

Модель (теория) «Смысл \leftrightarrow Текст»

Дискуссионный клуб NLP

Владимир Салий

Игорь Александрович Мельчук

(СССР и Канада, род. 1932)



Некоторые факты о МСТ

- Основной автор МСТ: И. А. Мельчук.
- Соавторы МСТ: А. К. Жолковский, Ю. Д. Апресян и др.
- «Наиболее популярное в России направление в формальной лингвистике» (из: Я. Г. Тестелец, Введение в общий синтаксис, 2001)
- «Наша теоретическая база — это „генеративно-трансформационное“ учение Н. Хомского, естественным развитием которого и является, по нашему мнению, данная модель» (писал Мельчук в 1974 г.)

Основная идея:

Лингвистическая модель языка должна обеспечивать переход от смыслов к текстам и от текстов к смыслам в идеале так же хорошо, как это делает носитель языка.

Центральный постулат МСТ в формулировке самого Мельчука

Центральный постулат теории моделей «Смысл – Текст» гласит: язык есть система, устанавливающая соответствия между любым заданным смыслом и всеми выражающими его текстами; соответственно, лингвистическое описание языка должно представлять собой, грубо говоря, множество правил, ставящих в соответствие всякому смыслу все тексты, несущие этот смысл (И. А. Мельчук, 1984 г., цит. по: Russian Language Journal / Русский язык, Winter-Spring 1984, Vol. 38, No. 129/130)

Уровни языка / уровни представления языка по Мельчуку

- фонологический
- семантический
- поверхностно-морфологический
- глубинно-морфологический
- поверхностно-синтаксический
- глубинно-синтаксический

Из трудов Мельчука

Поход Мельчука к морфологии: см.:

Курс общей морфологии. В 5 томах. Москва — Вена, 1997—2006.

Подход к лексикологии и семантике: см.:

И. А. Мельчук, А. К. Жолковский. Толково-комбинаторный словарь современного русского языка: Опыты семантико-синтаксического описания русской лексики. Вена, 1984.

2-е изд., испр.: Москва, 2016 (ISBN 978-5-94457-271-4).

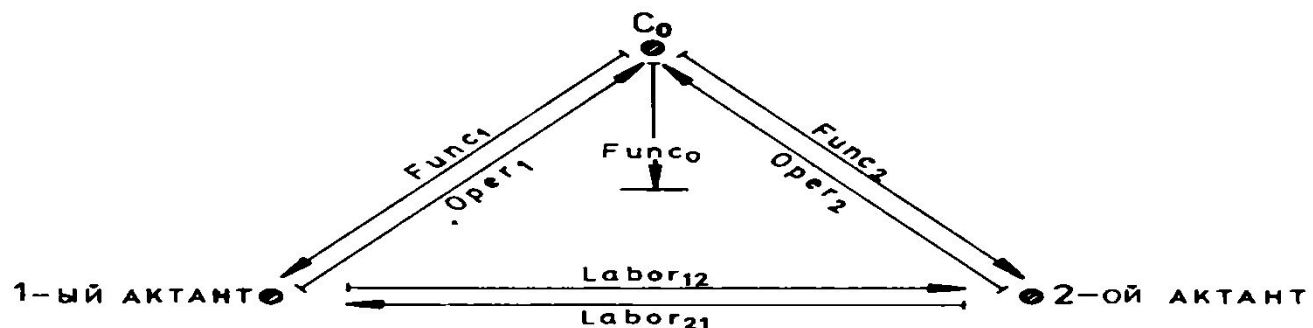
Как сделан толково-комбинаторный словарь

тцвление) = *оказывать*; **Oper₂**(*сопротивление*) = *встречать, наталкиваться* [на сопротивление].

