

Разрешение местоименной анафоры в системе ЭТАП-3

на примере возвратных местоимений и местоимений 3 лица

Евгения Иншакова
Лаборатория компьютерной
лингвистики ИППИ РАН
inshakova@iitp.ru

Система правил разрешения анафоры — один из модулей системы автоматического семантического анализа текста (см., например, [Богуславский И.М. и др. Семантический анализ и ответы на вопросы: система в стадии разработки // Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог». Вып. 14 (21): В 2 т. Т. 1: Основная программа конференции. — М.: Изд-во РГГУ, 2015, с. 62-79], [Boguslavsky, I. Semantic Analysis based on linguistic and ontological resources In Proceedings of the 5th International Conference on Meaning-Text Theory (MTT'2011). Barcelona, September 8—9, 2011. P. 25-36], [Boguslavsky, I. Semantic Analysis based on linguistic and ontological resources In Proceedings of the 5th International Conference on Meaning-Text Theory (MTT'2011). Barcelona, September 8—9, 2011. P. 25-36])

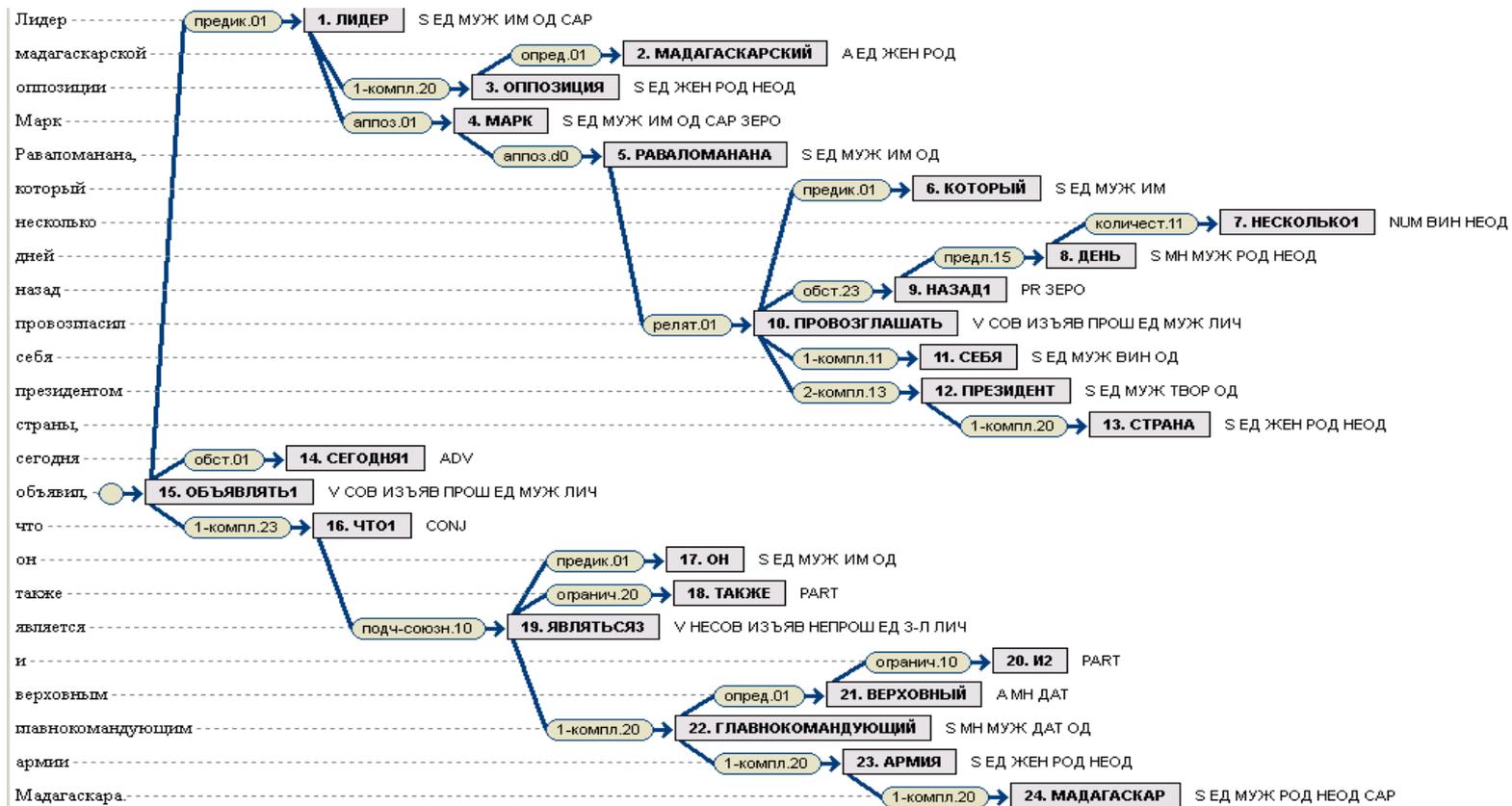
Местоимения, с которыми работают правила на данный момент:

- возвратные *себя* и *свой*
- относительные *который*, *кто*, *что*, *чей*
- местоимения 3 лица
- во многих случаях — нулевые местоимения

Виды данных, используемые ресурсы и инструменты разрешения анафоры

- Результаты работы парсера ЕТАР — деревья зависимостей

Лидер мадагаскарской оппозиции Марк Раваломанана, который несколько дней назад провозгласил себя президентом страны, сегодня объявил, что он также является и верховным главнокомандующим армии Мадагаскара.



- **Данные из статей русского комбинаторного словаря (КС):**
 - частеречная характеристика,
 - синтаксические признаки лексемы,
 - семантические признаки (дескрипторы)
 - онтокорреляты русских лексем (зона SEM-ONTO)
 - модель управления с семантическими ограничениями на заполнение некоторых валентностей
 - лексические функции

Пример статьи ТКС:

```
|16679 20:25:02 24-10-2016          ЗАВОЛЕВАТЬ
      POR:V      ЛИ
      //COB:ЗАВОЛЕТЬ
      DES:Event,AbstractObject,FCP-Process
      SEM-ONTO:PathologicProcess/EmotionalState
      SEM-TRANS:TakingIll/""
//Петя заболел ветрянкой / заболел этой мыслью
// У меня заболела нога - см. ЗАВОЛЕТЬ
      D1.1:Human,Animal
      D2.1:ТВОР,FCP-State,AbstractObject
      D3.1:OT
TRAF:1-КОМПЛ.20
TRAF:1-КОМПЛ.52
      *****
ZONE:EN
TRAF:RA-TRADUCT2.16
LA1:FALL2,LA2:ILL1
TRAF:RA-TRADUCT2.47
OPT:1
LA1:START1,LA2:ACHE2
TRAF:EXPANS.30
TRAF:EXPANS.94
LR:OT
      *****
ZONE:SEM
TRAF:SEM-CONV2.10          //EI
T1:ПРЕДИК,T2:hasObject
TRAF:SEM-CONV2.10
T1:1-КОМПЛ,T2:hasObject2
TRAF:SEM-CONV2.10
```

часть речи
зона дескрипторов
онтокорреляты

модель управления

зона семантического
перевода

Словарная статья сущ. ЗАПРЕТ

449 18:01:24 07-03-2017

ЗАПРЕТ

POR: S АЛ

SYNT: МУЖСК, ИСЧИСЛ, АГЕНС, ОБ-ИНФ

DES: FCP-Action, Event, AbstractObject

SEM-ONTO: Directing

SEM-ONTO: DeonticNecessity/Law/Ordering

SEM-TRANS: Prohibition/"/"/"

