УДК 519.711.3

АЛГЕБРАИЧЕСКИЙ ПОДХОД К РЕШЕНИЮ НЕКОТОРЫХ ЗАДАЧ РЕФАКТОРИНГА

С. Д. Махортов, В. А. Погореленко

Воронежский государственный университет

Поступила в редакцию 20.04.2008 г.

Аннотация. В статье предлагается алгебраический подход к решению некоторых задач рефакторинга объектно-ориентированных систем. Описана основанная на решетках алгебраическая модель, адекватно отражающая иерархии типов и взаимосвязи между ними. Показано, какие задачи рефакторинга могут быть решены с помощью данного подхода. Доказана теорема о замкнутости представленной модели. Эта теорема лежит в основе автоматизации эквивалентных преобразований и оптимизации иерархии типов.

Ключевые слова: алгебраическая система, иерархия типов, рефакторинг, решетки, логическое замыкание, эквивалентные преобразования.

Abstract. An algebraic approach to solving some problems of refactoring of object-oriented systems is proposed in this article. We give the description of lattice-based algebraic model which is adequately reflects type-hierarchies and their relationships. Some problems of the refactoring are shown which can be solved with the help of this approach. We prove closure theorem of the presented model. This theorem is at the core of equivalent transforms automation and type-hierarchy optimization.

Key words: algebraic system, type hierarchy, refactoring, lattices, logical closure, equivalent transforms.

1. РЕШАЕМЫЕ ЗАДАЧИ РЕФАКТОРИНГА

Рефакторинг — это изменение исходного кода программной системы с сохранением ее функциональности с целью улучшения качества и продления его жизнеспособности [1]. Одной из основных задач рефакторинга является устранение избыточности кода. Рассмотрим пример такой задачи (см. рис. 1, «до рефакторинга»). Здесь классы *Student* и *Teacher* имеют общий атрибут *name* типа *String*. Также *Student* и *Teacher* унаследованы от общего типа *Person*. Можно предположить, что изначально классы *Teacher* и *Student* разрабатывались независимо друг от друга, и в каждый из них был включен атрибут *name*. Однако в дальнейшем был создан общий тип *Person*, задающий общую функциональность классов *Student* и *Teacher*. На этапе слияния целесообразно провести рефакторинг и устранить дублирование атрибута *name* с помощью «подъема поля» [1]. В результате рефакторинга дублирование общего атрибута *name* устраняется (см. рис. 1., «после рефакторинга»), но функциональность классов сохраняется — им по-прежнему доступен атрибут *name*, только теперь уже через наследование.

Данный пример является простым и проведение рефакторинга в нем может быть сделано вручную. Однако когда в программной системе имеется большое количество типов (классов), стратегия проведения рефакторинга не является столь очевидной. Подобная ситуация возни-

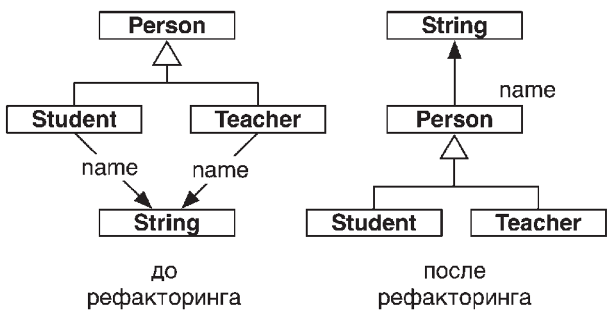

Рис. 1. Рефакторинг дублирующегося атрибута *name*

кает при реинжиниринге — модернизации устаревших программных продуктов. Также интересным является устранение дублирования в уже сложившихся библиотеках классов различных объектно-ориентированных языков программирования. Приведение большой иерархии типов в порядок вручную уже не является очевидной и простой задачей. Поэтому целесообразным является автоматизация этого процесса.

2. ОСОБЕННОСТИ ПОДХОДА

Рассматриваемый в данной статье подход дает математическую основу автоматизации устранения дублирования атрибутов в иерархии типов любой сложности. Предлагаемая алгебраическая модель отражает два вида связей между типами — наследование и агрегацию. Преимуществом данного подхода является полная формализация и широкий охват возможных ситуаций. В частности, типы атрибутов в свою очередь могут участвовать в общей иерархии. Необходимые для построения данной модели продукционно-логические структуры и соответствующая теоретическая база описаны в работах [2—3]. Основные идеи применения этих структур к иерархии типов изложены в статье [4]. Данную статью можно считать ее продолжением.

В основе используемого нами математического аппарата лежит теория решеток [5], бинарных отношений и основанные на них алгебраические структуры. Напомним некоторые основные понятия.

Решеткой называется полуупорядоченное множество $\mathbb{F}$, в котором наряду с отношением частичного порядка $\leq$ («не больше», «содержится») введены также два двуместных оператора $\wedge$ («пересечение») и $\vee$ («объединение»), дающие соответственно точную нижнюю и верхнюю грани любой пары $(a,b) \in \mathbb{F}$. Решетка $\mathbb{F}$ называется ограниченной, если она содержит общие верхнюю и нижнюю грани — такие два элемента O, I, что $O \leq a \leq I$ для любого $a \in \mathbb{F}$.

Бинарное отношение R на некотором множестве $\mathbb{F}$ называется рефлексивным, если для всех $a \in \mathbb{F}$ справедливо $(a,a) \in R$; транзитивным, если для любых $a,b,c \in \mathbb{F}$ из $(a,b),(b,c) \in R$ следует $(a,c) \in R$.

Известно, что существует замыкание R^* произвольного отношения R относительно свойств рефлексивности и транзитивности — рефлексивно-транзитивное замыкание. Пара элементов $a,c \in \mathbb{F}$ называется транзитивной в R, если $(a,c) \in R_1^*$, где R_1^* — транзитивное замыкание отношения $R_1 = R \setminus \{(a,c)\}$.