Func₁ – 1-ым глубинным актантом глагола (и его подлежащим) является имя ситуации *C₀*, а 2-ым глубинным актантом (и первым дополнением) – ее *i*-ый участник: **Func₁**(*удивление*) = *охватывать*; **Func₂**(*температура*) = *равняться*; **Func₁**(*предложение*) = *исходить* [от ково-л.]; **Func₂**(*предложение*) = *касаться* [чвео-л.]. Если **Func** вообще не имеет дополнения, то используется индекс **0**: **Func₀**(*дождь*) = *идти*.

Labor_{ij} – 1-ым глубинным актантом и подлежащим глагола является *i*-ый участник ситуации, 2-ым глубинным актантом *j*-ый участник ситуации, а 3-им глубинным актантом (и вторым дополнением) – имя самой ситуации *C₀*: **Labor₁₂**(*вопрос*) = *подвергать* [ково-л. вопросу]; **Labor₃₂**(*аренда*) = *сдавать* [что-л. в аренду].

Oper, **Func** и **Labor** соотносятся друг с другом как конверсивы, т.е. **Oper₁** = **Conv₂₁**(**Func₁**), **Labor₁₂** = **Conv₁₃₂**(**Oper₁**) и т.д. Эти отношения могут быть представлены с помощью следующей схемы:



Здесь ключевое слово обозначает ситуацию с двумя актантами. Каждая стрелка, соответствующая определенному полувспомогательному глаголу, направлена от

Поверхностно-син- Полу-таксические вспомога-роли тельные глаголы	Подлежащее	1-ое дополнение	2-ое дополнение
Oper _{1/2}	1-ый/2-ой актанта	C ₀	—
Func _{0/1/2}	C ₀	отсутствует/ 1-ый/2-ой актанта	—
Labor _{12/21}	1-ый/2-ой актанта	2-ой/1-ый актанта	C ₀

Приведем пример:



На этой схеме представлена ситуация 'А изменяет В'.

Вернемся к перечню ЛФ.

Involv – глагол, служащий для присоединения неучастника ситуации к имени ситуации, которая воздействует на него: **Conv₂₁Involv(ветер)** = *стоять [на*

БРУС, а, бя, бее, муж.

1. *Y-овый брус из X-а* – единица 3 вида Y-овых пиломатериалов из X-а – бревно, у которого с двух или четырех сторон опилены кромки [см. рисунок в статье ПИЛОМАТЕРИАЛЫ].

Ср. ДОСКА 1, РЕЙКА 1.

МУ – как у ПИЛОМАТЕРИАЛЫ

дубовый брус, брус из карельской сосны, еловый строгательный брус.

ЛФ Mult_c, Caus, предприятие-S₁^{usual}Caus, установка-S_{instr}Caus, человек-S₁^{usual}Real, F₂ – как у ПИЛОМАТЕРИАЛЫ; ЛФ S₀Caus, CausDegrad, F₁ – как у ДОСКА 1.

Syn _n	: спец. шпунтина
Dimun	: // брусок 1
Prepar	: обреза́ть, строгать, спец. шпунтовать [~]
Degrad	: треска́ться < ломаться
сделанный из	: брусяно́й

Виды брусьев

См. Виды пиломатериалов (1-3).

2. *X-овый брус* – единица 3 твердого материала X, которому придана форма бруса 1.

Ср. ДОСКА 2.

МУ – как у ДОСКА 2

металлический брус, брус из железобетона.

ВАР|ИТЬ, *врѣтит* <устар. *врѣтит*>, *варить*, прич страд прош неупотр, несов.

1.1a. сов нет. Разг. *X варит Y-у* = *X* создает достаточно простую горячую пищу [суп, кашу, ...], предназначенную для обеспечения *Y*-а рационом питания | узуально [= *Syn_c*{*готовить (пищу)*}]

1 = X [кто создает]	2 = Y [кого обеспечивает питанием]
1. <i>S_{ин}</i>	1. <i>S_{дат}</i> 2. <i>на Num S_{род}</i>

1) D_{2.2}: *S* – мы или собирают.

[*В экспедиции не было повара,*] и мы варили себе сами <на 10 человек, на (всю) группу>; Марфа одна в доме и стирала, и убиралась, и варила, и ходила по воду.

