Российская академия наук
Сибирское отделение
Институт систем информатики
имени А. П. Ершова

Д. К. Пономарёв

ПРОБЛЕМА РАЗЛОЖИМОСТИ ПРИ ФОРМАЛЬНОМ ОПИСАНИИ ЗНАНИЙ

Препринт
135

Новосибирск 2006

В данной работе под формальным описанием знаний понимается (элементарная) теория в логике первого порядка. Изучается свойство разложимости теории, которое на практике означает возможность разбить формальное описание интересующей предметной области на части, каждая из которых использует свой отдельный алфавит. Проблема разложимости элементарной теории сводится к вопросу о минимизации системы определяющих ее аксиом. Сформулирован критерий разложимости, установлена однозначность разложения сигнатуры разложимой элементарной теории и однозначность разложения самой теории с точностью до формул чистого равенства.

E-mail: ponom@iis.nsk.su

Siberian Division of the Russian Academy of Sciences
A. P. Ershov Institute of Informatics Systems

Denis Ponomaryov

FORMAL KNOWLEDGE REPRESENTATION AND THE DECOMPOSABILITY PROBLEM

Preprint
135

Novosibirsk 2006

In this paper we consider formal representation of knowledge as an (elementary) first-order logical theory. We study the decomposability property of theories, which means the possibility to split a formal representation of a considered subject domain into parts, each described by a separate set of terms. The decomposability problem is reduced to the question of minimization of system of axioms for a theory. We formulate the criterion of decomposability and prove the uniqueness of signature decomposition, as well as the uniqueness of theory decomposition up to equality formulas.
E-mail: ponom@iis.nsk.su

1. ВВЕДЕНИЕ

В работе[1] исследуются свойства элементарных теорий. При этом под *элементарной теорией* понимаем дедуктивно замкнутое непротиворечивое множество предложений (замкнутых формул) исчисления предикатов. Каждую теорию $\mathcal{T}$ определяем системой ее *аксиом* — такой непротиворечивой системой предложений T этой теории, из которых дедуктивно выводятся все предложения теории, $\mathcal{T}= \langle T\rangle$. Определения всех используемых понятий можно найти в монографии [3].

Сформулируем основную проблему, которую мы исследуем в данной работе.

Определение 1. *Рассмотрим сигнатуру Σ и теорию $\mathcal{T}$ этой сигнатуры.* ***Теория $\mathcal{T}$ — разложимая,*** *если можно подобрать такое дизъюнктное разложение сигнатуры $\Sigma = \Sigma_1 \cup \Sigma_2,\ \Sigma_1 \cap \Sigma_2 = \varnothing,\ \Sigma_1 \neq \varnothing \neq \Sigma_2$ и такую систему аксиом теории S, которая представляется объединением $S = S_1 \cup S_2$, в котором в формулах $S_i,\ i = 1, 2$ используются только символы из части Σ_i сигнатуры Σ. Будем далее обозначать разложение теории как $\mathcal{T}= S_1 \otimes S_2$, а разложение сигнатуры как $\Sigma = \Sigma_1 \coprod \Sigma_2$.*

Для произвольной теории $\mathcal{T}$ можно было бы рассматривать *тривиальное* разложение, в котором $\Sigma_1 = \varnothing$ или $\Sigma_2 = \varnothing$, а множество формул S_1 (соответственно S_2) образовано формулами чистого равенства теории $\mathcal{T}$, предложениями теории, не использующими сигнатурных символов. Однако такое разложение малосодержательно, поэтому, согласно введенному определению, далее под *разложимой теорией* будем понимать такую теорию, которая обладает нетривиальным разложением, в котором $\Sigma_1 \neq \varnothing \neq \Sigma_2$. Если имеются только тривиальные разложения, то теорию будем называть ***неразложимой.***

К примеру, если сигнатура Σ состоит только из одного предиката, то любая теория этой сигнатуры Σ (даже определенная пустым множеством аксиом) неразложима.

Проблема 1. *Рассмотрим теорию $\mathcal{T}$ сигнатуры Σ, определенную некоторым множеством замкнутых формул Φ сигнатуры Σ. Как исходя из системы аксиом Φ определить разложимость теории $\mathcal{T}$?*

[1]Работа выполнена при финансовой поддержке гранта РФФИ № 05-01-04003-ННИО_а (DFG project COMO, GZ: 436 RUS 113/829/0-1).

Проблема 2. *Рассмотрим конечно аксиоматизируемую теорию $\mathcal{T}$ сигнатуры Σ, определенную некоторым конечным множеством предложений Φ сигнатуры Σ. Как исходя из системы аксиом Φ определить разложимость теории $\mathcal{T}$?*

Проблемы разложимости элементарных теорий были поставлены Д. Е. Пальчуновым в работе [4], посвященной изучению формальных онтологий. Вопрос разложимости имеет важное значение, в частности, для формализации знаний, поскольку разложимость означает возможность разбить формальное описание интересующей предметной области на части, каждая из которых использует свой отдельный алфавит. При построении формального описания некоторой предметной области часто оказывается, что данные, полученные от экспертов (или извлеченные автоматически), представляют собой некоторый набор фактов, которые необходимо структурировать, чтобы получить из них адекватную формальную модель. В частности, может быть интересно, существуют ли части знания, которые независимы друг от друга. Это в точности соответствует вопросу разложимости, если рассматривать формальное описание предметной области как логическую теорию (скажем, в некотором подмножестве логики первого порядка). Разложимость важна и для выполнения логических операций над объемными декларативными базами знаний, онтологиями, в частности, для проверки их непротиворечивости. Если онтология может быть разбита на несколько частей, имеющих различные сигнатуры, то эти части возможно проверить на непротиворечивость по отдельности. Это также дает возможность распределенного выполнения данной операции. Можно привести и другие примеры приложений данного вопроса. Идеологически все они исходят из того простого факта, что в любой сфере знаний *декомпозиция* всегда означает *упрощение*.