Также важной задачей является нахождение транзитивной редукции: по данному R ищется минимальное отношение R' такое, что его транзитивное замыкание совпадает с транзитивным замыканием R. Как обычно, для частично-упорядоченных множеств мы будем различать понятия минимального элемента (для него нет меньшего элемента) и наименьшего элемента (он меньше всех). В [6] приведен алгоритм построения транзитивной редукции ориентированных графов; показано, что эта задача вычислительно эквивалентна построению транзитивного замыкания, и доказана единственность транзитивной редукции ациклического графа.

3. АЛГЕБРАИЧЕСКАЯ МОДЕЛЬ ИЕРАРХИИ ТИПОВ

Теперь рассмотрим, каким образом с помощью вышеуказанного алгебраического аппарата может быть смоделирована иерархия типов.

В качестве базового множества (решетки) берется множество типов (классов) объектно-ориентированной системы, т.е. элементом решетки является тип. Будем считать, что если тип b унаследован от типа a (тип a является родителем b), то $b \leq a$. Данное отношение задает частичный порядок на множестве типов. Визуально оно проиллюстрировано на рис. 2.

В качестве результата операции объединения $\vee$ берется наименьший общий родитель, т.е. родитель двух классов, находящийся как можно ниже по иерархии наследования. Для классов *Student* и *Teacher* наименьшим общим родителем является класс *Person*. Для операции пересечения $\wedge$ результатом является наибольший общий потомок двух классов. Она актуальна для систем с множественным наследованием. Данная операция проиллюстрирована на рис. 3. В данном случае пересечением типов *LibraryUser* и *InternetUser* является их непосредственный общий наследник *Employee*.

Для ограниченности полученной решетки добавим два специальных элемента: I — универсальный тип (общий предок, имеется во многих современных языках: под именем *Object* в языках Java, C#, Smalltalk или *TObject* в Object Pascal и Delphi) и O — фиктивный потомок всех типов (см. рис. 4). Фиктивный потомок

Рис. 2. Отношение порядка иерархии типов

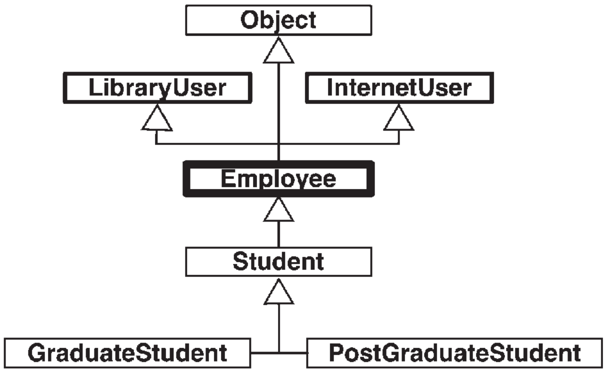

Рис. 3. Операция пересечения решетки типов

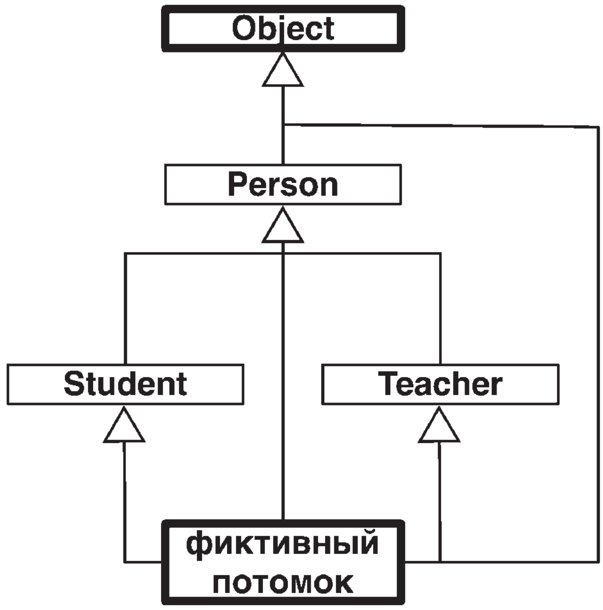

Рис. 4. Ограниченность решетки типов

всех типов семантически соответствует пустому типу. Например, пересечение двух любых типов, не имеющих общего родителя, дает в результате O. В системах с одиночным наследованием все пары типов, не связанные вышеуказанным отношением порядка, дают в пересечении O.

На решетке $\mathbb{F}$ рассмотрим второе (соответствующее агрегации) отношение R: если объект типа a в качестве атрибута содержится в типе b, то $(b,a)\in R$. Оба отношения ($\le$ и R) имеют общую семантику: в каждом случае, $b\le a$ или $(b,a)\in R$, тип b получает возможности типа a в виде доступа к его атрибутам. Семантически ясно, что это общее отношение «обладания набором возможностей», которое мы обозначаем $\xleftarrow{R}$, обязано быть рефлексивным и транзитивным.

Обсудим еще одно свойство введенных отношений. Пусть для элементов $a,b_1,b_2\in\mathbb{F}$ справедливо $b_1\le a$, $b_2\le a$. Тогда в соответствии с определением решетки имеем $b_1\vee b_2\le a$. Это естественное для отношения $\le$ свойство мы называем $\vee$-дистрибутивностью. Посмотрим, что будет означать обладание этим же свойством для отношения $\xleftarrow{R}$. Пусть $b_1\xleftarrow{R}a$ и $b_2\xleftarrow{R}a$, т.е. каждый тип b_1 и b_2 обладает возможностями типа a. Тогда в силу предполагаемой $\vee$-дистрибутивности имеем $b_1\vee b_2\xleftarrow{R}a$. Это значит, что тип $b_1\vee b_2$ также обладает возможностями типа a. С точки зрения проектирования типов это не обязательно. Однако в случае, если более одного наследника (b_1 и b_2) содержат одинаковые атрибуты, по принципам рефакторинга нужно «поднять» общие атрибуты, т.е. поместить один такой атрибут в общий тип-предок $b_1\vee b_2$, после чего каждый b_1 и b_2 получит возможности a в порядке наследования. В данном случае $\vee$-дистрибутивность отношения $\xleftarrow{R}$ означает решение важной задачи — устранение дублирования кода.

