М Е Т О Д Ы   А Н А Л И З А   Д А Н Н Ы Х
(Вычислительные системы)
1985 год   Выпуск III

УДК 519.812.2+519.7

КОНСТРУКТИВНОЕ ЧИСЛОВОЕ ПРЕДСТАВЛЕНИЕ ВЕЛИЧИН

Е.Е. Витяев

Исследования, проводимые в психологии, социологии, принятии решений, экспертном оценивании и других областях, показывают, что есть много сложных, структурных "нечисловых" величин (частичные порядки, толерантности, решетки и т.д.). Логический анализ таких величин, проведенный в теории измерений [1,2], теории принятия решений [3,4] и анализе нечисловой информации [5-7], показал, что формальные представления таких величин - эмпирические системы - являются такими алгебраическими структурами, которые нельзя сильным гомоморфизмом отобразить в поле вещественных чисел, т.е. для таких величин нельзя построить их числовые представления в теории измерений. С другой стороны, числовые представления величин обладают следующими достоинствами: они "удобны", по числовым значениям величин легко определяются исходные (в эмпирической системе) соотношения между значениями величин, для числовых величин разработано много математических методов их обработки. Поэтому, наряду с необходимостью разрабатывать "прямые" (например, логические) методы обработки структурных "нечисловых" величин, остается важной задача построения их числовых представлений. Именно этим мотивировалось включение этой задачи в программную систему логического анализа данных ЛАДА (см.наст.сб.).

Рассмотрим смысл и роль числового представления. Смысл состоит в том, чтобы значениям величины приписать числа так, что исходные отношения и операции преобразуются в некоторые "простые" и "удобные" числовые отношения и операции. В этом случае по значениям числовых отношений и операций легко определяются значения исходных отношений и операций.

В теории измерений числовое представление строится по следующей схеме. Предполагается, что величины измеряются некоторыми конкретными приборами. Свойства величин исследуются экспериментально и фиксируются в законах и гипотезах. Эмпирические системы величин являются идеализациями соответствующих приборов и процедур их использования. Величины известны нам в той мере, в какой известны соответствующие законы и гипотезы. Формально законы и гипотезы задаются системами аксиом, поэтому относительно эмпирических систем известно только, что они удовлетворяют соответствующей системе аксиом. Числовые представления величин получаются подбором такой числовой системы, в которую сильным гомоморфизмом отображается любая эмпирическая система, удовлетворяющая системе аксиом, описывающей данную величину.

Построенные по такой схеме числовые представления обладают следующими недостатками. В качестве числовых отношений и операций используется небольшое число обычных математических действий. Этого достаточно для числового представления большинства физических величин [2], но это препятствует числовому представлению многих других величин. Доказательство, что любая эмпирическая система, удовлетворяющая системе аксиом, сильным гомоморфизмом отображается в выбранную числовую систему, предъявляет чрезмерно сильные требования к системе аксиом. Приходится включать в нее аксиомы, не поддающиеся экспериментальной проверке, а также "чисто технические" аксиомы, не изменяющие множества экспериментально проверяемых следствий [I]. Это противоречит содержанию систем аксиом, как результатам экспериментального анализа свойств величин. Такие аксиомы часто отражают свойства числовой системы, а не свойства величин. Чрезмерным является также требование сильного гомоморфизма любой (?) алгебраической системы в числовую систему. Если теорию Т (систему аксиом) рассматривать как тип данных, а эмпирическую систему как структуру данных, то числовое представление строится для любой структуры данного типа Т . В программировании [8] тем не менее рассматриваются и другие способы задания семантики типа данных - инициальные, финальные, простые и другие структуры. Поскольку эти структуры определенным образом связаны со всеми остальными структурами типа Т, то числовое представление можно, в принципе, строить только для этих структур. Таким путем может быть получено конструктивное числовое представление экстенсивных величин. С практической точки зрения также имеет смысл получать числовые представления не для всех структур, а только для конечных, конечно-порожденных или конечных подструктур структур типа T. Отмеченные ограничения в постановке проблемы существования числового представления влекут за собой и соответствующие ограничения в постановках других проблем (проблемы единственности и адекватности).

