УДК 519.766.24 ISSN 1995-5499
DOI: https://doi.org/10.17308/sait/1995-5499/2023/3/156-166
Поступила в редакцию 28.03.2023
Подписана в печать 30.09.2023

ПРЕДСТАВЛЕНИЕ МНОГОЧЛЕНОВ МУЛЬТИМОДАЛЬНЫМИ КАТЕГОРИАЛЬНЫМИ ГРАММАТИКАМИ ЗАВИСИМОСТЕЙ

© 2023 Б. Н. Карлов✉

*Тверской государственный университет,
ул. Желябова, 33, 170100 Тверь, Российская Федерация*

Аннотация. Категориальные грамматики зависимостей являются обобщением классических категориальных грамматик. В отличие от контекстно-свободных грамматик, они описывают структуру предложения не с помощью составляющих, а посредством синтаксических зависимостей между словами. Эти грамматики весьма выразительны и позволяют порождать многие не контекстно-свободные языки. В статье изучаются выразительные возможности одного из вариантов этих грамматик — мультимодальных категориальных грамматик зависимостей. Ранее уже было установлено, что они способны порождать неполулинейные языки. В статье исследуются возможности задания многочленов с помощью таких грамматик. Доказано, что для любого полинома существует грамматика, порождающая множество положительных значений этого полинома, записанных в унарной системе. Аналогичным образом может быть представлено и множество абсолютных величин всех ненулевых значений полинома. Установлено, что можно построить грамматики, осуществляющие вычисление значения полинома на заданном аргументе и вычисление аргумента по заданному значению полинома.

Ключевые слова: категория, поляризованная валентность, мультимодальная категориальная грамматика зависимостей, непроективная зависимость, формальный язык, многочлен.

ВВЕДЕНИЕ

Категориальные грамматики являются одним из стандартных способов задания естественных и формальных языков. Они позволяют описывать структуру предложения в терминах синтаксических зависимостей между отдельными словами. Они сопоставляют каждому слову конечное множество синтаксических категорий, которые описывают синтаксические типы слов, то есть их роли в предложении. Простейший вариант таких грамматик (классические категориальные грамматики) эквивалентен контекстно-свободным грамматикам (см. [1]). Однако сейчас признано, что контекстно-свободные грамматики недостаточно выразительны для описания естественных языков. В частности, они неспособны обрабатывать так называемые непроективные зависимости, которые встречаются во многих естественных языках, в том числе в русском (см. [1]). Поэтому как с теоретической, так и с практической точки зрения большой интерес представляет разработка и исследование более выразительных моделей языка.

Существует большое количество типов грамматик, обобщающих контекстно-свободные: индексные, множественные контекстно-свободные, комбинаторные категориальные, различные варианты грамматик Ламбе-

✉ Карлов Борис Николаевич
e-mail: bnkarlov@gmail.com

ка и многие другие. Выразительные силы этих грамматик существенно различаются. К недавним достижениям в этой области относятся результаты о сильной эквивалентности двух вариантов слабо контекстных грамматик (см. [2–4]). Одним из обобщений классических категориальных грамматик являются категориальные грамматики зависимостей (КГЗ) и мультимодальные категориальные грамматики зависимостей (ммКГЗ), введённые А. Я. Диковским и М. И. Дехтярём (см. [5, 6]). Эти грамматики позволяют задавать непроективные зависимости, и в то же время для них существует алгоритм анализа, имеющий полиномиальную сложность при некоторых естественных ограничениях. В [7–10] описаны возможности практического применения КГЗ для анализа естественных языков, а в [11–13] получены некоторые теоретические результаты. В частности, в [11] построен пример ммКГЗ, порождающей язык, множество длин слов которого не содержит бесконечных арифметических прогрессий. Известно, что для контекстно-свободных грамматик и некоторых вариантов слабо контекстных грамматик множество длин слов является объединением конечного числа арифметических прогрессий, а образ Парика языка является полулинейным множеством (точные определения и доказательства этих фактов можно найти в [14, 15]). В связи с этим возникает вопрос, каким может быть множество длин слов языка, порождаемого ммКГЗ.

В настоящей статье мы изучаем выразительные возможности мультимодальных категориальных грамматик зависимостей. В разделе 1 приводятся основные определения. В разделе 2 мы доказываем, что для любого многочлена $f(n)$ с положительным старшим коэффициентом существует ммКГЗ G_f, порождающая язык $\{a^{f(n)} : n \ge n_0\}$, где n_0 — некоторая константа, зависящая от f. С помощью этого результата мы доказываем, что языки

$$L_1 = \{a^{f(n)} : n \ge 0, f(n) > 0\},$$

$$L_2 = \{a^{f(n)} : n \in Z, f(n) > 0\},$$

$$L_3 = \{a^{|f(n)|} : n \in Z, f(n) \ne 0\}$$

также порождаются ммКГЗ (здесь через Z обозначено множество целых чисел). Далее мы показываем, что существуют ммКГЗ, задающие языки

$$L_4 = \{b^n a^{f(n)} : n > 0, f(n) > 0\},$$

$$L_5 = \{b^{f^{-1}(m)} a^m : m > 0, f^{-1}(m) > 0\},$$

где $f^{-1}(m)$ — наибольшее натуральное число n, которое удовлетворяет неравенству $f(n) \le m$, то есть $f^{-1}(m)$ — некоторый вариант обратной функции для полинома f.

1. ОСНОВНЫЕ ОПРЕДЕЛЕНИЯ

Алфавит Σ — это произвольное конечное множество символов. *Слово* в алфавите Σ — это конечная последовательность символов из Σ. В лингвистических приложениях в роли Σ выступает множество словоформ, а их последовательности являются предложениями, но мы будем придерживаться терминологии теории формальных языков. *Длиной* слова w называется число символов в нём. Длина слова w обозначается $|w|$. *Пустое* слово ε — это слово длины 0. Через Σ^* обозначается множество всех слов в алфавите Σ, а через Σ^+ — множество всех непустых слов в Σ. *Конкатенацией* слов u и v (обозначается uv) называется слово, полученное приписыванием v в конец u. i-я степень слова w — это слово $w \ldots w$, где w повторяется i раз (обозначается w^i). В частности, $w^0 = \varepsilon$, $w^1 = w$.

