

Язык и логика: ризонинг в задачах обработки текста

Иван Рыгаев

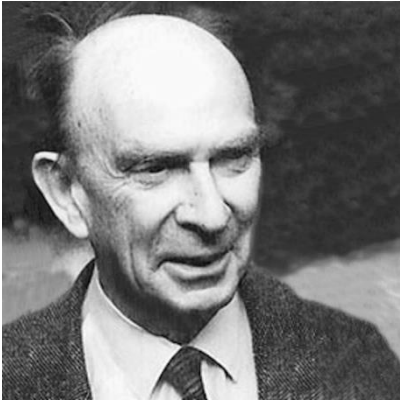
Лаборатория компьютерной лингвистики

Институт проблем передачи информации РАН

Притча

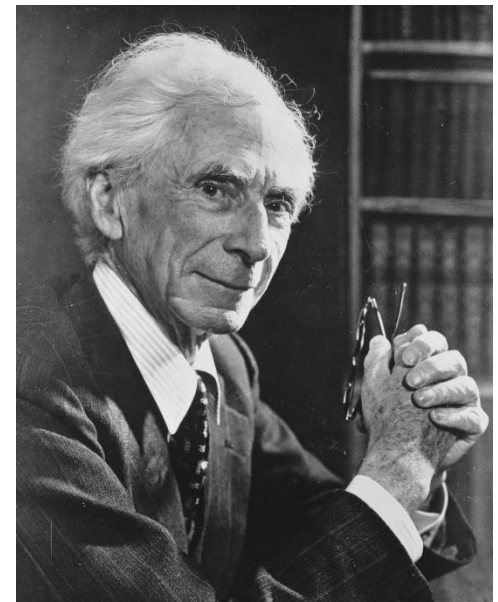


Стросон и Рассел



- Стросон, 1950:
 - *«Никакие правила ни Аристотеля, ни Рассела не могут предоставить чёткую логику для любого выражения обычного языка; ибо в обычном языке нет чёткой логики.»*

- Russell 1957:
 - *«Я абсолютно не в состоянии увидеть совсем никакой валидности ни в одном аргументе мистера Стросона. ... Однако я согласен с заявлением мистера Стросона, что в обычном языке нет чёткой логики.»*



Естественный язык и формальный

- [Джон] спит $\text{sleep}(\text{John})$
- Прямое соответствие между естественным языком и формальным
- [Все студенты] спят $\forall x[\text{student}(x) \rightarrow \text{sleep}(x)]$
- Какая часть высказывания справа соответствует именной группе «Все студенты»?
- Никакая.
- Вывод – естественный язык устроен нелогично.

Ричард Монтегю

- Монтегю занимался математикой и логикой и не участвовал в лингвистических войнах
- Но его раздражало, что столько шума происходит вокруг различия формальных и неформальных языков
- Поэтому однажды он занялся «довольно лёгкой и не очень важной» задачей – показать, что естественный язык можно рассматривать как формальный
- Ему это удалось, но это оказалось не так легко

Естественный язык и формальный

- [Все студенты] спят $\text{all_students}(\text{sleep})$
- $\text{all_students}(P) = \forall x[\text{student}(x) \rightarrow P(x)]$
- $\text{all_students}(\text{sleep}) = \forall x[\text{student}(x) \rightarrow \text{sleep}(x)]$
- Вывод – естественный язык допускает формальную интерпретацию.

- [Джон] спит $\text{John_property}(\text{sleep})$
- $\text{John_property}(P) = P(\text{John})$
- $\text{John_property}(\text{sleep}) = \text{sleep}(\text{John})$

Условия истинности и далее

- Значением предложения являются его условия истинности (как в математике)
- Интерпретация в модели:
 - John – индивид из множества дискурса
 - sleep – подмножество множества дискурса
- Однако:
 - Предложение – это больше, чем просто заявление об истинности высказывания
 - Предложение – это *сообщение, передача информации* от говорящего к слушающему

Снова язык и логика

- Логические операторы в естественном языке работают не так, как в формальной логике.
- Кванторы существования и всеобщности:
 - Некоторые студенты пришли (не все)
- Исключающее «или»:
 - У неё есть собака или кошка (но не собака и кошка)
- Двухсторонняя импликация:
 - Если пойдёт дождь, я возьму зонт (а если не пойдёт?)
- Некоммутативное «и»:
 - Она вышла замуж и родила ребёнка (а не наоборот)
- В языке другая логика (он нелогичен)?

