
УДК 004.823

М. И. СВЯТОЩИК

**МЕТОД РЕКУРСИВНОЙ
РЕКОНСТРУКЦИИ В
СИНТАКСИЧЕСКОМ АНАЛИЗЕ
ПРЕДЛОЖЕНИЙ**

В данной статье рассматривается метод рекурсивной реконструкции в комбинаторной семантике в рамках синтаксического анализа предложений. Метод рекурсивной реконструкции является важным инструментом для анализа и понимания языковых структур и семантики, а также устранения семантической неполноты предложения путём реконструкции её членов. Он основан на принципах комбинаторной семантики и позволяет эффективно анализировать структуру предложений, выявлять смысловые зависимости между словами и выражениями. В статье представлены основные принципы метода рекурсивной реконструкции, а также его применение на практике для анализа предложений.

Ключевые слова: комбинаторная семантика, рекурсия, роли индивидов, семантическая разметка, синтаксический анализ, модель мира, языковая картина мира.

M. I. SVYATOSHCHIK

**METHOD OF RECURSIVE RECONSTRUCTION IN THE
SYNTACTIC ANALYSIS OF SENTENCES**

This article discusses the method of recursive reconstruction in combinatory semantics within the framework of sentence syntactic analysis. The recursive reconstruction method is an important tool for analyzing and understanding language structures and semantics, as well as eliminating semantic incompleteness of a sentence by reconstructing its members. It is based on the principles of combinatory semantics and allows you to effectively analyze the structure of sentences, identify semantic dependencies between words and expressions. The article presents

the basic principles of the recursive reconstruction method, as well as its practical application for sentence analysis.

Keywords: combinatory semantics, recursion, roles of individuals, semantic markup, syntactic analysis, world model, linguistic picture of the world.

Синтаксический анализ предложений является одной из основных задач компьютерной лингвистики. Цель синтаксического анализа – определение структуры предложения, а также выявление связей между словами в предложении. Комбинаторная семантика подразумевает анализ семантического значения предложения, исходя из его структуры и синтаксиса. Метод рекурсивной реконструкции позволяет связать эти две области, что делает его важным инструментом для исследователей в области естественного языка.

Основы рекурсивной реконструкции, как метода синтаксического анализа, были впервые сформулированы А. Н. Гордеем и представляют собой дальнейшее развитие положений комбинаторной семантики. Комбинаторная семантика является лингвистической дисциплиной, которая изучает «отображение языком динамики ролей индивидов в событии» [1]. Положения комбинаторной семантики раскрываются в работах известных белорусских лингвистов, таких как В. В. Мартынов, А. Н. Гордей, Н. И. Руденко и других [2].

Рекурсивная реконструкция может быть использована при анализе рекурсивных предложений – это предложения, содержащие вложенные (рекурсивные) структуры, и для полного анализа необходима реконструкция членов предложения на глубинном уровне. В случае сложной (предикативной) рекурсии результатом такой реконструкции будет целое предложение. Этот метод также позволяет реконструировать пропущенные члены предложения на основе контекста или модели мира и устранить противоречия, которые часто возникают при линейном анализе предложений без учета уровней рекурсии [2].

При проведении рекурсивной реконструкции анализируются все уровни структуры предложения. Поверхностная синтаксическая структура представляет предложение на уровне поверхностного синтаксиса и основательно отражает лексический состав, синтаксические связи и порядок слов в данном предложении. Иногда термин «*поверхностная структура*» может использоваться для обозначения самого предложения, соответствующего данной структуре. Традиционно поверхностная структура подвергается анализу путем выделения составных частей предложения. Под глубинной структурой понимается «структура предложения до его трансформационной истории» [3]. В данном случае для анализа применяется трансформационный метод, который основывается на выводе сложных синтаксических структур из более простых при помощи набора правил преобразования, так называемых трансформаций [2].

