

И. А. ГОТЮР, А. В. КОСТРОМИТИНОВ

## ТЕХНОЛОГИЯ ОПИСАНИЯ И РАСКОДИРОВАНИЯ МЕТЕОРОЛОГИЧЕСКИХ ДАННЫХ НА ОСНОВЕ АЛГЕБРАИЧЕСКОГО ПОДХОДА

Предложен алгоритм описания и обработки метеорологических данных, основанный на алгебраическом подходе. Разработанный алгоритм позволяет ускорить обработку метеорологических данных и сократить затраты вычислительных ресурсов.

*Ключевые слова:* алгебраический подход, формальные языки, абстрактные типы данных, метеорология, буквенно-цифровые метеорологические коды, раскодирование.

**Введение.** Развитие гидрометеорологии в настоящее время невозможно без международного сотрудничества в глобальном масштабе. Международное сотрудничество, помимо всего прочего, заключается в осуществлении свободного и неограниченного обмена метеорологическими данными в реальном или близком к реальному масштабе времени. При этом необходимо, с одной стороны, по возможности сокращать объем передаваемых данных, а с другой — исключать неоднозначность их трактования при получении. Для решения этих задач используются буквенно-цифровые метеорологические коды (БЦМК) или традиционные коды. При разработке БЦМК прежде всего учитывалось, что для передачи метеорологических данных широко применяются телеграфные аппараты. Реализуемый в настоящее время обмен данными с помощью компьютерных сетей обусловил появление более совершенных алгоритмов сжатия данных. Это привело к созданию таблично-ориентированных кодов, которые должны со временем заменить традиционные коды. Тем не менее, в Глобальной сети телесвязи Всемирной метеорологической организации (ВМО) продолжают циркулировать сообщения, закодированные с помощью буквенно-цифровых кодовых форм.

**Алгебраический подход к описанию и раскодированию метеорологических данных.** Существующие алгоритмы раскодирования сообщений, представленных в виде БЦМК, являются более сложными по сравнению с алгоритмами обработки таблично-ориентированных кодов и требуют учитывать актуальную кодовую форму, а также факт наличия или отсутствия определенных групп данных и содержащихся в них значений метеорологических величин.

Традиционные коды составлены из набора кодовых форм и двоичных кодов, состоящих из букв (или групп букв), обозначающих метеорологические или, в определенных случаях, другие геофизические величины или явления. В сводках эти буквы (или группы букв) заменяются цифрами, обозначающими значение величины или состояние описываемых явлений. Для различных букв были выработаны спецификации, в соответствии с которыми эти буквы заменяются цифрами. В некоторых случаях спецификации букв позволяют непосредственно осуществлять их замену цифрами. В иных случаях необходимо использовать кодовые цифры, спецификации которых приводятся в кодовых таблицах. Кроме того, разработано определенное количество символических слов и символических цифровых групп для использования их в качестве названий кодов, кодовых слов, символических приставок или отличительных групп [1].

Правила, касающиеся выбора кодовых форм, используемых при международном обмене информацией, и выбора соответствующих символических слов, цифровых групп и букв, изложены в техническом регламенте (ВМО-№ 49) [1].

Анализ схем раскодирования БЦМК позволил составить общую схему раскодирования, которая заключается в выполнении следующих этапов:

- чтение закодированного метеорологического сообщения;
- определение кодовой формы;
- поиск в исходном сообщении информации о значениях гидрометеорологических величин или о служебных данных (индексе станции, времени наблюдения и пр.);
- преобразование по определенному для кодовой формы правилу (спецификации) информации о кодовом значении метеорологической величины в ее реальное значение.

Проведенное обобщение позволяет выделить две основные задачи раскодирования БЦМК:

- предварительный этап: синтаксический и лексический анализ;
- заключительный этап: непосредственно определение значений метеорологических величин.

Для автоматизации синтаксического и лексического анализа широко используется язык регулярных выражений, основу которого составляют теории автоматов и формальных языков [2, 3].

Регулярные выражения представляют собой формальный язык поиска и осуществления манипуляций с подстроками в тексте, основанный на использовании метасимволов. Метасимволом называется символ, обработка которого осуществляется по алгоритму, определяющему правилу обработки символов, предшествующих данному или следующих после него, либо его замещающих. Иными словами, регулярное выражение — это строка-шаблон, состоящая из символов и метасимволов, задающая правила поиска.

Для непосредственного определения значений гидрометеорологических величин на заключительном этапе раскодирования необходимо задать правила (спецификации) преобразования информации о кодовом значении метеорологической величины в ее реальное значение. В общем виде правилами являются следующие таблицы соответствия:

- для кодирования данных об облачности, о явлениях и пр.: кодовое значение величины  $\rightarrow$  ее реальное значение;
- для кодирования данных об атмосферном давлении, температуре и пр.: метеорологическая величина  $\rightarrow$  арифметические действия по преобразованию кодового значения в реальное.

