УДК 62.506.2

М. Ф. БОНДАРЕНКО, канд. техн. наук

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ СЛОВОИЗМЕНЕНИЯ ИМЕН ЧИСЛИТЕЛЬНЫХ

Цель данной статьи — описание процессов грамматической обработки русских числительных в терминах алгебры конечных предикатов [1]. Среди наиболее важных задач, составляющих понятие грамматической обработки числительных, можно указать следующие: синтез словоформ, их анализ, нормализацию (представление словоформ числительных в числовом виде). Использование алгебры конечных предикатов позволяет построить уравнения, описывающие какую-нибудь одну процедуру, и применять эти уравнения для выполнения других видов грамматической обработки. В этой статье в качестве первоначальной задачи выбран синтез словоформ числительных.

Входными данными в задаче синтеза является число в десятичном виде, т. е. в виде слова $C=c_1c_2 \ldots c_n$ в алфавите из десятичных цифр $0, 1, \ldots, 9$ (в записи слова меньшие индексы соответствуют старшим разрядам); другую часть входных данных составляет набор грамматических признаков $z_1, z_2, \ldots, z_6$. В результате синтеза выдается числительное, записанное в слове $Y=y_1y_2 \ldots y_m$ в русском алфавите.

На числительные, изучаемые в данной статье, накладываются некоторые ограничения: 1) не рассматриваются порядковые числительные, поскольку их проще рассматривать вместе с прилагательными; 2) не рассматриваются числительные, описывающие числа больше 1000, поскольку это привело бы лишь к увеличению формул, не добавляя ничего в идейном плане; 3) не рассматриваются составные числительные, т. е. состоящие из нескольких слов, поскольку это — самостоятельная следующая задача, которую нужно решать, опираясь на результаты данной работы.

Эти ограничения оставляют количественные числительные до тысячи включительно, простые (с одной основой) и сложные (с двумя основами), но зато их морфология изучается во всех подробностях. Поэтому в наших рассмотрениях $n=4$, а m с запасом принято равным 20; запас нужен, поскольку в ходе синтеза слова $Y=y_1y_2\ldots y_m$ понадобятся «рабочие ячейки». Если число не занимает всего слова C или числительное не занимает всего слова Y, то свободные места справа заняты пробелами (обозначаются «$\smile$»). Точнее, перед началом обработки все слова очищаются пробелами, а занесения производятся слева.

На первом шаге синтеза по числу C определяются цифровые значения основы $\tau_1, \tau_2, \ldots, \tau_n$ или σ. В τ_i записываются цифровые значения простых или препозиционных основ, а в σ — значения постпозиционных основ. Символы τ_i повторяют c_i, $i=1, 2, \ldots, n$. Например, число 1000 в терминах τ запишется так: $\tau_1=1$, $\tau_2=0$, $\tau_3=0$, $\tau_4=0$. В σ записывается число нолей в постпозиционной основе. Например, -десят в «пятьдесят» даст $\sigma=1$, а -сот в «пятьсот» даст $\sigma=2$.

На языке конечных предикатов нужные соотношения записываются следующим образом:

$$(c_1^5 \vee c_1^6 \vee c_1^7 \vee c_1^8)\, c_2^0 c_3^{\smile} \sim (\tau_1 = c_1)(\tau_2 = \smile)(\sigma = 1); \qquad (1)$$

$$A c_2^0 c_3^0 c_4^{\smile} \sim (\tau_1 = c_1)(\tau_2 = \smile)(\sigma = 2)(z_4 = \text{н})(z_2 = \text{ж}); \qquad (2)$$

$$A c_2^{\smile} \vee c_1^1 \vee (c_1^2 \vee c_1^3 \vee c_1^4 \vee c_1^9)\, c_2^0 c_4^{\smile} \sim (\tau_i = c_i)(\sigma = \smile). \qquad (3)$$

Здесь

$$A = c_1^2 \vee c_1^3 \vee c_1^4 \vee c_1^5 \vee c_1^6 \vee c_1^7 \vee c_1^8 \vee c_1^9,$$

в выражении (2) два последних члена справа характеризуют склонение препозиционной основы (см. выражения (6) и (8) ниже).