<i>Syn₃</i>	:	<i>готовить</i>
<i>Syn_n</i>	:	разг. <i>стряпать</i> , разг. <i>кашеварить</i>
<i>S₀</i> или <i>SingS₀</i>	:	<i>варка 1</i> [<i>Варка не отнимала у нее слишком много времени</i>]
человек- <i>S_{ic}^{usual}</i>	:	разг. <i>стряпуха</i> [женщина]; дорев. <i>кухарка</i> [женщина, работающая по найму в частном доме]; <i>повар</i> [на предприятии общественного питания или в частном доме]; <i>кок</i> [на корабле]; <i>кашевар</i> [в армии, в экспедиции, на полевом стане и т.п.]
сосуд- <i>S_{instr}^{usual}</i>	:	<i>кастрюля</i> , <i>котелок</i> [для варки на открытом огне]; <i>котел 1</i> [сосуд очень большого размера]; <i>чуунок</i> , <i>воршок</i> [для варки в печке]
устройство- <i>S_{loc-instr}^{usual}</i>	:	(<i>газовая</i> , <i>электрическая</i>) <i>плита</i> , (<i>электро</i>) <i>плитка</i> ,

Из трудов Мельчука (продолжение)

- *I. Mel'čuk, J. Milićević. An Advanced Introduction to Semantics: A Meaning-Text Approach. — Cambridge University Press, 2020.*
- *Л. Н. Иорданская, И. А. Мельчук. Смысл и сочетаемость в словаре. — М.: Языки славянских культур, 2007. ISBN 5-9551-0181-0*
- *I. A. Mel'čuk. Aspects of the Theory of Morphology. — Berlin; New York: Mouton de Gruyter, 2006. ISBN 3-11-017711-0, ISBN 978-3-11-017711-4*
- *I. A. Mel'čuk. Dependency syntax: Theory and practice. — Albany, NY: SUNY, 1988. ISBN 0-88706-450-7, ISBN 0-88706-451-5*

Применение МСТ

1) Машинный перевод:

само возникновение МСТ связано с темой машинного перевода (МП), над которой в составе научного коллектива работал Мельчук в свои ранние годы; МСТ – один из подходов к проблеме МП на основе правил (существует также принципиально иной, статистический подход); предполагалось, что программы МП будут опираться на МСТ; система ЭТАП (Ю. Д. Апресян и другие, СССР)

2) Другие системы автоматической обработки естественного языка (АОЕЯ):

в т. ч. современный лингвистический процессор ЭТАП-3 (разработан ИППИ им. Харкевича РАН; см. <http://iitp.ru/ru/science/works/452.htm>); морфологический и синтаксический анализаторы системы ЭТАП-3 используются в Национальном корпусе русского языка (НКРЯ; см. <https://ruscorpora.ru/new/instruction-syntax.html>); система ЭТАП-3 многофункциональна, на её основе создана в т. ч. одна из систем машинного перевода

Применение МСТ (продолжение)

- 3) В качестве теоретической основы для описательных работ по конкретным естественным языкам (подход получил определённое распространение в СССР/России, меньшее – в ряде зарубежных стран)

Некоторые источники по МСТ, автор которых не Мельчук

- Синтаксис в модели «Смысл \Leftrightarrow Текст» // Я. Г. Тестелец. Введение в общий синтаксис. Москва, 2001
- *Jasmina Milićević*. A short guide to the Meaning-Text Linguistic Theory. 2006. <http://olst.ling.umontreal.ca/pdf/IntroMTTJM.pdf>

Вместо заключения

В настоящее время подход к автоматической обработке естественного языка (NLP) на основе правил имеет тенденцию к выходу из употребления. Господствующим становится подход на основе статистики, в первую очередь — с использованием искусственных нейросетей.

Таким образом, и МСТ, оставаясь яркой страницей в истории русского и мирового языкознания, выходит из практического употребления в области NLP.