Пусть $\mathbf{C}$ — конечное непустое множество *элементарных категорий*. Его элементы соответствуют типам слов в предложении (подлежащее, сказуемое и т. п.). Полярностью называется элемент множества $\{\searrow, \swarrow, \nearrow, \nwarrow\}$. *Поляризованная валентность* — это строка вида vC, где v — полярность, C — элементарная категория. Пары валентностей вида $\swarrow C \nwarrow C$ и $\nearrow C \searrow C$ называются *правильными*. *Положительные* валентности $\nearrow C$ и $\nwarrow C$ задают начало непроективной зависимости между словами, а соответствующие им *отрицательные* валентности $\searrow C$ и $\swarrow C$ — её конец. *Категория* — это выражение вида

$$[L_k \backslash \ldots \backslash L_1 \backslash H / R_1 / \ldots / R_m]^P,$$

где H, L_i, R_j — элементарные категории, $k, m \ge 0$, P — конечная последовательность поляризованных валентностей. H называется *головой* категории, последовательности $L_k \backslash \ldots \backslash L_1$ и $R_1 / \ldots / R_m$ — левым и правым списком зависимостей соответственно, выражение $[L_k \backslash \ldots \backslash L_1 \backslash H / R_1 / \ldots / R_m]$ — *локальной частью* категории, а P — *потенциалом*. Множество всех категорий обозначается $\mathrm{Cat}(\mathbf{C})$. Если потенциал P содержит $\nearrow C$, то мы будем говорить, что он выпускает $\nearrow C$ направо, а если P содержит $\searrow C$, то он ловит $\searrow C$ слева.

Мультимодальная категориальная грамматика зависимостей (ммКГЗ) — это пятёрка $G = (\Sigma, \mathbf{C}, S, \delta, \pi)$, где

- Σ — алфавит терминальных символов;
- $\mathbf{C}$ — множество элементарных категорий;
- $S \in \mathbf{C}$ — *главная* категория;
- δ — *словарь*, отображение множества Σ в множество всех конечных подмножеств множества $\mathrm{Cat}(\mathbf{C})$;
- π — *функция запретов*, отображение множества $\mathbf{C}$ в множество всех своих подмножеств.

Если $\delta(a) = \{\gamma_1, \ldots, \gamma_n\}$, то мы будем записывать это в виде $a \mapsto \gamma_1, \ldots, \gamma_n$. Если $w = a_1 a_2 \ldots a_n$, то через $\delta(w)$ обозначается множество последовательностей категорий, которые можно приписать слову w согласно словарю δ:

$$\delta(w) = \{\gamma_1 \ldots \gamma_n : \gamma_i \in \delta(a_i) \text{ для } 1 \le i \le n\}.$$

На множестве конечных последовательностей категорий $\mathrm{Cat}(\mathbf{C})^*$ определяются следующие правила сокращения:

$\mathbf{L}^{\mathbf{l}} : \Gamma_1 [C]^{P_1} [C \backslash \beta]^{P_2} \Gamma_2 \vdash \Gamma_1 [\beta]^{P_1 P_2} \Gamma_2$;

$\mathbf{L}^{\mathbf{r}} : \Gamma_1 [\beta / C]^{P_1} [C]^{P_2} \Gamma_2 \vdash \Gamma_1 [\beta]^{P_1 P_2} \Gamma_2$;

$\mathbf{D}^{\mathbf{l}} : \Gamma_1 [\beta]^{P_1 \swarrow C P_2 \nwarrow C P_3} \Gamma_2 \vdash \Gamma_1 [\beta]^{P_1 P_2 P_3} \Gamma_2$, если P_2 не содержит валентностей $\swarrow C$ и $\nwarrow C$, а также валентностей вида vB для всех $B \in \pi(C)$;

$\mathbf{D}^{\mathbf{r}} : \Gamma_1 [\beta]^{P_1 \nearrow C P_2 \searrow C P_3} \Gamma_2 \vdash \Gamma_1 [\beta]^{P_1 P_2 P_3} \Gamma_2$, если P_2 не содержит валентностей $\nearrow C$ и $\searrow C$, а также валентностей вида vB для всех $B \in \pi(C)$.

В частности, если $B \in \pi(A)$, $A \in \pi(B)$, то это означает, что потенциал не может содержать подпоследовательностей валентностей вида $\nearrow A \nearrow B \searrow A \searrow B$, то есть непроективные зависимости с именами A и B не могут пересекаться.

Через $\vdash^*$ обозначается рефлексивное транзитивное замыкание отношения $\vdash$ Потенциал P называется *сбалансированным*, если его можно сократить до пустого слова, используя правила $\mathbf{D}^{\mathbf{l}}$ и $\mathbf{D}^{\mathbf{r}}$.

Язык, порождаемый ммКГЗ G, — это множество слов, которым можно приписать категории так, чтобы они сократились до главной категории $[S]$:

$$L(G) = \{a_1 \ldots a_n \in \Sigma^+ : \text{существует такая строка } \Gamma \in \delta(w), \text{ что } \Gamma \vdash^* [S]\}.$$

Язык L называется ммКГЗ-языком, если он порождается некоторой ммКГЗ.

Лингвистический смысл введённых понятий описан в книге [1] и в статьях [5, 12, 13].

2. ЗАДАНИЕ МНОГОЧЛЕНОВ МУЛЬТИМОДАЛЬНЫМИ КАТЕГОРИАЛЬНЫМИ ГРАММАТИКАМИ ЗАВИСИМОСТЕЙ

Пусть $f(n)$ — произвольный многочлен степени k с целыми коэффициентами. Через $f_i(n)$ мы обозначим многочлены, определяемые индукцией по i следующим образом:

$$f_0(n) = f(n), \qquad f_{i+1}(n) = f_i(n+1) - f_i(n).$$

Из определения непосредственно следует, что $f_i(n) + f_{i+1}(n) = f_i(n+1)$. В частности, $f(n) + f_1(n) = f(n+1)$. В следующих трёх леммах сформулированы некоторые простые свойства многочленов $f_i(n)$. Эти леммы выражают известные свойства конечных разностей и поведение многочленов в пределе при $n \to \infty$.

Лемма 1. *Если $f(n)$ — многочлен степени k, то $f_i(n)$ — многочлен степени $k - i$.*

Лемма 2. *Если $f(n)$ — многочлен степени k, то $f_k(n)$ — отличная от нуля константа.*

Лемма 3. *Для любого многочлена $f(n)$ с положительным старшим коэффициентом существует такое число n_0, что $f_0(n) > f_1(n) > \ldots > f_k(n) > 0$ при $n \ge n_0$.*

Теперь для каждого многочлена f с положительным старшим коэффициентом мы определим ммКГЗ G_f, порождающую язык $L(G_f) = \{a^{f(n)} : n \ge n_0 + 2\}$. Грамматика содержит единственный терминальный символ a и элементарные категории

$$\mathbf{C} = \{S, C, D, E, F, X, Y\} \cup$$
$$\cup \{T_i : 1 \le T_i \le f(n_0) - 1\} \cup \{A_i, B_i : 1 \le i \le k\},$$
для которых определены следующие ограничения:
$$\pi(A_i) = \{X\}, \pi(X) = \{A_i : 1 \le i \le k\},$$
$$\pi(B_i) = \{Y\}, \pi(Y) = \{B_i : 1 \le i \le k\}.$$

Для остальных категорий ограничений нет.