Будем обозначать через $x_1, x_2, \ldots, x_k$ буквенные выражения основ простых числительных или препозиционных основ сложных числительных, а через $u_1, u_2, \ldots, u_k$ — буквенные значения постпозиционных основ. Эти соотношения записываются следующей формулой:

$$x_1^0 x_2^{\text{д}} (x_3^{\text{и}} x_4^{\text{н}} \vee x_3^{\text{н}} x_4^{\smile})\, x_5^{\smile} \tau_1^1 \tau_2^{\smile} \vee x_1^{\text{д}} x_2^{\text{в}} x_3^{\smile} \tau_1^2 \tau_2^{\smile} \vee$$

$$\vee\, x_1^{\text{т}} x_2^{\text{р}} x_3^{\smile} \tau_1^3 \tau_2^{\smile} \vee x_1^{\text{ч}} x_2^{\text{е}} x_3^{\text{т}} x_4^{\text{ы}} x_5^{\text{р}} x_6^{\smile} \tau_1^4 \tau_2^{\smile} \vee x_1^{\text{п}} x_2^{\text{я}} x_3^{\text{т}} x_4^{\smile} \tau_1^5 \tau_2^{\smile} \vee$$

$$\vee\, x_1^{\text{ш}} x_2^{\text{е}} x_3^{\text{с}} x_4^{\text{т}} x_5^{\smile} \tau_1^6 \tau_2^{\smile} \vee x_1^{\text{с}} x_2^{\text{е}} x_3^{\text{м}} x_4^{\smile} \tau_1^7 \tau_2^{\smile} \vee x_1^{\text{в}} x_2^0 x_3^{\text{с}} (x_4^{\text{е}} \vee x_4^{\text{ь}})\, x_5^{\text{м}} x_6^{\smile} \tau_1^8 \tau_2^{\smile} \vee$$

$$\vee\, x_1^{\text{д}} x_2^{\text{е}} x_3^{\text{в}} x_4^{\text{я}} x_5^{\text{т}} x_6^{\smile} \tau_1^9 \tau_2^{\smile} \vee x_1^{\text{д}} x_2^{\text{е}} x_3^{\text{с}} x_4^{\text{я}} x_5^{\text{т}} x_6^{\smile} \tau_1^1 \tau_2^0 \tau_3^{\smile} \vee$$

$$
\begin{gathered}
\vee\, ((x_1^{\text{д}} x_2^{\text{в}} x_3^{\text{е}} \tau_2^{2} \vee x_1^{\text{т}} x_2^{\text{р}} x_3^{\text{и}} \tau_2^{\sqcup} \vee x_1^{\text{п}} x_2^{\text{я}} x_3^{\text{т}} \tau_2^{5} \vee \\
\vee\, x_1^{\text{с}} x_2^{\text{е}} x_3^{\text{м}} \tau_2^{7})\, x_4^{\text{н}} x_5^{\text{а}} x_6^{\text{д}} x_7^{\text{ц}} x_8^{\text{а}} x_9^{\text{т}} x_{10}^{\sqcup} \vee \\
\vee\, (x_1^{0} x_2^{\text{д}} x_3^{\text{и}} x_4^{\text{н}} \tau_2^{1} \vee x_1^{\text{ш}} x_2^{\text{е}} x_3^{\text{с}} x_4^{\text{т}} \tau_2^{6})\, x_5^{\text{н}} x_6^{\text{а}} x_7^{\text{д}} x_8^{\text{ц}} x_9^{\text{а}} x_{10}^{\text{т}} x_{11}^{\sqcup} \vee \\
\vee\, (x_1^{4} x_2^{\text{е}} x_3^{\text{т}} x_4^{\text{ы}} x_5^{\text{р}} \tau_2^{4} \vee x_1^{\text{в}} x_2^{0} x_3^{\text{с}} x_4^{\text{е}} x_5^{\text{м}} \tau_2^{8} \vee \\
\vee\, x_1^{\text{д}} x_2^{\text{е}} x_3^{\text{в}} x_4^{\text{я}} x_5^{\text{т}} \tau_2^{9})\, x_6^{\text{н}} x_7^{\text{а}} x_8^{\text{д}} x_9^{\text{ц}} x_{10}^{\text{а}} x_{11}^{\text{т}} x_{12}^{\sqcup} \tau_1^{1} \tau_3^{\sqcup} \vee \\
\vee\, ((x_1^{\text{д}} x_2^{\text{в}} x_3^{\text{а}} \tau_1^{2} \vee x_1^{\text{т}} x_2^{\text{р}} x_3^{\text{и}} \tau_1^{3})\, x_4^{\text{д}} x_5^{\text{ц}} x_6^{\text{а}} x_7^{\text{т}} x_8^{\sqcup} \vee \\
\vee\, x_1^{\text{с}} x_2^{0} x_3^{\text{р}} x_4^{0} x_5^{\text{к}} x_6^{\sqcup} \tau_1^{4} \vee x_1^{\text{д}} x_2^{\text{е}} x_3^{\text{в}} x_4^{\text{я}} x_5^{\text{н}} x_6^{0} x_7^{\text{с}} x_8^{\text{т}} x_9^{\sqcup} \tau_1^{9}) \wedge \\
\wedge\, \tau_2^{0} \tau_3^{\sqcup} \vee x_1^{\text{с}} x_2^{\text{т}} x_3^{\sqcup} \tau_1^{1} \tau_2^{0} \tau_3^{0} \tau_4^{\sqcup} \vee \tau_1^{1} \tau_2^{0} \tau_3^{0} \tau_4^{0} \tau_5^{\sqcup} x_1^{\text{т}} x_2^{\text{ы}} x_3^{\text{с}} \wedge \\
\wedge\, x_4^{\text{я}} x_5^{\text{ч}} x_6^{\sqcup} \vee u_1^{\text{д}} u_2^{\text{е}} u_3^{\text{с}} u_4^{\text{я}} u_5^{\text{т}} u_6^{\sqcup} \sigma^{1} \vee (u_1^{\text{с}} u_2^{0} u_3^{\text{т}} u_4^{\sqcup} \vee u_1^{\text{с}} u_2^{\text{т}} u_3^{\sqcup})\, \sigma^{2}. \qquad (4)
\end{gathered}
$$