Подчеркнем, что мы рассматриваем здесь обобщенную постановку в плане «распределения возможностей» между типами: вариант, когда b_1 и b_2 по каким-либо практическим соображениям содержат в виде атрибутов представителей типа a под различными идентификаторами, нами в расчет не принимается. На практике тип может содержать много атрибутов, но не все они одинаково существенны при построении иерархии.

Заметим далее, что отношение $\xleftarrow{R}$, обладая свойством транзитивности вне зависимости от контекста, не может аналогично во всех ситуациях удовлетворять $\vee$-дистрибутивности, иначе это приведет к некорректным результатам. Для иллюстрации рассмотрим пример $b\xleftarrow{R}a$ и $a\xleftarrow{R}a$. При дистрибутивности $\xleftarrow{R}$ выполнялось бы $b\vee a\xleftarrow{R}a$. По идеологии ООП тип $b\vee a$ не имеет права что-либо знать о своих наследниках. Поэтому в данной ситуации $b\vee a$, являясь общим предком типов

a и b, может обладать возможностями типа a лишь в случае, если он совпадает с a (т.е. $b \le a$). Востальных вариантах соотношение $b \vee a \xleftarrow{R} a$ будет некорректным.

Существуют ситуации, когда выполнение $\vee$-дистрибутивности отношения $\xleftarrow{R}$ теоретически возможно, но нецелесообразно с точки зрения качества кода. Пусть при $b_1 \xleftarrow{R} a$, $b_2 \xleftarrow{R} a$ элементы $b_1 \vee b_2$ и a имеют непустое пресечение: $(b_1 \vee b_2) \wedge a = d \ne O$, причем $d < b_1 \vee b_2$ и $d < a$. Если в данном случае допустить $(b_1 \vee b_2, a) \in R$, то окажется, что тип d обладает возможностями типа a одновременно по двум линиям: как его наследник и как наследник типа $b_1 \vee b_2$, также имеющего возможности a.

Другая подобная ситуация — наличие конфликта. Пусть имеет место $(b_1, a), (b_2, a), (b_3, a) \in R$, причем $b_1 \vee b_2 \ne b_2 \vee b_3$. Тогда пары $(b_1, a), (b_2, a)$ «конфликтуют» с парами $(b_2, a), (b_3, a)$: если в обоих случаях «поднять» атрибуты, то тип b_2 унаследует атрибут типа a одновременно от двух различных предков — $b_1 \vee b_2$ и $b_2 \vee b_3$, что также ухудшит код.

Принятая в настоящей работе стратегия предполагает отказ от «поднятия» общих атрибутов при наличии подобных ситуаций (невыполнение $\vee$-дистрибутивности). В дальнейшем возможны и другие подходы, более тонко учитывающие особенности конкретной системы.

На основе приведенных здесь соображений в следующем разделе дается определение продукционно-логического отношения на ограниченной решетке. Оно отражает свойство «обладания набором возможностей» в иерархии типов. Логическое замыкание произвольного отношения R на решетке $\mathbb{F}$ предоставляет все такие пары (b, a), что в типе b доступны возможности типа a. Логическая редукция дает иерархию с минимальным эквивалентным набором связей и в определенном смысле минимальным дублированием кода. Этот формализм позволит проводить автоматизированные исследования иерархий типов, включая эквивалентные преобразования и верификацию. Он может служить основой для практической реализации (или модернизации) типов.

В работах [2—3] были изучены отношения, обладающие без ограничений как транзитивностью, так и дистрибутивностью. Область их применения — продукционные (в частности, экспертные) системы с монотонным выводом. Понятие $\vee$-дистрибутивности отношения $\xleftarrow{R}$ задает семантику монотонного вывода в том смысле что из $b \xleftarrow{R} a$, $a \xleftarrow{R} a$ следует $b \vee a \xleftarrow{R} a$. В данной работе в силу вышеназванных соображений свойство дистрибутивности зависит от контекста. Это, в частности, порождает немонотонность логического вывода на рассматриваемых структурах.

4. ФОРМАЛИЗАЦИЯ И ИССЛЕДОВАНИЕ

В этом разделе мы рассматриваем бинарные отношения на ограниченной решетке. Вначале сформулируем некоторые определения и условия, связанные с ограничением свойства $\vee$-дистрибутивности продукционно-логических отношений. Как говорилось в п. 3, данное свойство алгебраически описывает «поднятие» общих атрибутов по иерархии типов. Проблема здесь состоит в попытке ввести статическое формальное условие для описания динамического процесса. Конкретнее, условие $\vee$-дистрибутивности отношения R на любых подходящих парах (b_1, a), $(b_2, a) \in R$ должно быть выполнено как до «подъема» атрибута типа a, так и после него. В дальнейшем, ссылаясь на примеры, мы для краткости будем говорить о «подъеме» атрибута a, или просто «подъеме» a, если это не вызовет противоречий. Тем более что формальные определения данного раздела связаны не с типами, а абстрактными элементами решетки.

Пара (b, a) элементов ограниченной решетки $\mathbb{F}$ называется простой, если $a \wedge b = O$ — нижняя грань $\mathbb{F}$.

Определение 4.1. Пусть R — отношение на решетке $\mathbb{F}$. Две пары вида (b_1, a), $(b_2, a) \in R$ называются $\vee$-совместимыми в R, если существуют такие $c_1, c_2 \in \mathbb{F}$, что $b_1 \vee b_2 \le c_1 \vee c_2$, причем пара $(c_1 \vee c_2, a)$ простая, а пары (c_1, a), $(c_2, a) \in R$ нетранзитивны в $R \cup \le$. При этом набор элементов $C = (b_1, b_2, c_1, c_2, a)$ будем называть $\vee$-дистрибутивным кортежем, а $T = (c_1, c_2, a)$ — $\vee$-дистрибутивной тройкой (в R).