Обсуждение этих проблем следует начать с такого замечания.

Замечание 1. *Если теорию $\mathcal{T}$ сигнатуры Σ можно определить такой системой аксиом, в которой используются только символы части сигнатуры $\Sigma' \subset \Sigma$, то такая теория — разложимая. Она разлагается на теорию меньшей сигнатуры Σ' и теории символов из дополнения $\Sigma \backslash \Sigma'$, определенные множествами тождественно истинных предложений.*

В условиях замечания проблема разложимости по существу сводится к проблеме разложимости теории меньшей сигнатуры. Это рассуждение приводит к таким понятиям.

Определение 2. *Рассмотрим сигнатуру Σ и теорию $\mathcal{T}$ этой сигнатуры.* ***Теория $\mathcal{T}$ — приводимая****, если можно подобрать такое подмножество $\Sigma' \subset \Sigma$ сигнатуры Σ и такую систему аксиом теории S, в формулах которой используются только символы из части Σ' сигнатуры Σ. В этом случае будем говорить точнее, что* ***теория $\mathcal{T}$ приводится к теории $\mathcal{T}'$ меньшей сигнатуры*** *Σ'.*

Если же такое невозможно и любая система аксиом теории $\mathcal{T}$ использует все сигнатурные символы, то теорию $\mathcal{T}$ называем ***неприводимой.***

Символы сигнатуры Σ теории $\mathcal{T}$, которые невозможно исключить из любой ее системы аксиом, называем ***существенными*** *(символами теории $\mathcal{T}$).*

Таким образом, решение проблемы разложимости должно включать предварительный этап. На нем следует определить приводимость или неприводимость теории и сигнатуру, к которой она сводится, выделить существенные символы теории и аксиомы приводимости. А уже далее устанавливать разложимость или неразложимость неприводимой теории меньшей сигнатуры и определить аксиомы разложимости.

В данной работе представляем решение проблемы разложимости следующим образом. В разд. 2 рассмотрим вопрос приводимости теории и докажем следующее утверждение:

Предложение 1. *Пусть $\mathcal{T}$ — теория сигнатуры Σ. Рассмотрим множество существенных символов Σ' сигнатуры Σ: $\Sigma' \subseteq \Sigma$.*

Тогда теорию $\mathcal{T}$ можно определить некоторой системой аксиом сигнатуры Σ' . При этом, данная система аксиом в сигнатуре Σ' определяет неприводимую теорию $\mathcal{T}'$.

В разд. 3 будем изучать уже неприводимые теории и сформулируем основной технический результат работы — предложение 3; установим разложимость и однозначность разложения неприводимой теории в произведение неразложимых теорий с точностью до формул чистого равенства (теорема 1). В заключение выведем правило преобразования аксиом $(Ax)'$ неприводимой теории $\mathcal{T}$ к такой системе аксиом $(Ax)''$, исходя из которой можно определить разложимость или неразложимость теории $\mathcal{T}$, а также выделить системы аксиом, представляющие разложение.

Наконец, в разд. 4 для случая конечной аксиоматизируемости теории $\mathcal{T}$ определим понятие веса ее системы аксиом и заметим, что указанные выше преобразования минимизируют вес. Таким образом, установим, что в этом случае для получения разложения теории можно использовать систему аксиом минимального веса.

Для точной формулировки перечисленных результатов расширим определение 1, сформулировав его не для систем аксиом, а на языке теорий.

Определение 3. *Рассмотрим теории $\mathcal{S}_1$, $\mathcal{S}_2$ дизъюнктных сигнатур Σ_1, Σ_2; $\Sigma_1 \cap \Sigma_2 = \varnothing$, $\Sigma_1 \neq \varnothing \neq \Sigma_2$.*

*Теория $\mathcal{T}$ сигнатуры $\Sigma = \Sigma_1 \coprod \Sigma_2$ **разлагается в произведение теорий** $\mathcal{S}_1$, $\mathcal{S}_2$ (обозначаем $\mathcal{T} = \mathcal{S}_1 \otimes \mathcal{S}_2$), если она получается дедуктивным замыканием всех предложений теории $\mathcal{S}_1$ (сигнатуры Σ_1) и всех предложений теории $\mathcal{S}_2$ (сигнатуры Σ_2) в исчислении предикатов расширенной сигнатуры $\Sigma : \mathcal{S}_1, \mathcal{S}_2 \vdash \mathcal{T}$.*

*При этом теории $\mathcal{S}_1$, $\mathcal{S}_2$ называем **компонентами (разложимости)** теории $\mathcal{T}$.*

Заметим, что произведение произвольных теорий $\mathcal{S}_1 \otimes \mathcal{S}_2$ может приводить к противоречивому множеству предложений, так как в компонентах могут определяться несовместные формулы чистого равенства. Поэтому эта операция является частичной, произведение некоторых теорий будет неопределено. Обратим внимание на следующие свойства произведения теорий. Они легко выводятся из свойств эквивалентных систем предложений.

Замечание 2. *Произведение теорий $\otimes$ удовлетворяет свойствам коммутативности и ассоциативности. Для разложения теории в произведение теорий попарно непересекающихся сигнатур выполняются равенства: $\mathcal{S}_1 \otimes \mathcal{S}_2 = \mathcal{S}_2 \otimes \mathcal{S}_1$ и $(\mathcal{S}_1 \otimes \mathcal{S}_2) \otimes \mathcal{S}_3 = \mathcal{S}_1 \otimes (\mathcal{S}_2 \otimes \mathcal{S}_3)$.*

Кроме того, произведение теорий монотонно. Если имеются разложения теорий $\mathcal{S}_1 \otimes \mathcal{S}_2$ и $\mathcal{T}_1 \otimes \mathcal{T}_2$, для которых выполнено включение теорий $\mathcal{S}_1 \subseteq \mathcal{T}_1$ и $\mathcal{S}_2 \subseteq \mathcal{T}_2$, то выполнено включение $\mathcal{S}_1 \otimes \mathcal{S}_2 \subseteq \mathcal{T}_1 \otimes \mathcal{T}_2$.

