АНАЛИЗ РАЗНОТИПНЫХ ДАННЫХ
(Вычислительные системы)
1983 год Выпуск 99

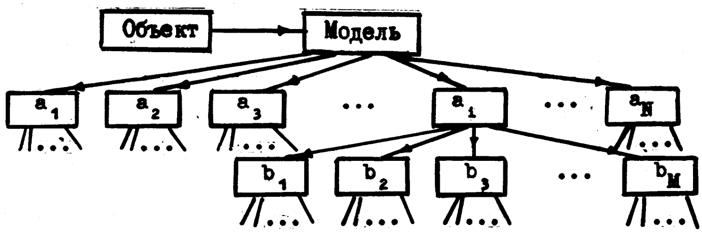
УДК 518.74

# ЛОГИЧЕСКАЯ ПРОСТОТА ИЕРАРХИЧЕСКИХ СТРУКТУР

Е.М.Черепанов

I. Во многих областях нашей деятельности приходится сталкиваться с объектами и явлениями, при описании которых приходится применять поэтапный метод структурного анализа. В начале выделяется некоторое конечное множество "опорных" и существенных в нашем рассмотрении элементов, затем устанавливаются отношения, которыми эти элементы между собой связаны. При этом полагается, что выделенные элементы в каком-то смысле независимы, автономны. Примером этому может служить программный комплекс, где каждый программный блок структурно и функционально автономен, как программная единица, но при этом связан с другими блоками определенными отношениями. Таким же образом можно рассматривать, например, крупный завод, в котором структурно и функционально автономные цеха связаны друг с другом определенными отношениями. Поэтому будем полагать, что существует некий, либо интуитивный, либо вполне точный критерий, позволяющий нам различать эти элементы и объявлять их на данном этапе рассмотрения примитивными и далее не анализируемыми. Такого сорта анализ объекта наблюдения приводит нас к построению конечной модели $W = \langle U, P_1^{m_1}, \dots, P_n^{m_n} \rangle$, где $U = \{a_1, a_2, \dots, a_N\}$ - конечное множество выделенных элементов, а $P_i^{m_i}$ - имя $i$-го $m_i$ -местного отношения, заданного на множестве $U$ .

Во многих случаях построение такого сорта модели не удовлетворяет нашим целям. В этих случаях приходится переходить к рассмотрению выделенного множества элементов, отношения на которых, как нам кажется, характеризуют явление или объект с нужной нам стороны. Может оказаться, как в приведенных выше примерах, что некоторые или все из этих элементов в свою очередь достаточно сложны и, следовательно,появляется возможность описания их логической структуры конечной моделью. Этот процесс "дробления" объекта наблюдения может идти, в принципе, сколь угодно долго. Как правило, критерием остановки служит достижение цели, которую мы преследуем, т.е. интуитивно понимаемая степень адекватности нашего описания наблюдаемому объекту. В конечном итоге мы получаем иерархическую древовидную структуру (см. рисунок).




Естественно, что возможен и прямо противоположный путь, приводящий к построению такого рода иерархических структур.Например, для всякой классификации типичными являются объединения множества объектов в группы по совокупности признаков. Полученные группы в свою очередь связаны между собой какими-то отношениями и могут быть объединены в группы более высокого ранга и т.д.

В дальнейшем нас будут интересовать лишь иерархические структуры вышеописанного типа и мы попытаемся найти метод оценки значения сложности такого рода структур.

Будем называть уровнем иерархии иерархической структуры совокупность всех моделей, находящихся на одной и той же глубине от вершины, таким образом построенного дерева.

Будем говорить, что иерархическая структура имеет ранг $k$, если она имеет $k$ уровней иерархии. Для того чтобы оценить значение сложности такого рода иерархических структур, необходимо определить влияние сложности структурной организации нижнего уровня на значение сложности соседнего верхнего уровня.

2. Допустим,для простоты, что рассматриваемый объект описывается иерархической структурой ранга 2. Пусть первый уровень представлен конечной моделью $W = \langle U, P_1^{m_1}, \ldots, P_n^{m_n} \rangle$ и множество $U = \{A_1, A_2, \ldots, A_N\}$ есть множество моделей второго уровня, где для всех $i = 1, 2, \ldots, N$ $A_i = \langle U_i, Q_{1i}^{m_{1i}}, \ldots, Q_{ni}^{m_{ni}} \rangle$. Таким образом, $n$-ки отношений модели $W$ определены на элементах множества $U$, которые в свою очередь являются моделями второго уровня иерархии.