Теперь определим словарь грамматики. Обозначим через P_0 потенциал
$$P_0 = (\nearrow A_1)^{f_1(n_0)} (\nearrow A_2)^{f_2(n_0)} \ldots (\nearrow A_k)^{f_k(n_0)}.$$

Перечислим сопоставления категорий символу a:
$$a \mapsto [S / T_1]^{\nearrow X P_0 \nearrow Y}, [T_1 / T_2], [T_2 / T_3], \ldots, [T_{f(n_0)-1} / C]; \quad (1)$$
$$a \mapsto [C / C]^{P_1}, [C / D]^{P_1 \searrow X \nearrow X}, [C / F]^{P_1 \searrow X}. \quad (2)$$
В категориях (2) потенциал P_1 принимает следующие значения:

$P_{1,1} = \searrow A_1 \nearrow B_1;$

$P_{1,2} = \searrow A_1 \searrow A_2 (\nearrow B_1)^2 \nearrow B_2;$

$P_{1,3} = \searrow A_1 \searrow A_2 \searrow A_3 (\nearrow B_1)^2 (\nearrow B_2)^2 \nearrow B_3;$

$\ldots$

$P_{1,k} = \searrow A_1 \ldots \searrow A_k (\nearrow B_1)^2 \ldots (\nearrow B_{k-1})^2 \nearrow B_k.$

Следующие сопоставления аналогичны категориям (2):
$$a \mapsto [D / D]^{P_2}, [D / C]^{P_2 \searrow Y \nearrow Y}, [D / E]^{P_2 \searrow Y}. \quad (3)$$
В категориях (3) потенциал P_2 принимает следующие значения:

$P_{2,1} = \searrow B_1 \nearrow A_1;$

$P_{2,2} = \searrow B_1 \searrow B_2 (\nearrow A_1)^2 \nearrow A_2;$

$P_{2,3} = \searrow B_1 \searrow B_2 \searrow B_3 (\nearrow A_1)^2 (\nearrow A_2)^2 \nearrow A_3;$

$\ldots$

$P_{2,k} = \searrow B_1 \ldots \searrow B_k (\nearrow A_1)^2 \ldots (\nearrow A_{k-1})^2 \nearrow A_k.$

Следующие сопоставления задают категории на конце слова:
$$a \mapsto [E / E]^{P_3}, [E]^{P_3 \searrow X}. \quad (4)$$
В категориях (4) потенциал P_3 принимает следующие значения:

$P_{3,1} = \searrow A_1;$

$P_{3,2} = \searrow A_1 \searrow A_2;$

$\ldots$

$P_{3,k} = \searrow A_1 \ldots \searrow A_{k-1} \searrow A_k.$

Сопоставления (5) аналогичны (4):
$$a \mapsto [F / F]^{P_4}, [F]^{P_4 \searrow Y}. \quad (5)$$
В категориях (5) потенциал P_4 принимает следующие значения:

$P_{4,1} = \searrow B_1;$

$P_{4,2} = \searrow B_1 \searrow B_2;$

$\ldots$

$P_{4,k} = \searrow B_1 \ldots \searrow B_{k-1} \searrow B_k.$

Сначала мы докажем, что построенная грамматика позволяет разбивать слова на блоки длины
$$f(n_0), f_1(n_0 + 1), f_1(n_0 + 2), \ldots, f_1(n)$$
для любого $n \ge n_0$ так, что все валентности типа $\nearrow A_i$ (соответственно $\nearrow B_i$) каждого блока сокращаются с валентностями $\searrow A_i$ (соответственно $\searrow B_i$) следующего блока.

Лемма 4. *Для любого $n \ge n_0$ существует строка категорий $\Gamma = [\alpha_1]^{P_1} \ldots [\alpha_r]^{P_r}$, где $r = f(n)$, такая, что:*

1) если число $n - n_0$ чётно, то $\alpha_r = [D / C]$ или $\alpha_r = [T_{f(n_0)-1} / C]$, а потенциал $P_1 \ldots P_r$ сокращается до
$$\nearrow X (\nearrow A_1)^{f_1(n)} \ldots (\nearrow A_k)^{f_k(n)} \nearrow Y;$$

2) если число $n - n_0$ нечётно, то $\alpha_r = [C / D]$, а потенциал $P_1 \ldots P_r$ сокращается до
$$\nearrow Y (\nearrow B_1)^{f_1(n)} \ldots (\nearrow B_k)^{f_k(n)} \nearrow X;$$

3) если $[\alpha_r{}']$ получается из $[\alpha_r]$ удалением зависимой категории, то
$$[\alpha_1] \ldots [\alpha_{r-1}][\alpha_r{}'] \vdash^* [S].$$

Доказательство. Индукция по n.

Базис индукции. Если $n = n_0$, то нужно взять строку, построенную из категорий (1):
$$[S / T_1]^{\nearrow X P_0 \nearrow Y} [T_1 / T_2][T_2 / T_3] \ldots [T_{f(n_0)-1} / C].$$
Правильность этой строки следует из определения P_0.