Герберт Пол Грайс

- Paul Grice. Logic and Conversation (1975)
- Значения операторов в естественном языке не отличаются от их значения в формальной логике
- Но надо различать то, что сказано (семантику), и какие выводы были сделаны из сказанного (прагматику)
- Грайс ввёл понятие импликатуры

Принцип кооперации

- Диалог не является набором несвязанных друг с другом реплик
- Как правило участники совместно следуют некоторой цели или направлению развития диалога
- Основной принцип: вноси такой вклад в диалог, который требуется на данном этапе его развития в соответствии с его целью
- Принцип реализуется четырьмя типами максим

Максимы Грайса

- Максима количества:
 - Будь настолько информативным, насколько это нужно
 - Не будь более информативным, чем нужно
- Максима качества:
 - Не говори того, что ты считаешь ложным
 - Не говори того, для чего у тебя нет достаточных оснований

Максимы Грайса

- Максима релевантности:
 - Не отклоняйся от темы (будь уместным)
- Максима способа:
 - Избегай непонятных выражений
 - Избегай неоднозначности
 - Будь краток (избегай ненужного многословия)
 - Говори по порядку
- Максимы работают не только в общении, но и в любой кооперации

Коммуникативные импликатуры

- Коммуникативные импликатуры возникают из-за того, что мы ожидаем от собеседника следования принципу кооперации
- Они не являются логическим выводом из сказанного
- Они могут быть отменены последующим дискурсом

Примеры

- Максима количества
 - Некоторые студенты пришли (а на самом деле все)
 - У неё есть собака или кошка (а может и то, и другое)
- Максима способа
 - Она вышла замуж и родила ребёнка (но не в этом порядке)

Нарушения максим

- Говорящий может:
 - Незаметно нарушать максимы. Тогда он берёт на себя ответственность за «обман».
 - Явно отказаться следовать им. «Я не могу сказать больше».
 - Нарушать одну из-за желания следовать другой, когда они входят в противоречие.
 - Явным образом эксплуатировать нарушение максимы, таким образом создавая новую импликатуру.

Ещё примеры

- – У меня кончился бензин.
 - За углом есть заправка (импликатура: она открыта)
- – Где живёт Мари?
 - Где-то на юге Франции (противоречие максим качества и кол-ва)
- – На войне как на войне (эксплуатация максимы количества)
- – Джон – отличный друг (ирония, эксплуатация максимы качества).

Схемы Винограда

- Предложены Гектором Левеком в 2011 г.
- Приз не помещается в чемодан, потому что **он** слишком *большой*. Кто *большой*?
 - приз
 - чемодан
- Маша поблагодарила Катю за всю ту помощь, что **она** *оказала*. Кто *оказал* помощь?
 - Маша
 - Катя
- Терри Виноград предложил первый пример в 1970 г.

Структура схем Винограда

- Проблема разрешения анафоры
- В предложении два потенциальных антецедента
- Грамматически признаки, статистика коллокаций и ограничения на тип аргументов не помогают
- Замена специального слова меняет правильный ответ (*большой -> маленький*)
- Приз не помещается в чемодан, потому что **он** *слишком маленький*. Кто *маленький*?
 - приз
 - чемодан

Здравый смысл

- Люди хорошо справляются со схемами Винограда
- Тесты показывают 91-92% правильных ответов
- Что требуется для правильного ответа?
- Понимать значение глагола 'помешаться'
 - Если А помещается в В, то А должен быть меньше В
- Понимать значение связки «потому что»
 - Замена на «несмотря на» тоже инвертирует ответ
- Для прохождения теста требуется извлечение имплицитной информации из текста

Результаты соревнования

- Было представлено 6 решений от 4 команд:

Contestant	Number correct	Percentage correct
Patrick <u>Dhondt</u>	27	45%
Denis Robert	19	31.666%
<u>Nicos Issak</u>	29	48.33%
<u>Quan Liu</u> (1)	28	46.9% (48.33)*
<u>Quan Liu</u> (2)	29	48.33% (58.33)*
<u>Quan Liu</u> (3)	27	45% (58.33)*

- Случайными ответам можно было получить 45%

Неопределённые выражения

- Являются ли неопределённые выражения референтными?
- Рассел использовал квантор существования
 - A dog came in. $\exists x(\text{dog}(x) \ \& \ \text{came-in}(x))$
- Аргументы против референтности:
 - John is friends with a dog, and Mary is friends with a dog
 - It is not the case that a dog came in
 - Every child owns a dog

Неопределённые выражения

- Но как быть с примерами:
 - A dog came in. *It* lay down under the table.
- Расширить действие квантора (Geach):
 - $\exists x(\text{dog}(x) \ \& \ x \text{ came in} \ \& \ x \text{ lay down})$
- Но:
 - A man fell over the edge. – *He* didn't fall; he jumped
 - A dog came in. – What did *it* do next?