Интуитивно ясно, что более простой иерархической структурой будет та, в которой все модели второго уровня в каком-то смысле подобны друг другу, а не та, в которой они резко отличаются одна от другой. Поэтому необходимо ввести какие-то формальные критерии, сокращающие значение сложности нашей структуры, в зависимости от того, насколько подобны друг другу модели второго уровня. Произвольное отношение $P$ модели $W$ есть $n$-ка моделей $\langle A_{i_1}, A_{i_2}, \ldots, A_{i_n} \rangle$. Ясно, что если любая другая $n$-ка этого отношения в каком-то подходящем смысле подобна этой, то данная структура нам кажется более простой, чем та, в которой этого подобия нет. Попытаемся формализовать это наше интуитивное предположение.

Для каждого отношения $P_i^{m_i}$ модели $W$ рассмотрим два множества $n$-ок:

$$\mathcal{H}^{m_i} = \{\langle x_1, \ldots, x_{m_i} \rangle \mid P_i^{m_i}(x_1, \ldots, x_{m_i})\}$$

и

$$\tilde{\mathcal{H}}^{m_i} = \{\langle x_1, \ldots, x_{m_i} \rangle \mid \sim P_i^{m_i}(x_1, \ldots, x_{m_i})\}.$$

Определим на каждом из множеств $\mathcal{H}^{m_i}$ и $\tilde{\mathcal{H}}^{m_i}$ отношение эквивалентности $\sigma$. Будем считать, что две $n$-ки $\sigma$-эквивалентны друг другу, если и только если $\sigma$-эквивалентны в этих $n$-ках пары моделей с одинаковыми номерами (какой может быть эта эквивалентность, будет рассмотрено ниже). Эквивалентность $\sigma$ порождает пару фактор-множеств $\mathcal{H}^{m_i}/\sigma$ и $\tilde{\mathcal{H}}^{m_i}/\sigma$. Пусть $\mathcal{H}^{m_i}/\sigma = \{E_{i_1}, \ldots, E_{i_k}\}$ и $\tilde{\mathcal{H}}^{m_i}/\sigma = \{\tilde{E}_{i_1}, \ldots, \tilde{E}_{i_l}\}$.

Теперь необходимо установить, каким образом мы можем оценивать значение сложности таким образом полученных фактор-множеств, состоящих из классов $n$-ок моделей. Для достижения этой цели введем полезный принцип, выраженный следующим постулатом:

PI. Пусть $A/\sigma$ – фактор-множество множества $n$-ок моделей по эквивалентности $\sigma$ и $A/\sigma = \{E_1, \ldots, E_n\}$, тогда значение сложности $v(A/\sigma) = \sum_{i=1}^{n} v(E_i)$.

Так как каждый класс фактор-множества содержит $\sigma$-эквивалентные $n$-ки моделей, будем считать значением сложности отдельного класса значение сложности наиболее сложного представителя этого класса.

Р2. Пусть $E \in A/\sigma$ - класс эквивалентности и $E = \{N_1,\dots,N_m\}$, тогда $v(E) = \max\{v(N_1),\dots,v(N_m)\}$.

Каждая n-ка в нашем рассмотрении есть n-ка моделей второго уровня иерархии. Будем полагать значением сложности n-ки моделей сумму значений сложности ее составляющих моделей.

Р3. Пусть $N = \langle A_1,\dots,A_n \rangle$ - n-ка моделей, тогда $v(N) = \sum_{i=1}^{n} v(A_i)$.

Таким образом, мы пришли к тому, что мы умеем делать, т.е. к вычислению значения сложности отдельной модели. По определению [I], если $W = \langle U, P_1^{m_1},\dots,P_n^{m_n} \rangle$ – модель конечной сигнатуры, то $v(W) = \sum_{i=1}^{n} v(P_i^{m_i})$, где $v(P_i^{m_i})$ - значение сложности $P_i^{m_i}$. О том, как осуществляется вычисление значения сложности каждого отношения модели, можно узнать в [I].