Смыслу числового представления точнее всего соответствует понятие конструктивизации [9] эмпирической системы. В этом случае значениям величины приписываются натуральные, рациональные или другие числа (или коды) так, чтобы значения отношений и операций в эмпирической системе можно было эффективно определить по этим числам. Такой способ получения числовых представлений не накладывает на числовые отношения и операции никаких ограничений, кроме эффективности, предъявляет более слабые требования к системе аксиом и не связан с требованием существования гомоморфизма в какие-то другие системы. Этот способ называется в работе конструктивным числовым представлением и может использоваться для числового представления структурных "нечисловых" величин. Заметим, что для конструктивных числовых представлений справедливы некоторые из замечаний, сделанных относительно постановки основных проблем (существования, единственности и адекватности) в теории измерений. Их учет открывает широкие возможности для дальнейших исследований.

I. Основные понятия теории измерений. Пусть знания о некоторой величине, свойстве, признаке сформулированы в виде некоторой теории T сигнатуры $\Omega = \langle P_0, P_1, \ldots, P_n, \rho_1, \ldots, \rho_m, c_0, c_1, c_2, \ldots \rangle$, где P_i, $i \leq n$, - предикатные символы; ρ_j, $j \leq m$, - символы операций; c_l, $l \in I$, - символы констант ($I = \emptyset$, I - начальная часть $\omega = \{0, 1, 2, \ldots\}$, $I = \omega$); P_0 - равенство. Величиной будем называть неприводимую [I] (равенство является единственным отношением конгруэнтности) систему $\mathfrak{M} = \langle A; \Omega_M \rangle$ сигнатуры Ω, удовлетворяющую теории T; A - множество значений величины, $\Omega_M = \{P_0^M, P_1^M, \ldots, P_n^M, \rho_1^M, \ldots, \rho_m^M, c_0^M, c_1^M, c_2^M, \ldots\}$ - множество отношений, операций и констант типа Ω, интерпретируемых в понятиях предметной области. Числовыми системами называются системы $\mathcal{R} = \langle Re^k, \Omega_R \rangle$ сигнатуры Ω, где k - размерность числового представления, Re - поле вещественных чисел, $\Omega_R = \{=, P_1^R, \ldots, P_n^R, \rho_1^R, \ldots, \rho_m^R, c_0^R, c_1^R, c_2^R, \ldots\}$ - множество отношений, операций и констант, определенных на Re или Re^k. Зафиксируем некоторую числовую систему $\mathcal{R}$. Шкалой [I] (числовым представлением) величины $\mathfrak{M} = \langle A; \Omega_M \rangle$ называется отображение (сильный гомоморфизм) μ: $A \to Re^k$, удовлетворяющее условиям:

1. $P_i^M(a_1, \ldots, a_{m_i}) \Leftrightarrow P_i^R(\mu a_1, \ldots, \mu a_{m_i})$, $i = 0, 1, \ldots, n$;

2. $\mu \rho_j^M(a_1, \ldots, a_{m_j}) = \rho_j^R(\mu a_1, \ldots, \mu a_{m_j})$, $j = 1, 2, \ldots, m$;

3. $\mu c_l^M = c_l^R$, $l \in I$.

Введем обозначения. $AC(T)$ – множество неприводимых (алгебраических) систем теории T; $AC_1(T), AC_0(T)$ – подмножества $AC(T)$, содержащие системы не более чем континуальной и счетной мощности соответственно; $F(\mathcal{M}, \mathcal{R})$ – множество шкал величины $\mathcal{M}$.

В теории измерений исследуются три основные проблемы [I,2].

I. Проблема существования. Для данной теории T величины найти достаточно простую и удобную числовую систему $\mathcal{R}$ (например, поле вещественных чисел) и доказать, что для любой величины $\mathcal{M} \in AC_1(T)$ существует шкала ($F(\mathcal{M}, \mathcal{R}) \neq \emptyset$).

Из формулировки проблемы существования следует, что знаний T должно быть достаточно для выбора числовой системы $\mathcal{R}$ и построения шкалы для любой системы $\mathcal{M} \in AC_1(T)$. Системы из $AC_1(T)$ являются величинами, которые удовлетворяют нашим знаниям T о них и для которых мы можем построить числовое представление. Решение проблемы существования должно, кроме того, давать метод шкалирования приборов, измеряющих эти величины. Этот метод обычно извлекается из доказательства теоремы существования.

