный узел или связь, которая позволяет распознать присутствие фрейма в тексте;
- *добавляется в сеть* – порождаемый узел (связь), вносящийся в граф семантической сети при срабатывании описания фрейма.

Метка – обозначение роли (строка из латинских символов) участника ситуации (имени слота фрейма), соответствующего данному узлу. Установка метки возможна только для целевых узлов.

Обязательная вершина – флаг указывает, является ли присутствие данного узла или связи в семантической сети текста необходимым для срабатывания описания фрейма. Снятие этого флага имеет смысл только для целевых узлов.

Связи

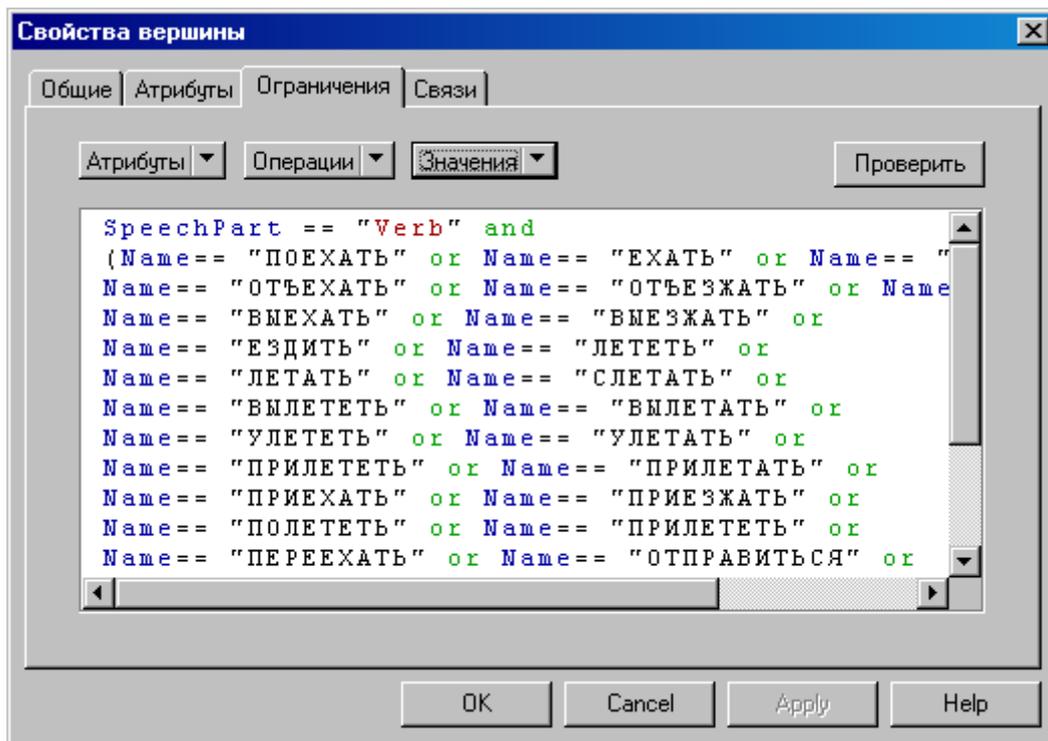
Свойства, содержащиеся в закладке «**Связи**», отражены на рисунке ниже.



Свойства этой группы присутствуют только у связей графа и позволяют задать тип узла, из которого исходит (*объект-источник*) и в который приходит связь (*объект-приемник*). Значениями этих свойств являются идентификаторы узлов графа, установленные в группе свойств «*Общие*».

Ограничения

В закладке «**Ограничения**» задается логическое выражение, определяющее условия, которым должен удовлетворять узел (связь) семантической сети, соответствующий данному элементу описания фрейма.



Логическое выражение содержит условия на отдельные атрибуты, объединенные логическими связками «and» и «or». Список predetermined атрибутов и их значений приведен в [Приложении 1](#). Синтаксис описания ограничений – в [Приложении 2](#).

Кроме того, в ограничениях можно использовать любые другие атрибуты, которые могли быть добавлены в элементы семантической сети текста за счет описаний фреймов, чьи элементы содержат добавляемые атрибуты.

Атрибуты

В закладке «Атрибуты» указываются атрибуты вершины или связи графа семантической сети текста, получаемые при срабатывании описания фрейма. Эти атрибуты могут быть заданы в дополнение к уже существующим атрибутам или взамен них. Атрибуты, установленные при срабатывании описания фрейма, используются в дальнейшем при проверке ограничений в других описаниях фреймов наравне с ранее существовавшими атрибутами.

Атрибуты могут задаваться для элементов графа любого типа – целевых, порождаемых, вспомогательных. Создание порождаемых узлов и связей без указания их атрибутов бессмысленно.

Анализ текста и тестирование описаний фреймов

Для анализа текста (отдельного предложения) введите его в *окно 2* и нажмите клавишу **F5**.