Используя это замечание, не будем расставлять скобок в произведениях $\otimes$.

Компоненты разложимости теории определены неоднозначно. Есть неопределенность в том, что в компоненты можно произвольным образом включать формулы чистого равенства теории. Для теории $\mathcal{T}$ через

$\mathcal{T}^{\#}$ будем обозначать теорию чистого равенства, содержащуюся в $\mathcal{T}$. Она определяется замкнутыми формулами чистого равенства, принадлежащими теории $\mathcal{T}$. Для произвольной теории $\mathcal{S}$ через $\langle\mathcal{S},\mathcal{T}^{\#}\rangle$ будем обозначать теорию $\mathcal{S}$ вместе с предложениями чистого равенства теории $\mathcal{T}$.

Следующее замечание легко следует из понятия выводимой секвенции исчисления предикатов.

Замечание 3. *Если теория $\mathcal{T}$ — конечно аксиоматизируема, определяется конечным набором аксиом, то она определяется одной формулой — конъюнкцией этих аксиом. В этом случае компоненты разложимости можно выбрать конечно аксиоматизируемыми теориями.*

Основной результат работы заключен в следующем утверждении об однозначности разложения неприводимой теории.

Теорема 1. *Для любой неприводимой теории $\mathcal{T}$ сигнатуры Σ однозначно определяется дизъюнктное разложение сигнатуры $\Sigma = \Sigma_1\coprod\ldots\coprod\Sigma_n$, которое с точностью до перестановки сомножителей отвечает любому разложению теории $\mathcal{T}{=}\mathcal{S}_1{\otimes}\ldots{\otimes}\mathcal{S}_n$ на неразложимые компоненты разложимости $\mathcal{S}_i$ — теории сигнатуры Σ_i, $i=1,\ldots,n$. Зафиксируем некоторое такое разложение.*

Для произвольного разложения теории на неразложимые компоненты $\mathcal{T} = \mathcal{T}_1\otimes\ldots\otimes\mathcal{T}_m$ имеем $n = m$ и после подходящей перенумерации компонентов каждая теория $\mathcal{S}_i$ отличается от теории $\mathcal{T}_i$ некоторыми предложениями чистого равенства из исходной теории $\mathcal{T}$: $\langle\mathcal{S}_i,\mathcal{T}^{\#}\rangle{=}\langle\mathcal{T}_i,\mathcal{T}^{\#}\rangle$.

Отметим, что главным теоретико-модельным инструментом исследования в статье служит интерполяционная теорема Крейга. Будем использовать формулировку этой теоремы из [3].

В заключение введения автор выражает искреннюю благодарность профессору Д. Е. Пальчунову за постановку задачи и за многочисленные советы по улучшению содержания статьи.

2. ПРИВОДИМЫЕ ТЕОРИИ

В данном разделе будем изучать предложения различных сигнатур и их расширений. При этом, без оговорок будем использовать такое известное свойство.

Замечание 4. *Рассмотрим сигнатуру Σ' и ее расширение Σ.*

Пусть ϕ и ψ — предложения меньшей сигнатуры Σ', причем в исчислении предикатов большей сигнатуры Σ выводима секвенция $\phi \vdash \psi$. Тогда эта секвенция выводима и в меньшей сигнатуре Σ'.

Действительно, в противном случае предложения ϕ и $\neg\psi$ меньшей сигнатуры Σ' образуют непротиворечивое множество предложений $T = \{\phi, \neg\psi\}$ меньшей сигнатуры Σ', которое противоречиво в расширении сигнатуры Σ. Тогда по теореме существования модели имеется модель теории $T = \{\phi, \neg\psi\}$ в сигнатуре Σ', а в расширенной сигнатуре эта теория модели не имеет.

Однако произвольную модель теории T сигнатуры Σ' всегда можно расширить до модели большей сигнатуры Σ, подбирая интерпретации символов из $\Sigma \backslash \Sigma'$ произвольным образом. Полученное противоречие доказывает замечание. $\square$

Предложение 2. *Рассмотрим расширение сигнатуры $\Sigma' \subseteq \Sigma$ и теорию $\mathcal{P}$ сигнатуры Σ' соответственно. Рассмотрим предложение φ сигнатуры Σ.*

Тогда если формула φ следует из теории $\mathcal{P}$, то найдется предложение θ теории $\mathcal{P}$ такое, что θ включает только те символы сигнатуры Σ', которые встречаются в формуле φ и из которого предложение φ следует: $\mathcal{P} \vdash \theta$, $\theta \vdash \varphi$.

Действительно, по определению выводимости найдется формула P теории $\mathcal{P}$, для которой $P \vdash \varphi$. Далее, по теореме Крейга найдется такое предложение θ расширенной сигнатуры Σ, которое использует только символы, присутствующие одновременно в формуле P и формуле φ, и для которого выполняются секвенции $P \vdash \theta$, $\theta \vdash \varphi$.

Таким образом, предложение θ принадлежит теории $\mathcal{P}$ и удовлетворяет утверждению. $\square$

Определение 4. *Рассмотрим теорию $\mathcal{T}$ и предложение теории $\varphi \in \mathcal{T}$.*

Предложение φ назовем ***приводимым в теории $\mathcal{T}$****, если найдется такое предложение θ теории $\mathcal{T}$, $\theta \in \mathcal{T}$, которое использует строго меньшее число сигнатурных символов формулы φ, и выполнена секвенция $\theta \vdash \varphi$. При этом, предложение θ назовем* ***сужением формулы φ в теории $\mathcal{T}$****.*

Если таких предложений θ найти нельзя, то формулу φ будем называть ***неприводимой*** *в теории $\mathcal{T}$.*

Замечание 5. *Любая неприводимая формула φ теории $\mathcal{T}$ включает только существенные символы теории $\mathcal{T}$.*

Необходимо отметить, что данное замечание следует из следствия теоремы Крейга, приведенного в [1]. В нем показано, что для каждого предложения логики предикатов первого порядка существует эквивалентное ему предложение, содержащее наименьшее по включению множество сигнатурных символов.