2. Проблема единственности. Для выбранной числовой системы $\mathcal{R}$ определить все шкалы $F(\mathcal{M}, \mathcal{R})$ величин $\mathcal{M} \in AC_1(T)$. Эти множества можно, в частности, определить, найдя группу допустимых преобразований [I].

Обычно требуется, чтобы не только числовая система, но и все множества $F(\mathcal{M}, \mathcal{R})$ были просты и удобны. Простота и удобство нужны для решения следующей проблемы.

3. Проблема адекватности. Числовые утверждения должны быть инвариантны относительно произвола в выборе шкал из $F(\mathcal{M}, \mathcal{R})$ (см. [I]).

Решение этих проблем позволяет корректно вводить числовые представления величин и в определенной степени корректно их использовать.

2. Конструктивные представления величин. При конструктивном представлении величин значения $a \in A$ величин $\mathcal{M} = \langle A; \Omega_M \rangle \in AC_0(T)$ нумеруются (кодируются). Нумерацией множества A называется отображение ν множества натуральных чисел $\omega = \{0, I, 2, \ldots\}$ на

A, $\nu: \omega \to A$ [9]. Пару $(\mathfrak{M},\nu)$ будем называть конструктивным представлением величины $\mathfrak{M}$ (конструктивной системой [9]), а нумерацию ν – конструктивным числовым представлением (конструктивизацией [9]), если существуют характеристические общерекурсивные функции $P_0^N, P_1^N, \ldots, P_n^N$ со значениями во множестве $\{0,I\}$, общерекурсивные функции $\rho_1^N, \ldots, \rho_m^N$ и натуральные числа $c_0^N, c_1^N, c_2^N, \ldots$ такие, что

1) $P_i^N(n_1, \ldots, n_{m_i}) = 1 \leftrightarrow P_i^M(\nu n_1, \ldots, \nu n_{m_i})$, $i = 1, \ldots, n$;

2) $\nu \rho_j^N(n_1, \ldots, n_{m_j}) = \rho_j^M(\nu n_1, \ldots, \nu n_{m_j})$, $j = 1, \ldots, m$;

3) $\nu c_l^N = c_l^M$, $l \in I$.

Конструктивное числовое представление ν аналогично шкале, только вместо числовых отношений, операций и констант используются общерекурсивные функции и натуральные числа. Конструктивной числовой системой является система $\mathfrak{N} = \langle \omega; P_0^N, P_1^N, \ldots, P_n^N, \rho_1^N, \ldots, \rho_m^N, c_1^N, c_1^N, \ldots \rangle$. Сформулируем проблемы существования, единственности и адекватности для конструктивного числового представления.

I. Проблема существования I. Доказать, что для любой системы $\mathfrak{M} \in AC_0(T)$ существует конструктивное числовое представление.

Данная формулировка предъявляет довольно сильные требования к теории T. Более слабой, но также практически интересной является следующая формулировка проблемы существования. Пусть Φ – система аксиом теории T, Φ^* – совокупность эрбрановых форм предложений Φ (скулемизация [9,с.345]), $f_1, f_2, \ldots$ – символы скулемовских функций. Определим сигнатуру $\Omega' \rightleftharpoons \Omega \cup \{f_1, f_2, \ldots\}$. Скулемовской оболочкой $\mathfrak{M}^*(X)$ подмножества $X \subset |\mathfrak{M}^*|$ системы $\mathfrak{M}^* \in AC(\Phi^*)$ сигнатуры Ω' называется подсистема системы $\mathfrak{M}^*$, порожденная подмножеством X. Можно доказать [9], что $\mathfrak{M}^*(X) \in AC(\Phi^*)$ для любого подмножества $X \subset |\mathfrak{M}^*|$.

Проблема существования 2. Доказать, что для любой величины $\mathfrak{M}^* \in AC(\Phi^*)$ сигнатуры Ω' и любого конечного подмножества $X \subset |\mathfrak{M}^*|$ скулемовская оболочка $\mathfrak{M}^*(X)$ имеет конструктивное числовое представление.