Приведем соотношения, которые характеризуют связь основ и окончаний количественных числительных в зависимости от грамматических признаков $z_1 \dots z_6$. При этом следует иметь в виду, что словоформы различных числительных характеризуются различными наборами грамматических признаков. Например, числительное «пять» изменяется только по падежам, а числительное «один» — по падежам, родам, числам и имеет противопоставление словоформ по признаку одушевленности или неодушевленности. Приводимые ниже уравнения учитывают всю эту пестроту морфологии числительных. Условно считается, что окончание любой словоформы числительного, т. е. окончание простого числительного или основы в препозиции в сложном числительном, состоит из трех символов s_1, s_2 и s_3. Например, для словоформы «двумя» — $s_1=$у, $s_2=$м, $s_3=$я, а для словоформы $s_1=$у, $s_2=$х, $s_3=\sqcup$ — «двухсот».

Символы, образующие окончание постпозиционных основ в сложных числительных, обозначим T_1, T_2, T_3. Например, для словоформы «двумястами» $T_1=$а, $T_2=$м, $T_3=$и, а для «двухсот» — $T_1=\sqcup$, $T_2=\sqcup$, $T_3=\sqcup$.

При формализации словоизменения числительных можно описывать условия выбора сразу трех символов окончания s_1, s_2, s_3 и/или T_1, T_2, T_3 в зависимости от лексического (числового) значения основы и грамматических признаков словоформы. Но оказалось, что, как и в других частях речи [2], описание будет более экономным, если описывать условия употребления той или иной буквы в отдельных позициях окончаний s_l и T_l ($l=1$, 2, 3). Приводимые ниже уравнения построены именно по этому принципу. Они связывают буквенные значения отдельных позиций в окончаниях числительных с выражаемыми ими грамматическими категориями и особенностями примыкающих основ.

Уравнения, описывающие области определения переменных для обозначения грамматических категорий числительных:

$$z_1^{\text{и}} \vee z_1^{\text{р}} \vee z_1^{\text{д}} \vee z_1^{\text{в}} \vee z_1^{\text{т}} \vee z_1^{\text{п}} \quad (5); \qquad z_2^{\text{м}} \vee z_2^{\text{ж}} \vee z_2^{\text{с}} \quad (6);$$

$$z_3^{\text{Е}} \vee z_3^{\text{м}}\ \ \mathbf{(7)};\quad z_4^{0} \vee z_4^{\text{н}}\ \ \mathbf{(8)};\quad z_5^{\text{с}} \vee z_5^{\text{н}}\ \ \mathbf{(9)}.$$

Здесь z_1 — переменная для обозначения падежа словоформ; z_2 — переменная для обозначения рода; z_3 — переменная для обозначения числа; z_4 — для обозначения признака одушевленности (неодушевленности); z_5 — для обозначения признака старой или новой формы слова.

Уравнения, описывающие третью букву s_3 или T_3 в окончаниях числительных s_1, s_2, s_3 или T_1, T_2, T_3:

$$s_3^{\smile} \vee s_3^{0} \vee s_3^{\text{у}} \vee s_3^{\text{я}} \vee s_3^{\text{и}}\ \ \mathbf{(10)};\quad T_3^{\text{и}} \vee T_3^{\smile}\ \ \mathbf{(11)};$$

$$s_3^{0} = \left(z_1^{\text{р}} z_2^{\text{м}} \vee z_1^{\text{в}} z_4^{0} z_2^{\text{м}} \vee z_1^{\text{р}} z_2^{\text{с}}\right) z_3^{\text{Е}} x_3^{\text{н}} \tau_1^{1} \tau_2^{\smile}; \tag{12}$$

$$s_3^{\text{у}} = z_1^{\text{д}} z_2^{\text{ж}} z_3^{\text{е}} x_3^{\text{н}} \tau_1^{1} \tau_2^{\smile}\ \ \mathbf{(13)};\quad s_3^{\text{я}} = z_1^{\text{т}} \left(\tau_1^{2} \vee \tau_1^{3}\right) \tau_2^{\smile}\ \ \mathbf{(14)};$$

$$s_3^{\text{и}} \sigma^{\smile} \vee T_3^{\text{и}} \sigma^{2} = z_1^{\text{р}} \left(\sigma^{2} \vee z_3^{\text{м}} \tau_4^{0}\right). \tag{15}$$

Не обязательно выписывать в явном виде уравнения, описывающие условия употребления пробела ($\smile$) в качестве значения s_3 или T_3, так как это условие может быть определено из равенств (10) и (11) в совокупности с (12) — (15).