Неопределённые выражения

- Местоимения – замаскированные определённые выражения (Evans E-type pronouns)
 - A dog came in. It [the dog that came in] lay down under the table.
- Подразумевается уникальность определённых выражений, но:
 - Everybody who bought *a sage plant* here bought *eight others* along with *it*.

Дискурсивные референты

- Карттунен предложил понятие дискурсивных референтов
- В отличие от реальных референтов эти конструкты возникают в процессе дискурса (в головах собеседников)
- У дискурсивного референта может отсутствовать реальный референт
- Одному реальному референту могут соответствовать разные дискурсивные референты

Теория «знакомства» (familiarity)

- Определённые выражения – не являются уникальными, а являются лишь знакомыми собеседникам
- Хайм привлекает понятие дискурсивных референтов в защиту этой теории:
 - John came, and so did Mary. *One of them* brought a cake.
- *One of them* – неопределённое выражение, но оно не вводит нового референта, а вводит лишь нового *дискурсивного* референта.

File Change Semantics

- Неопределённые выражения вводят новую переменную (дискурсивного референта)
- Определённые выражения ссылаются на уже существующую переменную (дискурсивного референта)
- Каждый дискурсивный референт соответствует «карточке» (File Card) во внутренней «картотеке» (File) собеседников

File Change Semantics

- *A woman was bitten by a dog*

1
– is a woman
– was bitten
by 2

2
– is a dog
– bit 1

File Change Semantics

- *A woman was bitten by a dog*
- *She hit it*

1
– is a woman
– was bitten
 by 2
– hit 2

2
– is a dog
– bit 1
– was hit by 1

File Change Semantics

- *A woman was bitten by a dog*
- *She hit it*
- *It jumped over the fence*

1

- is a woman
- was bitten
by 2
- hit 2

2

- is a dog
- bit 1
- was hit by 1
- jumped over 3

3

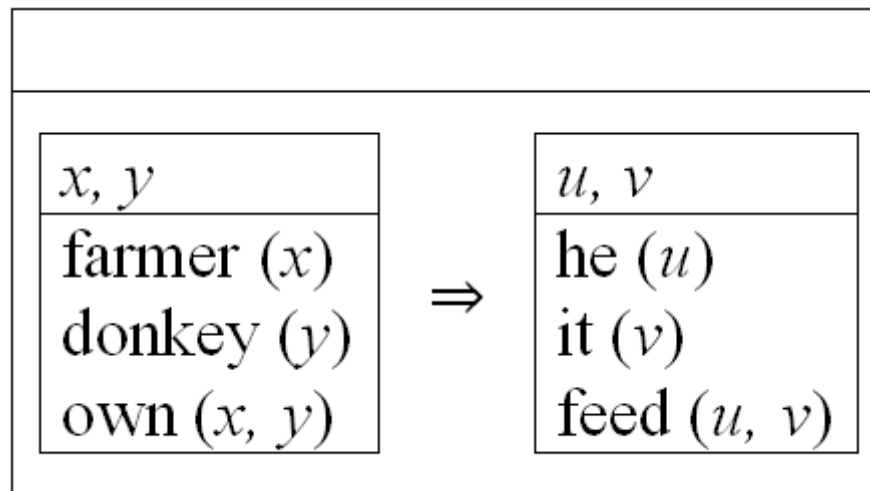
- is a fence
- was jumped
over by 2

Условия истинности

- Весь набор карточек является истинным если существует такой набор реальных референтов, которые соответствуют карточкам
- Отдельное предложение является истинным, если при добавлении его к истинному файлу получается истинный файл
- Но для предложения важнее не его условия истинности, а его потенциал изменения файла (file change potential)

