

Трансляция математических записей
с естественного языка при помощи
формальных грамматик

Михаил Пожидаев

10 октября 2019 г.

Три формы математической записи

- ▶ Интеграл в пределах от a до b от функции f от x по dx
- ▶ $\int_a^b f(x) dx$
- ▶ $\int_a^b f(x) dx$

Пример фрагмента грамматики Томита-парсера

- ▶ eq -> NormEnt interp(Equality.Name1)
Word<kwtype="равенство» ("также")
NormEnt<gram="твор» interp (Equality.Name2);
- ▶ diap -> "от"chislit "до"chislit;
- ▶ diap -> chislit "—"chislit;

Порядок работы

1. Добрый человек читает математические записи с печатного издания.
2. Записанная речь распознаётся при помощи движка распознавания речи.
3. Полученный текст подвергается обработке при помощи Томита-парсера с генерацией фактов.
4. Факты транслируются в записи в формате T_EX .

Спасибо за внимание!

Веб-сайт: <http://marigostra.ru/>

E-mail: msh@luwrain.org