Уравнения, описывающие вторую букву s_2 или T_2 в окончаниях числительных s_1, s_2, s_3 и T_1, T_2, T_3:

$$s_2^{\smile} \vee s_2^{\text{р}} \vee s_2^{\text{м}} \vee s_2^{\text{х}} \vee s_2^{\text{й}} \vee s_2^{\text{ю}} \vee s_2^{\text{в}}\ \ \mathbf{(16)};\quad T_2^{\text{м}} \vee T_2^{\text{х}} \vee T_2^{\text{ю}} \vee T_2^{\smile}\ \ \mathbf{(17)};$$

$$s_2^{\text{р}} = z_3^{\text{Е}} \tau_1^{1} \tau_2^{\smile} \left(z_2^{\smile} \left(z_2^{\text{м}} \left(z_1^{\text{р}} \vee z_1^{\text{в}} z_4^{0}\right) \vee z_1^{\text{р}} z_2^{\text{с}};\right.\right. \tag{18}$$

$$s_2^{\text{м}} = \left(z_3^{\text{Е}} \left(z_2^{\text{с}} \vee z_2^{\text{м}}\right) \left(z_1^{\text{д}} \vee z_1^{\text{т}} \vee z_1^{\text{п}}\right) \vee z_2^{\text{н}} z_3^{\text{м}} \left(z_1^{\text{д}} \vee\right.\right.$$
$$\left.\left.\vee\, z_1^{\text{т}}\right)\right) \tau_1^{1} \tau_2^{\smile} \vee \left(z_1^{\text{д}} \vee z_1^{\text{т}}\right) \left(\tau_1^{2} \vee \tau_1^{3} \vee \tau_1^{4}\right) \tau_2^{\smile} \vee \left(z_1^{\text{д}} \vee z_1^{\text{т}}\right) z_3^{\text{м}} \tau_4^{0}; \tag{19}$$

$$s_2^{\text{х}} = \left(z_1^{\text{р}} \vee z_1^{\text{п}}\right) z_3^{\text{м}} \tau_1^{1} \tau_2^{\smile} \vee \left(z_1^{\text{р}} \vee z_1^{\text{в}} z_4^{0} \vee z_1^{\text{п}}\right) \left(\tau_1^{2} \vee\right.$$
$$\left.\vee\, \tau_1^{3} \vee \tau_1^{4}\right) \tau_2^{\smile} \vee z_1^{\text{п}} \tau_4^{0} z_3^{\text{м}}; \tag{20}$$

$$s_2^{\text{й}} = \left(z_1^{\text{р}} \vee z_1^{\text{д}} \vee z_1^{\text{т}} z_5^{\text{н}} \vee z_1^{\text{п}}\right) z_2^{\text{ж}} z_3^{\text{Е}} \tau_1^{1} \tau_2^{\smile}; \tag{21}$$

$$s_2^{\text{ю}} = z_1^{\text{т}} \left(z_2^{\text{ж}} z_3^{\text{Е}} z_5^{\text{с}} \tau_1^{1} \tau_2^{\smile}\right) \vee \left(\tau_1^{5} \vee \tau_1^{6} \vee \tau_1^{7} \vee\right.$$
$$\left.\vee\, \tau_1^{8} \vee \tau_1^{9}\right) \tau_2^{\smile} \vee \tau_1^{1} \tau_3^{\smile} \left(\tau_2^{0} \vee \tau_2^{1} \vee \tau_2^{2} \vee \tau_2^{3} \vee\right.$$
$$\left.\vee\, \tau_2^{4} \vee \tau_2^{5} \vee \tau_2^{6} \vee \tau_2^{7} \vee \tau_2^{8} \vee \tau_2^{9}\right) \left(\tau_1^{2} \vee \tau_1^{3}\right) \tau_2^{0} \tau_3^{\smile}; \tag{22}$$

$$s_2^{\text{в}} = z_1^{\text{р}} z_3^{\text{м}} \tau_5^{0}. \tag{23}$$

Нет необходимости записывать специальное уравнение, которое фиксирует условия, когда вторая буква окончания s_2 равна пробелу ($\smile$). Это условие вполне однозначно определяется как дополнение к уравнениям (18) — (23) по отношению к полному перечню возможных значений, указанному в формуле (16):

$$T_2^{\text{м}} = \left(z_1^{\text{д}} \vee z_1^{\text{т}}\right) \sigma^{2} x_2^{\text{т}}\ \ \mathbf{(24)};\quad T_2^{\text{х}} = z_1^{\text{п}} \sigma^{2} x_2^{1}\ \ \mathbf{(25)};\quad T_2^{\text{ю}} = z_1^{\text{т}} \sigma^{1}\ \ \mathbf{(26)}.$$

Вторая буква постпозиционного окончания T_2 равна пробелу, когда она не равна «м», «х» или «ю».