Индукционный шаг. Рассмотрим случай, когда число $(n+1) - n_0$ нечётно. Тогда число $n - n_0$ чётно. По индукционному предположению существует строка категорий Γ_1 длины $f(n)$, потенциал которой сокращается до
$$\nearrow X (\nearrow A_1)^{f_1(n)} (\nearrow A_2)^{f_2(n)} \ldots (\nearrow A_k)^{f_k(n)} \nearrow Y.$$

Рассмотрим строку категорий

$$\Gamma = \Gamma_1([C/C]^{P_{1,1}})^{f_1(n)-f_2(n)}\ldots([C/C]^{P_{1,k-1}})^{f_{k-1}(n)-f_k(n)}$$
$$([C/C]^{P_{1,k}})^{f_k(n)-1}[C/D]^{P_{1,k}\searrow X\nearrow X}.$$

По индукционному предположению Γ_1 заканчивается на категорию с локальной частью $[D/C]$ или $[T_{f(n_0)-1}/C]$. Непосредственно проверяется, что после удаления D в последней категории локальная часть получившейся строки сокращается до $[S]$. Потенциал строки сокращается до

$$P' = \nearrow X(\nearrow A_1)^{f_1(n)}(\nearrow A_2)^{f_2(n)}\ldots(\nearrow A_k)^{f_k(n)}\nearrow Y$$
$$(P_{1,1})^{f_1(n)-f_2(n)}\ldots(P_{1,k-1})^{f_{k-1}(n)-f_k(n)}(P_{1,k})^{f_k(n)-1}P_{1,k} \quad (6)$$
$$\searrow X\nearrow X.$$

Число добавленных категорий равно

$$(f_1(n)-f_2(n))+(f_2(n)-f_3(n))+\ldots+$$
$$+(f_{k-1}(n)-f_k(n))+f_k(n)=f_1(n),$$

поэтому длина строки Γ равна

$$f(n)+f_1(n)=f(n+1).$$

Найдём число валентностей $\searrow A_i$ в потенциале P'. Эти валентности содержатся только в потенциалах $P_{1,i},P_{1,i+1},\ldots,P_{1,k}$. Значит, их количество равно

$$(f_i(n)-f_{i+1}(n))+(f_{i+1}(n)-f_{i+2}(n))+\ldots+$$
$$+(f_{k-1}(n)-f_k(n))+f_k(n)=f_i(n).$$

Поскольку валентности $\nearrow A_i$ и $\nearrow Y$ не блокируют сокращение друг друга, то все пары $\nearrow A_i\searrow A_i$ сокращаются. После этого между валентностями $\nearrow X$ и $\searrow X$ остаются только валентности вида $\nearrow B_i$. Так как $\nearrow X$ и $\nearrow B_i$ также не препятствуют сокращению, то пара $\nearrow X\searrow X$ сокращается. Остаётся найти количество валентностей вида $\nearrow B_i$. Валентность $\nearrow B_i$ встречается один раз в потенциале $P_{1,i}$ и по два раза в потенциалах $P_{1,i+1},\ldots,P_{1,k}$. Поэтому при $i<k$ количество $\nearrow B_i$ равно

$$(f_i(n)-f_{i+1}(n))+2(f_{i+1}(n)-f_{i+2}(n))+\ldots+$$
$$+2(f_{k-1}(n)-f_k(n))+2f_k(n)=$$
$$=f_i(n)+f_{i+1}(n)=f_i(n+1).$$

Поскольку $\nearrow B_k$ встречается только в $P_{1,k}$, то число вхождений этой валентности в потенциал равно $f_k(n)=f_k(n+1)$, так как $f_k(n)$ — константа. Итак, весь потенциал сокращается до

$$\nearrow Y(\nearrow B_1)^{f_1(n+1)}(\nearrow B_2)^{f_2(n+1)}\ldots(\nearrow B_k)^{f_k(n+1)}\nearrow X.$$

Случай чётного $(n+1)-n_0$ рассматривается аналогично. □

Теперь докажем, что описанное в предыдущей лемме разбиение является единственно возможным.

Лемма 5. *Пусть* $\Gamma=[\alpha_1]^{P_1}\ldots[\alpha_r]^{P_r}$ *— строка категорий, сокращающаяся до* $[S]$, *и пусть* $n\ge n_0$ *— такое натуральное число, что* $f(n)\le r$ *и при этом строка категорий* $\Gamma_1=[\alpha_1]^{P_1}\ldots[\alpha_{f(n)}]^{P_{f(n)}}$ *не содержит элементарных категорий* E *и* F. *Тогда для* Γ_1 *выполняется следующее:*

1) если число $n-n_0$ *чётно, то* $\alpha_r=[D/C]$ *или* $\alpha_{f(n)}=[T_{f(n_0)-1}/C]$, *а потенциал строки* Γ_1 *сокращается до*

$$\nearrow X(\nearrow A_1)^{f_1(n)}\ldots(\nearrow A_k)^{f_k(n)}\nearrow Y;$$

2) если число $n-n_0$ *нечётно, то* $\alpha_{f(n)}=[C/D]$, *а потенциал строки* Γ_1 *сокращается до*

$$\nearrow Y(\nearrow B_1)^{f_1(n)}\ldots(\nearrow B_k)^{f_k(n)}\nearrow X.$$

Доказательство. Индукция по n.

Базис индукции. Так как Γ сокращается до $[S]$, то единственным возможным началом строки Γ может быть

$$[S/T_1]^{\nearrow XP_0\nearrow Y}[T_1/T_2][T_2/T_3]\ldots[T_{f(n_0)-1}/C].$$

По построению P_0 имеет вид, указанный в пункте 1.

Индукционный шаг. Снова рассмотрим случай, когда число $(n+1)-n_0$ нечётно. По индукционному предположению строка Γ начинается на префикс Γ_1, потенциал которого сокращается до

$$\nearrow X(\nearrow A_1)^{f_1(n)}(\nearrow A_2)^{f_2(n)}\ldots(\nearrow A_k)^{f_k(n)}\nearrow Y,$$

а локальная часть заканчивается на категорию $\alpha_r=[D/C]$ или $\alpha_{f(n)}=[T_{f(n_0)-1}/C]$. Следовательно, сначала после Γ_1 стоят несколько категорий с локальной частью $[C/C]$, а затем стоит одна категория с локальной частью $[C/D]$. Поэтому потенциал всей строки категорий от начала Γ_1 до конца добавленной части сокращается до

$$\nearrow X(\nearrow A_1)^{f_1(n)}(\nearrow A_2)^{f_2(n)}\ldots$$
$$\ldots(\nearrow A_k)^{f_k(n)}\nearrow YP\searrow X\nearrow X \quad (7)$$

для некоторого P. Так как валентность типа X препятствует сокращению валентностей типов A_i, то потенциал P содержит ровно

$f_i(n)$ валентностей $\searrow A_i$ для каждого i. Валентность $\searrow A_k$ содержится только в потенциале $P_{1,k}$, поэтому он встречается в P ровно $f_k(n)$ раз. Докажем индукцией по i, что $P_{1,i}$ встречается в P ровно $f_i(n)-f_{i+1}(n)$ раз. Валентность $\searrow A_{k-1}$ встречается только в потенциалах $P_{1,k}$ и $P_{1,k-1}$. Так как $f_k(n)$ таких валентностей уже содержатся в потенциалах $P_{1,k}$, то потенциал $P_{1,k-1}$ встречается $f_{k-1}(n)-f_k(n)$ раз. Валентность $\searrow A_i$ встречается только в потенциалах $P_{1,i},\ldots,P_{1,k}$. По индукционному предположению потенциалы $P_{1,i+1},\ldots,P_{1,k}$ встречаются

$$f_{i+1}(n)-f_{i+2}(n),\ldots,f_{k-1}(n)-f_k(n),f_k(n)$$

раз соответственно, а значит, они содержат

$$(f_{i+1}(n)-f_{i+2}(n))+\ldots+$$
$$+(f_{k-1}(n)-f_k(n))+f_k(n)=f_{i+1}(n)$$

вхождений валентности $\searrow A_i$. Так как общее число вхождений этой валентности равно $f_i(n)$, то $P_{1,i}$ встречается в P ровно $f_i(n)-f_{i+1}(n)$ раз.