С учетом всего вышесказанного теперь можно определить значение сложности фактор-множеств $\mathcal{K}^{m_i}/\sigma$ и $\tilde{\mathcal{K}}^{m_i}/\sigma$ :

$$v(\mathcal{K}^{m_i}/\sigma) = \sum_{p=1}^{i_k} v(E_{ip}) = \sum_{p=1}^{i_k} \sum_{q=1}^{m_i} v(A^*_{pq}),$$

где $\langle A^*_{p1},\dots,A^*_{pm_i} \rangle$ - выделенный представитель класса $E_p$, поставляющий максимальное значение сложности. Аналогичным образом вычисляем и $v(\tilde{\mathcal{K}}^{m_i}/\sigma)$. Будем считать, что полное значение сложности, поставляемое носителем модели отношению $P_i^{m_i}$, есть: $v(U^{m_i}) = \max\{v(\mathcal{K}^{m_i}/\sigma), v(\tilde{\mathcal{K}}^{m_i}/\sigma)\}$. Тогда, обозначив через $v^{(2)}$ значение сложности иерархической структуры ранга 2, будем полагать, что $v^{(2)}(W) = \sum_{i=1}^{n} v(P_i^{m_i}) v(U^{m_i})$.

3. Теперь необходимо рассмотреть все те вопросы, которые связаны с построением фактор-множеств $\mathcal{K}^{m_i}/\sigma$ и $\tilde{\mathcal{K}}^{m_i}/\sigma$ для каждого отношения $P_i^{m_i}$ и сформулировать условия выбора представителя каждого класса эквивалентности. Ясно, что при сравнении двух произ-

вольных n-ок моделей, соответствующих отношению $P_i^{n_i}$, возможны два случая. Первый - когда модели с одинаковыми номерами внутри n -ок являются однотипными и второй - когда эти модели не однотипны. Для того чтобы в обоих случаях иметь возможность устанавливать подобие одной n-ки другой, рассмотрим два типа эквивалентности - сильную эквивалентность ($\approx$) и слабую эквивалентность ($\sim$).

ОПРЕДЕЛЕНИЕ I. Пусть $N = \langle A_1, \ldots, A_n \rangle$ и $M = \langle B_1, \ldots, B_n \rangle$ - n-ки отношения P. Будем считать, что $N \approx M$, если и только если для всех $i = 1, 2, \ldots, n$ существует изоморфизм $\varphi_i$ такой, что $\varphi_i : A_i \to B_i$.

В случае сильной эквивалентности все члены одного из классов полученного фактор-множества будут иметь одно и то же значение сложности и поэтому выбор представителя класса произволен. За деталями обоснования этого факта автор отсылает к работам [I,2].

ОПРЕДЕЛЕНИЕ 2. Пусть $N = \langle A_1, \ldots, A_n \rangle$ и $M = \langle B_1, \ldots, B_k \rangle$ - n-ки отношения P. Будем считать, что N изоморфно вложима в M, $N \subseteq M$, если и только если для всех $i = 1, \ldots, n$ существует изоморфизм $\varphi_i$ такой, что $\varphi_i : A_i \to B_i^*$, где $B_i^* \subset B_i$ есть модель с обедненной сигнатурой.

ОПРЕДЕЛЕНИЕ 3. Пусть $N = \langle A_1, \ldots, A_n \rangle$ и $M = \langle B_1, \ldots, B_n \rangle$ - n-ки отношения P. Будем считать, что $N \sim M$, если и только если существует n-ка отношения P $K = \langle C_1, \ldots, C_n \rangle$ такая, что $N \subseteq K$ и $M \subseteq K$.

Легко проверить, что отношение слабой эквивалентности $\sim$ в общем случае будет рефлексивно и симметрично, но не транзитивно и поэтому может называться эквивалентностью лишь условно, так как всегда можно подобрать такой случай, что $N \sim M$ по вложимости в $L_1$, а $M \sim K$ по вложимости в $L_2$, однако может оказаться, что $N \not\sim K$. Для того чтобы избежать такого рода неприятности, будем подразумевать, что сравниваемые нами n-ки моделей связаны отношением слабой эквивалентности по вложимости в определенную фиксированную n-ку моделей. Тогда все n-ки моделей, объединенные в классы эквивалентности по такому признаку, будут эквивалентны в строгом смысле, т.е. отношение эквивалентности $\sim$ будет также и транзитивно.

Теперь сформулируем условия выбора представителя класса из фактор-множества по эквивалентности $\sim$.

