

621.391.15:004.65

В.В. Кручинин, Б.А. Люкшин

Метод кодирования информационных объектов на основе деревьев И/ИЛИ

Рассматриваются вопросы применения метода построения алгоритмов комбинаторной генерации, основанного на деревьях И/ИЛИ для кодирования информационных объектов. Предлагается метод кодирования, основанный на представлении всего множества информационных объектов в виде дерева И/ИЛИ.

Ключевые слова: кодирование, информационный объект, методы генерации комбинаторных множеств, дерево И/ИЛИ.

Кодирование является одним из способов повышения информационной безопасности в системах хранения и передачи информации [1]. Зачастую информационные объекты в таких системах носят иерархическую или рекурсивную природу описания. В таком случае древовидная структура является естественной формой представления информационного объекта. Однако современные базы данных имеют реляционную модель хранения данных. Это приводит к коллизии между естественной структурой информационного объекта и табличной формой представления информации.

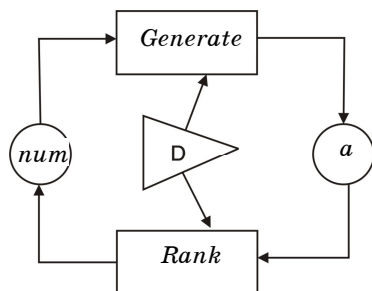


Рис. 1. Основная схема алгоритмов *Generate* и *Rank*

Для решения указанной проблемы предлагается использовать подход, основанный на применении алгоритмов комбинаторной генерации *Generate* и *Rank*. Обобщенная схема работы этих алгоритмов представлена на рис. 1.

В общем случае для получения процедур генерации *Generate* и нумерации *Rank* комбинаторного множества задается биективное отображение:

$$\text{Generate}: N_n \rightarrow A_n,$$

где N_n – конечное подмножество натуральных чисел; A_n – комбинаторное множество. Тогда обратное отображение $\text{Rank} \equiv \text{Gen}^{-1}$ задает процедуру нумерации:

$$\text{Rank}: A_n \rightarrow N_n.$$

Алгоритм *Generate* на основании числа $\text{num} \in N_n$ и некоторого описания комбинаторного множества D создает элемент $a \in A_n$. Алгоритм *Rank* производит обратное действие на основании $a \in A_n$ и некоторого описания D , формирует номер $\text{num} \in N_n$.

Рассмотрим методы построения алгоритмов генерации и нумерации комбинаторных объектов, основанные на представлении описания D в виде дерева И/ИЛИ [2]. Введем определения:

- 1) деревом И/ИЛИ называется дерево, которое содержит узлы двух типов: И-узел и ИЛИ-узел;
- 2) вариантом дерева И/ИЛИ назовем дерево, которое получается из данного путем отсечения всех дуг, кроме одной, у всех ИЛИ-узлов. На рис. 2 показан пример дерева И/ИЛИ и всех его вариантов.

Число вариантов ω_z для узла z дерева И/ИЛИ вычисляется по формуле

$$\omega_z = \begin{cases} 1, & \text{для листа,} \\ \sum_{i=1}^k \omega_i, & \text{для ИЛИ-узла,} \\ \prod_{i=1}^k \omega_i, & \text{для И-узла,} \end{cases}$$

где ω_i – число вариантов для узлов, являющихся сыновьями узла z .

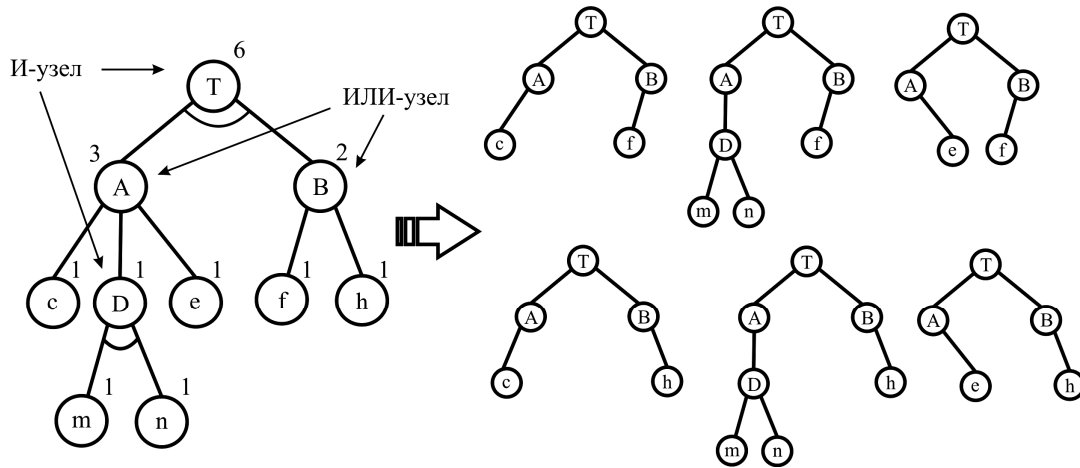


Рис. 2. Дерево И/ИЛИ и все его варианты

Зная число вариантов ω_i для всех узлов дерева И/ИЛИ, строятся алгоритмы генерации и нумерации вариантов.

Алгоритм генерации варианта следующий:

1. В стек помещается пара $(num, root)$ номер варианта и корень дерева.
2. Из стека вынимается очередная пара (l_z, z) .
3. Если z – это И-узел, то вычисляются l_i для всех сыновей s_i узла z , используя ω_i и схему Горнера $l_z = l_1 + \omega(s_1)(l_2 + \omega(s_2)(\dots\omega(s_{n-1})(l_n)\dots))$.

Все пары (l_i, s_i) заносятся в стек, переход на шаг 2.