Итак, мы доказали, что потенциал (7) равен потенциалу (6) из доказательства предыдущей леммы с точностью до порядка валентностей $\searrow A_i$. Следовательно, он также содержит ровно $f_i(n+1)$ валентностей $\nearrow B_i$, а соответствующая строка категорий имеет длину $f_1(n)$. Значит, общая длина строки категорий от начала слова до конца потенциала (7) равна $f(n)+f_1(n)=f(n+1)$.

Случай чётного $(n+1)-n_0$ рассматривается аналогично. □

Теперь мы можем доказать главный результат о грамматике G_f.

Теорема 1. *Для любого многочлена f с положительным старшим коэффициентом грамматика G_f порождает язык*

$$L(G_f)=\{a^{f(n)}:n\geq n_0+2\}.$$

Доказательство. Рассмотрим слово $w=a^{f(n)}$, где $n\geq n_0+2$. Пусть Γ — строка категорий длины $f(n-2)$ из леммы 4. Предположим для определённости, что число $(n-2)-n_0$ чётно, так что Γ заканчивается на $[D/C]$ или $[T_{f(n_0)-1}/C]$, а её потенциал сокращается до

$$\nearrow X(\nearrow A_1)^{f_1(n-2)}(\nearrow A_2)^{f_2(n-2)}\ldots(\nearrow A_k)^{f_k(n-2)}\nearrow Y.$$

Построим продолжение Γ_1 этой строки так же, как это было сделано в лемме 4, но в качестве последней категории используем не $[C/D]^{P_1\searrow X\nearrow X}$, а $[C/F]^{P_1\searrow X}$. Тогда потенциал строки $\Gamma\Gamma_1$ сократится до

$$\nearrow Y(\nearrow B_1)^{f_1(n-1)}(\nearrow B_2)^{f_2(n-1)}\ldots(\nearrow B_k)^{f_k(n-1)}.$$

Теперь построим продолжение Γ_2 получившейся строки, как в лемме 4, для случая нечётной разности, но заменим категории $[D/D]^{P_2}$ и $[D/C]^{P_2\searrow Y\nearrow Y}$ на категории $[F/F]^{P_4}$ и $[F]^{P_4\searrow Y}$ соответственно. Так как P_4 отличается от P_2 только тем, что не содержит положительных валентностей, то потенциал строки $\Gamma\Gamma_1\Gamma_2$ сбалансирован. Кроме того, непосредственно проверяется, что локальная часть последовательности $\Gamma\Gamma_1\Gamma_2$ сокращается до $[S]$. Так как $|\Gamma|=f(n-2)$, $|\Gamma_1|=f_1(n-2)$, $|\Gamma_2|=f_1(n-1)$, то

$$|\Gamma\Gamma_1\Gamma_2|=f(n-2)+f_1(n-2)+f_1(n-1)=$$
$$=f(n-1)+f_1(n-1)=f(n).$$

Это и значит, что $a^{f(n)}\in L(G_f)$. Случай нечётного $(n-2)-n_0$ рассматривается аналогично.

Теперь рассмотрим некоторое слово $w=a^r\in L(G_f)$, в котором $r\geq f(n_0+2)$. Тогда существует такая строка категорий $\Gamma\in\delta(w)$ длины r, что $\Gamma\vdash^*[S]$. Так как ни одна категория не имеет левого списка зависимостей, то все сокращения выполняются справа налево. Это значит, что последней категорией может быть только категория $[E]^{P_3\searrow X}$ или $[F]^{P_4\searrow Y}$. Предположим для определённости, что это $[F]^{P_4\searrow Y}$. Тогда перед последней категорией стоят p категорий вида $[F/F]^{P_4}$ для некоторого $p\geq 0$. Перед первой из этих категорий обязательно стоит категория $[C/F]^{P_1\searrow X}$, перед ней — q категорий вида $[C/C]^{P_1}$ для некоторого $q\geq 0$, а ещё раньше — категория $[D/C]^{P_2\searrow Y\nearrow Y}$. Обозначим через Γ_1 строку категорий от начала слова w до этой категории $[D/C]^{P_2\searrow Y\nearrow Y}$, через Γ_2 обозначим строку $([C/C]^{P_1})^q[C/F]^{P_1\searrow X}$, а через Γ_3 — строку $([F/F]^{P_4})^p[F]^{P_4\searrow Y}$. Заметим, что P_1 в строке Γ_2 обозначает различные потенциалы из определения грамматики G_f, и аналогично для P_4 в строке Γ_3. По лемме 5 потенциал строки Γ_1 сокращается до

$$\nearrow X(\nearrow A_1)^{f_1(n)}(\nearrow A_2)^{f_2(n)}\ldots(\nearrow A_k)^{f_k(n)}\nearrow Y$$

для некоторого $n \geq n_0$. Следовательно, строка Γ_2 имеет длину $f_1(n)$, а строка Γ_3 — длину $f_1(n+1)$ (это устанавливается так же, как в доказательстве леммы 5). Значит, общая длина строки Γ равна

$$f(n) + f_1(n) + f_1(n+1) =$$
$$= f(n+1) + f_1(n+1) = f(n+2),$$

поэтому $w = a^{f(n+2)}$. □

Грамматику G_f можно модифицировать так, чтобы она задавала все положительные значения произвольного многочлена.

Теорема 2. *Для любого многочлена $f(n)$ следующие языки являются ммКГЗ-языками (через Z обозначено множество целых чисел):*

1) $L_1 = \{a^{f(n)} : n \geq 0, f(n) > 0\}$;

2) $L_2 = \{a^{f(n)} : n \in Z, f(n) > 0\}$;

3) $L_3 = \{a^{|f(n)|} : n \in Z, f(n) \neq 0\}$.

