

По данной блок-схеме алгоритма составлена программа для ЭВМ «Минск-32». Данный алгоритм может быть использован для некоторых задач анализа и синтеза фраз и предложений русского языка, в целях обнаружения и исправления ошибок в написании порядковых числительных и т. д.

#### ЛИТЕРАТУРА

1. Грамматика русского языка, т. I. М., Изд-во АН СССР, 1960. 719 с.
2. Гвоздев А. Н. Современный русский литературный язык. М., «Просвещение», 1967. 432 с.
3. Розенталь Д. Э. Русский язык. Пособие для поступающих в вузы. М., Изд-во Моск. ун-та, 1967. 303 с.
4. Орфографический словарь русского языка. М., Гос. изд-во иностр. и нац. словарей, 1959. 1259 с.
5. Добромыслов В. А., Розенталь Д. Э. Трудные вопросы грамматики и правописания. М., Учпедгиз, 1955. 288 с.
6. Шабанов-Кушнаренко Ю. П., Якименко Л. И. Об одной тематической модели морфологической классификации множества имен существительных русского языка.— В сб.: Проблемы бионики. Вып. 6. Харьков, 1970, с. 104—107.

УДК 62. 506. 2

М. Ф. БОНДАРЕНКО, канд. техн. наук,  
Э. М. БУЗНИЦКАЯ, инж.

### АЛГОРИТМ МОРФОЛОГИЧЕСКОГО АНАЛИЗА ИМЕН ПРИЛАГАТЕЛЬНЫХ РУССКОГО ЯЗЫКА

Рассмотрим моделирование речевого поведения человека при решении задач морфологического анализа имен прилагательных.

*Задача 1. Выделение классов имен прилагательных в соответствии с наличием противопоставлений по различным грамматическим категориям.*

Требуется произвести разбиение исходного множества  $M$  имен прилагательных русского языка на непересекающиеся подмножества (классы)  $M_1, M_2, M_3$  по некоторому грамматическому показателю  $\Pi$ , формирующемуся на основании сведений об окончании и характере основы. Каждое такое разбиение определяет между элементами этого множества отношение  $R$  типа эквивалентности, т. е.  $R$  является одновременно рефлексивным, симметричным и транзитивным.

Лингвистическими предпосылками организации формальных классов  $M_1, M_2, M_3$  могут служить следующие соображения. В парадигме формального класса необходимо учитывать все противопоставления вариантов, свойственные членам этого класса. Поэтому парадигма может быть построена лишь на основании сопоставления парадигм отдельных лексем и выявления таким образом всех противопоставлений, встречающихся между элементами внутри данных частных парадигм.

У имен прилагательных парадигма легко выделяется по категориям рода, числа и падежа. Можно сказать, что все другие категории, приписываемые обычно прилагательному, не входят в состав парадигмы данного класса слов. Прежде всего это относится к полным кратким формам. Употребление этих форм там, где обе они возможны (т. е. в позиции именной части сказуемого), является свободным. Выбор в указанной позиции одной из двух форм факультативен и не определяется согласовательными связями. Можно равно сказать «он красив» и «он красивый». Обе данные формы, таким образом, следует рассматривать как разные слова, а не как члены одной парадигмы. В равной степени сказанное относится и к степеням сравнения прилагательных.

Классы  $M_1$ ,  $M_2$ ,  $M_3$  — характеризуются:  $M_1$  — противопоставлениями по роду, числу и падежу (это соответствует полным именам прилагательным);  $M_2$  — противопоставлениями по роду и числу (соответствует кратким формам имен прилагательных и превосходной степени сравнения — аналитической);  $M_3$  — нулевой парадигмой (сюда войдут неизменяемые прилагательные — слова иностранного происхождения и компаратив). Общую парадигму имени прилагательного можно представить в виде

$$D = (x_1 \vee \bar{x}_1) \cdot (x_2 \vee \bar{x}_2) \cdot (x_3 \vee \bar{x}_3) \cdot \\ \cdot (x_4 \vee \bar{x}_4) \wedge (x_5 \vee \bar{x}_5) \wedge (x_6 \vee \bar{x}_6), \\ (y_1 \vee \bar{y}_1) \wedge (y_2 \vee \bar{y}_2), (z_1 \vee \bar{z}_1) \wedge (z_2 \vee \bar{z}_2) \wedge (z_3 \wedge \bar{z}_3); \\ D = D_1 \cup D_2 \cup D_3,$$