Здесь мы приводим другое доказательство, использующее предложение 2. Докажем от противного. Пусть φ включает символы, несущественные для теории $\mathcal{T}$. Тогда можно подобрать такие аксиомы теории $\mathcal{T}$, которые используют только часть Σ' сигнатурных символов, не содержащую все символы формулы φ. Тогда по предложению 2 формула φ выводится из такой формулы θ теории $\mathcal{T}$, которая использует только символы из Σ'. Это противоречит неприводимости φ. $\square$

Из этого замечания следует, что существенные символы теории определяются по любой системе аксиом теории, образованной неприводимыми предложениями. Отметим к тому же, что любая система аксиом теории (Ax) может быть переработана в неприводимую систему аксиом $(Ax)'$, если ее предложения заменять их неприводимыми сужениями в данной теории. Это доказывает предложение 1.

Далее в разделе 3 будем изучать уже неприводимые системы аксиом.

3. РАЗЛОЖИМЫЕ ТЕОРИИ

Решающим шагом к определению компонентов разложимости служит следующий вариант интерполяционной теоремы Крейга.

Предложение 3. *Рассмотрим разложение сигнатуры $\Sigma = \Sigma_1 \coprod \Sigma_2$ и теории $\mathcal{P}$,$\mathcal{Q}$ сигнатур Σ_1, Σ_2 соответственно. Рассмотрим предложение φ сигнатуры Σ.*

Тогда если φ следует из объединения теорий $\mathcal{P},\mathcal{Q} \vdash \varphi$, то найдутся такие предложения θ и ϕ теорий $\mathcal{P}$ и $\mathcal{Q}$ соответственно, из которых следует предложение $\varphi : \mathcal{P} \vdash \theta,\ \ \mathcal{Q} \vdash \phi,\ \ \theta, \phi \vdash \varphi$, причем предложение θ включает только такие символы сигнатуры Σ_1, которые встреча-

ются в формуле φ, а предложение ϕ содержит только те символы сигнатуры Σ_2, которые используются в формуле φ.

Докажем это предложение. Будем предполагать выполненными все его посылки.

По условию $\mathcal{P},\mathcal{Q} \vdash \varphi$. По определению выводимости найдутся предложения $P \in \mathcal{P}$, $Q \in \mathcal{Q}$, для которых $P,\ Q \vdash \varphi$. Тогда $P \vdash Q \to \varphi$.

Поскольку P и Q — формулы сигнатур Σ_1, Σ_2, не имеющих общих символов, по интерполяционной теореме Крейга найдется предложение θ, содержащее только символы сигнатуры Σ_1, которые присутствуют в формуле φ. При этом $P \vdash \theta$ и $\theta \vdash Q \to \varphi$.

Выводим отсюда, что $Q \vdash \theta \to \varphi$. Также по интерполяционной теореме находим такое предложение ϕ, которое содержит только символы сигнатуры Σ_2, включенные в формулу φ. Для него также имеем $Q \vdash \phi$, $\phi \vdash \theta \to \varphi$.

Следовательно, получаем, что $\theta,\ \phi \vdash \varphi$. □

Из предложения 3 выводится способ установления разложимости теории $\mathcal{T}$. Для его описания определим понятие *разложимого предложения теории.*

Определение 5. *Рассмотрим теорию $\mathcal{T}$ и предложение теории $\varphi \in \mathcal{T}$.*

Предложение φ назовем ***разложимым в теории*** *$\mathcal{T}$, если найдутся такие предложения θ и ψ теории $\mathcal{T}$, $\theta \in \mathcal{T}, \psi \in \mathcal{T}$, которые включают символы только из φ, не содержат общих сигнатурных символов, не являются формулами чистого равенства и для которых выполнена секвенция $\theta,\psi \vdash \varphi$. При этом θ и ψ будем называть* ***фрагментами разложения*** *предложения φ.*

Если таких θ и ψ найти нельзя, то предложение φ будем называть ***неразложимым в теории*** *$\mathcal{T}$.*

Замечание 6. *Если предложение φ теории $\mathcal{T}$ неприводимо, то фрагменты его разложения θ и ψ также являются неприводимыми предложениями теории.*

Действительно, если бы θ, ψ были приводимы, то из этого бы следовала приводимость предложения φ, так как $\theta \wedge \psi \vdash \varphi$.

Лемма 1. *При любом разложении теории $\mathcal{T} = \mathcal{S}_1 \otimes \mathcal{S}_2$ в произведение теорий сигнатур Σ_1, Σ_2 соответственно ($\Sigma_1 \neq \varnothing \neq \Sigma_2$) каждое неразложимое предложение φ теории $\mathcal{T}$, включающее сигнатур-*

ные символы, выводится либо только из теории $\langle\mathcal{S}_1,\mathcal{T}^{\#}\rangle$, либо только из теории $\langle\mathcal{S}_2,\mathcal{T}^{\#}\rangle$.

Если формула φ, к тому же, неприводима, то она содержится либо в теории $\langle\mathcal{S}_1,\mathcal{T}^{\#}\rangle$ сигнатуры Σ_1, либо в теории $\langle\mathcal{S}_2,\mathcal{T}^{\#}\rangle$ сигнатуры Σ_2. В частности, она содержит символы либо только из сигнатуры Σ_1, либо только из сигнатуры Σ_2.

Действительно, рассмотрим неразложимое предложение φ. По определению разложения теории $\mathcal{T}$ это предложение выводится из формул теории $\mathcal{S}_1$ и формул теории $\mathcal{S}_2$. Из предложения 3 следует, что найдутся такие предложения $\theta \in\mathcal{S}_1$ и $\psi \in\mathcal{S}_2$, для которых $\theta, \psi \vdash \varphi$. Тогда, в силу неразложимости φ, одна из формул θ, ψ не использует сигнатурных символов. Отсюда следует первое утверждение леммы.

