

Управление, вычислительная техника и информатика

УДК 004.8

РАЗРАБОТКА МЕТОДА БИНАРНОЙ ДЕКОМПОЗИЦИИ ФУНКЦИОНАЛЬНЫХ ПРЕДИКАТОВ

М.Н. Рудометкина

Томский политехнический университет
Карагандинский государственный технический университет
E-mail: monika_7272@mail.ru

Проведен анализ единственного обоснованного метода бинарной декомпозиции конечных предикатов – метода декартовой декомпозиции. Впервые разработан и обоснован частный метод бинарной декомпозиции для функциональных предикатов, основанный на анализе зависимостей, при этом количество значений вспомогательной переменной сведено к минимуму. На основе предложенного метода создан алгоритм бинарной декомпозиции функциональных предикатов.

Ключевые слова:

Бинарная декомпозиция, предикат, функциональные предикаты, декартова декомпозиция.

Key words:

Binary decomposition, predicate, function predicate, Descartes decomposition.

Средства декомпозиции предикатов разрабатывались в работах Ю.Н. Павловского, М.Ф. Бондаренко, Ю.П. Шабанова-Кушнаренко, А.И. Мальцева [1–4]. Однако обзор работ этих авторов показал, что в области аппарата декомпозиции обнаруживается недостаток обоснованных, не требующих проверки эквивалентности полученного результата исходному предикату, и в то же время рациональных методов, позволяющих получать компактное бинарное представление. Следует отметить, что почти все имеющиеся формальные средства декомпозиции предикатов в указанных работах не учитывают зависимости между переменными. Учет таких особенностей в строении предиката позволяет выполнять его декомпозицию более рациональным путем. Между тем теория реляционных баз данных предлагает средства декомпозиции отношений в виде утверждений о зависимостях между атрибутами. Если взять во внимание наличие связи между реляционными отношениями и предикатами, то можно использовать средства декомпозиции отношений для декомпозиции предикатов. Как показал анализ литературы, такие средства в аппарате алгебры предикатов на данный момент отсутствуют.

Целью данной статьи является разработка бинарной конъюнктивной декомпозиции функциональных предикатов.

Функциональным (однозначным) назовем предикат $P \in \text{Pre}(S)$ с функциональной зависимостью $X \rightarrow Y$, для которого $X \cup Y = V$. Иначе говоря, в функциональном предикате должна содержаться функ-

циональная зависимость, которая охватывает все его переменные. Предикат $P \in \text{Pre}(S)$ может содержать функциональную зависимость $X \rightarrow Y$, но при этом не быть функциональным. Это возможно в случае, когда $X \cup Y \neq V$, т. е. когда функциональная зависимость охватывает не все переменные предиката P .

Для постановки задачи вводим обозначения, характеризующие предметное пространство. Пусть задан остов $T \subseteq V \times U$, а конечное множество V состоит из двух непересекающихся групп переменных: $X = \{x_1, x_2, \dots, x_m\}$ и $Z = \{z_1, z_2, \dots, z_n\}$, т. е. $V = \{x_1, x_2, \dots, x_m, z_1, z_2, \dots, z_n\}$. В качестве универсума предметов выбирается произвольное непустое не более чем счетное множество U , элементы которого называются предметами. Пусть предметное пространство S размерности n образуется с помощью остова T как декартово произведение областей изменения переменных из V .

Теперь задан произвольный конечный предикат $P \in \text{Pre}(S)$ с функциональной зависимостью $X \rightarrow Z$. Необходимо выполнить бинарную декомпозицию предиката $P(X, Z)$ с помощью одной вспомогательной переменной.

С точки зрения дальнейшей реализации модели важно, чтобы результат декомпозиции получился как можно более компактным (формулы полученных предикатов должны быть простыми, как и выражаемые ими отношения). Для этого надо, чтобы число значений вспомогательной переменной в пределах предикатов, полученных в результате декомпозиции, было минимальным.