где  $x_1 - x_6$  — набор бинарных признаков, характеризующих противопоставления соответственно по именительному, родительному, дательному, винительному, творительному, предложному падежам;

$y_1, y_2$  — противопоставления по числу (единственное, множественное);

$z_1, z_2, z_3$  — противопоставления по роду (мужской, женский, средний).

Грамматическая информация для элементов  $M_1$  получается как объединение грамматических признаков из  $D_1, D_2, D_3$ ; для  $M_2$  — из  $D_2, D_3$ . Для элементов  $M_3$  цепочку грамматических признаков можно записать в виде

$$f = \{x_1 x_2 x_3 x_4 x_5 x_6, y_1 y_2, z_1 z_2 z_3\}.$$

### Работа алгоритма

1. Проверить, входит ли анализируемое слово в словарь 1 (неизменяемые слова). Если ответ «да», то слово — элемент из  $M_3$ .

2. Сравнить окончание слова \* с нулевым окончанием; если они совпадают, перейти к 3, если нет — к 5.

3. Сравнить две последние буквы основы с *ин, ын, ов, ев*; если ответ «да», то перейти к 4, если «нет», — слово из  $M_2$ .

4. Проверить анализируемое слово на принадлежность словарю 2\*\*; если вхождение имеет место, то слово из  $M_2$ , если нет — из  $M_1$ .

5. Сравнить окончание входного слова с однобуквенными окончаниями *-а, -о, -ы*; если они совпадают, перейти к 6, если нет — к 7.

6. Отбросить окончание входного слова, перейти к 3.

7. Сравнить окончание входного слова с однобуквенными окончаниями *-я, -и*; если ответ «да», то перейти к 8, если «нет» — к 9.

8. Сравнить предпоследнюю букву основы с *ь*; если имеет место совпадение, то слово — элемент из  $M_1$ , если нет — из  $M_2$ .

9. Проверить на совпадение окончание основы с *ь*; если они совпадают, то слово из  $M_2$ , если нет — перейти к 10.

10. Сравнить окончание анализируемого слова с однобуквенным окончанием *е*; если ответ «да», то перейти к 11, если «нет» — к 16.

11. Сравнить предпоследнюю букву слова с *ц*; если это *ц*, то перейти к 12, если нет — к 13.

12. Проверить слово на вхождение в словарь 3; если слово принадлежит словарю, то это элемент из  $M_3$ , если нет — из  $M_2$ .

13. Сравнить предпоследнюю букву слова с *ж*; если они совпадают, то перейти к 14, если нет — к 15.

14. Проверить, входит ли анализируемое слово в словарь 4; если ответ «да», то оно принадлежит  $M_2$ , если «нет» —  $M_3$ .

15. Сравнить предпоследнюю букву основы с *н*; если совпадение имеет место, то входное слово — элемент из  $M_2$ , если нет — из  $M_3$ .

16. Сравнить окончание анализируемого слова с двухбуквенными окончаниями *-еа, -ей*; если оно совпадает, то перейти к 17, если нет — к 19.

17. Отбросить окончание слова, перейти к 18.

18. Проверить три последние буквы основы на совпадение с *айш* и *ейш*; если ответ «да», то слово из  $M_2$ , если «нет» — из  $M_1$  или  $M_3$ .

19. Проверить, совпадает ли окончание входного слова с *-ий, -ая, -ие*; если ответ «да», то перейти к 17, если «нет» — анализируемое слово — элемент из  $M_1$ .

---

\* Алгоритм выделения окончания слова в работе не приводится.

\*\* 2,3,4 — словари исключений.