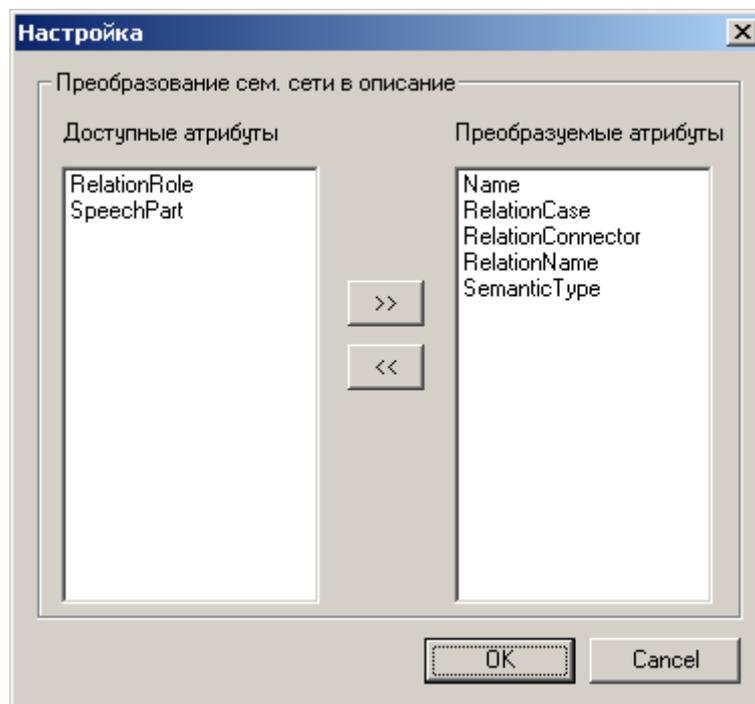
В результате анализа в *окне 3* отобразится граф, представляющий семантическую сеть проанализированного текста, а в *окне 4* в виде таблицы отобразится список выделенных фреймов и участников соответствующих ситуаций.

Преобразование предложения в описание фрейма

После анализа текста семантическая сеть автоматически преобразуется во временное описание с идентификатором *source*, которое может служить прототипом для создания соответствующего описания фрейма.

При преобразовании семантической сети в описание фрейма атрибуты каждого узла и связи преобразуются в условия, представляющие собой возможные **ограничения** на элементы описания.

Перечень преобразуемых атрибутов регулируется при помощи диалога настройки, вызываемого выбором пункта меню «**Анализ**→**Настройки**».



Используйте кнопки “>>” и “<<” для перемещения атрибутов из списка доступных в список преобразуемых и обратно. Сохраненные настройки вступают в силу при следующем запуске процедуры анализа текста.

Сценарий создания описаний фреймов

Типичный сценарий создания и настройки системы описаний целевых фреймов включает в себя следующие фазы:

1. Ввод типового предложения, представляющего один из вариантов описания ситуации в тексте, в *окно 2* и его анализ программой – клавиша **F5**;
2. Просмотр результатов выделения фреймов на основе уже существующих описаний – списка найденных фреймов в *окне 4*. Если соответствующий фрейм корректно выделен (нужное описание уже существует), можно перейти к шагу 1 или завершить работу. Иначе – шаг 3;
3. Просмотр графа семантической сети, построенного в результате анализа предложения – временного описания *source* в списке описаний (*окно 1*). Создание прототипа нового описания – копирование временного описания в рабочие описания с указанием **свойств** нового описания;
4. **Удаление/добавление** элементов нового описания (при необходимости). Изменение требуемых **свойств элементов** нового описания;
5. Тестирование нового описания на различных лексико-синтаксических трансформациях типового предложения (синонимические замены, изменение порядка слов, осложнения и т.п.) – ввод предложения, нажатие клавиши **F5**. Просмотр результата выделения фреймов в *окне 4*. Если соответствующий фрейм не выделен в соответствии с новым описанием или выделены не все участники – возврат к шагу 4;
6. Новое описание фрейма готово. Переход к шагу 1 или конец.

Приложение 1.

Описание атрибутов

Ниже приведены атрибуты узлов и связей, используемые в графах описаний фреймов.