Исходя из этого значение  $v$  метеорологической величины целесообразно ассоциировать с правилами  $r$  синтаксического и лексического анализа и спецификациями  $s$ . Таковую ассоциативную конструкцию можно объединить в структуру  $M = (r, s, v)$ . В связи с тем, что в структуре  $M$  могут храниться данные о любой метеорологической величине, для однозначной ее идентификации предлагается использовать множество уникальных числовых дескрипторов  $D = \{d_i\}, i = \overline{1(1), N}$ , где  $N$  — количество раскодируемых метеорологических величин.

Для хранения произвольного количества пар  $\langle d_i; M \rangle, d_i \in D, i = \overline{1(1), N}$ , в современных методах обработки данных применяются абстрактные типы данных — словари. Словари обладают широкими возможностями доступа к произвольным элементам (словарным единицам) и характеризуются незначительными затратами вычислительных ресурсов на операции поиска и добавления нового элемента, которые наиболее часто выполняются при раскодировании метеорологических данных.

На множестве структур определяются нуль-арные операции получения регулярного выражения  $g_1 : 0 \rightarrow r$ , получения спецификации  $g_2 : 0 \rightarrow s$  и определения значения метеорологической величины  $g_3 : 0 \rightarrow v$ . Тогда операцию раскодирования метеорологических данных можно определить как последовательно выполняемые операции поиска в исходном сообще-

нии закодированной информации о значении  $f_1$  данных и определении их непосредственно значения  $f_2$ . Сигнатура операции  $f_1$  имеет следующий вид:

$$f_1 : b \times r \rightarrow k,$$

где  $b$  задается на множестве БЦМК,  $r$  — на множестве регулярных выражений,  $k$  — на множестве закодированных значений метеорологической величины.

Сигнатура операции  $f_2$  имеет вид

$$f_2 : s \times k \rightarrow v,$$

где  $v$  задается на множестве значений метеорологической величины.

**Алгоритм раскодирования метеорологических данных.** На рис. 1 представлена схема алгоритма раскодирования метеорологических данных, закодированных с помощью буквенно-цифровых форм.

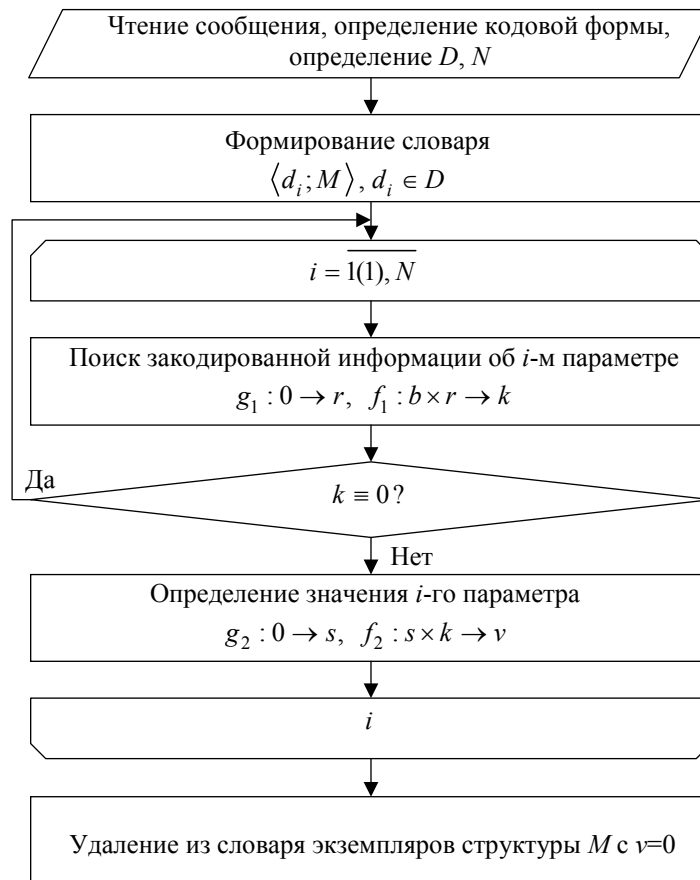


Рис. 1

Рассмотрим алгоритм более подробно.

**Шаг 1.** Производится чтение метеорологического сообщения  $b$  и по заголовку определяется кодовая форма, в которой оно закодировано. По кодовой форме определяется множество дескрипторов  $D = \{d_i\}$ ,  $i = \overline{1(1), N}$ , метеорологических величин, данные о которых могут содержаться в сообщении.