Поясним это определение на примере. Предположим, что типы b_1 и b_2 содержат атрибут a и пара $(b_1 \vee b_2, a)$ является простой. Тогда, в соответствии с обсуждением в п.3, общий атрибут может быть «поднят», т.е. перемещен в тип $b_1 \vee b_2$. До этого действия имеем $c_1 = b_1, c_2 = b_2$. После «подъема» атрибута условие определения 4.1 остается выполненным, но с другими промежуточными элементами: $c_1 = c_2 = b_1 \vee b_2$.

Требуемая в определении нетранзитивность пар (c_1, a), (c_2, a) обеспечивает «подъем» непосредственного атрибута типов b_1 и b_2, а не косвенного, связанного с ним цепочкой наследований и агрегаций. Конечно, «поднимаемый» атрибут неявно перемещается вместе со своими «возможностями», но именно он сам должен удовлетворять условиям определения 4.1.

Следующее определение формально описывает возможные варианты конфликтов между кортежами (точнее — их тройками), претендующими на свойство $\vee$-дистрибутивности (см. также о конфликте пар в п. 3).

Определение 4.2. Рассмотрим $\vee$-дистрибутивную тройку $T = (c_1, c_2, a)$, а также тройку $T' = (c_1', c_2', a')$, проверяемую на аналогичное свойство. Тройка T называется нейтрализующей для T' (обозначим $T' \prec T$), если выполнено одно из условий:

— $a = a'$, $c_1 \vee c_2 \neq c_1' \vee c_2'$ и справедливо хотя бы одно из неравенств $c_i' < c_1 \vee c_2$;

— $a < a'$, $c_1 \vee c_2 \neq c_1' \vee c_2'$ и выполнено хотя бы одно из неравенств $c_i' \leq c_1 \vee c_2$; $a < a'$, ; $c_1 \vee c_2 = c_1' \vee c_2'$.

Говоря неформально, это определение описывает ситуации, когда наличие нейтрализующей тройки T «угрожает» свойству $\vee$-дистрибутивности тройки T'. Рассмотрим в этом плане пункты 1)—3) определения 4.2.

Пусть имеет место 1). Тогда, после возможного «подъема» a из c_1 и c_2 в $c_1 \vee c_2$, в силу соотношений $c_i' < c_1 \vee c_2$, $(c_1 \vee c_2, a) \in R$, $a = a'$ пара (c_i', a') окажется транзитивной в $R \cup \leq$, что сделает невозможной $\vee$-дистрибутивность тройки T'. П. 1) определения 4.2, в частности, содержит вариант конфликта, о котором говорилось в примере предыдущего раздела работы. Для этого примера справедливы оба соотношения $T' \prec T$ и $T \prec T'$.

Если для T, T' выполнен вариант 2), то после «подъема» a в $c_1 \vee c_2$ получим соотношения $c_i' \leq c_1 \vee c_2$, $(c_1 \vee c_2, a) \in R$, $a < a'$, что вновь означает транзитивность в $R \cup \leq$ пары (c_i', a'). В случае же 3) после «подъема» атрибута a в $c_1 \vee c_2 = c_1' \vee c_2'$ получим, что обе пары (c_i', a') транзитивны в $R \bigcup \leq$: $c_i' \leq c_1' \vee c_2'$, $(c_1' \vee c_2', a) \in R$, $a < a'$.

Определение 4.3. Тройку $T' = (c_1', c_2', a')$ будем называть неконфликтной, если для нее не существует ни одной нейтрализующей $\vee$-дистрибутивной тройки.

Понятия, введенные в определениях 4.2—4.3 для троек вида $T = (c_1, c_2, a)$ и $T' = (c_1', c_2', a')$, автоматически распространяются и на связанные с ними кортежи $C = (b_1, b_2, c_1, c_2, a)$ и $C' = (b_1', b_2', c_1', c_2', a')$.

Определение 4.4. Две $\vee$-совместимые пары $(b_1, a), (b_2, a)$ называются неконфликтно $\vee$-совместимыми, если для них существует неконфликтный $\vee$-дистрибутивный кортеж (b_1, b_2, c_1, c_2, a).

Определение 4.5. Отношение R на решетке $\mathbb{F}$ называется ограниченно $\vee$-дистрибутивным, если для любых неконфликтно $\vee$-совместимых в R пар $(b_1, a), (b_2, a)$ справедливо $(b_1 \vee b_2, a) \in R$.

Определение 4.6. Отношение называется логическим с ограничением объединений (в данной работе — просто логическим), если оно содержит $\leq$, транзитивно и ограниченно $\vee$-дистрибутивно. Логическим замыканием отношения R называется наименьшее логическое отношение, содержащее R и его множество неконфликтно совместимых пар.

Определение 4.7. Два отношения R_1 и R_2, определенные на общей решетке, называются эквивалентными ($R_1 \sim R_2$), если их логические замыкания совпадают. Логической редукцией отношения R называется эквивалентное ему минимальное (в смысле теории множеств) отношение R_0.

Для выяснения вопроса о существовании логического замыкания введем следующее определение.

Определение 4.8. Пусть задано произвольное отношение R на решетке $\mathbb{F}$. Будем говорить, что упорядоченная пара $b, a \in \mathbb{F}$ логически связана отношением R ($b \xleftarrow{R} a$), если выполнено одно из следующих условий:

$$(b, a) \in R; \; b \leq a;$$

— существуют такие $b_1, b_2 \in \mathbb{F}$, что $b_1 \vee b_2 = b$, причем $b_1 \xleftarrow{R} a, b_2 \xleftarrow{R} a$ и пары $(b_1, a), (b_2, a)$ неконфликтно $\vee$-совместимы;

— существует элемент $c \in \mathbb{F}$ такой, что $b \xleftarrow{R} c$ и $c \xleftarrow{R} a$.

Условия 1)—4) определения 4.8 мы будем также называть правилами (вывода). Справедлива следующая.

Теорема 4.1. Для произвольного отношения R на решетке $\mathbb{F}$ логическое замыкание существует и совпадает с множеством $\xleftarrow{R}$ всех упорядоченных пар, логически связанных отношением R.