## Окончание работы алгоритма.

	<i>Словарь 1</i>			<i>Словарь 2</i>
ампир	демисезон	минимум	перманент	нов
апаш	декольте	маренго	реглан	здоров
апплике	дезабилье	массака	рондо	суров
беж	дяз	мокко	сомон	готов
бемоль	ин-кварто	мажор	соте	
бордо	ин-октава	минор	трико	<i>Словарь 3</i>
брутто	ин-фолио	модерн	томат	гуще
буфф	кlesh	моль		проще
валансьен	комильфо	неглиже		слаще
волан	коми	пикколо		толще
галифе	фантази	плаке		чаще
гала	люкс	пике		чище
гротеск	майноз	плиссе		
гольф	максимум	пик		<i>Словарь 4</i>
				свеже
				несвеже

Пункт 18 данного алгоритма предполагает неоднозначный ответ:  $M_1$  или  $M_3$ . Данную дизъюнкцию можно определить из контекста. Критерием для установления степени имени прилагательного может служить слово, следующее за определяемым. Если это существительное, то прилагательное получает признак положительной степени. Следовательно, ответом будет «элемент из  $M_1$ ». При отсутствии глагола в личной форме или любого другого сказуемого в предложении прилагательное получает признак «сравнительная степень, элемент из  $M_3$ ».

Решение данной задачи позволяет в дальнейшем анализировать имена прилагательные отдельно для каждого класса.

**Задача 2.** Приведение имени прилагательного к словарной форме (лемме слова).

Требуется определить вид отображения  $F$ , устанавливающего соответствие между множеством словоформ одной и той же парадигмы словоизменения из  $M_1$ , которые обладают различными наборами грамматических признаков и исходной словоформой той же парадигмы.

Отображение  $F$  ищется в виде следующего алгоритма.

1. Сравнить окончание анализируемого слова с нулевым окончанием; если ответ «да», то исходное слово совпадает с его словарной формой (обозначим словарную форму через СФ), если «нет», то перейти к 2.

2. Сравнить окончание входного слова с однобуквенными окончаниями *-а*, *-о*, *-у*, *-ы*; если они совпадают, то перейти к 3, если нет — к 4.

3. Отбросить окончание слова на выходе СФ.

4. Сравнить окончание анализируемого слова с однобуквенными окончаниями *-я*, *-ю*, *-и*; если ответ «да», то перейти к 5, если «нет» — к 7.

5. Отбросить окончание слова, перейти к 6.

6. Заменить *ь* на *ий*, на выходе СФ.

7. Сравнить окончание входного слова с двубуквенными окончаниями *-ый, -ий*; если ответ «да», то на выходе СФ, если «нет» — перейти к 8.

8. Сравнить окончание анализируемого слова с двубуквенными окончаниями *-ой, -ом, -ое, -ых, -их, -ым*; если они совпадают, то перейти к 9, если нет — к 20.

9. Отбросить окончание слова, перейти к 10.

10. Сравнить две последние буквы слова с *ов, ев*; если ответ «да», то перейти к 11, если «нет» — к 13.

11. Проверить, входит ли анализируемое слово в словарь 1\*; если оно входит, то на выходе — СФ, если нет — перейти к 12.

12. Присоединить окончания *-ый* или *-ой*; на выходе — СФ.

13. Сравнить две последние буквы основы с *ин, ын*; если они совпадают, то перейти к 14, если нет — к 18.

14. Проверить слово на вхождение в словарь 2; если оно входит, то это — СФ, если нет — перейти к 15.

15. Сравнить входное слово со словом из словаря 3; если они совпадают, то перейти к 16, если нет — к 17.

16. Присоединить окончание *-ой*; на выходе — СФ.

17. Присоединить окончание *-ый*; на выходе — также СФ.

18. Сравнить последнюю букву основы с *г, к, х*; если это *г, к, х*, то перейти к 19, если нет — к 12.

19. Присоединить окончания *-ий* или *-ой*; на выходе — СФ.

20. Сравнить окончание анализируемого слова с двубуквенными окончаниями *-их, -им, -ие*, если ответ «да», то перейти к 21, если «нет» — перейти к 25.

21. Отбросить окончание слова и перейти к 22.

22. Сравнить последнюю букву основы с *ь*; если это *ь*, то перейти к 6, если нет — к 23.

23. Сравнить последнюю букву основы с *г, к, х, ж, ш*; если они совпадают, то перейти к 19, если нет — к 24.