Используя методы трансформационной грамматики, восстанавливаем пропущенные члены предложения с опорой на модель мира и языковую картину мира, чтобы фразы соответствовали требованиям очередности, наглядности и простоты. Согласно алгоритму ролей индивидов (список ролей индивидов Теории автоматического порождения архитектуры знаний (ТАПАЗ)), где разграничиваются физические и информационные процессы, а также при

поддержке модели мира и языковой картины мира, мы можем восстановить пропущенные члены предложения. Таким образом, исходное предложение «*Стул сломан*» реконструируется в «*Некто сломал стул*», где «*некто*» – реконструированное подлежащее, отображающее живого индивида в роли субъекта, «*сломан*» – сказуемое, отображающее акцию, а «*стул*» – прямое дополнение, отображающее объект.

В работе А. Н. Гордея «Семантическая разметка события и её отражение средствами китайского и русского языков» (2021) подробно можно ознакомиться и рассмотреть определения ролей индивидов. Автор также представляет анализ семантической разметки события в контексте русского и китайского языков. Исследование включает основные аспекты выявления и описания ролей индивидов в событийных контекстах [4, с. 15–16].

В комбинаторной семантике применяется понятие «*свёртки*», которая означает пропуск составного элемента единицы на определенном языковом уровне [4]. Говоря о свёртке знаков в языке, которая предполагает наличие вложенных конструкций, следует упомянуть о принципе экономии. Этот принцип был сформулирован французским структуралистом Андре Мартине в работе «Принцип экономии в фонетических изменениях» [5, с. 126]. В ней он указывал, что регулирующим фактором в развитии языка является «*принцип наименьшего усилия*» или экономии [там же]. Как отмечает В. В. Мартынов, принцип экономии определяет эллиптичность выражений в естественном языке. До начала речевой коммуникации известно, что собеседник в некоторой степени знаком с ситуацией, о которой будет говориться, поэтому говорящий сокращает текст высказывания (компрессирует код), в надежде, что семантическая нагрузка сохранится в ожидаемом объеме [6, с. 179].

В своих работах о структуре предложения В. В. Мартынов и А. Н. Гордей указывают на неполноту двухчленной структуры предложения (подлежащее + сказуемое), ссылаясь на исследования З. Харриса (1964). Они отмечают, что минимальная ядерная цепочка должна состоять из трех структурных компонентов: субъекта, акции и объекта [5, с. 43]. В связи с этим для полного выражения смысла необходима процедура реконструкции сверхглубинной структуры, которая предполагает восстановление пропущенных элементов предложения на основе расширенной ядерной семантической цепочки "субъект – инструмент – акция – медиатор – объект – продукт" [7, с. 57].

В настоящее время Теория автоматического порождения архитектуры знаний (ТАПАЗ-2) располагает полным набором алгебраически исчисленных ролей индивидов. В декодировании некоторых алгебраических формул и представления их значений в терминосистемах английского и русского языков мы оказали сильное содействие.

Ролевой лист ТАПАЗ-2 составляет следующий кортеж: «*субъект (инициатор → вдохновитель → распространитель → вершитель) → инструмент (активатор → суппрессор → усилитель → преобразователь) → медиатор (ориентир → локус → транспортёр → адаптер → акцептор → запас → сепаратор → материал → макет → фиксатор → ресурс → стимул → регулятор → хронотон → источник → индикатор) → объект (покрытие →*

корпус → прослойка → сердцевина) → **продукт** (заготовка → полуфабрикат → прототип → изделие)» [3, с. 10].

На примере предложения рассмотрим процесс перехода от лексически размеченного текста к его синтаксической структуре, где П – подлежащее, СК – сказуемое, ОБ_{причины} – обстоятельство причины, S – субъект, А – акция, О – объект, М – медиатор, Р – продукт, I – инструмент. Переход может осуществляться в два этапа: сначала происходит переход от исходного варианта текста к восстановленному, а затем – от восстановленного текста к семантике.

Восстановление текста происходит путем реконструкции пропущенных членов предложения с помощью модели мира или языковой картины мира. После этого происходит нормализация синтаксической структуры путем переписывания членов предложения.

Рассмотрим предложение: тоска вызывает заболевание.

Синтаксическая формула исходного предложения примет вид: П+СК+ПД.