Для доказательства этой теоремы нам потребуются некоторые свойства $\vee$-дистрибутивных троек.

Лемма 4.1. Пусть $T=(c_1,c_2,a)$ — неконфликтная $\vee$-дистрибутивная тройка и элемент $c_3\in\mathbb{F}$ таков, что пары $(c_1\vee c_3,a),(c_2\vee c_3,a)$ простые, а пара $(c_3,a)\in R$ нетранзитивна в $R\cup\le$. Тогда $c_1\vee c_2\vee c_3=c_1\vee c_2=c_2\vee c_3=c_1\vee c_3$.

Доказательство. Сразу заметим, что интерес представляет лишь случай, когда $c_3\ne c_1$, $c_3\ne c_2$, иначе утверждение леммы тривиально. Рассмотрим тройку $T'=(c_2,c_3,a)$. В силу условия леммы она $\vee$-дистрибутивна. Кроме того, поскольку $c_2\le c_2\vee c_3$ и пары $(c_2,a),(c_3,a)$ нетранзитивны в $R\cup\le$, то $c_2<c_2\vee c_3$. Последнее неравенство означает, что тройка T' может оказаться нейтрализующей для $T=(c_1,c_2,a)$. Чтобы этого не случилось (ведь по условию T неконфликтна!) в силу п. 1) определения 4.2 необходимо выполнение равенства $c_1\vee c_2=c_2\vee c_3$, из которого сразу следуют соотношения $c_1\vee c_2\vee c_3=c_1\vee c_2=c_2\vee c_3$. Недостающее равенство $c_1\vee c_2\vee c_3=c_1\vee c_3$ можно получить, если вместо T' аналогично рассмотреть тройку (c_1,c_3,a). □

Лемма 4.2. Пусть R — произвольное бинарное отношение на $\mathbb{F}$ и $T=(c_1,c_2,a)$ — неконфликтная $\vee$-дистрибутивная тройка. Тогда множество неконфликтно $\vee$-совместимых пар отношения $R\cup\{(c_1\vee c_2,a)\}$ совпадает с множеством неконфликтно $\vee$-совместимых пар исходного отношения R.

Доказательство. Вначале докажем, что в результате указанной операции — добавления $(c_1\vee c_2,a)$ к R — исходное множество не сужается. Из определения 4.1 следует, что множество $\vee$-совместимых пар полностью определяется множеством нетранзитивных пар отношения $R\cup\le$. Таким образом, при доказательстве достаточно рассмотреть те пары нового отношения $R\cup\{(c_1\vee c_2,a)\}$, нетранзитивность которых может нарушиться в результате присоединения к R пары $(c_1\vee c_2,a)$ (например, (c',a') при $c'\le c_1\vee c_2$ и $a\le a'$).

Покажем, что любые неконфликтно $\vee$-совместимые пары $(d_1,a'),(d_2,a')$ останутся таковыми после добавления к R указанной в лемме пары $(c_1\vee c_2,a)$. Соответствующий парам $(d_1,a'),(d_2,a')$ неконфликтный $\vee$-дистрибутивный кортеж обозначим $C'=(d_1,d_2,c_1',c_2',a')$. Согласно определению 4.1, справедливо $d_1\vee d_2\le c_1'\vee c_2'$, пара $(c_1'\vee c_2',a')$ простая и пары $(c_1',a'),(c_2',a')\in R$ нетранзитивны в $R\cup\le$. Их нетранзитивность может нарушиться в результате добавления пары $(c_1\vee c_2,a)$ к отношению R. Покажем, как в этом случае заменить соответствующие элементы c_1',c_2' на другие $\tilde{c}_1,\tilde{c}_2$, чтобы для пар $(d_1,a'),(d_2,a')$ вновь выполнялись все требования определения 4.1.

Рассмотрим вначале случай $c_1\vee c_2=c_1'\vee c_2'$. Поскольку обе тройки $T=(c_1,c_2,a)$ и $T'=(c_1',c_2',a')$ неконфликтны, то согласно п. 3) определения 4.2 имеем $a=a'$. Пусть при этом $c_i'<c_1\vee c_2$ ($i=1$ или $i=2$). Тогда после добавления $(c_1\vee c_2,a)$ пара (c_i',a') становится транзитивной в $R\cup\le$. Обозначив $\tilde{c}_i=c_1\vee c_2$, заметим, что пара $(\tilde{c}_i,a')$ нетранзитивна в $R\cup\le$. Если же неравенство $c_i'<c_1\vee c_2$ не выполнено, то пара (c_i',a') остается нетранзитивной, и мы выбираем $\tilde{c}_i=c_i'$. Проверим, что выбранные элементы $\tilde{c}_1,\tilde{c}_2$ наряду с нетранзитивностью пар $(\tilde{c}_1,a'),(\tilde{c}_2,a')\in R$ обеспечивают и остальные требования $\vee$-совместимости пар $(d_1,a'),(d_2,a')$.

Во-первых, по построению $\tilde{c}_i$ справедливо $c_i'\le\tilde{c}_i$, отсюда имеем $c_1'\vee c_2'\le\tilde{c}_1\vee\tilde{c}_2$. Поэтому из соотношения $d_1\vee d_2\le c_1'\vee c_2'$ следует $d_1\vee d_2\le\tilde{c}_1\vee\tilde{c}_2$, что и требуется в определении 4.1. Покажем далее, что $\tilde{c}_1\vee\tilde{c}_2=c_1'\vee c_2'$. Поскольку рассматривается вариант $c_1\vee c_2=c_1'\vee c_2'$, то из построения $\tilde{c}_i$ ($\tilde{c}_i=c_i'$ либо $\tilde{c}_i=c_1\vee c_2$) следует, что $\tilde{c}_i\le c_1'\vee c_2'$, $i=1,2$. Отсюда вытекает соотношение $\tilde{c}_1\vee\tilde{c}_2\le c_1'\vee c_2'$. С другой стороны, как было замечено в начале абзаца, $c_1'\vee c_2'\le\tilde{c}_1\vee\tilde{c}_2$. В итоге имеем $\tilde{c}_1\vee\tilde{c}_2=c_1'\vee c_2'$. Это равенство означает, что пара $(\tilde{c}_1\vee\tilde{c}_2,a')$ является простой, т.к. пара $(c_1'\vee c_2',a')$ простая по определению. Таким образом, для элементов $\tilde{c}_1,\tilde{c}_2$ выполнены все требования определения 4.1.