24. Присоединить окончание *-ий*; на выходе — СФ.

25. Проверить на совпадение окончание анализируемого слова с двубуквенными окончаниями *-ем, -ей*; если они совпадают, то перейти к 26, если нет — 29.

26. Отбросить окончание слова, перейти к 27.

27. Сравнить последнюю букву основы с *ц*; если это *ц*, то перейти к 17, если нет — к 28.

28. Сравнить последнюю букву основы с *ь*; если ответ «да», то переход к 6, если «нет» — к 24.

29. Сравнить окончание анализируемого слова с двубуквенными окончаниями *-ее, -яя, -юю*; если они совпадают, то перейти к 30, если нет — к 31.

30. Отбросить окончание слова, перейти к 27.

---

\* Словари 1, 2, 3 — исключения, связанные с нерегулярностью русского языка.

31. Установить, совпадает ли окончание входного слова с окончанием *-ые*; если ответ «да», то перейти к 32, если «нет» — к 33.

32. Отбросить окончание *-ые*, перейти к 12.

33. Сравнить окончание анализируемого слова с трехбуквенными окончаниями *-ого*, *-ому*; если они совпадают, то перейти к 34, если нет — к 35.

34. Отбросить окончание слова, перейти к 18.

35. Проверить, совпадает ли окончание входного слова с трехбуквенными окончаниями *-его*, *-ему*; если ответ «да», то перейти к 36, если «нет» — к 37.

36. Отбросить окончание входного слова, перейти к 27.

37. Установить, является ли окончание анализируемого слова окончанием *-ыми*; если ответ «да», то перейти к 38, если «нет» — к 39.

38. Отбросить окончание *-ыми*, перейти к 10.

39. Проверить, совпадает ли окончание входного слова с трехбуквенным окончанием *-ими*; если совпадает, то перейти к 40, если нет, то следует проверить входное слово.

40. Отбросить окончание *-ими*, перейти к 22\*.

Окончание работы алгоритма.

#### Словарь 1

кесарев  
царев  
дедов  
крокодилов  
буйволов  
фараонов  
драконов  
воронов  
петров  
чертов

христов  
сизифов  
отцов



#### Словарь 2

сердечкин  
ласточкин  
бабушкин  
батюшкин  
филькин  
маменькин  
мамин  
женин  
мужнин  
дочернин

братнин  
папин  
дочерин  
курицын

Словарь 3  
свиной

В некоторых случаях формирования именительного падежа серьезные затруднения вызывают блоки 12, 19 приведенного выше алгоритма. Это случаи, когда вопрос о выборе необходимого окончания нельзя решить однозначно. Однако дальнейший анализ имен прилагательных, создание специальных словарей и отыскание новых формальных признаков, на наш взгляд, поможет разрешить эту задачу.

Результатом решения предыдущей задачи явилось формирование элементов  $\underline{a}_i$ , обладающих набором признаков  $\underline{f}_i = \{x_1, x_2, x_3, x_4, x_5, x_6, y_1, y_2, z_1, z_2, z_3\}$ , из которых может быть образовано некоторое множество  $M'$ . Требуется произвести разбиение множества  $M' = \{a'_1, a'_2, \dots, a'_n\}$  на классы по отношению подобия системы падежных окончаний.

\* Словари 1,2,3 — исключения, связанные с нерегулярностью русского языка.

На выходе вырабатывается алгоритмический признак  $p$ , принимающий значения 1—15. Это соответствует номеру класса, в который будут включены прилагательные одного типа склонения. В результате непосредственного склонения слов, относящихся к множеству  $M^1$ , было установлено 16 типов склонения полных имен прилагательных, которые вместе с определяющими их формальными признаками приведены в таблице.