Пусть теперь формула φ неприводима. Будем для определенности считать, что $\theta \vdash \varphi$, где $\theta \in\langle\mathcal{S}_1,\mathcal{T}^{\#}\rangle$.

По интерполяционной теореме Крейга из неприводимости φ следует, что все символы в ее записи принадлежат сигнатуре Σ_1. Значит эта формула принадлежит теории $\langle\mathcal{S}_1,\mathcal{T}^{\#}\rangle$. $\Box$

Докажем усиление второй части этой леммы, применяя индукцию по числу сомножителей.

Следствие 1. *Рассмотрим произвольное нетривиальное разложение теории $\mathcal{T} = \mathcal{S}_1 \otimes \ldots \otimes \mathcal{S}_n$. Тогда любое неприводимое и неразложимое предложение теории $\varphi \in \mathcal{T}$ принадлежит некоторой теории $\langle\mathcal{S}_i,\mathcal{T}^{\#}\rangle$.*

Для произвольного $i = 1,\ldots,n$ обозначим $\mathcal{S}'_i = \langle\mathcal{S}_i,\mathcal{T}^{\#}\rangle$, тогда $\mathcal{T} = \mathcal{S}'_1 \otimes\ldots\otimes \mathcal{S}'_n$, и требуется показать, что любое неприводимое и неразложимое предложение теории $\varphi \in\mathcal{T}$ принадлежит некоторому $\mathcal{S}'_i$.

Действительно, база индукции для $n = 2$ следует из леммы 1. Перейдем к шагу индукции и предполагаем, что утверждение выполнено для числа сомножителей строго меньше n. Установим его для n сомножителей.

Обозначим $\mathcal{P} = \mathcal{S}'_1\otimes\ldots\otimes\mathcal{S}'_{n-1}$, $\mathcal{Q} = \mathcal{S}'_n$. В силу ассоциативности, по замечанию 2 имеем $\mathcal{T} = \mathcal{S}'_1\otimes\ldots\otimes\mathcal{S}'_n = \mathcal{P} \otimes \mathcal{Q}$.

Рассмотрим неприводимое и неразложимое предложение теории $\varphi \in\mathcal{T}$ и по лемме 1 заметим, что либо $\varphi \in \mathcal{P}$, и утверждение следует из предположения индукции, либо $\varphi \in \mathcal{Q} = \mathcal{S}'_n$. В любом случае, следствие доказано. $\Box$

Данное следствие показывает, что для любого разложения теории любое неприводимое и неразложимое предложение теории содержит сигнатурные символы только из какого-то одного компонента разложимости теории. Это позволяет определить разложение сигнатуры и компоненты разложимости теории $\mathcal{T}$.

Следует формализовать полученные результаты с помощью следующего ниже понятия. Приведем его для произвольной системы аксиом Φ некоторой теории сигнатуры Σ.

Определение 6. *Два сигнатурных символа P, $Q \in \Sigma$ назовем* ***непосредственно связанными*** *(системой аксиом Φ теории $\mathcal{T}$), если P и Q входят в одну аксиому $\psi \in \Phi$.*

Символы P, Q назовем ***связанными*** *(системой аксиом Φ), если найдется такая последовательность символов $P = T_1, \dots, T_q = Q$, в которой любая соседняя пара T_i, T_{i+1} связана непосредственно.*

Отношение связанности является отношением эквивалентности для неприводимой теории $\mathcal{T}$, так что множество Σ разбивается на компоненты связности — эквивалентности. По сути, получаем неориентированный помеченный граф (с петлями в вершинах), вершинами которого служат символы сигнатуры, а аксиомы определяют связывающие их ребра.

Условие неприводимости теории гарантирует отсутствие в таком графе изолированных вершин без петель — синглетонов. Каждый сигнатурный символ включается в некоторую аксиому неприводимой теории $\mathcal{T}$, так что в этом случае каждая вершина графа инцидентна некоторому ребру. При этом условии отношение связности можно определить двойственным образом через отношения не вершин, а ребер графа инцидентности.

Определение 7. *Две аксиомы $\varphi, \psi \in \Phi$ назовем* ***непосредственно связанными****, если они содержат некоторый общий символ сигнатуры Σ.*

Аксиомы φ, ψ назовем ***связанными*** *(системой аксиом Φ), если найдется такая последовательность аксиом из $\Phi : \varphi = \phi_1, \dots, \phi_q = \psi$, в которой любая соседняя пара ϕ_i, ϕ_{i+1} связана непосредственно.*

Компоненты связности однозначно определяются либо перечислением входящих в них сигнатурных символов (вершин графа), либо перечислением аксиом (ребер графа), включающих эти символы.

Отметим, что определения 6 и 7 являются обобщением понятия синтаксической сети предложений логики предикатов первого порядка, введенного в [2].

Завершим нашу конструкцию разложения теории и приступим к доказательству теоремы 1. Вначале докажем следующее предложение.

Предложение 4. *Любая теория $\mathcal{T}$ имеет систему аксиом, состоящую из неприводимых и одновременно неразложимых предложений.* Действительно, согласно предложению 1 для теории $\mathcal{T}$ существует система аксиом $(Ax)'$, состоящая из неприводимых предложений. Применим к этой системе следующие преобразования.

Каждую разложимую аксиому $\varphi \in (Ax)'$ заменим на ее фрагменты θ, ψ, каждый из которых выводится из $(Ax)'$ и для которых $\theta, \psi \vdash \varphi$. Неразложимые аксиомы не изменяем. Обозначим полученную после преобразования систему аксиом как $T(Ax)'$.

По построению система аксиом $T(Ax)'$ также определяет теорию $\mathcal{T}$. Будем повторять это преобразование и построим последовательность систем аксиом $T(Ax)'$, $T^2(Ax)', \ldots$.