ОПРЕДЕЛЕНИЕ 4. Пусть $E_p$ - класс фактор-множества по эквивалентности $\sim$. Будем называть n-ку $N_1$ представителем этого класса, если и только если для любой произвольной n-ки $N_i \in E_p$, $N_i \subseteq N_1$.

Условимся, что для каждого класса эквивалентности $E_p = \{N_1, \dots, N_k\}$ на первом месте стоит представитель этого класса. Для всех $i = 1, \dots, k$, $v(N_1) \geq (N_i)$. Построим теперь алгоритм образования фактор-множества множества n-ок отношения по слабой эквивалентности, индуцированной изоморфной вложимостью.

Пусть $\mathcal{H}^n = \{\langle x_1, \dots, x_n \rangle \mid P(x_1, \dots, x_n)\} = \{N_1, \dots, N_M\}$. Выберем из $\mathcal{H}^n$ $N_{i_1}$ такую, что $v(N_{i_1}) = \max\{v(N_1), \dots, v(N_M)\}$. Теперь из множества $\mathcal{H}^n \setminus N_{i_1}$ выберем все n-ки, изоморфно вложимые в $N_{i_1}$, и обозначим этот класс $E_1 = \{N_{i_1}, \dots, N_{i_k}\}$. Допустим, что $k \leq M$, тогда в множестве $\mathcal{H}^n \setminus E_1$ снова выберем n-ку с максимальным значением сложности и, проделав ту же процедуру, построим класс $E_2$ и т.д. Процесс построения классов оборвется тогда и только тогда, когда $\mathcal{H}^n \setminus \bigcup\limits_p^L E_p = \emptyset$. Допустим, что на каком-то шаге k $v(N_i) = v(N_j)$ и $N_i \not\sim N_j$, тогда, проделав процедуру построения класса $E_k$ для обоих представителей, из двух классов выберем класс наибольшей мощности. Ясно, что таким образом построенное разбиение будет удовлетворять всем требованиям определенной выше слабой эквивалентности.

Покажем теперь, что такое разбиение единственно. Допустим, что это не так. Тогда существует такое разбиение $\mathcal{H}^{*n}/\sim = \{E_1^*, \dots, E_k^*\}$, где хотя бы один класс $E_p^*$ не совпадает полностью ни с одним из классов фактор-множества $\mathcal{H}^n/\sim$. Пусть $N_{1p}^*$ – представитель этого класса, т.е. $v(E_p^*) = v(N_{1p})$. Допустим, $v(E_p^*) > v(E_1)$, но этого не может быть, так как по условию построения в $E_1$ в качестве представителя мы выбираем n-ку с максимальным значением сложности. Допустим, что $v(E_p^*) = v(E_1)$, тогда если мощности этих классов совпадают (в противном случае мы должны были выбрать класс большей мощности), то возможны два случая: а) $E_p^* \cap E_1 \neq \emptyset$ и б) $E_p^* \cap E_1 = \emptyset$. В случае "а" n-ки, принадлежащие $E_p^* \cap E_1$, изоморфно вложимы в представителей обоих классов, а так как отношение $\sim$ – эквивалентность, то n-ки множества $E_1 \setminus (E_p^* \cap E_1)$ изоморфно вложимы в представителя $E_p^*$, а n-ки множества $E_p^* \setminus (E_p^* \cap E_1)$ изоморфно вложимы в представителя $E_1$, что противоречит тому, что мы выбираем все n-ки, вложимые в представителей этих классов. В случае "б" $E_p^*$ может частично совпадать с одним из классов $\mathcal{H}^n/_\sim \setminus E_1$.

Допустим, что $v(E_p^*) \geq v(E_2)$. Повторяя все рассуждения, приходим к выводу, что и $E_p^* \cap E_2 = \emptyset$ и т.д. Перебор всех классов из $\mathcal{R}^n/\sim$ показывает, что $E_p^* \cap E_i = \emptyset$ для всех $i = 1,2,\ldots,L$. Значит, наше допущение о том, что такое разбиение не единственно, неверно.

## Литература


I. GOODMAN N. The Structure of Appearance.- Harvard:University Press,1951. - 319 p.

2. КЕЙСЛЕР Г., ЧЭН Ч.Ч. Теория моделей.М.: Мир,1977.-614 с.

3. МАЛЬЦЕВ А.И. Алгебраические системы.М.: Наука,1970.-392 с.



Поступила в ред.-изд.отд.
20 июня 1983 года