Тип склонения	Формальные признаки слов
1	Окончание <i>-ый</i> , основа, оканчивающаяся не на <i>ц</i>
2	« <i>-ый</i> , основа на <i>ц</i>
3	« <i>-ий</i> , « « <i>г, к, х</i>
4	« <i>-ий</i> , « « <i>б, в, з, л, м, с</i>
5	Слова, входящие в словарь 1
6	Окончание <i>-ий</i> , основа на <i>ж</i>
7	Слова, входящие в словарь 2
8	Окончание <i>-ий</i> , основа на <i>н</i>
9	Слова, входящие в словарь 3
10	Окончание <i>-ий</i> , основа на <i>ч</i>
11	Слова, входящие в словарь 4
12	Окончание <i>-ий</i> , основа на <i>ш</i>
13	« <i>-ой</i> , « « <i>г, к, х, ж, ш</i>
14	Окончание <i>-ой</i>
15	«Нулевое» окончание, основа на <i>ов, ев</i>
16	» » » » <i>ин, ын</i>

### Работа алгоритма

1. Сравнить окончание анализируемого слова с нулевым окончанием; если они совпадают, то перейти к 2, если нет — к 3.
2. Сравнить две последние буквы с *ов, ев*, если ответ «да», то на выходе  $p = 15$ , если «нет» —  $p = 16$ .
3. Сравнить окончание входного слова с двубуквенным окончанием *-ий*; если совпадение имеет место, то перейти к 4, если нет — к 14.
4. Отбросить окончание *-ий*, перейти к 5.
5. Проверить на совпадение последнюю букву основы с *г, к, х*; если ответ «да», то на выходе  $p = 3$ , если «нет» — перейти к 6.
6. Сравнить последнюю букву основы с *ж*; если они совпадают, то перейти к 7, если нет — к 8.
7. Проверить анализируемое слово на вхождение в словарь 1; если ответ «да», то  $p = 5$ , если «нет» —  $p = 6$ .
8. Сравнить последнюю букву основы с *н*; если они совпадают, то перейти к 9, если нет — к 10.
9. Проверить принадлежность входного слова словарю 2; если слово из словаря, то  $p = 7$ , если нет —  $p = 8$ .

10. Проверить, совпадает ли последняя буква основы с *ч*; если ответ «да», то обратиться к 11, если «нет» — к 12.

11. Проверить, принадлежит ли слово словарю 3; если ответ «да», то  $p = 9$ , если «нет» —  $p = 10$ .

12. Сравнить последнюю букву основы с *ш*; если это *ш*, то перейти к 13, если нет, то значение признака  $p$  на выходе равно 4.

13. Проверить, входит ли анализируемое слово в словарь 4; если входит, то  $p = 11$ , если нет —  $p = 12$ .

14. Сравнить окончание входного слова с двубуквенным окончанием *-ой*; если это *-ой*, то перейти к 15, если нет — к 17.

15. Отбросить окончание *-ой*, перейти к 16.

16. Проверить, совпадает ли последняя буква основы с *г*, *к*, *х*, *ж*, *ш*; если ответ «да», то  $p = 13$ , если «нет» —  $p = 14$ .

17. Сравнить окончание анализируемого слова с *ый*; если они совпадают, то перейти к 18, если нет, то проверить входное слово.

18. Отбросить окончание *-ый*, перейти к 19.

19. Проверить совпадение последней буквы основы с *ц*; если это *ц*, то  $p = 2$ , если нет —  $p = 1$ .

Окончание работы алгоритма.

Словарь 1

наважий  
медвежий  
божий  
бульдожий  
белужий  
верблюжий  
севрюжий  
лебяжий  
стерляжий  
княжий

Словарь 2

кабаний  
фазаний  
сазаний  
тараканий  
бараний  
олений  
селезений  
тюлений  
павлиний  
бизоний  
вороний  
обезьяний

Словарь 3

водчий  
лежащий  
широкоплечий  
печный  
равноплечий  
рабочий  
прочий  
охотный  
неохотный  
писчий  
зыбучий  
певучий

живучий  
сидячий  
бродячий  
ходячий  
стоячий  
зрячий  
горячий  
висячий  
колючий  
воноющий  
горючий

Словарь 4

росомашный  
черепашный  
пастуший  
старуший  
мышный

ЛИТЕРАТУРА

1. Грамматика современного русского языка. М., «Наука», 1970. 726 с.
2. Bielfeldt. H. H. Rückläufiges Wörterbuch der russischen Sprache der gegenwart. Berlin, Akademie Verlag, 1965. 392S.