4. Если z – это ИЛИ-узел, то определяется пара (l_k, s_k) из условия:

если $\omega(s_1) \leq l_z < \omega(s_1) + \omega(s_2)$, то выбирается второй сын и $l_2 := l_z - \omega(s_1)$;

если $\sum_{i=1}^k \omega(s_i) \leq l_z < \sum_{i=1}^{k+1} \omega(s_i)$, то выбирается $k+1$ сын и $l_{k+1} := l_z - \sum_{i=1}^k \omega(s_i)$.

Найденная пара (l_k, s_k) заносится в стек, переход на шаг 2.

5. Если z – лист, то переход на шаг 2.

6. Если стек пуст, то останов.

Представленный алгоритм задает биективное отображение $Generate: N_n \rightarrow W$, где N_n – подмножество натуральных чисел; W – множество вариантов. Из этого следует, что для любого варианта дерева И/ИЛИ можно найти единственный номер i . Таким образом, можно однозначно задать алгоритм нумерации варианта $Rank$. Временная сложность алгоритмов генерации и нумерации вариантов в среднем зависит от структуры дерева И/ИЛИ и пропорциональна среднему числу узлов в варианте.

Рассмотрим представление информационного объекта в виде дерева И/ИЛИ. Пусть все множество информационных объектов из некоторой предметной области можно представить пятеркой (O, S, T, P, O_n) , где O – множество объектов, имеющих некоторую структуру; O_n – обозначение множества объектов, S – множество структур объектов; T – множество элементарных объектов; P – множество правил перехода вида:

- 1) $O_i \Rightarrow T_{1i} | T_{2i} | \dots | T_{ni}$.
- 2) $O_i \Rightarrow S_{1i} | S_{2i} | \dots | S_{ki}$.
- 3) $S_{ij} \Rightarrow O_{i1} O_{i2} \dots O_{im}$.

Правила типа (1) гласит, что объект имеет одну из возможных элементарных реализаций. Правила типа (2) означает, что объект имеет одну из возможных структур, и правила типа (3), что структура S_{ij} содержит некоторое подмножество объектов. Такое представление информационного объекта позволяет построить для него дерево И/ИЛИ:

- 1) для правил типа (1) записывается ИЛИ – узел O_i , у которого будут сыновья $T_{1i}, T_{2i}, \dots, T_{ni}$ – листья;

- 2) для правил типа (2) записывается ИЛИ – узел O_i , у которого будут сыновья $S_{1i}, S_{2i}, \dots, S_{ki}$ – И-узлы;

- 3) для правил типа (3) записывается И – узел S_{ij} , у которого будут сыновья $O_{i1} O_{i2} \dots O_{im}$ – ИЛИ-узлы.

В качестве примера рассмотрим информационный объект, имеющий атрибуты стандарта сертификата x.509 (stdat). Перечислим их: С (страна); L (размещение); ST (штат или провинция); O (организация); OU (подразделение); CN (фамилия, имя, отчество); STREET (адрес); E (электронный адрес). На основе данного стандарта каждый удостоверяющий центр формирует свои собственные атрибуты полей. Например, в поле CN могут быть занесены паспортные данные, номер свидетельства ИНН, номер пенсионного свидетельства и т.д. Тогда атрибуты stdat можно представить деревом И/ИЛИ (рис. 3).

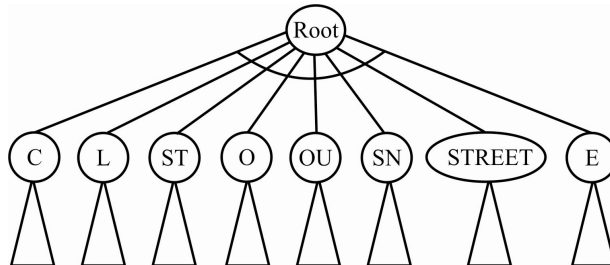


Рис. 3. Дерево И/ИЛИ для атрибутов сертификата

Далее строятся деревья И/ИЛИ для конкретных атрибутов. Например, атрибут «адрес».

Если для множества информационных объектов получено дерево И/ИЛИ, то алгоритм Rank по описанию конкретного информационного объекта получает уникальный номер, а алгоритм Generate по номеру получает описание этого объекта.

Такой метод кодирования информационных объектов позволяет:

1. Повысить информационную безопасность системы хранения данных, поскольку внутри системы информационные объекты явно не хранятся, а хранятся их номера и дерево И/ИЛИ.

2. Обеспечить быстрый поиск информационных объектов, поскольку в базе данных хранятся только номера.

3. Сократить объем базы данных.

Литература

1. Белов Е.Б. Основы информационной безопасности / Е.Б. Белов, В.П. Лось, Р.В. Мещеряков, А.А. Шелупанов. – М.: Горячая линия – Телеком, 2006. – 544 с.
2. Кручинин В.В. Методы построения алгоритмов генерации и нумерации комбинаторных объектов на основе деревьев И/ИЛИ. – Томск: В-Спектр, 2007. – 200 с.

Кручинин Владимир Викторович

Канд. техн. наук, зам. директора института инноватики ТУСУРа
Тел.: (+7 382-2) 41-50-00
Эл. адрес: kru@tdec.ru

Люкшин Борис Александрович

Доктор техн. наук, профессор, зав. каф. механики, графики и управления качеством ТУСУРа
Тел.: (+7 382-2) 51-05-30
Эл. адрес: lba2008@yandex.ru

V.V. Kruchinin, B.A. Lukschin

Method of Coding of Information Objects on the Basis of Trees And-Or

Questions of application of a method of construction of algorithms of the combinatory generation, based on trees And-Or for coding of information objects are considered. The method of coding based on representation of all set of information objects in the form of a tree And-Or is offered.

Keywords: coding, information object, methods of generation of combinatory sets, a tree And-Or.