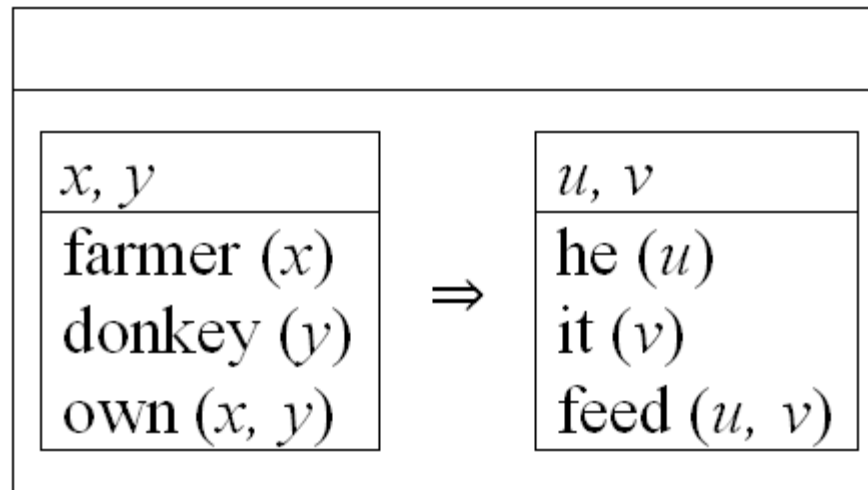
Discourse Representation Theory

- DRT – теория динамической семантики (Kamp 1981)
 - Использует семантические представления – DRS
 - DRS состоит из дискурсивных референтов и условий
 - Для сложных предложений DRS может включать под-DRS
- *If a farmer owns a donkey, he feeds it*