Уравнения, описывающие первую букву s_1 или T_1 в окончаниях числительных $s_1s_2s_3$ или $T_1T_2T_3$:

$$s_1^{\frown} \vee s_1^{0} \vee s_1^{\text{и}} \vee s_1^{\text{а}} \vee s_1^{\text{у}} \vee s_1^{\text{Е}} \vee s_1^{\text{ь}} \vee s_1^{\text{ы}}; \tag{27}$$

$$T_1^{\text{и}} \vee T_1^{\text{ь}} \vee T_1^{\text{а}} \vee T_1^{\frown}\left(z_1^{\text{и}}\bar{\sigma} \vee z_1^{\text{в}}z_4^{\text{н}}\bar{\sigma} \vee z_1^{\text{р}}\sigma^{2}\right) \vee \sigma^{\frown}); \tag{28}$$

$$\begin{gathered} s_1^{0} = \left(z_1^{\text{р}} \vee z_1^{\text{д}} \vee z_1^{\text{п}} \vee z_1^{\text{в}}z_4^{\text{с}}z_2^{\text{м}} \vee z_1^{\text{т}}z_2^{\text{ж}} \vee z_1^{\text{в}}z_2^{\text{с}}\right) z_3^{\text{Е}}\tau_1^{1}\tau_2^{\frown} \vee \\ \vee \left(z_1^{\text{и}} \vee z_1^{\text{в}}\right)\left(\tau_1^{1}\tau_3^{0}\tau_4^{\frown} \vee \tau_1^{9}\tau_3^{\frown}\right)\tau_2^{0} \vee \left(z_1^{\text{т}}z_3^{\text{Е}} \vee z_1^{\text{р}}z_3^{\text{м}}\right)\tau_3^{0}; \end{gathered} \tag{29}$$

$$\begin{gathered} s_1^{\text{и}} = \left(z_3^{\text{м}} \vee z_1^{\text{т}}\left(z_2^{\text{м}} \vee z_2^{\text{с}}\right) z_2^{\text{Е}}\right)\tau_1^{1}\tau_2^{\frown} \vee \left(z_1^{\text{и}} \vee z_1^{\text{в}}z_4^{\text{н}}\right)\tau_1^{3}\tau_3^{\frown} \vee \\ \vee \left(z_1^{\text{р}} \vee z_1^{\text{Д}} \vee z_1^{\text{п}}\right)\beta \vee \left(z_1^{\text{р}}z_3^{\text{Е}} \vee z_1^{\text{и}}z_3^{\text{м}} \vee z_1^{\text{в}}z_3^{\text{м}}\right)\tau_4^{0}\tau_5^{\frown}. \end{gathered} \tag{30}$$

Здесь

$$\begin{gathered} \beta = \left(\tau_1^{5} \vee \tau_1^{6} \vee \tau_1^{7} \vee \tau_1^{8} \vee \tau_1^{9}\right)\tau_2^{\frown} \vee \left(\tau_2^{0} \vee \tau_2^{1} \vee \right. \\ \left. \vee\, \tau_2^{2} \vee \tau_2^{3} \vee \tau_2^{4} \vee \tau_2^{5} \vee \tau_2^{6} \vee \tau_2^{7} \vee \tau_2^{8} \vee \tau_2^{9}\right)\tau_1^{1}\tau_3^{\frown} \vee \left(\tau_1^{2} \vee \right. \\ \left. \vee\, \tau_1^{3}\right)\tau_2^{0}\tau_3^{\frown}; \quad s_1^{\text{а}} = z_1^{\text{и}}z_2^{\text{ж}}z_3^{\text{Е}}\tau_1^{1}\tau_2^{\frown} \vee \left(z_1^{\text{и}} \vee z_1^{\text{в}}z_4^{\text{н}}\right) z_2^{\text{м}}\tau_1^{2}\tau_2^{\frown} \vee \\ \vee\, z_1^{\text{в}}z_4^{\text{н}} \vee \left(z_1^{\text{р}} \vee z_1^{\text{д}} \vee z_1^{\text{т}} \vee z_1^{\text{п}}\right)\sigma \vee \left(\left(z_1^{\text{р}} \vee z_1^{\text{т}} \vee z_1^{\text{п}}\right) z_3^{\text{м}} \vee \right. \\ \left. \vee \left(z_1^{\text{м}}z_2^{\text{х}}\tau_5^{\frown} \vee z_1^{\text{р}}z_2^{\text{м}}\tau_5^{0}\right) z_3^{\text{Е}}\right)\tau_4^{0}, \end{gathered}$$