При восстановлении текста путём реконструкции пропущенных членов предложения с опорой на модель мира и языковую картину мира предложение приобретает вид:

1. Когда я тоскую, я болею.
2. Болеть – процесс денормализации.
3. Выхлащивание – уничтожения ядра.
4. Некто \ нечто выхлащивает\разрушает меня.
5. Недостаток фасцинационного\информационного ресурса приводит к физической и\или информационной\психической денормализации.
6. Источник – информационный ресурс.
7. Источник – правило, предписание.
8. Вызвать тоску – изъять регулятор – изъять источник.
9. Некто\нечто изымает регулятор, в результате чего происходит моя физическая и\или информационная\психическая денормализация.

Синтаксическая разметка исходного текста в технологии ТАПАЗ:

Некто\нечто (П) изымает (СК) регулятор (ПД), затем (ЗАС) происходит (СК₀) моя физическая и\или информационная\психическая денормализация (П₀).

ОБ_{причины} [П+СК+ПД]  П₀+СК₀

Результат восстановления семантической структуры исходного рассматриваемого текста:

Некто (S₁=вершитель) изымает (A₁) регулятор (O₁) у меня (M₁=ориентир) при помощи чего-то (I₁), затем некто (S₂) затемняет (A₂) источник (M₂=источник) в подсознании (O₂) при помощи чего-то (I₂), далее я (S₄=вершитель) тоскую (M₃=индикатор), я (S₄) не реагирую(A₃) на тоску (M₃=индикатор), затем происходит (A₄) моя физическая и\или информационная\психическая денормализация (P₁).

Вывод: индикатор (тоска) сигнализирует о том, что некто\нечто затемнил(о) источник в объекте. Наблюдается, это приводит к психической (информационной) и\или физической денормализации объекта.

Семантическая формула исходного рассматриваемого текста:

$S_1 S_2 S_3 S_4 I_1 I_2 A_1 A_2 A_3 A_4 M_1 = \text{ориентир} M_2 = \text{источник} M_3 = \text{индикатор} O_1 O_2 \longrightarrow P_1$

Поэтому путём использования метода рекурсивной реконструкции стало возможным преодолеть непосредственные противоречия, которые обычно возникают в процессе синтаксического анализа предложений естественного языка, и достичь необходимой полноты и глубины синтаксического описания.

ЛИТЕРАТУРА

1. *Гордей, А. Н.* Реконструкция и рекурсия в синтаксическом анализе предложения / А. Н. Гордей // Пятые научные чтения, посвященные памяти профессоров В. А. Карпова и С. М. Прохоровой, Минск, ГУ, 18-19 марта 2011 г. / редкол. : А. И. Головня [и др.]. – Минск : Бел. Дом печати, 2011. – С. 18–26.
2. *Лещенко, К. И.* Рекурсивная реконструкция в методологическом аспекте / К. И. Лещенко // Вестник МГЛУ. Сер. 1, Филология. – Минск : МГЛУ, 2019. – № 3 (100). – С. 69–75.
3. *Гордей, А. Н.* Теоретическая грамматика восточных языков: лекционный курс / А. Н. Гордей. – Минск, 2007 (CD-ROM).
4. *Гордей, А. Н.* Иностранные языки в высшей школе / А. Н. Гордей / ФГБОУВО «Рязанский государственный университет имени С. А. Есенина». – Рязань : Редакц.-изд. центр РГУ им. С. А. Есенина, 2021. – Вып. 2 (57). – С. 5–26.
5. *Мартине, А.* Принцип экономии в фонетических изменениях / А. Мартине. – М. : ИЛ, 1960. – 263 с.
6. *Мартынов, В. В.* Категории языка: семиологический аспект / В. В. Мартынов. – М. : Изд-во «Наука», 1982. – 192 с.
7. *Гордей, А. Н.* Теория автоматического порождения архитектуры знаний (ТАПАЗ-2) и дальнейшая минимизация семантических исчислений / А. Н. Гордей // Открытые семантические технологии проектирования интеллектуальных систем = Open Semantic Technologies for Intelligent Systems (OSTIS-2014) : материалы IV Междунар. науч.-техн. конф. (Минск, 20-22 февраля 2014 года) / редкол. : В. В. Голенков (отв. ред.) [и др.]. – Минск : БГУИР, 2014. – С. 49–64.

Информация об авторе:

Святошиц Марина Игоревна – младший научный сотрудник Национальной академии наук Республики Беларусь, г. Минск, Республика Беларусь.