D1.1: ТВОР, Agent

D1.2: РОД, Agent

D2.1: РОД

D2.2: НА1

D2.3: ИНФ

D3.1: ДАТ, Agent

_SYN: ЗАПРЕЩЕНИЕ //LF1_CIN

_A2: ЗАПРЕТНЫЙ

_ANTI: РАЗРЕШЕНИЕ2

_V0: ЗАПРЕЩАТЬ

_MAGN: ПОЛНЫЙ/ СТРОГИЙ/ КАТЕГОРИЧЕСКИЙ

_BON: РАЗУМНЫЙ

_ANTIBON: ВЕССМЫСЛЕННЫЙ

_OPER2: БЫТЬ<ПОД2>

_INCEOPER2: ПОПАДАТЬ1<ПОД1>/ ПОДВЕРГАТЬСЯ

_FUNC2: ЛЕЖАТЬ<НА2>

_CAUSEFUNC2: НАЛАГАТЬ<НА1>/ НАКЛАДЫВАТЬ<НА1>/ ВВОДИТЬ<НА1>/ УСТАНОВЛИВАТЬ2<НА1>

_SO_CAUSEFUNC2: НАЛОЖЕНИЕ<НА1>/ ВВЕДЕНИЕ1<НА1>/ УСТАНОВЛЕНИЕ<НА1>

_LIQUFUNC2: СНИМАТЬ1<С1>/ ОТМЕНЯТЬ

_CAUSOPER2: СТАВИТЬ<ПОД1>

_REAL3-М: СОВЛЮДАТЬ/ ПОДЧИНЯТЬСЯ

_SO_REAL3-М: СОВЛЮДЕНИЕ/ ПОДЧИНЕНИЕ|

ANTIREAL3-М: НАРУШАТЬ

• **Онтология** (сейчас используются не все ее возможности)

The screenshot shows the Protégé web interface. At the top, there is a browser address bar with the URL `jabberwocky:8080/webprotege/#Edit:projectId=9df7424f-424b-44d7-9b89-310573cb4a88` and a search bar. Below the browser bar is the Protégé logo and a tab labeled `WebProtege` with a sub-tab `ontoetap`. A navigation bar contains tabs for `Classes`, `Properties`, `Individuals`, `Changes By Entity`, and `Project Dashboard`.

The main interface is divided into two panels:

- Classes Panel:** Located on the left, it features a search bar with the text `Search: walkin`. Below it is a tree view of classes. The class `Constructing` is selected and highlighted in blue. Other visible classes include `Keeping`, `LieDown`, `Listening`, `Looking`, `Maintaining`, `Making`, `Carpentry`, `Cooking`, `Copying`, `DevelopingWeaponOfMassDestructor`, `Manufacture`, `Maneuver`, `Ambush`, `Attack`, `BaseballManeuver`, `Blockade`, `Catching`, `DirectActionActivity`, `GameShot`, `GolfManeuver`, `MilitaryManeuver`, `Score`, `Modeling`, and `OpeningEyes`.
- Class description for Constructing Panel:** Located on the right, it displays the details for the selected class. It includes:
 - Display name:** `Constructing`
 - IRI:** `http://proling.iitp.ru/ontologies/ontoetap.owl#Constructing`
 - Annotations:**
 - `rdfs:comment` with source `SUMO, etaponto`
 - `rdfs:comment` with definition `The &subclass of &Making in which a &StationaryArtifact is built.` and language `en`
 - `rdfs:comment` with source `file: Merge.kif`
 - Properties:**
 - `hasInstrument` with value `Device`
 - `hasObject` with value `StationaryArtifact`

- **Программа, автоматически строящая кореферентные связи**, сравнивающая их с эталонной анафорической разметкой и вычисляющая значения точности (число построенных правильных связей / число всех построенных связей) и полноты (число построенных правильных связей / число эталонных связей)

Функционал для разметки анафорического корпуса и тестирования на нем правил разрешения анафоры

Coreferences

5

6

Manual addition

Pronoun [dropdown] Add

Antecedent

Begin [dropdown]

Sent. number [input]

End [dropdown]

0 [dropdown]

Automatic addition

Run coreference rules Show Listing

Refresh Call tracer

View struct Cmp all

Previous sentences

Семен Еремеевич снова задумался. 1

Может быть, ему вспомнился давний случай с женщиной, которую долго не принимали на работу, потому что муж ее оказался замешанным в какую-то непонятную историю и был арестован. 2

(3:ему,-1;1:Семен)
(15:ее,8:женщиной) 3

Sentences

Show current only

Show previous

Show two previous

4

Cancel

OK

I am not sure

Clear

Go to next sentence

- Корпуса групп предложений из корпуса SynTagRus в формате .tgt с анафорической разметкой

500-Copy1.tgt - Structure Editor

File Edit Sentence Tools Options View Help

Sentence	ID	Class	Comment	Coref	MicroSynt
Не уверен, что я согласен с гайдаровской исторической философией.	[285]		traced		
Но уверен, что если он окажется прав, то это будет в моих интересах.	[286]		traced		
=====	[287]				
Этот объемистый меморандум фактически уже был черновиком будущего хрущевского доклад...	[288]	LF			
Окончательное решение о том, что на закрытом заседании XX съезда будет произнесен антист...	[289]	LF	traced	(16:его,12:доклад)	
=====	[290]				
А то, что Хрущев сохранил многие низовые уравнилельные мифы, хотел и мог быть публичной...	[291]		traced	(22:нему,4:Хру...	
Исходя из этого, вполне естественными были два проекта, которые Хрущев предложил номен...	[292]			(8:которые,7:пр...	
Во-первых, снять с нее ответственность за террор, переложив ее на Сталина, а попутно и все...	[293]	LF		(7:ее,4:ответст...	
=====	[294]				
Впрочем, проверить это нелепо, так что настаивать не будем...	[295]		traced		
Выстраивается очередь и из тех стран, которые объявляют, что лунные исследования станов...	[296]		traced	(7:которые,6:ст...	
=====	[297]				
А далее: либо она замерла из-за неблагоприятных условий, либо она там присутствует, но не ...	[298]			(9:она,3:она)	
А в низших формах могла приспособиться и сохраниться.	[299]		traced		
Ее поиски, ее изучение, ее познание - все это стоит того, чтобы летать на Марс, на другие план...	[300]		traced	(1:Ее,-2;3:она)(3...	

Маркировка знаком “!” предложений, в которых автоматически построенная разметка расходится с эталоном

Как выглядят кореферентные связи в системе ЭТАП?

- Все элементы кореферентной цепочки объединяются в гиперузел класса COREF-GROUP с порядковым номером;
- антецедент получает признак ANTEC;
- если в предложении оказывается несколько кандидатов на роль антецедента, все они получают признак POS-ANTEC, который после прохождения ими правил-«фильтров» должен либо удаляться, либо заменяться на ANTEC.

Порядок работы алгоритма разрешения анафоры

- I. Автоматическое разбиение текста на предложения и его конвертация в формат .tgt (tagged tree), с которым будет работать редактор синтаксических структур
- II. Синтаксический анализ
- III. Построение анафорических связей с оценкой качества

Построение анафорических связей

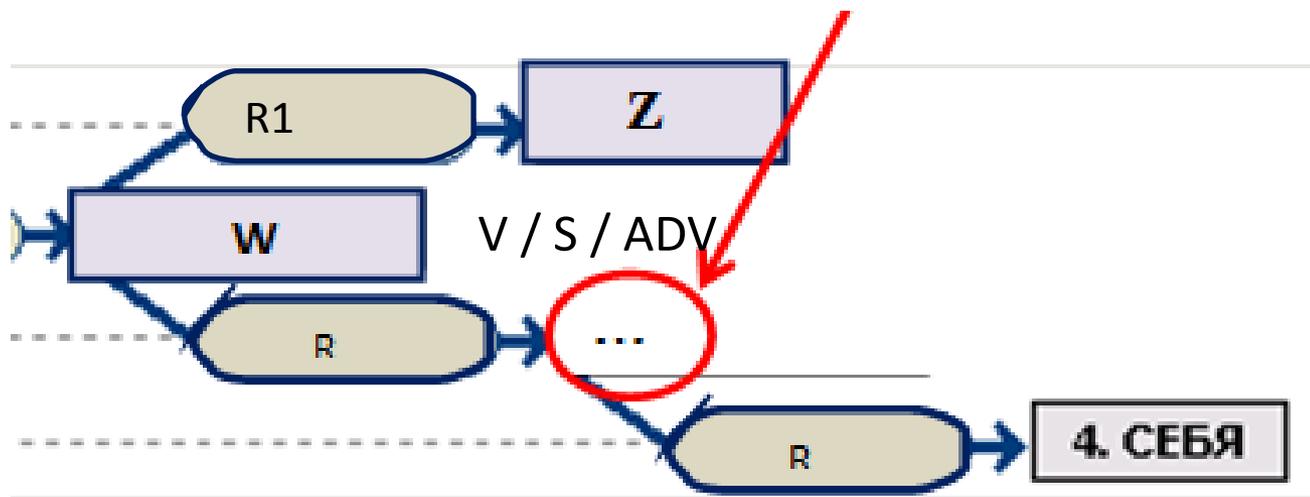
На этапе постсинтаксической обработки предложения:

1. Установление антецедентов возвратных местоимений в финитных клаузах и именных группах (для правильного перевода на английский местоимений СЕБЯ и СВОЙ)

Информация, которая используется в этих правилах

1) лемма местоимения и структура, в которой нужно искать его антеcedент:

ничего / любое слово или цепочка слов,
кроме глаголов в личной форме, зависят
по любой связи, кроме сочинительной



R = любое синтаксическое отношение

R1 = отношение ПРЕДИК / ДАТ-СУБЪЕКТ / КВАЗИАГЕНТ / АГЕНТ

Принцип А: анафор (возвратное или взаимно-возвратное местоимение) должен быть связан в своей локальной области.