Доказательство. Если старший коэффициент многочлена отрицателен, то язык L_1 конечен, так как все значения $f(n)$ отрицательны, начиная с некоторого значения аргумента n_0. Если старший коэффициент положителен, то пусть n_0 — число из леммы 3. Тогда по теореме 1 язык

$$L_1{}' = \{a^{f(n)} : n \geq n_0 + 2\}$$

является ммКГЗ-языком. Обозначим через $L_1{}''$ множество оставшихся слов, принадлежащих языку L_1:

$$L_1{}'' = \{a^i \in L_1 : 1 \leq i \leq f(n_0 + 2) - 1\}.$$

Так как язык $L_1{}''$ конечен, то он порождается некоторой ммКГЗ. Но класс ммКГЗ-языков замкнут относительно объединения (см. [6, 12]), поэтому язык $L_1 = L_1{}' \cup L_1{}''$ также является ммКГЗ-языком.

Язык L_2 можно представить в виде

$$L_2 = \{a^{f(n)} : n \geq 0, f(n) > 0\} \cup$$
$$\cup \{a^{f(-n)} : n < 0, f(-n) > 0\}.$$

Оба члена объединения являются ммКГЗ-языками согласно пункту 1 настоящей теоремы, а значит, L_2 также является ммКГЗ-языком в силу замкнутости класса ммКГЗ-языков относительно объединения.

Язык L_3 можно представить в виде объединения

$$L_3 = \{a^{f(n)} : n \in Z, f(n) > 0\} \cup$$
$$\cup \{a^{-f(n)} : n \in Z, -f(n) > 0\}.$$

Тогда L_3 является ммКГЗ-языком, так как согласно предыдущему пункту оба члена объединения — ммКГЗ-языки. □

Построенные грамматики позволяют задавать только значения многочлена. В следующей теореме показано, как с помощью ммКГЗ задать слова, хранящие как аргумент, так и соответствующее значение многочлена.

Теорема 3. *Для любого многочлена $f(n)$ язык $L = \{b^n a^{f(n)} : n > 0, f(n) > 0\}$ является ммКГЗ-языком.*

Доказательство. Если старший коэффициент многочлена $f(n)$ отрицателен, то L — конечный язык, а значит, и ммКГЗ-язык. Рассмотрим случай положительного старшего коэффициента. Достаточно построить ммКГЗ для языка L с дополнительным ограничением $n \geq n_0 + 2$. Множество элементарных категорий грамматики содержит все элементарные категории грамматики G_f, а также три новые категории S_0, R и V. Главной категорией является S_0. Сопоставим символу b следующие категории:

$$b \mapsto [S_0 / R]^{\nearrow V}, [R / R]^{\nearrow V}, [R / S]^{\nearrow V}, [S_0 / S]^{\nearrow V}.$$

Символу a сопоставим категории из грамматики G_f со следующими изменениями (здесь n_0 — число из леммы 3):

• категория $[S / T_1]^{\nearrow X P_0 \nearrow Y}$ заменяется на категорию $[S / T_1]^{(\searrow V)^{n_0} \nearrow X P_0 \nearrow Y}$;

• перед каждой валентностью $\searrow X$ и $\searrow Y$ добавляется валентность $\searrow V$. Например, если в G_f есть потенциал $\searrow A_1 \nearrow B_1 \searrow X \nearrow X$, то он заменяется на $\searrow A_1 \nearrow B_1 \searrow V \searrow X \nearrow X$.

Любое слово, порождаемое этой грамматикой, имеет вид $b^n a^m$ для некоторых n и m. Категории для символов a сокращаются так же, как в грамматике G_f, но они дополнительно содержат отрицательные валентности $\searrow V$. Как мы уже показали в доказательстве лемм 4 и 5, слово a^m разбивается на блоки с длинами $f(n_0)$, $f_1(n_0)$, $f_1(n_0 + 1), \ldots, f_1(n_0 + p)$ для некоторого p. В новой грамматике первый блок ловит n_0 валентностей типа V, а остальные $p+1$ блоков — по одной такой валентности. Так как категории слова b^n выпускают направо в сумме n валентностей типа V, то $n = n_0 + p + 1$, откуда $p = n - n_0 - 1$. Следовательно,

$$m = f(n_0) + f_1(n_0) + f_1(n_0+1) + \ldots + f_1(n-1) = \\ = f(n).$$

Наоборот, рассмотрим слово $b^n a^{f(n)}$. Припишем символам b строку категорий

$$[S_0 / R]^{\nearrow V}([R/R]^{\nearrow V})^{n-2}[R/S]^{\nearrow V},$$

а символам a — строку категорий из доказательства теоремы 1 с заменой категорий на соответствующие им в новой грамматике. Валентности $\nearrow V$ и $\searrow V$ сокращаются, так как их количество одинаково, а оставшаяся часть строки категорий сокращается до $[S]$, как это было показано в доказательстве теоремы 1. □

Наконец, покажем, что ммКГЗ позволяют задавать и некоторый вариант функций, являющихся обратными к многочленам. Для многочлена $f(n)$ с положительным старшим коэффициентом обозначим через $f^{-1}(m)$ функцию, определяемую следующим образом:

$$f^{-1}(m) = \max\{n \geq 0 : f(n) \leq m\}.$$

Теорема 4. *Для любого многочлена с положительным старшим коэффициентом* $f(n)$ *язык* $L = \{b^{f^{-1}(m)} a^m : m > 0, f^{-1}(m) > 0\}$ *является ммКГЗ-языком.*

Доказательство. Пусть n_0 — число из леммы 3. При необходимости увеличим его так, чтобы при всех $n \geq n_0$ функция $f(n)$ строго возрастала. Любое слово языка L имеет вид $b^n a^{f(n)+p}$, где $p \geq 0$ — такое целое число, что $f(n) + p < f(n+1)$. Мы модифицируем грамматику из теоремы 3 так, чтобы в конце слова мог находиться ещё один блок, длина p которого удовлетворяет двойному неравенству $0 \leq p \leq f_1(n) - 1$. Для этого удалим из грамматики все категории вида $[E/E]^{P_3}, [E]^{P_3 \searrow X}, [F/F]^{P_4}, [F]^{P_4 \searrow Y}$ и добавим следующие категории:

$$a \mapsto [E/E]^{P_3'}, [E]^{P_3'' \searrow X}, [E/E_1]^{P_3'' \searrow X}, \\ [E_1/E_1]^{\searrow B_1}, [E_1]^{\searrow B_1};$$

$$a \mapsto [F/F]^{P_4'}, [F]^{P_4'' \searrow Y}, [F/F_1]^{P_4'' \searrow Y}, \\ [F_1/F_1]^{\searrow A_1}, [F_1]^{\searrow A_1}.$$