Единственным обоснованным методом бинар-

ной декомпозиции предикатов, гарантирующим корректность результата, является, декартова декомпозиция [1], которая позволяет выполнять декомпозицию любых конечных предикатов, а не только функциональных. Хотя недостатки этого метода и обсуждались ранее, рассмотрим декартову декомпозицию как отправную точку для поиска более рационального метода в смысле экономики аппаратных ресурсов.

Для выполнения декартовой декомпозиции предиката $P \in \text{Pre}(S)$ необходимо в явном виде представить его область истинности. В качестве такого вида можно использовать совершенную нормальную дизъюнктивную форму, предиката P . Пусть P_T – область истинности предиката P , $P_T \subseteq S$. Тогда

$$P(X, Z) = \bigvee_{(a_1, a_2, \dots, a_m, c_1, c_2, \dots, c_n) \in P_T} x_1^{a_1} x_2^{a_2} \dots x_m^{a_m} z_1^{c_1} z_2^{c_2} \dots z_n^{c_n} = \\ = \bigvee_{i=1}^r x_1^{a_{i1}} x_2^{a_{i2}} \dots x_m^{a_{im}} z_1^{c_{i1}} z_2^{c_{i2}} \dots z_n^{c_{in}},$$

где r – количество кортежей отношения P_T (которое будет конечно в силу конечности предиката P).

Согласно методу декартовой декомпозиции вводится вспомогательная переменная y со значениями $B = \{b_1, b_2, \dots, b_r\}$, которые интерпретируются как уникальные имена кортежей отношения P_T . Каждому кортежу из P_T произвольным образом взаимно однозначно сопоставляется уникальное имя:

$$x_1^{a_{11}} x_2^{a_{12}} \dots x_m^{a_{1m}} z_1^{c_{11}} z_2^{c_{12}} \dots z_n^{c_{1n}} = y^{b_1}, \\ x_1^{a_{21}} x_2^{a_{22}} \dots x_m^{a_{2m}} z_1^{c_{21}} z_2^{c_{22}} \dots z_n^{c_{2n}} = y^{b_2}, \\ \dots \\ x_1^{a_{r1}} x_2^{a_{r2}} \dots x_m^{a_{rm}} z_1^{c_{r1}} z_2^{c_{r2}} \dots z_n^{c_{rn}} = y^{b_r}. \quad (1)$$

На основе системы (1) определяется вспомогательный предикат P_y :

$$P_y(X, Z, y) = \bigvee_{i=1}^r y^{b_i} x_1^{a_{i1}} x_2^{a_{i2}} \dots x_m^{a_{im}} z_1^{c_{i1}} z_2^{c_{i2}} \dots z_n^{c_{in}}. \quad (2)$$

Так как предикат $P_y(X, Z, y)$ выражает взаимно однозначное соответствие между каждым кортежем (X, Z) и его именем y , то он удовлетворяет паре функциональных зависимостей $y \leftrightarrow \{x_1, x_2, \dots, x_m, z_1, z_2, \dots, z_n\}$. Тогда в силу следствия из теоремы Хеза предикат $P_y(X, Z, y)$ удовлетворяет зависимости конъюнкции $\wedge \{\{x_1, y\}, \{x_2, y\}, \dots, \{x_m, y\}, \{z_1, y\}, \{z_2, y\}, \dots, \{z_n, y\}\}$. Бинарная декомпозиция предиката P_y выполняется с помощью кванторов существования в соответствии с последней зависимостью.

Предикат P_y связан с исходным P равенством

$$P(X, Z) = \exists y P_y(X, Z, y). \quad (3)$$

Действительно, последовательно применяя к правой части (3) тождество (2) и равенство (1), имеем

$$\exists y P_y(X, Z, y) = \bigvee_{b \in B} y / b(P_y(X, Z, y)) = \\ = \bigvee_{i=1}^r (X, Z, b_i) = \\ = \bigvee_{i=1}^r x_1^{a_{i1}} x_2^{a_{i2}} \dots x_m^{a_{im}} z_1^{c_{i1}} z_2^{c_{i2}} \dots z_n^{c_{in}} = P(X, Z).$$

Таким образом, декартова декомпозиция $P(X, Z)$ выполняется путем построения вспомогательного предиката $P_y(X, Z, y)$ в соответствии с равенствами (1), (2) и последующей бинарной декомпозицией $P_y(X, Z, y)$ с помощью кванторов существования. Равенство (3) позволяет вернуться к исходному предикату $P(X, Z)$.