Иван_i говорит о себе_i.

*Иван_i слышит, что Петр_j говорит о себе_{*i/j}.*

Локальная область – некоторая составляющая (обычно клауза определенного типа или именная группа). В русском языке локальной областью для возвратных местоимений является финитная клауза, поэтому у них может быть несколько возможных антецедентов в одном предложении (по одному в каждой клаузе):

Он_i не разрешает мне_j [PRO_j производить опыты над собой_{i/j}]

- 2) семантическая категория антецедента местоимения — дескриптор Agent, т.к. известно, что возвратные местоимения чаще всего ссылаются на одушевленные объекты;
- 3) списки конструкций, в которых не должна проводиться антецедентная связь, т.к. в них сочетания местоимений СЕБЯ и СВОЙ и их синтаксических «хозяев» идиоматизировались и местоимения уже не отсылают ни к какому объекту в действительности

СЕБЯ

*представлять² себе, представлять² собой, вести³ себя, чувствовать себя [как], приходить в себя, приводить¹ в себя, включать¹ в себя, уходить¹ в себя, выходить¹ из себя, выводить⁴ из себя, строить из себя, содержать¹ в себе, нести¹ в себе, представлять¹ из себя, отказывать¹ себе, замыкаться² в себе, брать, забирать, нести¹, приносить, уносить, везти¹, привозить, увозить¹ с собой, брать на себя, напускать на себя, находить¹ себя, терять себя, изживать себя, посвящать¹ себя, усилие над собой, так себе, ничего² себе, не по себе, не в себе, само собой, сам по себе, себе на уме, глаголы речи и мыслительной деятельности с сочетанием *про себя* 'не вслух', зависимость от глаголов по ограничительному отношению (*А он спит себе*), от прилагательных эстетической оценки (*хорош/нехорош/красив/некрасив собой*).*

СВОЙ

в свою очередь, в свое время, в своем роде, в своем уме, своего рода, в большинстве своем, со своей стороны.

Работа с установленной анафорической связью на этапе русско-семантического перевода

- возвратное местоимение удаляется,
- синтаксическое отношение, ведущее к местоимению СЕБЯ, перенаправляется к его антецеденту,
- синтаксическое отношение ОПРЕД, ведущее к местоимению СВОЙ, переводится по словарному правилу, которое опирается на следующую иерархию синтаксических отношений, на которые может быть заменена эта связь:

1) существительное — хозяин местоимения СВОЙ управляет род. п. или тв. п. по 1-й валентности (*своя победа, свой дом, свое звание генерала, свой анализ проблемы / анализ проблемы Ивановым*) → ОПРЕД заменяется на КВАЗИАГЕНТ или АГЕНТ:

Перевод на семантический язык предложения *Я не вижу свой нос.*

Translation

```
hasObject (Negation_1_1, Seeing_1_1)
hasExperiencer (Seeing_1_1, UtteranceSpeaker_1_1)
hasStimulus (Seeing_1_1, Nose_1_1)
isPartOf (Nose_1_1, UtteranceSpeaker_1_1)
```

Семантический перевод предложения
Вася вспомнил о своей встрече с Машей

Translation

```
hasGivenName (Human_1_1, "Vasya"^^xsd:string)  
hasAboutness (Remembering_1_1, Meeting_1_1)  
hasExperiencer (Remembering_1_1, Human_1_1)  
hasAgent (Meeting_1_1, Human_1_1)  
hasCoagent (Meeting_1_1, Human_1_2)  
hasGender (Human_1_2, Female_1_1)  
hasGivenName (Human_1_2, "Masha"^^xsd:string)
```

2) существительное — хозяин местоимения СВОЙ управляет род. п. и тв. п. по 1-й валентности и управляет род. п. по 2-й валентности (*свое обучение / обучение сотрудников*) →

ОПРЕД заменяется на 1-КОМПЛ

Начало перевода предложения *Иван закончил свое обучение*

```
X: 3.2 СВОЙ
SEM-CONV1.D0.2      ЗАМЕНА НА АНТЕЦЕДЕНТ
***** Удаление пустого гиперузла :COREF-GROUP-01 *****
```

Гипотезы 3

2.1 ЗАКАНЧИВАТЬ	1.1 ИВАН
	3.2 ОБУЧЕНИЕ
3.2 ОБУЧЕНИЕ	1.1 ИВАН

ПРЕДИК
1-КОМПЛ
1-КОМПЛ

3) существительное — хозяин местоимения СВОЙ не управляет род. п. ни по 1-й, ни по 2-й валентности (*свой стол*) →

ОПРЕД заменяется на АТРИБ

Начало перевода предложения *Иван сел за свой стол.*

X: 4.2 СВОЙ

SEM-CONV1.DO.3 ЗАМЕНА НА АНТЕЦЕДЕНТ

***** Удаление пустого гиперузла :COREF-GROUP-01 *****

Гипотезы 4

2.1 САДИТЬСЯ	1.1 ИВАН	▲	ПРЕДИК
	3.1 ЗА1	▲▲	1-КОМПЛ
3.1 ЗА1	4.2 СТОЛ	▲▲	ПРЕДЛ
4.2 СТОЛ	1.1 ИВАН		АТРИБ

На этапе перевода русской поверхностно-синтаксической структуры в семантическую структуру:

1. восстановление синтаксических «нулей» — невыраженных субъектов инфинитивных, причастных, деепричастных оборотов и придаточных предложений, установление antecedентов относительных местоимений (см. [Маракасова 2016]):

1) у финитных сентенциальных актантов и сирконстантов [Shushurin 2016], а также матричных предикатов:

Нарушитель не подозревал, что [стражи порядка], садились в машину, чтобы PRO; отправиться на его поиски.

SEM-conv1.02.10 ВОССТАНОВЛЕНИЕ СУБЪЕКТА В ПРИЧАСТНЫХ, ДЕЕПРИЧАСТНЫХ, ИНФИНИТИВНЫХ ОБО:
X: 10.1 ОТПРАВЛЯТЬСЯ

Гипотезы 13

3.1 ПОДОЗРЕВАТЬ	1.1 НАРУШИТЕЛЬ	▲▲	ПРЕДИК
	2.1 НЕ1		ОГРАНИЧ
	6.2 САДИТЬСЯ		1-КОМПЛ
4.1 СТРАЖ	5.2 ПОРЯДОК	▲▲	1-КОМПЛ
6.2 САДИТЬСЯ	4.1 СТРАЖ	▲	ПРЕДИК
	7.14 В1	▲▲	1-КОМПЛ
	9.1 ЧТОБЫ		ОБСТ
7.14 В1	8.4 МАШИНА	▲▲	ПРЕДЛ
9.1 ЧТОБЫ	10.1 ОТПРАВЛЯТЬСЯ		ИНФ-СОЮЗН
10.1 ОТПРАВЛЯТЬСЯ	4.1 СТРАЖ		ПРЕДИК
	11.1 IN1	▲	2-КОМПЛ
11.1 IN1	13.2 ПОИСК	▲	ПРЕДЛ
13.2 ПОИСК	12.1 ОН		1-КОМПЛ

Поскольку он ; считал, что гора Хайыракан является одним из символов шаманизма, то в своем завещании PRO ; попросил, чтобы его ; прах развеяли над этой горой.