Имя атрибута	Семантика атрибута	Основные полезные значения
	Атрибуты узлов графа	
Name	Строка текста, соответствующая узлу. В узле может быть несколько строк, каждая из которых соответствует полному словосочетанию, образованному от ключевого существительного в узле, например: «НОВЫЙ УКАЗ ПРЕЗИДЕНТА», «УКАЗ ПРЕЗИДЕНТА», «УКАЗ». Рекомендуется использовать для сравнения содержимого узла с конкретными словами при написании схемы фрейма.	Любая текстовая строка.
StrictName	Строка текста, соответствующая полному словосочетанию, образованному от ключевого существительного в узле. Может быть несколько, например: «ИТАЛЬЯНСКАЯ ПОГРУЗОЧНАЯ ТЕХНИКА», «НЕМЕЦКАЯ ПОГРУЗОЧНАЯ ТЕХНИКА». Эти строки выдаются как значение атрибута при выделении фрейма.	
SpeechPart	Часть речи (английское название), соответствующая ключевому слову в узле.	SpeechPart=="Noun", SpeechPart=="Verb", SpeechPart=="Adjective"
SemanticType	Семантическая категория, соответствующая ключевому слову в узле. Представляет альтернативу части речи в ряде случаев.	SemanticType=="Organization" – любая организация; SemanticType=="Organization: Name" – именованная организация; SemanticType=="Person" – любая персона; SemanticType=="Person" – именованная персона; SemanticType=="Geoplace" – географическое место; SemanticType=="Event" – событие (действие или состояние)

	Атрибуты связей графа	
RelationName	Название синтактико-семантического отношения, соответствующего связи.	<p>«<i>аргумент</i>» – связь предиката с аргументом, устанавливаемая на основании словаря моделей управления. Конкретный тип связи описывается указанными ниже атрибутами связи (Иванов <- подписал, подписание -> указа);</p> <p>«<i>принадлежность</i>» – связь существительного с подчиненным генитивом, не описанная в словаре моделей управления (указ -> президента);</p> <p>«<i>тождество</i>» – связь существительного с приложением или связь в конструкции с предикатом тождества (директор <- Иванов, Иванов -> это директор);</p> <p>«<i>атрибут</i>» – связь существительного с прилагательным (президентский <- указ);</p> <p>«<i>атрибут-наречие</i>» – связь с наречием (уволить -> нехотя);</p> <p>«<i>количество</i>» – связь существительного с числительным (пять <- человек);</p> <p>«<i>инфинитив</i>» – связь в глагольной группе (начал -> скупать);</p> <p>«<i>обстоятельство</i>» – связь существительного или глагола с сирконстантом (написал -> в Америке, цветок ->на окне);</p> <p>«<i>квантор</i>» – частица, местоимение, прилагательное (этот <- человек, не <- поддержал).</p>
RelationConnector	Коннектор отношения – предлог, при помощи которого устанавливается синтаксическая связь.	Любые предлоги или пустое поле, например: RelationConnector==“” RelationConnector==“ПРИ ПОМОЩИ”, RelationConnector==“ПРИ”
Relation Case	Семантический падеж аргумента предиката в синтактико-семантическом отношении, соответствующем связи. Соответствует грамматическому падежу в модели управления при употреблении в простом неосложненном предложении с активной конструкцией (без причастных оборотов и т.п.). Пример: подписать -> указ –	И, Р, Д, В, Т, П или пусто, например: RelationCase == “И”, RelationCase == “”

	Винительный гр. и Винительный сем.; указ <- подписан – Именительный гр. но! Винительный сем.	
Relation Role	Семантическая роль аргумента предиката в синтактико-семантическом отношении, соответствующем связи. Извлекается из словаря моделей управления, но ввиду высокой неоднозначности вместо нее в настоящее время рекомендуется использовать атрибуты RelationCase и RelationConnector.	« <i>субъект</i> » – это исключение. В конструкциях при существительном со строевым глаголом типа «Вася осуществил продажу» служебные описания фреймов могут достраивать прямую связь (Вася -> продажа), приписывая ей роль «субъект». Этим можно и нужно пользоваться в тех описаниях фреймов, где есть семантический субъект при предикате, выраженном существительным.

Приложение 2.

Синтаксис описания ограничений

При задании ограничений в узлах и связях описаний фреймов используется следующий синтаксис.

Логические операции:

- and (И), например: `RelationCase == "P" and RelationConnector == "ВТЕЧЕНИЕ"`;
- or (ИЛИ), например: `Name == "АКЦИЯ" or Name == "КОНТРОЛЬНЫЙ ПАКЕТ"`;
- ! (НЕ), например: `!Name == "АКЦИЯ"`.

Условие на атрибут:

- `<атрибут> <операция> <значение>`.

Операции с атрибутами:

- `==` (равно);
- `=~` (регулярное выражение);
- `>` (больше);
- `<` (меньше);
- `>=` (больше либо равно);
- `<=` (меньше либо равно).

Имена атрибутов:

- предопределенные – см. [Приложение 1](#);
- пользовательские – произвольная строка, состоящая из латинских символов.

Значения атрибутов:

- строки – в двойных кавычках;
- целые числа – последовательности цифр.

Приоритет операций:

- условия на атрибуты;
- операции в скобках `()`;
- ! – относится либо к ближайшему справа выражению в скобках, либо к ближайшему условию на атрибут;
- and, or – слева направо.