Недостатком такой декомпозиции является то, что переменная y принимает столько же значений в рамках предиката P_y и его проекций, сколько кортежей содержит область истинности предиката P . Будем считать, что это максимальное количество значений вспомогательной переменной, которое требуется для бинарной декомпозиции предиката, поскольку такого количества всегда достаточно. Кроме того, при декартовой декомпозиции не учитываются характерные особенности строения предиката для того, чтобы уменьшить число значений вспомогательной переменной.

Задача состоит в том, чтобы, используя функциональную зависимость (и другие зависимости при их наличии в предикате), выполнить бинарную декомпозицию либо непосредственно с помощью кванторов существования, либо с предварительным введением одной вспомогательной переменной с минимальным количеством значений.

Введем несколько обозначений. Через P_T обозначим область истинности предиката $P \in \text{Pre}(S)$, $P_T \subseteq S$. Предикат $P(X, Z)$ с зависимостью $X \rightarrow Z$ можно рассматривать как сюръективную функцию $p: A \rightarrow C$, где

$$A = P_T[X] \subseteq \text{Dom} X, C = P_T[Z] \subseteq \text{Dom} Z. \quad (4)$$

Пусть r – число кортежей отношения A , l – число кортежей отношения C . Т. к. для любого $\bar{c} \in C$ найдется $\bar{a} \in A$ такой, что $\bar{c} = p(\bar{a})$, то $p: A \rightarrow C$ – сюръекция. Для любого предиката $P(X, Z)$ с зависимостью $X \rightarrow Z$ справедливо неравенство $r \geq l$, поскольку в разных кортежах $(\bar{a}, \bar{c}) \in P_T$ значения c могут повторяться, а значения $\bar{a}$ – нет. Учитывая зависимость $X \rightarrow Z$, число кортежей в отношении P_T будет равно r , т. е. числу кортежей отношения A .

При декартовой декомпозиции предиката $P(X, Z)$ число значений вспомогательной переменной y в рамках полученных формул будет равно количеству всех кортежей отношения P_T , которое совпадает с количеством элементов области определения функции $p: A \rightarrow C$, т. е. $k = r$.

В силу принципа однозначности [3] все морфологические модели, построенные на его основе, являются функциями [2, 3]. При этом оказывается, что количество элементов области определения таких функций сильно превосходит количество элементов области значений.

Рассмотрим модель флексии как функцию $p: A \rightarrow C$. Область определения такой функции – это различные наборы признаков, характеризующие влияние контекста на флексию (окончание), область значений – множество всех флексий. Количество элементов множества A лежит в пределах от 10^3 до 10^4 , тогда как количество элементов множества C составляет от 10^1 до 10^2 .

Смысл предлагаемого ниже метода состоит в том, чтобы соотнести вспомогательную переменную, которая вводится с целью бинарной декомпозиции предиката $P(X, Z)$ не с областью определения функции $p: A \rightarrow C$ (как это происходит при декартовой декомпозиции), а с областью ее значений или со вспомогательной областью, которая строится на ее основе. Кроме того, предлагаемый метод учитывает функциональную зависимость в исходном предикате. Тем самым предполагается сократить количество значений вспомогательной переменной до минимума.

В предлагаемом методе будут повсеместно использоваться зависимости конъюнкции вида $X \rightarrow x_1 | x_2 | \dots | x_m$, которые являются мощным инструментом декомпозиции предикатов. Непосредственными операторами декомпозиции в данном методе являются кванторы существования.