2. Проблема единственности. Ее можно разбить на две части. Первая связана с существованием не сводимых друг к другу посредством эффективного отображения (неавтоэквива -

лентных [9]) конструктивных числовых представлений. В [10] показано, что число неавтоэквивалентных конструктивизаций может быть любым. Неавтоэквивалентные конструктивные числовые представления принципиально различны, поэтому необходимо учитывать возможный произвол в выборе одной из них.

<u>Проблема единственности А.</u> Для каждой величины $\mathcal{M} \in AC_0(T)$ определить число неавтоэквивалентных конструктивных числовых представлений.

Вторая часть проблемы единственности так же, как и в случае числовых представлений, связана с произволом в выборе одного из автоэквивалентных конструктивных числовых представлений.

<u>Проблема единственности Б.</u> Для каждой величины $\mathcal{M} \in AC_0(T)$ определить все классы автоэквивалентных конструктивных числовых представлений.

3. П р о б л е м а а д е к в а т н о с т и. Она также разбивается на две части в зависимости от того, какой произвол в выборе конструктивных числовых представлений нужно учитывать.

<u>Проблема адекватности А.</u> Выбор класса автоэквивалентных числовых представлений должен учитывать имеющиеся знания Т.

<u>Проблема адекватности Б.</u> Числовые утверждения должны быть инвариантны относительно выбора одного из автоэквивалентных конструктивных числовых представлений.

3. <u>Взаимосвязь конструктивного и числового представлений.</u> Предположим, что для некоторой величины, описываемой теорией T, решены проблемы существования как для числового, так и для конструктивного числового представлений. Пусть $\mathcal{R}$ – выбранная числовая система, $\mathcal{M} \in AC_1(T)$ – величина и $\mu : \mathcal{M} \to \mathcal{R}$ – шкала. Из решения проблемы существования конструктивного числового представления следует, что для любой не более чем счетной величины $\mathcal{N} \in AC_0(T)$, являющейся подсистемой величины $\mathcal{M}$, $\mathcal{N} \subset \mathcal{M}$, существует конструктивное числовое представление $\nu : \omega \to |\mathcal{N}|$. Рассмотрим образ $\mu(\mathcal{N})$ величины $\mathcal{N}$ при его отображении в числовую систему $\mathcal{R}$. Так как подсистема $\mathcal{N}$ содержит все константы c_i^M, $i \in I$, и замкнута относительно операций, то из определения шкалы следует, что отображение $\mu : \mathcal{N} \to \mathcal{R}$ также является шкалой величины $\mathcal{N}$. Поэтому для каждой подсистемы $\mathcal{N} \in AC_0(T)$ любой из величин $\mathcal{M} \in AC_1(T)$, $\mathcal{N} \subset \mathcal{M}$ существуют как конструктивное ν, так и числовое μ представления. Рассмотрим отображение $\mu\nu : \omega \to \mu(\mathcal{N})$. Из определений шкалы и конструктивного числового представления следует, что пара $(\mu\mathcal{N}, \mu\nu)$ яв-

ляется конструктивным представлением числового представления $\mu\mathfrak{N}$ величины $\mathfrak{N}$. Поэтому для величин $\mathfrak{N} \in AC_0(T)$, $\mathfrak{N} \subset \mathfrak{M}$, $\mathfrak{M} \in AC_1(T)$ существуют конструктивное числовое представление ν, числовое представление μ и конструктивное представление $\mu\nu$ числового представления $\mu\mathfrak{N}$.

4. Примеры конструктивных представлений величин. Рассмотрим линейный порядок. Знания T о линейном порядке содержат аксиомы антисимметричности, полноты и транзитивности. Линейными порядками являются, например, балльные величины и величины типа "число": число рабочих на предприятии, число автокатастроф, число браков или разводов и т.д. Значениями многих таких величин являются натуральные числа, поэтому их естественным числовым представлением может быть конструктивное числовое представление. Такие величины удовлетворяют также аксиоме I.

АКСИОМА I. Любая ограниченная сверху (снизу) последовательность $a_1 < a_2 < a_3 < \ldots < a$ $(a < \ldots < a_3 < a_2 < a_1)$ конечна.