SEM-conv1.02.1

ВОССТАНОВЛЕНИЕ СУБЪЕКТА В ПРИЧАСТНЫХ, Деепричастных, инфинитивных оборотах, придаточных

X: 15.1 ПРОСИТЬ1

Гипотезы 21

1.1 ПОСКОЛЬКУ	3.1 СЧИТАТЬ2	ПОДЧ-СОЮЗН
	11.1 ТО2	СООТНОС
3.1 СЧИТАТЬ2	2.1 ОН	ПРЕДИК
	6.1 ЯВЛЯТЬСЯ1	1-КОМПЛ
4.1 ГОРА	5.1 ФРМ-ЗАГЛАВН(Хайыракан)	КВАЗИАГЕНТ
6.1 ЯВЛЯТЬСЯ1	4.1 ГОРА	ПРЕДИК
	7.1 ОДИН1	1-КОМПЛ
7.1 ОДИН1	8.1 ИЗ	ЭЛЕКТИВ
8.1 ИЗ	9.1 СИМВОЛ	ПРЕДЛ
9.1 СИМВОЛ	10.1 ШАМАНИЗМ	1-КОМПЛ
12.1 В2	14.1 ЗАВЕЩАНИЕ	ПРЕДЛ
14.1 ЗАВЕЩАНИЕ	13.1 СВОЙ	ОПРЕД
15.1 ПРОСИТЬ1	1.1 ПОСКОЛЬКУ	ОБСТ
	2.1 ОН	ПРЕДИК
	12.1 В2	ОБСТ
	18.1 РАЗВЕИВАТЬ	2-КОМПЛ
17.1 ПРАХ	16.1 ОН	КВАЗИАГЕНТ
18.1 РАЗВЕИВАТЬ	17.1 ПРАХ	1-КОМПЛ
	19.1 НАД	2-КОМПЛ
19.1 НАД	21.1 ГОРА	ПРЕДЛ
21.1 ГОРА	20.1 ЭТОТ	ОПРЕД

2) при сочинительном и сравнительном сокращении:

Попыталась ей объяснить, она ; мало что поняла, но PRO ; посмеялась.

SEM-conv1.02.3 ВОССТАНОВЛЕНИЕ СУБЪЕКТА В ПРИЧАСТНЫХ, ДЕЕПРИЧАСТНЫХ, ИНФИНИТИВНЫХ ОБОРОТА
X: 9.1 ПОСМЕЯТЬСЯ

Гипотезы 11

1.1 ПЫТАТЬСЯ	3.1 ОБЪЯСНЯТЬ	▲	1-КОМПЛ
3.1 ОБЪЯСНЯТЬ	2.1 ОНА	▲	2-КОМПЛ
	7.1 ПОНИМАТЬ		1-КОМПЛ
6.3 ЧТОЗ	5.2 МАЛО		ВСПОМ
7.1 ПОНИМАТЬ	4.1 ОНА	▲	ПРЕДИК
	6.3 ЧТОЗ	▲	1-КОМПЛ
	8.1 НО		СОЧИН
8.1 НО	9.1 ПОСМЕЯТЬСЯ		СОЧ-СОЮЗН
9.1 ПОСМЕЯТЬСЯ	4.1 ОНА		ПРЕДИК
10.1 SetAnd	7.1 ПОНИМАТЬ		hasElement
	9.1 ПОСМЕЯТЬСЯ		hasElement

3) в причастных и деепричастных оборотах:

Спектакль, PRO ; решённый в стилистике условного театра, получился неожиданным.

SEM-conv1.02.6 ВОССТАНОВЛЕНИЕ СУБЪЕКТА В ПРИЧАСТНЫХ, ДЕЕПРИЧАСТНЫХ, ИНФИНИТИВНЫХ ()
X: 2.1 РЕШАТЬ1

Гипотезы 7

2.1 РЕШАТЬ1	1.1 СПЕКТАКЛЬ		1-КОМПЛ
	3.15 В2		ОБСТ
3.15 В2	4.2 СТИЛИСТИКА	▲▲	ПРЕДЛ
4.2 СТИЛИСТИКА	6.1 ТЕАТР	▲	1-КОМПЛ
6.1 ТЕАТР	5.1 УСЛОВНЫЙ	▲▲	ОПРЕД
7.1 ПОЛУЧАТЬСЯ	1.1 СПЕКТАКЛЬ	▲	ПРЕДИК
	8.1 НЕОЖИДАННЫЙ	▲	1-КОМПЛ

4) в инфинитивных оборотах:

По словам генерального менеджера Dimension Data Стива Росса, задача курса — научить его слушателей, " PRO ; думать, как хакеры".

SEM-CONV1.02.11
X: 15.1 ДУМАТЬ

ВОССТАНОВЛЕНИЕ СУБЪЕКТА В ПРИЧАСТНЫХ, Деепричастных, инфинитивных оборотах, ПРИДАТОК

Гипотезы 17

1.1 ПО1	2.1 СЛОВО	▲▲▲	ПРЕДЛ
2.1 СЛОВО	4.1 МЕНЕДЖЕР	▲	КВАЗИАГЕНТ
4.1 МЕНЕДЖЕР	3.1 ГЕНЕРАЛЬНЫЙ	▲▲	ОПРЕД
	5.1 ФИКТ-ЛИЧИМЯ(Dimension)		АППОЗ
	7.1 СТИВ1		АППОЗ
5.1 ФИКТ-ЛИЧИМЯ(Dimension)	6.1 ФИКТ-ФАМИЛИЯ(Data)	▲	АППОЗ
7.1 СТИВ1	8.1 РОСС	▲	АППОЗ
9.1 ЗАДАЧА	10.1 КУРС1	▲	1-КОМПЛ
11.1 БЫТЬ	1.1 ПО1	▲	ВВОДН
	9.1 ЗАДАЧА		ПРИСВЯЗ
	12.1 НАУЧАТЬ	▲	ПРЕДИК
12.1 НАУЧАТЬ	14.2 СЛУШАТЕЛЬ	▲	1-КОМПЛ
	15.1 ДУМАТЬ		2-КОМПЛ
14.2 СЛУШАТЕЛЬ	13.1 ОН		1-КОМПЛ
15.1 ДУМАТЬ	14.2 СЛУШАТЕЛЬ		ПРЕДИК
	16.1 КАК1		СРАВНИТ
16.1 КАК1	17.1 ХАКЕР	▲	СРАВН-СОЮЗН

5) у событийных существительных – аргументов предикатов с объектным контролем (признаками ОБ-ИНФ/ОБ-СУЩ):

Я бы простила авторам ; PRO ; недоработку сценария.

SEM-CONV1.02.15 ВОССТАНОВЛЕНИЕ СУБЪЕКТА В ПРИЧАСТНЫХ, ДБЕПРИЧАСТНЫХ, ИНФИНИТИВНЫХ ОБОРОТ
X: 4.1 НЕДОРАБОТКА

Гипотезы 5

2.1 ПРОЩАТЬ	1.2 Я		ПРЕДИК
	3.1 АВТОР		2-КОМПЛ
	4.1 НЕДОРАБОТКА		1-КОМПЛ
4.1 НЕДОРАБОТКА	3.1 АВТОР		КВАЗИАГЕНТ
	5.1 СЦЕНАРИЙ		1-КОМПЛ

Их ; обвиняют в PRO ; нарушении общественного порядка.

SEM-CONV1.02.14 ВОССТАНОВЛЕНИЕ СУБЪЕКТА В ПРИЧАСТНЫХ, ДБЕПРИЧАСТНЫХ, ИНФИНИТИВНЫХ ОБОРОТ
X: 3.1 НАРУШЕНИЕ

Гипотезы 5

2.1 ОБВИНЯТЬ	1.2 ОНИ	▲	1-КОМПЛ
	3.1 НАРУШЕНИЕ	▲	2-КОМПЛ
3.1 НАРУШЕНИЕ	1.2 ОНИ		АГЕНТ
	5.2 ПОРЯДОК	▲	1-КОМПЛ
5.2 ПОРЯДОК	4.1 ОБЩЕСТВЕННЫЙ	▲▲	ОПРЕД

- 6) у событийных существительных – аргументов лексических функций с субъектным контролем (_OPER1, _INCEPOPER1, _FINOPER1, _OPER1-NS, _SO_OPER1, _REAL1-M, _INCEPREAL1-M, _FINREAL1-M, _ANTIREAL1-M, _MANIF, LABOR1-2, _INCEPLABOR1-2, _FINLABOR1-2)

Также восстанавливаются нулевые дополнения в таких же контекстах (восстанавливается отношение 1-КОМПЛ/2-КОМПЛ):

Только [Красный Крест]_i выразил Сахарову_j PRO_i благодарность OBJ_j.