Каждая формула содержит лишь конечное число сигнатурных символов, поэтому каждая аксиома φ из $(Ax)'$ может быть последовательно разложена только конечное число раз (оговоримся, что рассматримаемые сигнатуры конечные). Таким образом, начиная с некоторого n в системе аксиом $T^m(Ax)'$, $m \geqslant n$, будут только неразложимые фрагменты аксиомы φ.

Обозначим через $(Ax)''$ такие предложения φ, каждое из которых встречается во всех членах построенной последовательности кроме, может быть, конечного их числа. Это неразложимые фрагменты аксиом из $(Ax)'$. По замечанию 6 система $(Ax)''$ состоит из неразложимых и неприводимых предложений теории $\mathcal{T}$ и представляет систему аксиом теории $\mathcal{T}$.$\square$

Обозначим через $\Psi = \Psi(\mathcal{T})$ все неприводимые и неразложимые предложения неприводимой теории $\mathcal{T}$. Заметим, что множество Ψ образует систему аксиом теории. Оно включает все неразложимые фрагменты неприводимых сужений предложений теории $\mathcal{T}$. Поэтому любое предложение теории $\mathcal{T}$ выводится из предложений системы Ψ. Кроме того, эта система аксиом максимальна — любая система неприводимых и неразложимых аксиом теории $\mathcal{T}$ включается в Ψ.

Применим введенные выше определения 6,7 к максимальной системе неприводимых и неразложимых аксиом $\Psi(\mathcal{T})$ неприводимой теории $\mathcal{T}$. Определим *стандартные компоненты связности сигнатуры* $\Sigma = \Sigma_1 \coprod \ldots \coprod \Sigma_n$ относительно этой системы аксиом. Построим *стандартное разложение теории* на компоненты разложимости, отвечающее этому разложению сигнатуры.

По построению каждая аксиома из системы $\Psi(\mathcal{T})$ содержит символы только из какой-то одной дизъюнктной части сигнатуры Σ_i. Обозначим для каждого $i = 1, \ldots, n$ через S_i все аксиомы, содержащие символы из Σ_i. Поскольку теория неприводимая, каждый символ сигнатуры присутствует в некоторой аксиоме из Ψ. Значит, S_i — множество аксиом соответствующего компонента связности системы аксиом. Вместе эти множества аксиом определяют *стандартное разложение теории* $\mathcal{T}{=}S_1 \otimes \ldots \otimes S_n$.

Заметим, что аксиомы $S_i,\ i = 1, \ldots, n$ определяют неразложимые теории. Действительно, если бы некоторая система S_i определяла разложимую теорию, то по лемме 1 сигнатура Σ_i и система аксиом S_i была бы несвязна. А по построению это компонент связности.

Лемма 2. *Любые системы неприводимых и неразложимых аксиом Φ неприводимой теории $\mathcal{T}$ определяют одно и то же отношение связности на символах сигнатуры Σ теории $\mathcal{T}$. В частности, они определяют одни и те же компоненты связности $\Sigma_i,\ i = 1, \ldots, n$ сигнатуры Σ.*

Покажем, что любая система неприводимых и неразложимых аксиом Φ определяет то же отношение связности, что и максимальная система $\Psi(\mathcal{T})$.

В самом деле, поскольку $\Psi(\mathcal{T})$ — максимальная система неприводимых и неразложимых аксиом, то использование меньшей системы Φ может нарушить некоторые отношения связности. Поэтому следует только проверить, что символы любого неприводимого и неразложимого предложения φ, связанные относительно $\Psi(\mathcal{T})$, являются связанными и относительно системы аксиом Φ.

Разложение на компоненты связности системы аксиом Φ, как и выше, определяет разложение теории $\mathcal{T}$ на компоненты разложимости. По следствию 1 φ содержит символы только одного компонента связности системы Φ. Поэтому эти символы принадлежат одному компоненту связности и являются связанными.$\square$

Таким образом, любая система неприводимых и неразложимых аксиом неприводимой теории приводит к стандартному разложению сигнатуры.

Лемма 3. *Пусть* $\mathcal{T}$ *— неприводимая теория сигнатуры* Σ, *а* $\Sigma = \Sigma_1 \coprod \ldots \coprod \Sigma_n$ *— есть стандартное разложение сигнатуры на компоненты связности.*

Рассмотрим произвольное разложение теории на компоненты $\mathcal{T}{=}\mathcal{T}_1 \otimes \ldots \otimes \mathcal{T}_m$. *Тогда разложение сигнатуры, отвечающее этому разложению теории, получается из разложения* $\Sigma = \Sigma_1 \coprod \ldots \coprod \Sigma_n$ *объединением некоторых компонентов* Σ_i.

В частности, из связности некоторой системы неприводимых и неразложимых аксиом теории $\mathcal{T}$ *следует неразложимость теории* $\mathcal{T}$.

Используем максимальную систему аксиом теории $\Psi = \Psi(\mathcal{T})$. В силу следствия 1, каждая аксиома $\theta \in \Psi$ включается в некоторую теорию $\mathcal{T}_i$, $i = 1, \ldots, m$.

Поскольку теории $\mathcal{T}_j$ используют символы различных дизъюнктных частей сигнатуры, теория $\mathcal{T}_i$ включает и все аксиомы, непосредственно связанные с θ. Значит, она включает и весь компонент связности — все аксиомы связанные с θ. Но тогда каждая теория $\mathcal{T}_i$ использует все символы некоторых компонентов сигнатуры.

Заключительное утверждение леммы следует из леммы 2 и доказанного выше. Стандартные компоненты связности определяются по любой системе неприводимых и неразложимых аксиом. Если имеется только один компонент связности сигнатуры, то любое нетривиальное разложение теории состоит из одного компонента. $\square$

Продолжим доказательство теоремы 1. Присоединение к теории $\mathcal{T}$ предложений теории чистого равенства не изменяет используемую сигнатуру, поэтому достаточно доказать теорему для таких разложений теории $\mathcal{T}$, компоненты которых содержат все предложения чистого равенства теории $\mathcal{T}$.

