

# О странных деревьях

Semyon Grigorev

20 декабря 2016 г.

Все знают про контекстно-свободные грамматики и некоторые даже любят их. Некоторые что-то слышали про конъюнктивные и даже булевы грамматики и побаиваются их. И есть за что. Давайте рассмотрим простой язык  $L_0 = \{\omega\omega \mid \omega \in \{a,b\}^*\}$ . Данный язык не является контекстно-свободным. Но для него в работе Александра Охотина “Boolean Grammars” предложена замечательная булева грамматика (обозначим её  $G_0$ ):

$$\begin{aligned} S &\rightarrow -AB \ \& \ -BA \ \& \ C \\ A &\rightarrow XAX \mid a \\ B &\rightarrow XBX \mid b \\ C &\rightarrow XXC \mid \varepsilon \\ X &\rightarrow a \mid b \end{aligned}$$

Очень просто, правда? Задание для самопроверки — убедитесь, что данная грамматика действительно описывает язык  $L_0$ . Теперь давайте рассмотрим язык  $L_1 = \{\omega c \omega \mid \omega \in \{a,b\}^*\}$ , который получается из языка  $L_0$  простым добавлением разделителя ‘с’, что наводит на мысль, что грамматики этих языков должны быть очень похожи. Давайте попробуем получить грамматику для  $L_1$  из грамматики  $G_0$ . Не уверен, что это очень простая задача, поэтому предъявим грамматику для этого языка, приводимую в работе “”. Любопытный читатель может попробовать построить дерево вывода цепочки aabсаab в данной грамматике, а так же получить из неё грамматику для языка  $L_0$ .

$$\begin{aligned}
S &\rightarrow C \ \& \ D \\
C &\rightarrow aCa \mid aCb \mid bCa \mid bCb \mid c \\
D &\rightarrow aA \ \& \ aD \mid bB \ \& \ bD \mid cE \\
A &\rightarrow aAa \mid aAb \mid bAa \mid bAb \mid cEa \\
B &\rightarrow aBa \mid aBb \mid bBa \mid bBb \mid cEb \\
E &\rightarrow aE \mid bE \mid \varepsilon
\end{aligned}$$

Не самое очевидный результат, не правда ли? Некоторые трудности в работе с булевыми грамматиками отмечает и Andrew Stevenson в своей диссертации в разделе “Managing Grammar Complexity”.

Давайте вместо букв оперировать более сложными объектами (предпосылки к этому есть в гиперграмматиках, где оперируют множествами). Деревья. Строгое расширение КС. Видна лемма о накачке. Структура грамматики более простая. Примеры для приведённых ранее языков. Ещё для некоторых.

## Список литературы

- [1] Alexander Okhotin, “Parsing by matrix multiplication generalized to Boolean grammars”, *Theoretical Computer Science*, V. 516, p. 101–120, January 2014