SEM-conv1.02.18 ВОССТАНОВЛЕНИЕ СУБЪЕКТА В ПРИЧАСТНЫХ, Деепричастных, инфинитивных OI
X: 6.2 БЛАГОДАРНОСТЬ

Гипотезы 7

2.1 КРАСНЫЙ	1.1 ТОЛЬКО1		ОГРАНИЧ
3.1 КРЕСТ	2.1 КРАСНЫЙ	▲▲	ОПРЕД
4.1 ВЫРАЖАТЬ	3.1 КРЕСТ	▲▲	ПРЕДИК
	5.1 САХАРОВ	▲	2-КОМПЛ
	6.2 БЛАГОДАРНОСТЬ	▲▲	1-КОМПЛ
6.2 БЛАГОДАРНОСТЬ	3.1 КРЕСТ		КВАЗИАГЕНТ
	5.1 САХАРОВ		1-КОМПЛ

7) у событийных существительных – аргументов адвербиальной лексической функции _ADV1-UN, а также у адвербиалов с предлогами С,БЕЗ,ОТ,ПРИ

Всё это PRO ; замирающая от PRO ; страха Маргарита ; разглядела в коварных тенях от свечей кое-как.

SEM-CONV1.02.20
X: 5.1 СТРАХ

ВОССТАНОВЛЕНИЕ СУБЪЕКТА В ПРИЧАСТНЫХ, Деепричастных, инфинитивных оборота

Гипотезы 12

2.2 ЭТО	1.2 ВЕСЬ		ОПРЕД
3.1 ЗАМИРАТЬ	4.1 ОТ		ОБСТ
	6.1 МАРГАРИТА		ПРЕДЖК
4.1 ОТ	5.1 СТРАХ	▲▲	ПРЕДЛ
5.1 СТРАХ	6.1 МАРГАРИТА		КВАЗИАГЕНТ
7.1 РАЗГЛЯДЕТЬ	2.2 ЭТО		1-КОМПЛ
	6.1 МАРГАРИТА	▲▲	ПРЕДЖК
	8.15 В2		ОБСТ
	12.1 КОЕ-КАК		ОБСТ
8.15 В2	10.1 ТЕНЬ1	▲▲	ПРЕДЛ
10.1 ТЕНЬ1	9.3 КОВАРНЫЙ	▲▲	ОПРЕД
	11.1 СВЕЧА	▲	1-КОМПЛ

2. работа правил, восстанавливающих разного рода имплицитные связи:

- 1) Установление квазиагентивной, агентивной, 1-й комплетивной или атрибутивной связи существительных с некоторыми типами внешних поссессоров
- 2) Канонизация ЛФ _OPER1, _OPER2 и _FUNC0
- 3) Построение предикативных связей между 2-м и 1-м дополнениями предикатов с признаком СОГЛАКТ-3-2
- 4) Установление квазиагентивной, агентивной, 1-й комплетивной или атрибутивной связи существительных с притяжательными прилагательными
- 5) Коррекция релятивных связей и установление антецедентов относительных местоимений [Маракасова 2016]

3. Установление анафорических связей возвратных местоимений с восстановленными синтаксическими «нулями»

За вчерашнее халатное PRO_i отношение к своим_i обязанностям ему_i (фамилия, имя, отчество прописными буквами) объявлялся строгий выговор.

SEM-CONV1.A4.1
X: 5.1 СВОЙ

УСТАНОВЛЕНИЕ АНТЕЦЕДЕНТНОЙ СВЯЗИ С ВОССТАНОВЛЕННЫМ ПОДЛЕЖАЩИМ В КОНСТ

Гипотезы 19

1.1 ЗА1	4.1 ОТНОШЕНИЕ2	ПРЕДЛ
4.1 ОТНОШЕНИЕ2	2.1 ВЧЕРАШНИЙ	ОПРЕД
	3.1 ХАЛАТНЫЙ	ОПРЕД
	6.1 ОБЯЗАННОСТЬ	1-КОМПЛ
	7.1 ОН	КВАЗИАГЕНТ
6.1 ОБЯЗАННОСТЬ	5.1 СВОЙ	ОПРЕД
	7.1 ОН	КВАЗИАГЕНТ
7.1 ОН	12.1 БУКВА	ПРИМЬКАТ
8.1 ФАМИЛИЯ	9.1 ИМЯ	СОЧИН
9.1 ИМЯ	10.1 ОТЧЕСТВО	СОЧИН
12.1 БУКВА	11.1 ПРОПИСНОЙ	ОПРЕД
13.1 ВЫГОВАРИВАТЬ1	1.1 ЗА1	2-КОМПЛ
	7.1 ОН	2-КОМПЛ
		1-КОМПЛ
	8.1 ФАМИЛИЯ	FICTIT
	14.1 СТРОГИЙ	ОБСТ
	15.1 ВЫГОВОР1	1-КОМПЛ

4. Установление именных групп, которые не могут быть antecedентами местоимений 3 лица (они получают признак NON-ANTEC)

Это именные группы, которые:

- Относятся к определенному классу лексем:
ВСЕ1, ВСЯ, ГЛАВНОЕ2, НЕКОГО, НЕЧЕГО1, НИЧТО, СЕЗ,
ТО1, ТО-ТО2, ЧТО2, ЧТО-ЛИБО, ЧТО-НИБУДЬ,
ЧТО\$БЫ\$ТО\$НИ\$БЫЛО, ЭТО
- зависят по отношениям:
 - ПРИСВЯЗ (*Петя был студентом*),
 - АППОЗ (*студент Петя*),
 - ОБ-АППОЗ (*Петя, студент МГУ*),
 - НУМ-АППОЗ (*дом номер восемь*),

*Дима_i жил в квартире номер_j пятнадцать, я хорошо его_{i/*j} помню.*

- КОМПЛ-АППОЗ (вес (в) сто килограммов),
- по 2-КОМПЛ от слова с синтаксическими признаками СОГЛАКТ-3-1, СОГЛАКТ-3-2 или СОГЛАКТ-2-1 (репутация Петра как умного человека; признание Мадагаскара независимым государством).

*Миша_i притворялся дурачком_j, чтобы к нему_{i/*j} не приставали с вопросами.*

Исключение — зависящие по присвязочному отношению местоимения 3 лица в твор. п., которые могут отсылать к существительным с предикатным статусом (*Ивана принимали за директора_i, когда он им_i еще не был*);

- сами управляют чем-то по отношению ПРЕДИК (*Петя студент*):

*Если Маша_i — высокая блондинка_j, значит, та девушка — не она_{i/*j}.*

- Зависят по предикативному отношению (т.е. стоят в позиции подлежащего), но фактически являются именными сказуемыми (см. критерии, предложенные в [Падучева, Успенский])

*Но главное его богатство_j - вода_i. Она_i.../Оно*_j...*

*И главная причина_j - именно недостатки_i в работе с кадрами. Они_i.../Она*_j...*

*Десятая часть_j рабочей силы страны - ремонтники_i. Они_i... / Она*_j...*

Но ср. пример анафоры к предикату:

Микросекунды - это, в общем, довольно уже далеко отстоящий от обычной повседневной жизни диапазон_j времен, но, тем не менее, его_i можно наблюдать и с помощью обычных видеокамер.

- относятся к классу Event (события) или AbstractObject (например, количества и временные интервалы) и при этом зависит от предлога, который является его значением лексической функции *_ADV1-UN* (по аналогии, на бегу, в аспекте), *_ADV2-UN* (по ошибке, под влиянием, под руководством) или *_LOC* (в апреле, на ветру)

*Если авария_i произойдет с вероятностью_j $1/100000$, то никто ее_{i/*j} не боится.*

- относятся к классу Object (материальные объекты) и зависят от предлога, который является их значением лексической функции $_LOC$ и зависит по обстоятельству отношению.