**Шаг 2.** Для каждого элемента  $d_i$  множества  $D$  создается структура  $M = (r, s, v)$  с заданными элементами  $r$  и  $s$ . Далее формируется словарь  $\langle d_i; M \rangle$ ,  $d_i \in D, i = \overline{1(1), N}$ .

**Шаг 3.** Для  $i$ -го экземпляра структуры  $M$  выполняется операция получения регулярного выражения  $g_1 : 0 \rightarrow r$ . Полученное регулярное выражение используется для поиска и извлечения закодированной информации об  $i$ -й метеорологической величине:  $f_1 : b \times r \rightarrow k$ .

*Шаг 4.* Если  $k \equiv 0$  (данные об  $i$ -й метеорологической величине отсутствуют), полагается  $i = i + 1$ , и переход к шагу 3.

*Шаг 5.* Для  $i$ -го экземпляра структуры  $M$  выполняется операция получения спецификации  $g_2 : 0 \rightarrow s$ . Полученная спецификация используется для раскодирования  $i$ -й метеорологической величины:  $f_2 : s \times k \rightarrow v$ , значение которой сохраняется в  $i$ -м экземпляре структуры  $M$ .

Шаги 3—5 выполняются в цикле для всех  $i = \overline{1(1), N}$ .

*Шаг 3* ( $N+1$ ). Из словаря удаляются элементы, для которых на шаге 3 не была найдена информация о метеорологической величине.

Отметим, что реализация операции  $f_1$  (работа с регулярными выражениями) имеется практически во всех современных языках программирования. Для реализации операции  $f_2$  — определения значения  $v$  метеорологической величины — предлагается алгоритм, схема которого представлена на рис. 2.

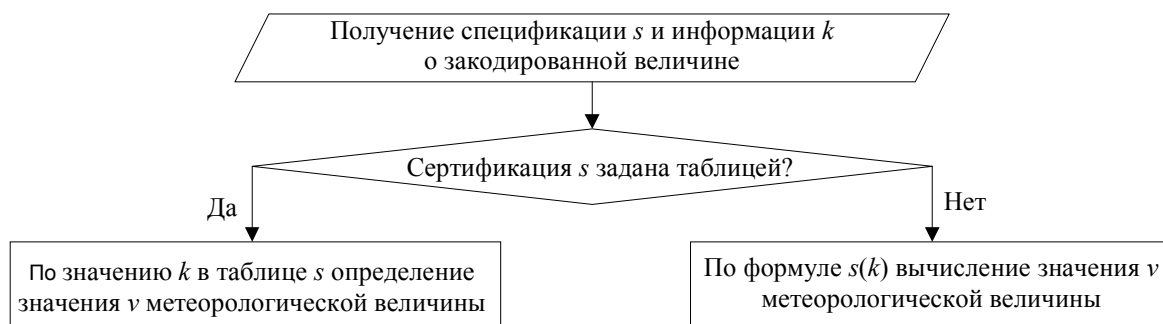


Рис. 2

**Заключение.** Представленная технология реализована в программном обеспечении автоматизированного рабочего места военного метеоролога, которое позволяет производить обработку метеорологических сообщений, циркулирующих в Глобальной сети телесвязи ВМО, а также их визуализацию. Результаты эксплуатации свидетельствуют о том, что использование таких современных способов представления данных и их обработки, как абстрактные типы данных и регулярные выражения позволяет значительно упростить алгоритмы обработки метеорологической информации, а также сократить затраты вычислительных ресурсов.

#### СПИСОК ЛИТЕРАТУРЫ

1. Наставление по кодам: Международные коды. Т. I.1; Ч. А — Буквенно-цифровые коды [Электронный ресурс]: <[http://fip.rshu.ru/2/doc/WMO306\\_Vol\\_I.1\\_2011\\_ru.pdf](http://fip.rshu.ru/2/doc/WMO306_Vol_I.1_2011_ru.pdf)>.
2. Хопкрофт Дж., Мотвани Р., Ульман Дж. Введение в теорию автоматов, языков и вычислений. М.: Изд. дом „Вильямс“, 2002.
3. Пентус А. Е., Пентус М. Р. Математическая теория формальных языков: Учеб. пособие. М.: Интернет-университет информ. технологий; Бином, Лаборатория знаний, 2006.

#### Сведения об авторах

**Иван Алексеевич Готюр**

— канд. техн. наук; Военно-космическая академия им. А. Ф. Можайского, кафедра технологий и средств геофизического обеспечения войск, Санкт-Петербург; E-mail: gotur@newmail.ru

**Алексей Валерьевич Костромитинов**

— Военно-космическая академия им. А. Ф. Можайского, кафедра технологий и средств геофизического обеспечения войск, Санкт-Петербург; инженер; E-mail: alexakos@yandex.ru

Рекомендована кафедрой технологий и средств геофизического обеспечения войск

Поступила в редакцию 21.10.13 г.