Обозначим через $L(\Sigma)$ множество всех предложений, использующих символы только из сигнатуры Σ.

Оказывается, если компоненты разложимости включают все предложения чистого равенства теории $\mathcal{T}$, то они однозначно определяются по соответствующим им сигнатурам.

Лемма 4. *Пусть теория $\mathcal{T}$ сигнатуры $\Sigma = \Sigma_1 \cup \Sigma_2$ разлагается в произведение теорий $\mathcal{S}_1$ и $\mathcal{S}_2$ дизъюнктных сигнатур Σ_1, Σ_2, $\Sigma_1 \cap \Sigma_2 = \varnothing, \Sigma_1 \neq \varnothing \neq \Sigma_2$. Предположим, что эти теории содержат все формулы чистого равенства теории $\mathcal{T}$: $\mathcal{S}_i \supseteq \mathcal{T}^{\#}$, $i = 1, 2$.*

Тогда каждая теория $\mathcal{S}_i$ состоит из всех предложений теории $\mathcal{T}$, которые включают только символы из части сигнатуры Σ_i, $\mathcal{S}_i = \{\varphi \in L(\Sigma_i) |\ \varphi \in \mathcal{T}\}$, $i = 1, 2$. В частности, такое разложение определяется по разложению сигнатуры однозначно.

Введем $T_i = \{\varphi \in L(\Sigma_i) |\ \varphi \in \mathcal{T}\}$, $i = 1, 2$. По определению разложения в произведение теорий имеем включение $\mathcal{S}_i \subseteq T_i$, $i = 1, 2$. Поэтому следует доказать только обратное включение $\mathcal{S}_i \supseteq T_i$, $i = 1, 2$.

Пусть для определенности $\varphi \in L(\Sigma_1)$ и $\varphi \in \mathcal{T}$. Тогда $\mathcal{S}_1$, $\mathcal{S}_2 \vdash \varphi$. Пусть $\sigma_1 \in \mathcal{S}_1$ и $\sigma_2 \in \mathcal{S}_2$ — такие формулы, для которых σ_1, $\sigma_2 \vdash \varphi$. Тогда $\sigma_2 \vdash \sigma_1 \to \varphi$. По интерполяционной теореме Крейга найдется промежуточное предложение чистого равенства θ, для которого $\sigma_2 \vdash \theta$, $\theta \vdash \sigma_1 \to \varphi$. В частности, предложение θ принадлежит теории $\mathcal{T}$ и $\theta, \sigma_1 \vdash \varphi$.

Поскольку $\mathcal{S}_1$ включает все предложения чистого равенства теории $\mathcal{T}$, то $\mathcal{S}_1 \vdash \varphi$. $\square$

Замечание 7. *Утверждение леммы 4 легко обобщается на случай произвольного числа сомножителей в разложении теории.*

Завершим доказательство теоремы 1. Для произвольного $i = 1, \ldots, n$ обозначим через $\mathcal{S}_i$ теорию сигнатуры Σ_i, порожденную системой аксиом S_i, образующей компонент связности системы Ψ. Также обозначим через $\mathcal{S}'_i = \langle \mathcal{S}_i, \mathcal{T}^{\#} \rangle$ теорию $\mathcal{S}_i$ вместе с предложениями чистого равенства теории $\mathcal{T}$. Имеем разложение теории на неразложимые компоненты $\mathcal{T} = \mathcal{S}'_1 \otimes \ldots \otimes \mathcal{S}'_n$.

Рассмотрим произвольное разложение теории на неразложимые компоненты $\mathcal{T} = \mathcal{T}_1 \otimes \ldots \otimes \mathcal{T}_m$. По лемме 3 сигнатура каждого компонента $\mathcal{T}_j$ получается объединением некоторых компонентов Σ_{i_j}, $i_j = 1, \ldots, n$ стандартного разложения общей сигнатуры.

Дополним каждую теорию формулами чистого равенства: $\mathcal{T}'_i = \langle \mathcal{T}_i, \mathcal{T}^{\#} \rangle$. Тогда также имеем $\mathcal{T} = \mathcal{T}'_1 \otimes \ldots \otimes \mathcal{T}'_m$.

В силу леммы 4 и замечания 7, заключаем, что каждая теория $\mathcal{T}'_j$ образована всеми формулами теории $\mathcal{T}$, которые включают символы из соответствующей ей сигнатуры Δ_j. Однако по лемме 3 Δ_j получается объединением некоторых $\Sigma_i,\ i=1,2,\ldots,n$.

Для произвольного j рассмотрим произведение теорий $\mathcal{S'}_i$ сигнатур $\Sigma_i \subseteq \Delta_j$. Оно содержится в теории $\mathcal{T}$, является непротиворечивым и определеным. Применяя замечание 7, получаем, что теория $\mathcal{T'}_j$ совпадает с этим произведением. В силу неразложимости теории $\mathcal{T'}_j$, она совпадает с некоторой теорией $\mathcal{S'}_i$, использующей те же символы сигнатуры.

Этим построено отображение $j \to i$. Компоненты разложения $\mathcal{T'}_j = \mathcal{S'}_i$ по замечанию 7 однозначно определяются используемыми символами сигнатуры — компонентами стандартного разложения Σ_i. Поэтому это отображение разнозначно. Сюръективность отображения очевидна, так как в силу разложимости любой символ сигнатуры должен использоваться некоторой теорией $\mathcal{T'}_j$.

Теорема 1 доказана. □

Из леммы 2 следует, что для определения компонентов связности сигнатуры достаточно использовать любую систему $(Ax)''$ аксиом теории, образованную неприводимыми и неразложимыми предложениями. Ее нетрудно получить, исходя из неприводимых аксиом $(Ax)'$ теории $\mathcal{T}$, построенных в разделе 2, с помощью преобразований из доказательства предложения 4.