*В деревне_i Ольга_j рано вставала и работала в саду, а вечером она*_{i/j} ходила в гости к соседям.*

*В институте_i Андрей_j изучал право, а работает он*_{i/j} продавцом.*

При этом существительное X не управляет указательными и относительными местоимениями и элективными прилагательными или номером по отношению НУМ-АППОЗ (*в этом институте; в институте, который он окончил; в доме номер восемь*);

- относятся к классу Event и зависят от предлога, который близок к значению лексической функции `_ADV1-UN` или `_ADV2-UN`, но не указан в этом качестве в их словарных статьях (*при наличии, с благодарностью, в присутствии, на вкус, по привычке, от нетерпения, по вине, без сожаления, с непривычки, за неимением, на радость*):

*В ответ_i Коля_j сказал, что он^{*_{i/j}} уходит.*

- относятся к классу AbstractObject и зависят от предлога, который близок к их значению лексической функции _ADV1-UN или _ADV2-UN, но не указан в этом качестве в словарной статье существительного (*в количестве, за/на время, с января по май, со временем, на днях*), при этом существительное не управляет указательными, притяжательными и относительными местоимениями или существительными по отношению КВАЗИАГЕНТ (*в этом количестве, в вашем присутствии, с его благодарностью, с благодарностью Пети*):

*Машина_i ехала на большой скорости_j, и мы не разглядели, кто ее_{i/*j} ведет.*

- зависят по отношениям:

- ОБСТ (*Они шли дворами*),

*Осенью_i Таня_j уехала, и Коле было одиноко без нее^{*_{i/j}}.*

- ДЛИТЕЛЬН (*Весь день он работал*),

- КРАТНО-ДЛИТЕЛЬН (*Два раза они возвращались*),

- ДИСТАНЦ (*идти два километра*),

- ОБ-ОБСТ (*пить водку стаканами*),

- АТРИБ в качестве твор. п. способа (*плавать брассом*).

Исключения — случаи, когда непосредственно справа от существительного — предложная группа *из них*, а справа от нее — числительное (*проехать на лыжах 982 км_i (из них_i 285 км с малокалиберным карабином за спиной)*);

- являются количественными существительными с признаком ЧИСЛ-ОГР — *десяток, сотня, тысяча* и т.д. (*Призовой фонд равен миллиону*) и при этом не зависят от предиката из класса ChangeOfPossession (*Вася дал Пете тысячу_i рублей, и он ее_i потратил*) или не зависят от предиката с дескриптором 'ДЕЙСТВИЕ' и не управляют существительным из класса денежных единиц (*Вася потерял тысячу_i рублей, а Петя нашел ее/их_i*):

*Книга_i стоила тысячу_j рублей, так что я решила не покупать ее_{i/*j}.*

- являются именами параметров и зависят от существительных по атрибутивной связи (*дети в возрасте от пяти лет, книги по цене от 100 рублей*);

*Там была стена_i высотой_j 10 метров, и на нее_{i/*j} можно было залезть.*

- являются именами единиц измерения и зависят по отношению 1-КОМПЛ от существительного или глагола со значением параметра (*вес в сто тонн, стоит пятьсот долларов*), предикатов *составлять¹, достигать, равняться¹, равный, длиться, вмещать*;

- являются именами единиц измерения и зависят от предлогов из списка: *в1, в2, до1, за1, к, на1, от, около, порядка, по1, под1, при, с1, с2, через, назад1, спустя, тому\$назад*, которые сами зависят по отношению ОБСТ или АППОЗ (*в 2 метрах, до 6 часов, за час, к 5 часам, на 20 лет, порядка 50 %, по 2 штуки, под 90 лет, при 10 градусах, с 2 раз, человек с 10, 5 дней спустя, через год*), или зависит от предлога *в1* или *на1*, который зависит от прилагательного или наречия в сравнительной степени (*в три раза меньше*), или являются единицами измерения времени и зависят по отношению АТРИБ (*человек лет сорока*):

*Вчера зрителей_i было на десять человек_j больше, и они_{i/*j} слушали внимательно.*

- непосредственно или опосредованно зависят по отношению ВВОДН: *Проблема, как считает Иван, существует; Коля, закройте дверь*).

- включены в состав фразеологических/микросинтаксических единиц (их полный список не установлен):
Дело в том, что..., Это другое дело; Какая разница?; не может быть и речи; о том и речь; что толку, если...; слава богу, черт возьми, курам на смех, Но ведь ржавчина ржавчине рознь
и многие другие.

5. Установление антецедентов и/или множества кандидатов в антецеденты местоимений 3 лица в пределах одного или нескольких предложений

Типы антецедентов, которые обрабатываются правилом

на данный момент — только существительные и элективные группы, не связанные сочинительной связью и находящиеся слева от местоимения

Правило включает:

- 1) список лемм местоимений, к которым оно применяется (*он, она, оно, они, него1, него2, нее, них*);
- 2) зону поиска существительного Z — кандидата на роль контролера местоимения: слева от него по всему предложению,
- 3) его морфологическую категорию:
 - существительное (кроме лексем Я, МЫ, ТЫ, ВЫ, СЕБЯ, КОТОРЫЙ),
 - прилагательное с признаком ЭЛЕКТ (*один/трое из мальчиков*),
- 4) условие, что существительное Z не имеет признаков POS-АНТЕС, NON-АНТЕС;

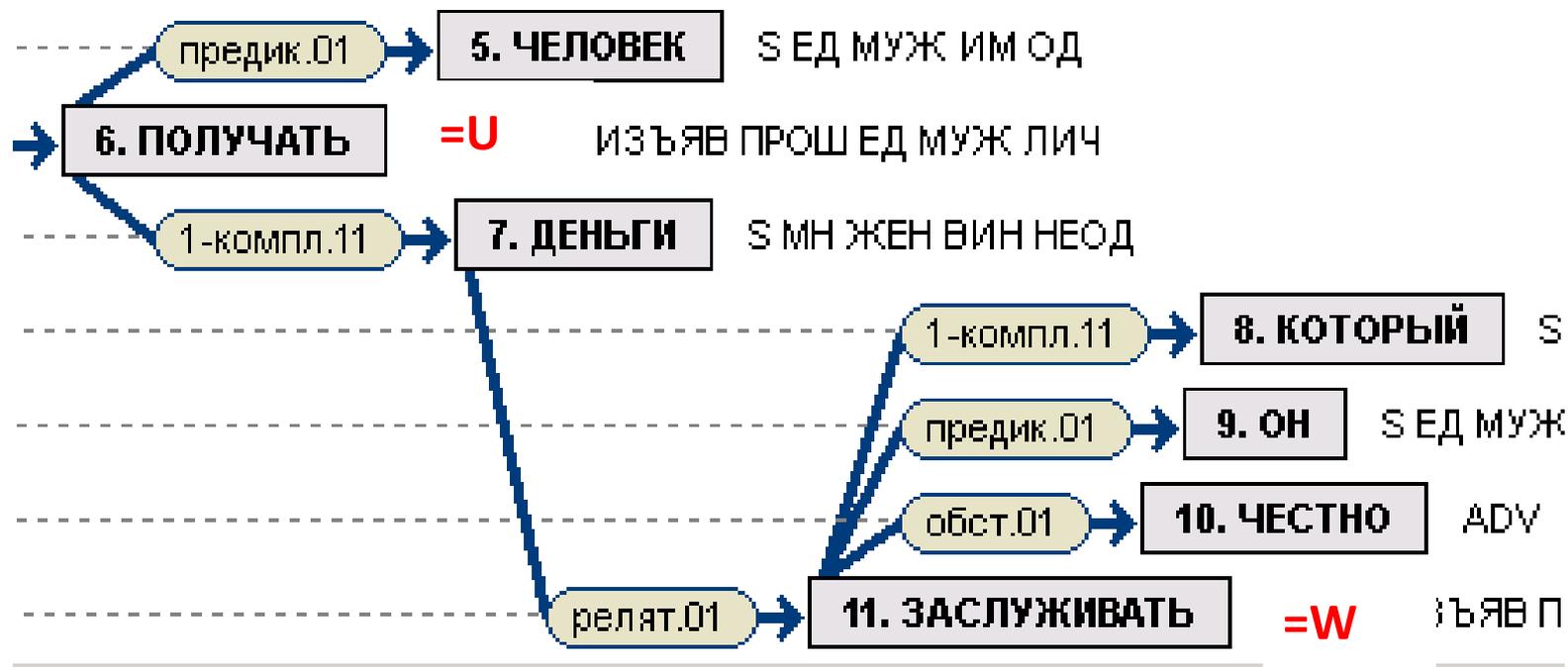
5) морфологические характеристики местоимения и его потенциального контролера:

- либо они согласуются по роду и числу;
- либо местоимение *они/них* контролируется количественной группой с числительным *полтора / два / три / четыре*, и признаком ед. ч. у существительного (*два мальчика*);
- либо местоимение — *он* в косв. падеже или *него¹*, а существительное — среднего рода (*государство; в лице принадлежащих; ему монополий*)

6) путь по синтаксической структуре от местоимения к предикату *W* — глаголу, существительному, прилагательному или наречию, от которого оно зависит (зависимость непосредственная, через предлог и/или элективную конструкцию (*один из них*));

7) путь по синтаксической структуре от существительного/прилагательного Z (кандидата на роль контролера местоимения) к предикату U — глаголу, существительному, прилагательному или наречию, от которого оно зависит (зависимость непосредственная, через предлог и/или количественную конструкцию (*много людей, у десятка человек*), а также редкие контексты, в которых Z зависит от того же предиката W, что и местоимение, которое в таких случаях управляет интенсификатором *сам* (*Иван говорил с Машей; о ней; самой*);

То, что предикаты U и W, за редкими исключениями, не совпадают, обеспечивает соблюдение принципа В теории связывания в большинстве случаев (отсутствие кореферентной связи в контекстах типа *Саша_i видит / хочет видеть его_{j/*i}, рассказ Саша_i о нем_{j/*i}*).



VS.



8) проверку семантического согласования по модели управления для предиката W и потенциального антецедента местоимения (Z). Например, если местоимение зависит от W по квазиагентивной связи, то нужно проверить, что Z тоже может занимать эту позицию, т.е. что в 1-м столбце модели управления предиката W содержится хотя бы один из тех семантических дескрипторов или онтологических концептов, которые имеются у Z в зоне дескрипторов:

*Там сначала, наверное, страниц на тридцать следовал
трактат о том, что Дарвин с его теорией
естественного отбора был неправ.*

```
29:36 23-07-2015          ТЕОРИЯ
POR: S      АЧ
SYNT: ЖЕНСК, ИСЧИСЛ, АГЕНС
DES: 'ИНФОРМАЦИЯ', 'ФАКТ', 'АВСТРАКТ'
SEM-ONTO: Proposition
D1.1: РОД, 'ЛИЦО'
D2.1: РОД, 'ФАКТ', 'ВЕЩЕСТВО', 'СИСТЕМА', 'АВСТРАКТ', 'МЕХА
D2.2: O2
D2.3: ЧТО1
D2.3: OU    //ТЕОРИЯ. КАК ПОСТУПАТЬ В ТАКИХ СЛУЧАЯХ
```

```
4:02 07-05-2015          ТРАКТАТ
POR: S      НФ
SYNT: МУЖСК, ИСЧИСЛ, АГЕНС
DES: 'ИНФОРМАЦИЯ', 'АВСТРАКТ', 'ФАКТ', 'ПРЕДМЕТ'
SEM-ONTO: Book/TextDocument
D1.1: РОД
D2.1: O2
D2.2: ПРО1
D2.3: НА+ТЕМУ
```

После того, как сработало правило:

Фраза

1.1	ТАМ	ADV, CAP
2.1	СНАЧАЛА [ЗПТ]	ADV
3.1	НАВЕРНОЕ [ЗПТ]	ADV
4.1	СТРАНИЦА	S, РОД, МН, ЖЕН, НЕОД
5.1	НА1	PR
6.2	ТРИДЦАТЬ	NUM, ВИН
7.1	СЛЕДОВАТЬ1	V, ПРОШ, ЕД, ЛИЧ, ИЗЪЯВ, МУЖ, НЕОД
8.1	ТРАКТАТ	S, ИМ, ЕД, МУЖ, НЕОД
9.1	ТО1 [ЗПТ]	S, ПР, ЕД, СРЕД, НЕОД, PREP
10.1	ЧТО1	CONJ
11.1	ДАРВИН1	S, ИМ, ЕД, МУЖ, ОД, POS-ANTEC, C
12.16	СЗ	PR
13.1	ОН	S, РОД, ЕД, МУЖ, ОД
14.1	ТЕОРИЯ	S, ТВОР, ЕД, ЖЕН, НЕОД, NON-ANTEC
15.1	ЕСТЕСТВЕННЫЙ	A, РОД, ЕД, МУЖ
16.1	ОТБОР	S, РОД, ЕД, МУЖ, НЕОД
17.1	НЕПРАВ1	A, ПРОШ, ЕД, ЛИЧ, ИЗЪЯВ, КР, МУЖ
18.1	TimeInterval	
19.1	SpeechTimeInterval	
20.1	TimeInterval	
21.1	SpeechTimeInterval [ТЧК]	
-----Hypertext:-----		
22.1	FICT-HN (gr: COREF-GROUP-01)	

Групп кореферентности - 1

Group info

22 :COREF-GROUP-01

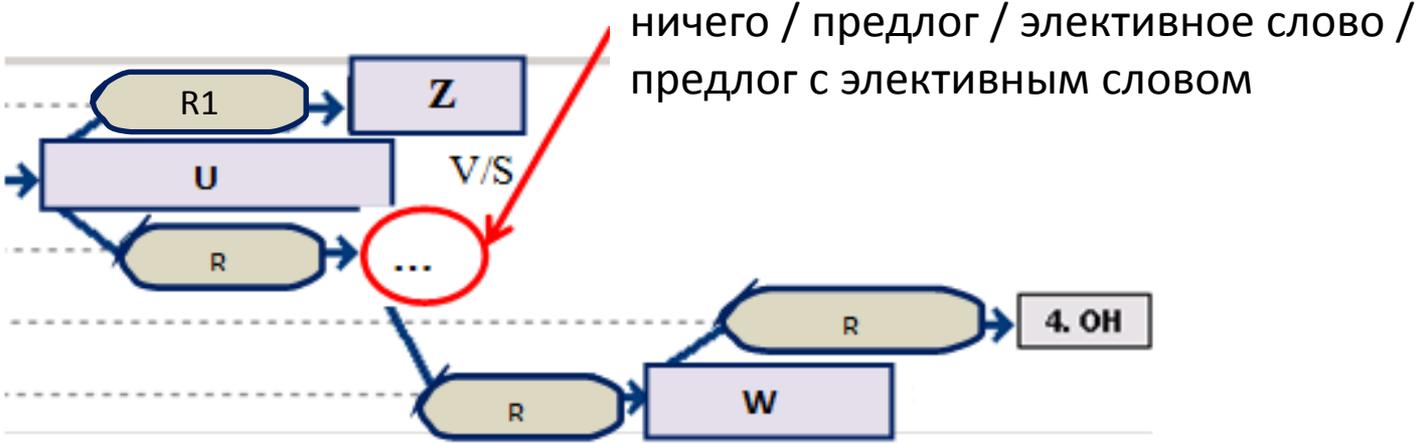
11.1 ДАРВИН1,

13.1 ОН

9) запрет на анафорическую связь в конструкции, где местоимение зависит от существительного W, его контролер (Z) является подлежащим клаузы или субъекта ИГ и при этом W и Z зависят от одного и того же глагольного или именного предиката U (непосредственно, через предлог и/или элективную конструкцию):

*Петя_i рассказал[U] некоторым из его_{j/*i} знакомых[W] о его_{j/*i} работе[W];*

*Петин_i разговор[U] с его_{j/*i} начальником[W]);*



R = любое синтаксическое отношение

R1 = отношение ПРЕДИК / ДАТ/СУБЪЕКТ / КВАЗИАГЕНТ / АГЕНТ

10) запрет на контроль местоимения одним из элементов сочиненной группы:

*Петя_i и Вася_j гуляли с его_{k/*i/*j} собакой;*

Исключения — контексты, где образуются пары антецедентов и местоимений: *Я долго смотрел, как удалялись физик_i и Таня_j, в конце улицы под фонарём он_i взял её_j под руку.*

11) запрет на то, чтобы предикат U управлял существительным W и при этом был значением его лексической функции с субъектным контролем существительного W :

*Петя_i получил_{_OPER2} от него_{j/*i} выговор.*

12) Запрет на то, чтобы местоимение и
антецедент находились по разные стороны
границы прямой речи

*«Петя_i скоро придет», – сказал он*_i*

Данное правило приписывает
существительному, для которого выполняются
все эти условия, признак POS-АНТЕС и
связывает с ним местоимение кореферентной
связью.