где

$$\sigma = \left(\left(\tau_1^{4} \vee \tau_1^{9}\right)\tau_3^{\frown} \vee \tau_1^{1}\tau_3^{0}\tau_4^{\frown}\right)\tau_2^{0}; \tag{31}$$

$$\begin{gathered} s_1^{\text{у}} = \left(z_1^{\text{р}} \vee z_1^{\text{д}} \vee z_1^{\text{т}} \vee z_1^{\text{п}}\right)\tau_1^{2}\tau_2^{\frown} \vee z_1^{\text{в}}\left(z_2^{\text{ж}}z_3^{\text{Е}}\left(\tau_1^{1}\tau_2^{\frown} \vee \right.\right. \\ \left.\left. \vee\, \tau_4^{0}\tau_5^{\frown}\right) \vee z_2^{\text{м}}z_4^{0}\tau_1^{2}\tau_2^{\frown}\right) \vee z_1^{\text{р}}z_2^{\text{м}}z_3^{\text{Е}}\tau_5^{0}; \end{gathered} \tag{32}$$

$$\begin{gathered} s_1^{\text{Е}} = \left(\left(z_1^{\text{и}} \vee z_1^{\text{в}}\right) z_2^{\text{ж}}\tau_1^{2} \vee \overline{z_1^{\text{и}}\left(z_1^{\text{в}}z_4^{\text{н}}\right)}\,\tau_1^{3} \vee z_1^{\text{т}}\tau_1^{4}\right)\tau_2^{\frown} \vee \\ \vee \left(z_1^{\text{д}} \vee z_1^{\text{т}} \vee z_1^{\text{п}}\right) z_2^{\text{ж}}z_3^{\text{Е}}\tau_4^{0}\tau_5^{\frown} \vee z_1^{\text{п}}z_2^{\text{м}}z_3^{\text{Е}}\tau_5^{0}; \end{gathered} \tag{33}$$

$$s_1^{\text{ь}} = z_1^{\text{т}}\tau_1^{4}\tau_2^{\frown} \vee \left(z_1^{\text{п}} \vee z_1^{\text{в}} \vee z_1^{\text{т}}\right)\beta; \tag{34}$$

$$s_1^{\text{ы}} = \left(z_1^{\text{и}} \vee z_1^{\text{в}}\right) z_2^{\text{м}}z_3^{\text{м}}\tau_5^{0}. \tag{35}$$

Условие, когда первая буква окончания s_1 равна пробелу ($\frown$), не записывается тогда, когда равны нулю правые части уравнений (29) — (35):

$$T_1^{\text{и}} = \left(z_1^{\text{р}} \vee z_1^{\text{д}} \vee z_1^{\text{п}}\right)\tau_0^{1} \vee \left(z_1^{\text{и}} \vee z_1^{\text{в}}\right)\sigma^{2}x_2^{\text{т}}\tau_1^{2}\tau_2^{\frown}; \tag{36}$$

$$T_1^{\text{а}} = \left(z_1^{\text{д}} \vee z_1^{\text{т}} \vee z_1^{\text{п}}\right) D \vee \left(z_1^{\text{и}} \vee z_1^{\text{в}}\right)\sigma^{2}x_2^{\text{г}}\tau_2^{\frown}\left(\tau_1^{3} \vee \tau_1^{1}\right); \tag{37}$$

$$D = \sigma^{2}x_2^{\text{т}}\tau_2^{\frown}\left(\tau_1^{2} \vee \tau_1^{3} \vee \tau_1^{4} \vee \tau_1^{5} \vee \tau_1^{6} \vee \tau_1^{7} \vee \tau_1^{8} \vee \tau_1^{9}\right)$$

$$T_1^{\text{ь}} = z_1^{\text{т}}\sigma^{1}\left(\tau_1^{5} \vee \tau_1^{6} \vee \tau_1^{1} \vee \tau_1^{8}\right)\tau_2^{\frown}. \tag{38}$$

Условие, когда $T_1 = \frown$, не записывается. Это условие выполняется, когда правые части уравнений (36) — (38) равны нулю.

Подставляя известные значения переменных в уравнения (1) — (38) и решая их, мы сможем получить значения переменных

$x_1, \ldots, x_k, s_1, s_2, s_3, u_1, \ldots, u_k, T_1, T_2, T_3$. Некоторые значения переменных будут равны пробелу ($\sqcup$), а некоторые отличны от пробела. Теперь необходимо объединить отличные от пробела символы и словоформу $y_1, \ldots, y_m$, что и является конечной задачей синтеза словоформы.