С учетом леммы 2 система неприводимых и неразложимых аксиом $(Ax)''$ теории $\mathcal{T}$ определяет разложение сигнатуры, а с учетом леммы 3 и самой теории. Вместе с заключительным утверждением леммы 3 это приводит нас к критерию разложимости элементарной теории $\mathcal{T}$.

Критерий разложимости. *Теория $\mathcal{T}$ сигнатуры Σ является разложимой тогда и только тогда, когда соответствующая ей система неприводимых и неразложимых аксиом $(Ax)''$ индуцирует более чем один компонент связности на сигнатуре Σ.*

4. КОНЕЧНО АКСИОМАТИЗИРУЕМЫЕ ТЕОРИИ

Отдельно рассмотрим проблему разложимости в случае конечно аксиоматизируемой теории $\mathcal{T}$ и сведем ее к к задаче минимизации системы аксиом следующим образом.

Каждой формуле ψ сигнатуры Σ можно сопоставить множество символов $supp(\psi)$, использованных в записи этой формулы. Обозначим их количество через $np(\psi) = \#(supp(\psi))$.

Определим *вес* формулы ψ как натуральное число

$$w(\psi) = 3^{np(\psi)} = 3^{\#(supp(\psi))}.$$

Замечание 8. *Основание степени 3 выбрано так, чтобы для любых натуральных чисел n, m выполнялось неравенство $3^{n+m} > 3^n + 3^m$ (в частности, чтобы суммарный вес фрагментов разложения формулы был всегда меньше, чем вес самой формулы).*

Определение 8. *Весом системы аксиом Ψ теории $\mathcal{T}$ назовем сумму весов составляющих ее аксиом:*

$$w(\Psi) = \sum_{\psi \in \Psi} w(\psi).$$

Минимальной системой аксиом теории $\mathcal{T}$ будем называть такую систему аксиом Ψ, вес которой минимален.

Поскольку веса образуют множество натуральных чисел, всегда можно выбрать систему аксиом Ψ минимального веса (из известного множества систем аксиом). Далее будем предполагать, что система аксиом Ψ имеет минимальный вес. Возможность использования такой системы для решения проблемы разложимости объясняет следующее утверждение.

Предложение 5. *Пусть Ψ — минимальная система аксиом теории $\mathcal{T}$. Тогда она образована такими предложениями теории $\mathcal{T}$, которые являются одновременно неразложимыми и неприводимыми.*

Пусть ψ — предложение системы аксиом Ψ минимального веса.

Проверим вначале неразложимость этого предложения. Определим такие предложения θ, ϕ теории $\mathcal{T}$, для которых $\theta, \phi \vdash \psi$, причем θ и ϕ включают разные наборы сигнатурных символов, встречающихся в формуле ψ. Из замечания 8 следует, что $w(\psi) > w(\theta) + w(\phi)$. Поэтому если в системе аксиом Ψ аксиому ψ заменить парой аксиом θ, ϕ, то получим систему аксиом меньшего веса. Получаем противоречие с тем, что система аксиом Ψ имеет наименьший вес. Выводим из этого, что ψ — неразложимое предложение.

Также от противного докажем неприводимость ψ. Предположим, найдется сужение этой формулы в теории $\mathcal{T}$: $\mathcal{T} \vdash \theta$, $\theta \vdash \psi$. При этом, в формуле θ присутствует только часть сигнатурных символов формулы ψ. Тогда $w(\psi) > w(\theta)$, и если заменить в системе аксиом формулу ψ на формулу θ меньшего веса, то получим систему аксиом меньшего веса. Следовательно ψ неприводимо. $\square$

5. ЗАКЛЮЧЕНИЕ

В данной работе мы сформулировали критерий разложимости элементарной теории, доказали однозначность разложения сигнатуры в дизъюнктное объединение, отвечающее разложению теории, и однозначность разложения самой теории с точностью до формул чистого равенства.

Мы рассмотрели тот факт, что *несущественные* символы, используемые в предложениях теории (символы, которые на самом деле можно исключить из ее аксиом), влияют на то, в какие компоненты разложимости распадается теория. Поэтому важным шагом к определению разложимости является устранение таких символов из аксиом теории, т. е. *приведение* теории.

В случае конечной аксиоматизируемости теории мы обобщили операции приведения и разложения, введя определение *веса* системы аксиом. Таким образом, на практике ключевыми шагами к определению разложимости являются следующие операции: определение систем аксиом по данному набору предложений (теории) и нахождение для данной теории системы аксиом с минимальным весом.

СПИСОК ЛИТЕРАТУРЫ

1. Пальчунов Д. Е. Алгебраическое описание смысла высказываний естественного языка // Модели когнитивных процессов. — Новосибирск, 1997. — Вып. 158: Вычислительные системы. — С. 127–148.
2. Пальчунов Д. Е. Синтаксическая близость предложений языка первого порядка // Измерение и модели когнитивных процессов. — Новосибирск, 1998. — Вып. 162: Вычислительные системы. — С. 58–80.
3. Кейслер Г.,Чэн Ч. Ч. Теория моделей. — М.: Мир, 1977.
4. Palchunov D. E. GABEK for Ontology Generation // Integration and Learning in Organizations / Ed. by J. Zelger, Ph. Herdina, A. Oberprantacher. — GABEK. Contributions to Knowledge Organization, Bd. II. — Wien, LIT-publishing Company, 2005.

Д. К. Пономарёв

ПРОБЛЕМА РАЗЛОЖИМОСТИ ПРИ ФОРМАЛЬНОМ ОПИСАНИИ ЗНАНИЙ

Препринт
135

Рукопись поступила в редакцию 26.08.2006
Рецензент Н. В. Шилов
Редактор З. В. Скок

Подписано в печать 05.09.2006
Формат бумаги 60×84 1/16 Объем 1,2 уч.-изд.л., 1,25 п.л.
Тираж 60 экз.

Центр оперативной печати "Оригинал 2", г. Бердск, 49-а, оф. 7
тел./факс 8 (241) 5 38 77