*В результате, к примеру, утверждение "Мы
проголосовали за политика X, потому что
он хороший" замещается нами в ходе
избавления от когнитивного диссонанса
другим...*

1.15	B2	PR, CAP
2.1	РЕЗУЛЬТАТ (ЗПТ)	S, PR, ЕД, МУЖ, НЕОД, NON-ANTEC
3.1	К	PR
4.1	ПРИМЕР (ЗПТ)	S, ДАТ, ЕД, МУЖ, НЕОД, NON-ANTEC
5.1	УТВЕРЖДЕНИЕ 1 (КАВ)	S, ИМ, ЕД, СРЕД, НЕОД
6.1	МЫ	S, ИМ, МН, ОД, IN-QUOT, CAPIT, CA
7.1	ГОЛОСОВАТЬ	V, ПРОШ, МН, ЛМЧ, МЗЪЯВ, СОВ, IN-
8.1	ЗА1	PR, IN-QUOT
9.2	ПОЛИТИК	S, ВМН, ЕД, МУЖ, ОД, IN-QUOT, POS
10.1	ФРИ (X) (ЗПТ)	S, IN-QUOT, NON-ANTEC CAP
11.1	ПОТОМУЧТО	CONJ, IN-QUOT
12.1	БМТЬ	V, НАСТ, ЛМЧ, МЗЪЯВ, НЕСОВ
13.1	ОН	S, ИМ, ЕД, МУЖ, ОД, IN-QUOT
14.2	ХОРОШОЙ (КАВ)	A, ВМН, ЕД, МУЖ, НЕОД, IN-QUOT
15.1	ЗАМЕЩАТЬ	V, НЕПРОШ, ЕД, ЛМЧ, МЗЪЯВ, 3-Л, Н
16.1	МЫ	S, ТВОР, МН, ОД
17.1	ВХОДЕ	PR
18.1	ИЗЪЯВЛЕНИЕ	S, РОД, ЕД, СРЕД, НЕОД, NON-ANTEC
19.1	КОГНИТИВНЫЙ	A, РОД, ЕД, МУЖ
20.1	ДИССОНАНС	S, РОД, ЕД, МУЖ, НЕОД, PREP
21.1	ДРУГОЙ	A, ТВОР, ЕД, МУЖ
22.1	TimeInterval	
23.1	SpeechTimeInterval	
24.1	TimeInterval	
25.1	SpeechTimeInterval	
26.1	TimeInterval	
27.1	SpeechTimeInterval (ТЧК)	

-----Hypernodes:-----

28.1 FICT-HN (gr: COREF-GROUP-01)

Группа кореферентности - 1
 Group info
 28 :COREF-GROUP-01
 9.2 ПОЛИТИК,
 13.1 ОН

- Данное правило должно применяться циклически к одному и тому же местоимению, пока не будут отображены все возможные кандидаты на роль его антецедента.

**6. Выбор из нескольких возможных
антецедентов (стирание неверных
коррелентных связей)**

1. Правила, использующие семантическую и онтологическую информацию

- 1) согласование по МУ, если местоимение зависит от нескольких предикатов

- 2) местоимение – субъект предиката, относящегося к классу «человеческих» действий и свойств – IntentionalProcess, SocialRole, SocialRelation и др., а потенциальный antecedent не относится к классу лиц – Agent или Human

Он доставал из-под террасы дрова, грохал возле печки, шуршал бумагой, разжигал, кряхтел, а ей не хотелось ничего, и она была не рада, что поехала с ним в этот раз.

```
Группы кореферентности - 2
  Group info
44 :COREF-GROUP-01
    4.1 ТЕРРАСА,
    8.1 ПЕЧКА,
    10.1 БУМАГА,
    14.1 ОНА,
    19.1 ОНА
45 :COREF-GROUP-02
    1.1 ОН,
    5.1 ДРОВА,
    24.1 НЕГО1

  Group info
```

```
SEM-COREF.C7.1      ПРОВЕРКА СОГЛАСОВАНИЯ ВОЗМ. АНТЕЦЕДЕНТОВ МЕСТ. З Л. С ОНТОКОРРЕЛЯТАМИ НЕКОТОРЫХ ПРЕДИКАТОВ, ОТ
X: 19.1      ОНА
```

```
Группы кореферентности - 2
  Group info
44 :COREF-GROUP-01
    14.1 ОНА,
    19.1 ОНА
```

2. Правила, использующие синтаксическую информацию

- 1) о конструкциях, в которых происходит поддержание или переключение референции

С позиции этого подхода очень важен учебник_i - он_i и несет в себе Знание.

Такая способность вовсе не делает попугая обладателем речи_i (даже в ее_i зачаточном состоянии).

Он даже позволил себе опасную фразу: "Если какая-либо точка мира может быть названа его (мира) центром, то это центр_i обращений небесных тел; а в нём_i, как известно всякому, кто разбирается в этих вопросах, находится Солнце, а не Земля".

- 2) о кореферентных группах, в которых нарушаются принципы связывания анафорических местоимений

- 3. Правила, использующие информацию о структуре кореферентной группы**
(например, что в ней была стерта какая-то кореферентная связь)

4. Правила, использующие дискурсивную информацию

1) роль протагониста в тексте или некотором его фрагменте

На данный момент определяется с ошибками, т.к. не устанавливаются кореф. связи между неместоименными ИГ

Протагонист – самый выделенный участник всего текста или его фрагмента, проходящий через предложения «из темы в тему»

Вермален_i вновь получил травму

Главный тренер "Арсенала" Арсен Венгер_j подчеркнул, что защитник Томас Вермален_i вновь получил повреждение и не сможет выйти на поле в ближайшее время. Бельгиец_i травмировался на тренировке "канониров", у него_{i/*j} снова проблемы с ахилловым сухожилием. <...>

Виктор Васин_i: «Не вижу себя в "Зените"»

Защитник нальчикского "Спартака" Виктор Васин_i, который недавно был вызван Диком Адвокатом_j в сборную России на товарищеский матч с бельгийцами, признался: приглашение в первую команду страны стало для него_{i/*j} сюрпризом. Наряду с этим 22-летний футболист_i в интервью пресс-службе южан подчеркнул: ему_{i/*j} нужно демонстрировать более стабильную игру.<...>

2) риторическое расстояние до антецедента

Это длина цепочки т.н. риторических отношений, последовательно связывающих клаузы (например «последовательность», «причина», «результат», «уступка» и др.) в которых упомянут один и тот же объект.

3) одушевленность

4) нахождение местоимения и существительного в пределах 1 абзаца

5) позиция подлежащего

6) семантическая роль субъекта

7) линейное расстояние, измеряемое в количестве клауз.

Дальнейшие задачи

- более качественное разрешение дискурсивной анафоры;
- разрешение более сложных случаев синтаксической анафоры:
 - местоимения *друг друга*,
 - установление антецедента местоимения ОНИ/НИХ, который расщеплен, т.е. представляет собой несколько именных групп (*Медвежонок_i прыгнул к Ёжику_j, они_{i+j} обхватили друг друга и так несколько мгновений стояли, трясясь; Пчёлы_i, как и кошки_j, не замечают красный цвет_k: он_k для них_{i+j} всё равно, что чёрный*), или является кванторной группой (*Понятно каждое слово в отдельности, но непонятно, что они означают в комбинации; И мы каждому_i помогаем развить свой_i потенциал*);
- создание корпуса с анафорической разметкой, чтобы на его материале провести оценку качества разрешения анафоры;
- восстановление эллипсиса (например, для предложений типа *Вася_j может поднять свой_j вес, а Петя_i не может поднять свой_i вес*).