В этих категориях потенциалы P_3' и P_4' принимают следующие значения:

$$P_{3,1}' = \searrow A_1 (\nearrow B_1)^s;$$
$$P_{3,2}' = \searrow A_1 \searrow A_2 (\nearrow B_1)^t;$$
$$P_{3,3}' = \searrow A_1 \searrow A_2 \searrow A_3 (\nearrow B_1)^t;$$
$$\ldots$$
$$P_{3,k}' = \searrow A_1 \ldots \searrow A_{k-1} \searrow A_k (\nearrow B_1)^t;$$
$$P_{4,1}' = \searrow B_1 (\nearrow A_1)^s;$$
$$P_{4,2}' = \searrow B_1 \searrow B_2 (\nearrow A_1)^t;$$
$$P_{4,3}' = \searrow B_1 \searrow B_2 \searrow B_3 (\nearrow A_1)^t;$$
$$\ldots$$
$$P_{4,k}' = \searrow B_1 \ldots \searrow B_{k-1} \searrow B_k (\nearrow A_1)^t.$$

Здесь $s \in \{0,1\}, t \in \{0,1,2\}$. Потенциалы P_3'' и P_4'' строятся аналогично, но s и t принимают значения $s = 0$, $t \in \{0,1\}$.

Так как в потенциалах допустим случай $s = 0$, $t = 0$, то категории, содержащие E_1 или F_1, могут вообще не использоваться. Это позволяет порождать слова вида $b^n a^{f(n)}$ при $p = 0$. Пусть $p > 0$. Тогда заключительный блок, содержащий категории E или F, может породить от 1 до $f_1(n) - 1$ валентностей вида $\nearrow A_1$ или $\nearrow B_1$, так как потенциалы порождают таких валентностей не больше, чем в «обычном» блоке, а потенциал последней категории порождает хотя бы на одну валентность меньше. Все эти валентности будут пойманы категориями с головами E_1 или F_1, поэтому потенциал оказывается сбалансированным. Такой способ сопоставления категорий символам является единственно возможным, поэтому любое слово, порождаемое грамматикой, имеет вид $b^n a^{f(n)+p}$. □

ЗАКЛЮЧЕНИЕ

В работе были исследованы выразительные возможности мультимодальных категориальных грамматик зависимостей. Было доказано, что для любого многочлена $f(n)$ существует ммКГЗ, порождающая в точности те слова, длины которых являются положительными значениями этого многочлена. Также было показано, что существует ммКГЗ, выполняющая вычисление значения многочлена на заданном аргументе, а также ммКГЗ, выполняющая вычисление обратной к многочлену функции. Дальнейшие исследования в этом направлении могут включать поиск более широких классов функций, представимых с помощью ммКГЗ, а также исследование выразительных возможностей КГЗ без ограничений на пересечение непроективных зависимостей.

КОНФЛИКТ ИНТЕРЕСОВ

Автор декларирует отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

СПИСОК ЛИТЕРАТУРЫ

1. *Гладкий А. В.* Формальные грамматики и языки / А. В. Гладкий. – М. : Наука, 1973. – 368 с.

2. *Kuhlmann M.* The Tree-Generative Capacity of Combinatory Categorial Grammars / M. Kuhlmann, A. Maletti, L. K. Schiffer // Proc. Foundations of Software Technology and Theoretical Computer Science. – Vol. 150 of LIPIcs. – Dagstuhl, Germany : Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2019. – P. 44:1–44:14.

3. *Kuhlmann M.* The Tree-Generative Capacity of Combinatory Categorial Grammars / M. Kuhlman, A. Maletti, L. K. Schiffer // Journal of Computer and System Sciences. – 2022. – Vol. 124. – P. 214–233. doi: https://doi.org/10.1016/j.jcss.2021.10.005

4. *Schiffer L. K.* Strong Equivalence of TAG and CCG / L. K. Schiffer, A. Maletti // Transactions of the Association for Computational Linguistics. – 2021. – Vol. 9 – P. 707–720. doi: https://doi.org/10.1162/tacl_a_00393

5. *Dekhtyar M.* Generalized Categorial Dependency Grammars / M. Dekhtyar, A. Dikovsky // Pillars of Computer Science. Lecture Notes in Computer Science, vol 4800. – Berlin, Heidelberg : Springer, 2008. – P. 230–255. doi: https://doi.org/10.1007/978-3-540-78127-1_13

6. *Dikovsky A.* Multimodal Categorial Dependency Grammars / A. Dikovsky // Proc. of the 12th Conference on Formal Grammar (FG 2007), Dublin, Ireland. – 2007. – P. 1–12.

7. *Béchet D.* Incremental learning of iterated dependencies / D. Béchet, A. Foret // Machine Learning. – 2021. https://doi.org/10.1007/s10994-021-05947-2

8. *Dikovsky A.* Categorial Dependency Grammars: from Theory to Large Scale Grammars / A. Dikovsky // Proc. of the 1st Intern. Conf. on Dependency Linguistics (Depling 2011), Barcelona, Spain. – 2011. – P. 262–271.

9. *Foret A.* On Categorial Grammatical Inference and Logical Information Systems / A. Foret, D. Béchet // Logic and Algorithms in Computational Linguistics 2018. – Cham : Springer, 2020. – P. 125–153. doi: https://doi.org/10.1007/978-3-030-30077-7_6

10. *Foret A.* Iterated Dependencies in a Breton Treebank and Implications for a Categorial Dependency Grammar / A. Foret, D. Béchet, V. Bellynck // Proc. of the 4th Celtic Language Technology Workshop within LREC2022, Marseille, France. – Paris : European Language Resources Association, 2022. – P. 40–46.

11. *Карлов Б. Н.* О свойствах языков, задаваемых мультимодальными категориальными грамматиками зависимостей / Б. Н. Карлов // Вестник ТвГУ. Серия: Прикладная математика. – 2011. – № 22. – С. 91–110.