Для удобства записи уравнений это объединение проведем поэтапно. Вначале объединим препозиционную основу $x_1, \ldots, x_k$ с окончанием $s_1s_2s_3$, в результате чего будет получена промежуточная последовательность символов $h_1, \ldots, h_p$. Затем объединим постпозиционную основу $u_1, \ldots, u_k$ со своим окончанием $T_1T_2T_3$, в результате чего будет получена последовательность символов $l_1, \ldots, l_p$. Словоформа $y_1; \ldots, y_m$ будет получена в результате присоединения справа к отличным от пробела символам из последовательности $h_1, \ldots, h_p$ символов последовательности $l_1, \ldots, l_p$. Таким образом, схема синтеза словоформы $y_1, \ldots, y_m$ на последнем этапе выглядит так:

$$x_1, \ldots, x_k + s_1s_2s_3 \to h_1, \ldots, h_p;$$

$$u_1, \ldots, u_k + T_1T_2T_3 \to l_1, \ldots, l_p;$$

$$h_1, \ldots, h_p + l_1, \ldots, l_p \to y_1, \ldots, y_m (m = 2p).$$

Запишем уравнения, которые объединяют $x_1, \ldots, x_k$ с $s_1s_2s_3$ в последовательности $h_1, \ldots, h_p$:

$$\overline{x_i^{\sqcup}} \supset (h_i = x_i); \ (i = 1, k) \ \mathbf{(39)}, \quad x_i^{\sqcup} \supset h_i^{\sqcup} \ \mathbf{(40)},$$

$$\overline{x_2^{\sqcup}}\, x_{2+1}^{\sqcup} \supset (h_{i+1} = s_1)(h_{i+2} = s_2)(h_{i+3} = s_3). \tag{41}$$

Уравнения, объединяющие $u_1, \ldots, u_k$ с окончанием $T_1T_2T_3$:

$$\overline{u_i^{\sqcup}} \supset (l_i = u_i) \ \mathbf{(42)}, \quad u_i^{\sqcup} \supset l_i^{\sqcup} \ \mathbf{(43)},$$

$$\overline{u_i^{\sqcup}}\, u_{i+1}^{\sqcup} \supset (l_{i+1} = T_1)(l_{i+2} = T_2)(l_{i+3} = T_3). \tag{44}$$

Уравнения, объединяющие $h_1, \ldots, h_p, l_1, \ldots, l_p$;

$$h_i^{\sqcup} \supset (y_i = h_i), \tag{45}$$

$$\overline{h_i^{\sqcup}}\, h_{i+1} \supset (y_{i+1} = l_1)(y_{i+2} = l_2) \ \ldots \ (y_{i+p} = l_p). \tag{46}$$

Рассмотрим на примерах, как действует предложенная система уравнений при синтезе словоформ числительных.

Пусть необходимо получить словоформу числительного с числовым значением «1» и такими значениями грамматических признаков: мужской род, единственное число, винительный падеж, одушевленный. Закодируем эти данные в соответствии с введенными обозначениями:

$$e_1 = 1, e_2 = \sqcup, \ldots, e_n = \sqcup; \ z_1 = \text{В}, \ z_2 = \text{М}; \ z_3 = \text{Е}; \ z_4 = 0.$$

Необходимо получить $y_1, \ldots, y_m$.

Исходные данные следует подставить во все 46 уравнений, формализующих морфологию численных. Но не во всех уравнениях имеются эти переменные. Кроме того, объем вычислений в значительной мере зависит от того, в какой последовательности будут поставлены заданные значения переменных в уравнения. Поэтому, рассматривая пример, будем подставлять заданные значения в систему уравнений в последовательности, которая позволяет получить $y_1, y_2, \ldots, y_m$ наиболее коротким путем.

Прежде всего подставим $c_1 = 1$, $c_2 = \smile$, ..., $c_n = \smile$ в формулы (1) — (3) для получения числового значения $\sigma, \tau_1, \ldots, \tau_n$ основы (или двух основ), входящих в состав искомой словоформы. $c_1 = 1$ удовлетворяет левой части только уравнения (3), из которого мы находим $\tau_1 = c_1 = 1$, $\tau_2 = c_2 = \smile, \ldots, \tau_n = c_n = \smile$ (σ по-прежнему равно пробелу).