Обозначим теорию линейного порядка T вместе с аксиомой I через T_1. Известно, что любой линейный, но не более чем счетный порядок, удовлетворяющий теории T_1, вложим в модель $\langle\omega; \leq\rangle$. Отсюда следует решение проблемы существования конструктивного числового представления для таких линейных порядков.

ПРЕДЛОЖЕНИЕ I. Любой линейный порядок $\mathfrak{M} \in AC_0(T_1)$ имеет конструктивное числовое представление.

Конструктивными числовыми представлениями могут служить обычные способы нумерации значений этих величин.

Рассмотрим линейные порядки, удовлетворяющие аксиоме полноты.

АКСИОМА II.

$$\forall a, b\, \exists c (a < c < b).$$

Обозначим через T_2 теорию линейного порядка T вместе с аксиомой II. Примерами полных линейных порядков, удовлетворяющих T_2, являются физические величины, используемые в нефизических областях, например, температура, давление, вес человека, рассматриваемые с медицинской точки зрения, или температура, освещенность, влажность почвы, рассматриваемые с сельскохозяйственной точки зрения. Для этих величин операция сложения (имеющая смысл в физике) смысла не имеет. Осмысленно только отношение порядка, являющееся полным линейным порядком. Такой порядок естественно пред - ставлять не действительными, а рациональными числами. Получим конструктивное числовое представление полных линейных порядков, используя рациональные числа. Известно, что любой полный, не более чем счетный линейный порядок $\mathfrak{M} \in AC_0(T)$ изоморфен одному из интервалов (0,I), [0,I), (0,I], [0,I] множества рациональных чисел.

ПРЕДЛОЖЕНИЕ 2. Любой полный линейный порядок $\mathfrak{M} \in AC_0(T_2)$ имеет конструктивное числовое представление.

Конструктивными числовыми представлениями могут служить градации шкал приборов, измеряющих эти величины.

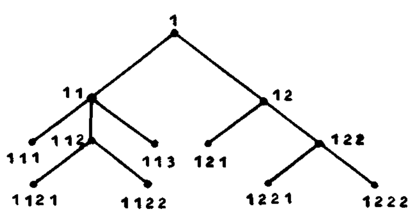
Рассмотрим деревья - рефлексивные, антисимметричные, транзитивные порядки, удовлетворяющие следующей аксиоме.

АКСИОМА Ш.

$$\forall a,b,c\ (c \leq a \ \&\ c \leq b \Rightarrow a \leq b \vee b \leq a).$$

Обозначим теорию деревьев через T_3. Конечными деревьями описываются такие величины,как должность, подчиненность, занимаемое место (в дереве рабочих мест некоторой организации), иерархиче - ские величины и т.д. Конечные деревья всегда конструктивизируемы, поэтому решение проблемы существования конструктивного числового представления сводится к построению простой и удобной конструктивизации, применимой к любому конечному дереву. Пример такого конструктивного числового представления приведен на рисунке.

Если у дерева несколько корневых вершин, то они нумеруются числами I,2,3,... Вершинам дерева (значениям величины) сопоставляются наборы натуральных чисел $a = \nu(\langle n_1^a,\dots,n_k^a\rangle)$, $b = \nu(\langle n_1^b,\dots,n_l^b\rangle)$. По числам из набора легко определяется отношение порядка между a и b.



Закономерные связи между конструктивными числовыми представлениями также должны задаваться общерекурсивными функциями.Например, закономерная связь между древовидными величинами - занимае - мое место, должность, подчиненность; балльными величинами - сте - пень, звание; полно линейно упорядоченной величиной - зарплата,определяется законодательством и может быть представлена общерекурсивной функцией.

5. <u>Конструктивные измерительные процедуры, тесты и анкеты.</u> Шкалы μ: $\mathfrak{M} \to \mathfrak{R}$ практически реализуются в виде шкал приборов (весов, линейки, термометра). Конструктивизации ν также могут реализовываться как показания некоторых измерительных процедур, в частности тестирования, анкетирования, обследования и т.д. Предположим, что нас интересует отношение предпочтения некоторой величины $\mathfrak{M} = \langle A; \leq \rangle$ (коэффициента интеллектуальности, удовлетворенности работой, темперамента) и способ прямого измерения отношения предпочтения $\leq$ дорог, неудобен, требует большого времени и т.д. Для более простого и быстрого измерения этого отношения разрабатывается и используется тест (анкета, обследование). Применение теста к испытуемому (респонденту, больному) дает для некоторого значения $a \in A$ величины $\mathfrak{M}$ набор ответов в виде некоторой последовательности натуральных или рациональных чисел $\langle n_1^a, \ldots, n_k^a \rangle$. Если по результатам теста для любых двух значений $a, b \in A$ величины $\mathfrak{M}$ можно эффективно определить отношение предпочтения