12. *Dekhtyar M.* Iterated Dependencies and Kleene Iteration / M. Dekhtyar, A. Dikovsky, B. Karlov // Proc. of the 15th Conference on Formal Grammar (FG 2010), Copenhagen, Denmark. Series: LNCS. 2012. Vol. 7395. – Berlin, Heidelberg : Springer, 2012. – P. 66–81. doi: https://doi.org/10.1007/978-3-642-32024-8_5

13. *Dekhtyar M.* Categorial Dependency Grammars / M. Dekhtyar, B. Karlov, A. Dikovsky // Theoretical Computer Science. – 2015. – Vol. 579. – P. 33–63. doi: https://doi.org/10.1016/j.tcs.2015.01.043

14. *Ginsburg S.* The Mathematical Theory of Context-Free Languages. – New York : McGraw-Hill, Inc., 1966. – 232 p.

15. *Joshi A. K.* Tree-Adjoining Grammars / A. K. Joshi, Y. Schabes // Handbook of Formal Languages. – Berlin, Heidelberg : Springer, 1997. – P. 69–123. doi: https://doi.org/10.1007/978-3-642-59126-6_2

Карлов Борис Николаевич — канд. физ.-мат. наук, доц., доцент кафедры информатики Тверского государственного университета. E-mail: bnkarlov@gmail.com
ORCID: 0000-0002-4340-2435

DOI: https://doi.org/10.17308/sait/1995-5499/2023/3/156-166 ISSN 1995-5499
Received 28.03.2023
Accepted 30.09.2023

REPRESENTATION OF POLYNOMIALS BY MULTIMODAL CATEGORIAL DEPENDENCY GRAMMARS

© 2023 B. N. Karlov✉

Tver State University,
33, Zhelyabova Street, 170100 Tver, Russian Federation

Annotation. Categorial dependency grammars are generalization of classical categorical grammars. Unlike context-free grammars, they describe the sentence structure not by using constituents, but by using syntactic dependencies between words. These grammars are very expressive, they allow to generate many non-context-free languages. The expressive power of one of the variants of these grammars is studied in this paper – multimodal categorical dependency grammars. It was established earlier that they can generate non-semilinear languages. The possibility of representation of polynomials by such grammars is studied in this paper. It is proved that for every polynomial there exists a grammar generating the set of all positive values of this polynomial encoded in unary system. The set of absolute values of all non-zero values of an arbitrary polynomial can be represented in an analogous way. It is established that it is possible to construct grammars performing calculation of the value of a polynomial by a given argument, and calculation of the argument by a given value of a polynomial.

Keywords: category, polarized valency, multimodal categorical dependency grammar, non-projective dependency, formal language, polynomial.

CONFLICT OF INTEREST

The author declare the absence of obvious and potential conflicts of interest related to the publication of this article.

REFERENCES

1. *Gladkiy A. V.* (1973) Formal'nye grammatiki i yazyki [Formal grammars and languages]. Moscow, Nauka. 368 p. (in Russian)

2. *Kuhlmann M., Maletti, A. and Schiffer L. K.* (2019) The Tree-Generative Capacity of Combinatory Categorial Grammars. *In Proc. Foundations of Software Technology and Theoretical Computer Science*. 150 of LIPIcs. Dagstuhl, Germany, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, P. 44:1–44:14.

3. *Kuhlmann M., Maletti A. and Schiffer L. K.* (2022) The Tree-Generative Capacity of Combinatory Categorial Grammars. *Journal of Computer and System Sciences*. 124. P. 214–233. doi: https://doi.org/10.1016/j.jcss.2021.10.005

4. *Schiffer L. K. and Maletti A.* (2021) Strong Equivalence of TAG and CCG. *Transactions of the Association for Computational Linguistics*. 9. P. 707–720. doi:https://doi.org/10.1162/tacl_a_00393

5. *Dekhtyar M. and Dikovsky A.* (2008) Generalized Categorial Dependency Grammars. *Pillars of Computer Science. Lecture Notes in Computer Science*. Vol 4800. Berlin, Heidelberg, Springer. P. 230–255. doi: https://doi.org/10.1007/978-3-540-78127-1_13

6. *Dikovsky A.* (2007) Multimodal Categorial Dependency Grammars. *In Proc. of the 12th Conference on Formal Grammar (FG 2007), Dublin, Ireland*. P. 1–12.

7. *Béchet D. and Foret A.* (2021) Incremental learning of iterated dependencies. *Machine Learning*. https://doi.org/10.1007/s10994-021-05947-2

8. *Dikovsky A.* (2011) Categorial Dependency Grammars: from Theory to Large Scale Grammars. *In Proc. of the 1st Intern. Conf. on Dependency Linguistics (Depling 2011), Barcelona, Spain*. P. 262–271.

9. *Foret A. and Béchet D.* (2020) On Categorial Grammatical Inference and Logical Information

✉ Karlov Boris N.
e-mail: bnkarlov@gmail.com

Systems. *Logic and Algorithms in Computational Linguistics 2018*. Cham, Springer. P. 125–153. doi: https://doi.org/10.1007/978-3-030-30077-7_6

10. *Foret A., Béchet D. and Bellynck V.* (2022) Iterated Dependencies in a Breton Treebank and Implications for a Categorial Dependency Grammar. *In Proc. of the 4th Celtic Language Technology Workshop within LREC2022, Marseille, France*. Paris, European Language Resources Association. P. 40–46.

11. *Karlov B. N.* (2011) O svoistvakh yazykov, zadavaemykh mul'timodal'nymi kategorial'nymi grammatikami zavisimostei [On properties of languages generated by multimodal categorical dependency grammars]. *Vestnik TvGU. Seria: Prikladnaya matematika*. (22). P. 91–110. (in Russian)

12. *Dekhtyar M., Dikovsky A. and Karlov B.* (2012) Iterated Dependencies and Kleene Iteration. *In Proc. of the 15th Conference on Formal Grammar (FG 2010), Copenhagen, Denmark. Series: LNCS*. 2012. Vol. 7395. Berlin, Heidelberg, Springer, P. 66–81. doi: https://doi.org/10.1007/978-3-642-32024-8_5

13. *Dekhtyar M., Karlov B. and Dikovsky A.* (2015) Categorial Dependency Grammars. *Theoretical Computer Science*. 579. P. 33–63. doi: https://doi.org/10.1016/j.tcs.2015.01.043

14. *Ginsburg S.* (1966) The Mathematical Theory of Context-Free Languages. New York, McGraw-Hill, Inc. 232 p.

15. *Joshi A. K. and Schabes Y.* (1997) Tree-Adjoining Grammars. *Handbook of Formal Languages*. Berlin, Heidelberg, Springer. P. 69–123. doi: https://doi.org/10.1007/978-3-642-59126-6_2

Karlov Boris N. — PhD, associate professor of the department of computer science, Tver State University.
E-mail: bnkarlov@gmail.com
ORCID: 0000-0002-4340-2435