Рассмотрим теперь вариант $c_1\vee c_2\ne c_1'\vee c_2'$ и в нем потенциально опасный случай $a\le a'$. В нем из-за неконфликтности тройки T' в силу п.п. 1)—2) определения 4.2 невозможно выполнение неравенств вида $c_i'\le c_1\vee c_2$. Рассматривая п.п. 1)—2) отдельно, замечаем, что в каждом из этих вариантов сохраняется нетранзитивность пар (c_i',a'), $i=1,2$, а вместе с ней и $\vee$-совместимость пар $(d_1,a'),(d_2,a')$. Таким образом, при исходном неравенстве $c_1\vee c_2\ne c_1'\vee c_2'$ кортеж C' не меняет своих свойств, т.е. можно выбрать $\tilde{c}_i=c_i'$, $i=1,2$.

Мы показали, что после добавления $(c_1\vee c_2,a)$ к R пары $(d_1,a'),(d_2,a')$ сохраняют свойство $\vee$-совместимости. Для доказательства требуемого вложения множеств осталось проверить, что сохраняется и их свойство неконфликтности. Достаточно показать, что для тройки $\tilde{T}=(\tilde{c}_1,\tilde{c}_2,a')$ нет ни одной нейтрализующей $\vee$-дистрибутивной тройки $N=(n_1,n_2,p)$. С этой целью предположим противное — что она существует ($\tilde{T}\prec N$), и при этом проанализируем п.п. 1)—3) определения 4.2. Учитывая полученные выше соотношения $c_i' \le \tilde{c}_i$ и $\tilde{c}_1\vee\tilde{c}_2=c_1'\vee c_2'$, нетрудно увидеть, что для $T'=(c_1',c_2',a')$ также выполнено $T'\prec N$. Если при этом тройка N (вместе со своими свойствами) существовала до добавления пары $(c_1\vee c_2,a)$ к R, то мы имеем противоречие, т.к. тройка T' изначально была неконфликтной. Рассмотрим ситуацию, когда сама тройка (n_1,n_2,p) или ее свойства (пара $(n_1\vee n_2,p)$ простая, а пары $(n_1,p),(n_2,p)\in R$ нетранзитивны в $R\cup\le$) появились в результате включения $(c_1\vee c_2,a)$ в отношение R. Очевидно, что свойство нетранзитивности любой пары не может возникнуть в результате добавления к отношению другой пары, равно как и свойство простоты фиксированной пары элементов решетки. Таким образом, единственно возможным вариантом остается лишь появление самой тройки (n_1,n_2,p), когда $p=a$ и $n_j=c_1\vee c_2$ ($j=1$ либо $j=2$). Положим для определенности $j=1$, т.е. $N=(c_1\vee c_2,n_2,a)$ (альтернативный случай является симметричным). Далее заметим, что поскольку пара $(c_1\vee c_2\vee n_2,a)$ простая, то для тройки $T=(c_1,c_2,a)$ и пары (n_2,a) выполнены условия леммы 4.1. В силу этой леммы $c_1\vee c_2\vee n_2=c_1\vee c_2$. С учетом последнего равенства рассмотрим применительно к $\tilde{T}\prec N$ каждый из п.п. 1)—3) определения 4.2 в отдельности.

Пусть $a=a'$, $c_1\vee c_2\vee n_2\ne\tilde{c}_1\vee\tilde{c}_2$ и справедливо одно из неравенств $\tilde{c}_i<c_1\vee c_2\vee n_2$, т.е. имеет место п.1). Тогда выполнены и соотношения $c_i'<c_1\vee c_2\vee n_2=c_1\vee c_2$. Кроме того, в рассматриваемом варианте $c_1\vee c_2\ne\tilde{c}_1\vee\tilde{c}_2=c_1'\vee c_2'$. Мы получили, что изначально тройка $T=(c_1,c_2,a)$ была нейтрализующей для тройки $T'=(c_1',c_2',a')$, что противоречит условию неконфликтности T'.

Рассмотрим следующий вариант, когда тройка $N=(c_1\vee c_2,n_2,a)$ является нейтрализующей для $\tilde{T}$ в силу п. 2) определения 4.2: $a<a'$, $c_1\vee c_2\vee n_2\ne\tilde{c}_1\vee\tilde{c}_2$ и выполнено хотя бы одно из неравенств $\tilde{c}_i\le c_1\vee c_2\vee n_2$. Тогда, по аналогии с предыдущим случаем, справедливо $c_i'\le c_1\vee c_2$ и $c_1\vee c_2\ne c_1'\vee c_2'$, т.е. $T'\prec T$ — аналогичное противоречие.

Наконец, если $a<a'$, $c_1\vee c_2\vee n_2=\tilde{c}_1\vee\tilde{c}_2$, то и $c_1\vee c_2=c_1'\vee c_2'$. И в данном случае оказывается, что $T'\prec T$ было выполнено до добавления $(c_1\vee c_2,a)$ к R, чего не могло быть по условию леммы.

Полученные противоречия показывают, что в результате добавления $(c_1\vee c_2,a)$ к R произвольные $\vee$-совместимые пары $(d_1,a'),(d_2,a')$ сохраняют свое свойство неконфликтности.