$$a \leq b \leftrightarrow P^N(\langle n_1^a, \ldots, n_k^a \rangle, \langle n_1^b, \ldots, n_l^b \rangle)$$

(например, подсчитывая сумму баллов, взвешенное среднее, кодируя ответы и т.д.), то отображение ν: $\langle n_1^a, \ldots, n_k^a \rangle \to a$, осуществляемое тестом, и будет конструктивным числовым представлением величины $\mathfrak{M} = \langle A; \leq \rangle$. Сама процедура тестирования (анкетирования, обследования) будет конструктивной измерительной процедурой со значениями вида $\langle n_1, \ldots, n_k \rangle$.

Примером такого отношения предпочтения и соответствующего теста является предпочтение между односемейными домами, разобранное в [II, с.243].

Использование теста для конструктивного измерения некоторой величины может быть обосновано решением следующей задачи. Пусть задано некоторое отношение предпочтения $\leq$ величины $\mathfrak{M}_0 = \langle A; \leq \rangle$, обладающее свойствами T (удовлетворяющее аксиомам частичного порядка, толерантности, решеток). Требуется решить проблему существования конструктивного числового представления ν для любой величины $\mathfrak{M} \in AC_0(T)$ и затем для данной нам величины $\mathfrak{M}_0$ разработать тест измеряющий конструктивное числовое представление ν величины $\mathfrak{M}_0$. Мы не можем сразу строить конструктивное числовое представление ν для исходной величины $\mathfrak{M}_0$, так как она известна нам только своими аксиомами, содержащимися в T. Поэтому решить проблему существования конструктивного числового представления нужно опираясь на $AC_0(T)$.

В работе получены конструктивные числовые представления лишь для простейших величин. Конструктивные числовые представления конечных решеток получены В.Ф.Новиковым. Дальнейшее построение конструктивных числовых представлений может быть стимулировано применением программной системы логического анализа данных ЛАДА.

Литература

1. ПФАНЦАГЛЬ И. Теория измерений. -М.: Мир, 1976. - 248 с.

2. KRANTZ D.H., LUCE R.D., SUPPES P., TVERSKY A. Foundations of measurement. V.1. - New York and London: Academic press,1971.- 570 p.

3. ФИШБЕРН П. Теория полезности для принятия решений. -М.: Наука, 1978. - 352 с.

4. КИНИ Р.Л., РАЙФА Х. Принятие решений при многих критериях: предпочтения и замещения. -М.: Радио и связь, 1981. - 560 с.

5. Анализ нечисловой информации /Тюрин Ю.Н., Литвак Б.Г.,Орлов А.И. и др. - М. Б.и., 1981. - 80 с. (Препринт/Научный совет по Комплексной проблеме "Кибернетика").

6. П Всесоюзная школа-семинар "Программно-алгоритмическое обеспечение прикладного многомерного статистического анализа".(Тезисы докладов). - М., 1983. - 374 с.

7. Всесоюзная конференция "Нечисловая статистика, экспертные оценки и смежные вопросы" (Тезисы докладов). -М.-Таллин, 1984, - [illegible] с.

8. Данные в языках программирования. Абстракция и типология. -М.: Мир, 1982. - 327 с.

9. ЕРШОВ Ю.Л. Проблемы разрешимости и конструктивные модели. -М.: Наука, 1980. - 415 с.

10. ГОНЧАРОВ С.С. Проблема числа неавтоэквивалентных конструктивизаций. -Алгебра и логика, 1980, т.19, №6, с.621-639.

11. КОЗЕЛЕЦКИЙ Ю. Психологическая теория решений. -М.: Прогресс, 1979. - 504 с.

Поступила в ред.-изд.отд.
10 сентября 1985 года