Подставим полученные значения $\tau_1, \ldots, \tau_n$ в равенство (4) для получения буквенного состава основы синтезируемой словоформы. $\tau_1 = 1$, $\tau_2 = \smile$ обращают в единицу единственное слагаемое уравнения (4):

$$x_1^{0} x_2^{д} x_3^{н} x_4^{н} x_5^{\smile} \vee x_1^{0} x_2^{д} x_3^{н} x_4^{\smile} x_5^{\smile}.$$

На следующем шаге определим s_3 и T_3 с помощью уравнений (10) — (15). Для этого подставим исходные и полученные значения переменных в правые части (12) — (15). $\tau_1 = 1$, $\tau_2 = \smile$ удовлетворяют только уравнениям (12) и (13), но значения $z_1 = в$, $z_2 = м$, $z_3 = Е$, $z_4 = 0$ удовлетворяют только (12). Из этого уравнения находим, что $s_3 = 0$, и одновременно устанавливаем, что $x_3 = н$. Переменная T_3 входит в (11) и (15). При подстановке известных и полученных значений переменных правая часть равенства (15) обращается в нуль, следовательно, $T_3 \neq п$. При этом условии из (11) определяется, что $T_3 = \smile$.

Определим s_2 и T_2 с помощью формул (16) — (26). Значения $\tau_1 = 1$ и $\tau_2 = \smile$ входят в уравнения (18) — (22). Назначения $z_1 = в$, $z_2 = М$, $z_3 = Б$, $z_4 = 0$ удовлетворяют только правой части уравнения (18), из которого определяем, что $s_2 = г$. Значения T_2 присутствуют в (17) — (26). При подстановке известных значений правые части (25) ÷ (26) обращаются в ноль, следовательно, $T_2 \neq м$, $T_2 \neq ю$ и $T_2 \neq х$. Из равенства (17) при этом условии следует, что $T_2 = \smile$.

Определим s_1 и T_1 с помощью уравнений (27) — (38); $\tau_1 = 1$ и $\tau_2 = \smile$ входит в формулы (29) — (32). Но значения грамматических признаков удовлетворяют только уравнению (29). Из него мы находим, что $s_1 = 0$. При подстановке известных данных в (36) — (38) их правые части обращаются в ноль. Это означает, что $T_1 \neq п$, $T_1 \neq а$ и $T_1 \neq ь$. При этом условии из (28) следует, что $T_1 = \smile$.

Таким образом, мы установили, что элементы словоформы имеют следующие значения: $x_1 = 0$, $x_2 = д$, $x_3 = н$, $x_4 = \smile$, $x_5 = \smile$, $s_1 = 0$, $s_2 = г$, $s_3 = 0$, $u_1 = \smile$, $T_1 T_2 T_3$. Значения u_i равны пробелу, так как они не изменились в процессе решения уравнений. Теперь необходимо объединить эти элементы в одну словоформу.

С помощью равенств (39) — (41) объединим значения $x_1, x_2, \ldots, x_{11} \subset s_1 s_2 s_3$. Согласно (39) $h_1 = x_1 = 0$, $h_2 = x_2 =$ д, $h_3 = x_3 =$ н. Из уравнения (40) следует, что остальные h принимают значения, равные пробелу (—). Но при решении (41) переменные h_4, h_5 и h_6 изменят свои значения: $h_4 = s_1 = 0$, $h_5 = s_2 =$ г, $h_6 = s_3 = 0$.

С помощью формул (42) — (44) объединим $x_1, \ldots, x_{11}$ с $T_1 T_2 T_3$. При подстановке известных значений только в (43) левая часть не равна нулю. Из этого уравнения определяем, что $l_1, \ldots, l_k$ равны пробелу (—), остальные $l_{k+1}, \ldots, l_p$ равны пробелу в результате первоначальной очистки.

С помощью равенств (45), (46) определим значение $y_1, \ldots, y_m$. Из уравнения (45) следует, что $y_1 = 0$, $y_2 =$ д, $y_3 =$ н, $y_4 = 0$, $y_5 =$ г, $y_6 = 0$. Согласно (46) $y_7 =$ —, $y_8 =$ —, ..., $y_{20} =$ —. В результате мы получили искомую словоформу: ОДНОГО.

Аналогичным образом можно решать задачи анализа, синтеза словоформ сложных числительных и др. Предложен и реализован на ЕС-1022 алгоритм решения уравнений алгебры конечных предикатов [3].

Список литературы: 1. *Шабанов-Кушнаренко Ю. П.* О теории интеллекта.— Проблемы бионики, 1979, вып. 22, с. 3—10. 2. *Шабанов-Кушнаренко Ю. П., Бондаренко М. Ф.* Математическая модель склонения полных непритяжательных имен прилагательных.— НТИ, сер. 2, 1979, № 6, с. 10—13. 3. *Бондаренко М. Ф., Бондарев В. М.* Программа решения системы уравнений алгебры конечных предикатов с ярусной структурой.— РФАП АН УССР, Справка № 111, 1980.

Поступила в редколлегию 05.03.81.