Таким образом, мы доказали, что в результате указанной операции исходное множество неконфликтно $\vee$-совместимых пар не сужается. Для полного доказательства леммы 4.2 осталось установить, что оно и не расширяется. С этой целью вначале покажем, что расширение было бы возможным лишь за счет единственной новой $\vee$-дистрибутивной тройки $T_\vee=(c_1\vee c_2,c_1\vee c_2,a)$. Предположим противное, что после добавления $(c_1\vee c_2,a)$ в R появились новые неконфликтно $\vee$-совместимые пары $(d_1,a'),(d_2,a')$ с соответствующей им тройкой $T'=(c_1',c_2',a')$, не совпадающей с $T_\vee$. Она не могла существовать до добавления $(c_1\vee c_2,a)$ к R. При этом, как и выше, единственно возможным вариантом является лишь появление самой тройки (c_1',c_2',a'), когда $a'=a$ и $c_j'=c_1\vee c_2$ ($j=1$ или $j=2$). Снова не ограничивая общности можно считать, что $j=1$, т.е. $T'=(c_1\vee c_2,c_2',a)$. Кроме того, в силу сделанного предположения имеем $c_1\vee c_2\ne c_2'$. Так как пара $(c_1\vee c_2\vee c_2',a)$ простая, то для тройки $T=(c_1,c_2,a)$ и пары (c_2',a) выполнены условия леммы 4.1, по которой $c_1\vee c_2\vee c_2'=c_1\vee c_2$, т.е. $c_2'<c_1\vee c_2$. Последнее неравенство приводит к противоречию, т.к. из него следует транзитивность пары (c_2',a) в $R\cup\le$, чего не может быть по определению дистрибутивной тройки T'.

Таким образом, в результате описанной в лемме операции появилась единственная новая $\vee$-дистрибутивная тройка $T_\vee=(c_1\vee c_2,c_1\vee c_2,a)$. Однако, ее появление не порождает новых дистрибутивно совместимых пар $(b_1,a),(b_2,a)$, поскольку неравенство $b_1\vee b_2\le(c_1\vee c_2)\vee(c_1\vee c_2)=c_1\vee c_2$ было бы выполнено и ранее для тройки $T=(c_1,c_2,a)$.

Итак, лемма 4.2 полностью доказана. □

Замечание 4.1. Из определений 4.1—4.2 следует, что множество $\vee$-совместимых пар (с их конфликтами) определяется множеством нетранзитивных пар отношения $R\cup\le$. Отсюда следует инвариантность множества неконфликтно $\vee$-совместимых пар отношения R относительно еще одной операции — расширения R за счет транзитивных пар отношения $R\cup\leq$.

Лемма 4.3. Множество неконфликтно $\vee$-совместимых пар отношения $\xleftarrow{R}$ совпадает с таковым множеством базового отношения R.

Доказательство. Достаточно установить, что при построении отношения $\xleftarrow{R}$ из отношения R за счет правил 2)—4) определения 4.8 указанное в лемме множество сохраняется. Для правил 2) и 4) требуемое утверждение сразу следует из замечания 4.1. Покажем, что аналогичное верно и для правила 3).

Пусть пары $(b_1,a),(b_2,a)$ неконфликтно $\vee$-совместимы. Рассмотрим соответствующий им неконфликтный кортеж $C=(b_1,b_2,c_1,c_2,a)$. Как указывалось в начале раздела, пары $(c_1,a),(c_2,a)$ также неконфликтно $\vee$-совместимы, поэтому отношение $\xleftarrow{R}$ должно содержать и пару $(c_1\vee c_2,a)$. При этом, по лемме 4.2, операция добавления к R пары $(c_1\vee c_2,a)$ сохраняет его множество неконфликтно совместимых пар.

Рассматривая далее пары $(b_1\vee b_2,c_1\vee c_2)$, $(c_1\vee c_2,a)$, замечаем, что пара $(b_1\vee b_2,a)$ транзитивна в $R\cup\leq$ при новом (расширенном) отношении R. Эта пара также должна быть добавлена к R по правилам 2) и 4), при этом, как указано в начале доказательства, исходное множество неконфликтно совместимых пар расширяемого отношения R не изменится. □

Лемма 4.4. При произвольном R отношение $\xleftarrow{R}$ является логическим.

Доказательство. Требуется проверить, что данное отношение содержит $\leq$, транзитивно и ограниченно $\vee$-дистрибутивно. Первые два указанных свойства непосредственно следуют из п.2) и п.4) определения 4.8. Свойство ограниченной $\vee$-дистрибутивности требует более подробного рассмотрения.

Пусть $b_1,b_2,a\in\mathbb{F}$, $b_1\xleftarrow{R}a, b_2\xleftarrow{R}a$ и пары $(b_1,a),(b_2,a)$ неконфликтно $\vee$-совместимы в $\xleftarrow{R}$. Требуется доказать, что тогда $b_1\vee b_2\xleftarrow{R}a$. С этой целью покажем, что при сделанных предположениях пары $(b_1,a),(b_2,a)$ неконфликтно $\vee$-совместимы в R. Тогда требуемое утверждение сразу будет следовать из п.3) определения 4.8.

Итак, согласно определению 4.1 имеем: существуют такие $c_1,c_2\in\mathbb{F}$, что $b_1\vee b_2\leq c_1\vee c_2$, пара $(c_1\vee c_2,a)$ простая, $c_1\xleftarrow{R}a, c_2\xleftarrow{R}a$ и пары $(c_1,a),(c_2,a)$ нетранзитивны в $\xleftarrow{R}$ (поскольку $(R\cup\leq)\subseteq\xleftarrow{R}$, то и в $R\cup\leq$). Последнее означает, что соотношения $c_1\xleftarrow{R}a, c_2\xleftarrow{R}a$ были получены без применения п.4) определения 4.8. Простота пары $(c_1\vee c_2,a)$ означает также, что не использовался и п.2). Таким образом, из п.1) и п.3) определения 4.8 следует, что для пар $(c_1,a),(c_2,a)$ остается лишь такой вариант, когда $c_1=c_1^1\vee\ldots\vee c_1^{m_1}$, $c_2=c_2^1\vee\ldots\vee c_2^{m_2}$, где $(c_j^{k_j},a)\in R$, $k_j=1,\ldots,m_j$, $j=1,2$, причем все пары $(c_j^{k_j},a)$ нетранзитивны в $R\cup\leq$.