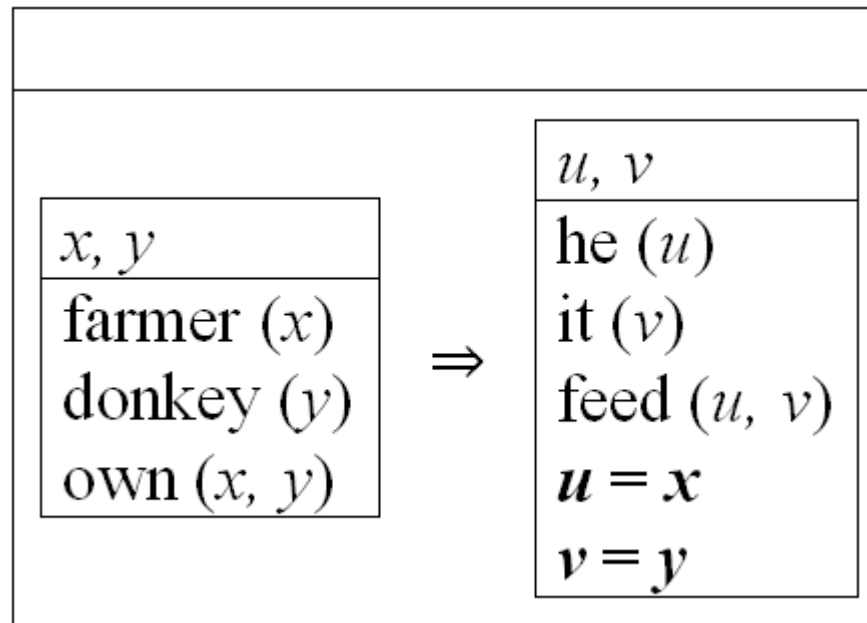
Разрешение анафоры

- *If a farmer owns a donkey, he feeds it*



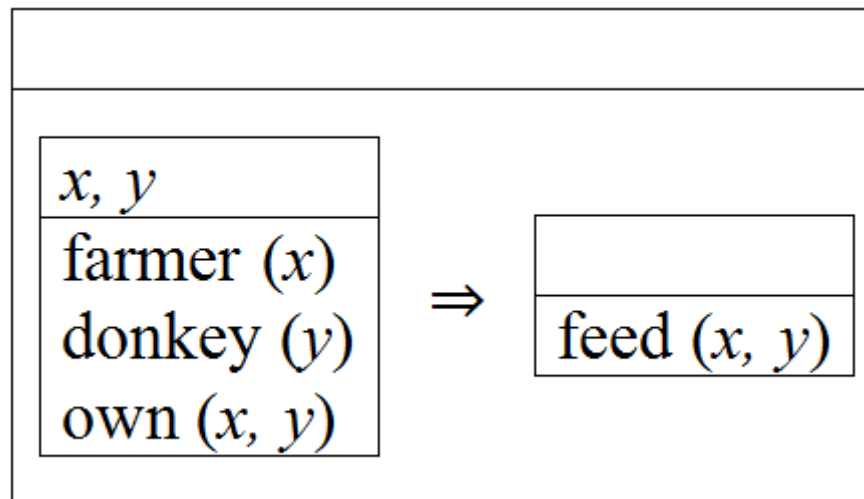
Разрешение анафоры

- *If a farmer owns a donkey, he feeds it*



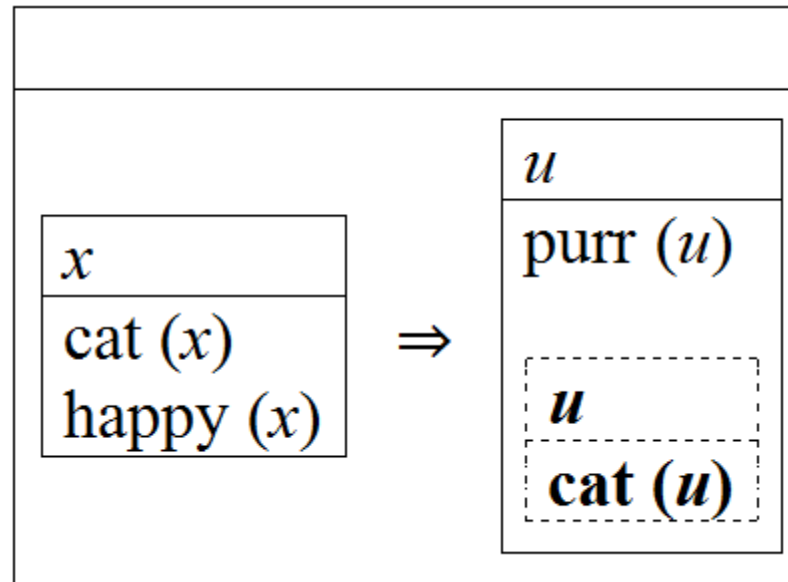
Разрешение анафоры

- *If a farmer owns a donkey, he feeds it*



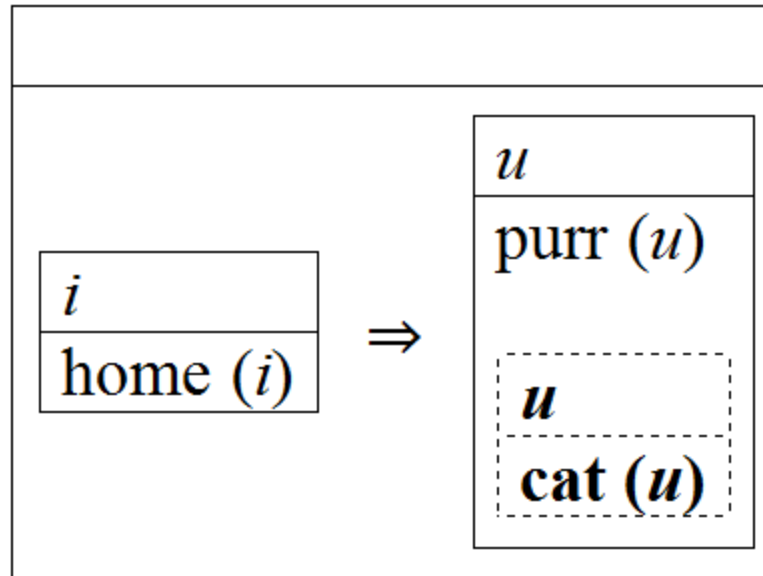
Проекция пресуппозиций

- Связывание пресуппозиций (van der Sandt 1992)
 - A-DRS хранит содержимое пресуппозиции
 - Предварительный DRS – DRS с непустыми A-DRS
- *If a cat is happy, **the cat** purrs*



Аккомодация пресуппозиций

- Аккомодация
 - Если антецедент не найден, его можно добавить
 - Это (вспомогательная) стратегия восстановления
- *When I am at home, the cat purrs*



Аккомодация пресуппозиций

- Связывание и аккомодация
 - Связывание идёт снизу вверх
 - Аккомодация идёт сверху вниз
- *When I am at home, the cat purrs*