Рассмотрим вначале случай $m_1=m_2=1$. Тогда $(c_1,a),(c_2,a)\in R$, и лемма уже доказана (неконфликтность тройки $T=(c_1,c_2,a)$ в R следует из ее неконфликтности в большем множестве $\xleftarrow{R}$).

Предположим теперь, что $m_1>1$. В этом случае дистрибутивная тройка $T'=(c_1^1,c_1^2,a)$ является неконфликтной в R. Поскольку пара $(c_1\vee c_2,a)$ простая, то для тройки T и любой оставшейся пары вида $(c_j^{k_j},a)$ выполнены условия леммы 4.1. По этой лемме справедливо равенство $c_1^1\vee c_1^2\vee c_j^{k_j}=c_1^1\vee c_1^2$. Складывая (в смысле операции объединения $\vee$) такие равенства по всем $k_j=1,\ldots,m_j$, $j=1,2$, получим $c_1\vee c_2=c_1^1\vee c_1^2$. В итоге мы пришли к следующему факту: существуют такие $c_1^1,c_1^2\in\mathbb{F}$, что $b_1\vee b_2\leq c_1^1\vee c_1^2$, пара $(c_1^1\vee c_1^2,a)$ простая, пары $(c_1^1,a),(c_1^2,a)\in R$ нетранзитивны в $R\cup\leq$, тройка $T'=(c_1^1,c_1^2,a)$ неконфликтна в R. Это означает, что пары $(b_1,a),(b_2,a)$ неконфликтно $\vee$-совместимы в R. Отсюда в силу п.3) определения 4.8 следует $b_1\vee b_2\xleftarrow{R}a$, т.е. отношение $\xleftarrow{R}$ ограниченно $\vee$-дистрибутивно. □

Лемма 4.5. Пусть R — некоторое логическое отношение на решетке $\mathbb{F}$ и $a,b\in\mathbb{F}$. Тогда, если справедливо $b\xleftarrow{R}a$, то $(b,a)\in R$.

Доказательство. Проведем его с помощью индукции по m — верхней оценке уровня рекурсии в соотношении $b\xleftarrow{R}a$. При $m=0$ имеет место одно из условий 1)—2) определения 4.8. Случай 1) означает требуемое утверждение. Если же справедливо 2), то и тогда $(b,a)\in R$, поскольку логическое отношение R содержит и $\leq$.

Предположим далее, что лемма верна для некоторого $m\geq 0$ и докажем ее утверждение при уровне рекурсии $m+1$. В этом случае новые для рассмотрения варианты могут дать условия 3)—4) определения 4.8.

Рассмотрим вариант, когда соотношение $b\xleftarrow{R}a$ происходит из условия 3) определения 4.8. При этом по предположению индукции соотношения $b_1\xleftarrow{R}a, b_2\xleftarrow{R}a$ имеют уровень рекурсии $\le m$, т.е. $(b_1,a),(b_2,a)\in R$. Тогда, в силу ограниченной дистрибутивности R, получим $(b,a)\in R$.

Наконец, пусть справедливо 4). И в этом случае по предположению индукции базовые соотношения $b\xleftarrow{R}c$ и $c\xleftarrow{R}a$ имеют уровень рекурсии $\le m$. Следовательно, $(b,c),(c,a)\in R$. Отсюда, в силу транзитивности логического отношения R, имеем $(b,a)\in R$. □

Доказательство теоремы 4.1. В силу п.1) определения 4.8 отношение $\xleftarrow{R}$ содержит R. По лемме 4.3 оно также содержит и множество неконфликтно $\vee$-совместимых в R пар. Из леммы 4.4 следует, что $\xleftarrow{R}$ является логическим. Осталось показать, что это — наименьшее из таковых отношений.

Пусть $R^{'}$ — другое логическое отношение, содержащее R и его множество неконфликтно совместимых пар. Тогда очевидно, что если $b\xleftarrow{R}a$, то $b\xleftarrow{R^{'}}a$. Отсюда по лемме 4.5 имеем $(b,a)\in R^{'}$. Следовательно, отношение $\xleftarrow{R}$ содержится в произвольно выбранном $R^{'}$, и, таким образом, является наименьшим логическим отношением, содержащим R и его множество неконфликтно совместимых пар. □

СПИСОК ЛИТЕРАТУРЫ

1. *Фаулер М., Скотт К. UML.* Основы: Пер. с англ. — СПб: Символ-Плюс, 2002. — 192 с.

2. *Махортов С.Д.* Логические отношения на решетках // Вестник ВГУ. Серия физика, математика. — Воронеж, 2003, 2. — С. 203—209.

3. *Махортов С.Д.* О редукции логических отношений на решетках // Вестник факультета ПММ: Вып. 5. — Воронеж: ВГУ, 2004. — с. 172—179.

4. *Махортов С.Д.* Методы исследования и преобразования иерархий типов на основе логических структур // Вестник ВГУ. Серия системный анализ и информационные технологии. — Воронеж, 2006, 2. — С. 24—27.

5. *Биркгоф Г.* Теория решеток: Пер. с англ. — М.: Наука, 1984. —568с.

6. *Aho A.V., Garey M.R., Ulman J.D.* The transitive reduction of a directed graph. SIAM J. Computing 1:2, P. 131—137 (1972).

Махортов Сергей Дмитриевич — канд. физ.-мат.наук, доцент, зав. каф. Математического обеспечения ЭВМ ф-та ПММ, Воронежский государственный университет, Тел. (4732) 208-698, E-mail: sd@expert.vrn.ru.

Погореленко Владимир Аркадьевич — аспирант ф-та ПММ, Воронежский государственный университет. E-mail: vladimir@livesystems.ru.

Makhortov Sergey Dmitrievich - Candidate of physics-math. Sciences, Associate Professor, the dept. of the Software for electronic-calculating machines, Voronezh State University. Тел. (4732) 208-698, E-mail: sd@expert.vrn.ru.

Pogorelenko Vladimir Arkadyevich - Post-Graduate Student the dept. of the Software for electronic-calculating machines, Voronezh State University. E-mail: vladimir@livesystems.ru.