Перейдем к описанию метода бинарной декомпозиции предиката $P(X, Z)$ с функциональной зависимостью $X \rightarrow Z$, где $X = \{x_1, x_2, \dots, x_m\}$, $Z = \{z_1, z_2, \dots, z_n\}$. Прежде всего, рассмотрим тривиальную ситуацию — когда группа переменных состоит только из одной переменной, т. е. $m=1$. Тогда на основании следствия из теоремы Хеза функциональная зависимость $x \rightarrow \{z_1, z_2, \dots, z_n\}$ влечет зависимость конъюнкции $\wedge \{\{x, z_1\}, \{x, z_2\}, \dots, \{x, z_n\}\}$. Бинарная декомпозиция $P(x, Z)$ выполняется на основе последней зависимости с помощью кванторов существования. Вспомогательная переменная не требуется. Но это тривиальный случай, не характерный для задач реальной сложности, поэтому будем полагать, что $m \geq 2$.

С точки зрения строения исходного функционального предиката возможны три разные ситуации, которые определяют последовательность действий для достижения бинарной декомпозиции минимальной ценой. Рассмотрим каждый случай в порядке возрастания сложности.

При бинарной декомпозиции функционального предиката $P(X, Z)$ прежде всего надо проверить, являются ли аргументы $\{x_1, x_2, \dots, x_m\}$ описываемой функции $p: A \rightarrow C$ независимыми друг от друга. Взаимная независимость переменных $x_1, x_2, \dots, x_m$ в предикате $P(X, Z)$ означает выполнение зависимости $Z \rightarrow x_1 | x_2 | \dots | x_m$. Для строгой проверки предиката $P(X, Z)$ на наличие зависимости $Z \rightarrow x_1 | x_2 | \dots | x_m$ необходимо выполнить дизъюнктивное разложение предиката $P(X, Z)$ по переменным Z . После этого с помощью утверждения этого разложения можно точно установить, выполняется зависимость $Z \rightarrow x_1 | x_2 | \dots | x_m$ в предикате $P(X, Z)$ или нет. Если предикат $P(X, Z)$ удовлетворяет зависимости $Z \rightarrow x_1 | x_2 | \dots | x_m$, то бинарная декомпозиция выполняется достаточно просто. При этом возможны две ситуации.

Первая ситуация имеет место, когда группа переменных Z состоит из одной переменной — z . Тогда вспомогательная переменная не требуется: бинарная декомпозиция $P(X, z)$ выполняется с помощью кванторов существования непосредственно на основе данной зависимости конъюнкции, которую можно представить в виде $\wedge \{\{x_1, z\}, \{x_2, z\}, \dots, \{x_m, z\}\}$. Это самый простой случай. Чуть более сложный слу-

чай — когда группа Z состоит из двух и более переменных; тогда имеет место вторая ситуация.

Во второй ситуации надо заменить группу переменных Z одной вспомогательной переменной y , после чего осуществить переход к двум вспомогательным предикатам — $P_1(X, y)$ и $P_2(y, Z)$, — которые являются результатом промежуточной декомпозиции предиката $P(X, Z)$. Бинарная декомпозиция предикатов $P_1(X, y)$ и $P_2(y, Z)$ осуществляется с помощью кванторов существования на основе зависимостей конъюнкции $\wedge \{\{y, x_1\}, \{y, x_2\}, \dots, \{y, x_m\}\}$ и $\wedge \{\{y, z_1\}, \{y, z_2\}, \dots, \{y, z_n\}\}$ соответственно, которые, как будет показано далее, выполняются в силу построения этих предикатов.

Если $P(X, Z)$ не удовлетворяет зависимости $Z \rightarrow x_1 | x_2 | \dots | x_m$, то имеет место самая сложная ситуация из трех возможных. Здесь так же происходит переход к вспомогательным предикатам $P_1(X, y)$ и $P_2(y, Z)$ с последующей бинарной декомпозицией последних. Однако способ образования данных вспомогательных предикатов в этом случае гораздо сложнее.

В первой ситуации ясно как проводить бинарную декомпозицию функционального предиката $P(X, Z)$.

Ситуация, когда предикат $P(X, Z)$ удовлетворяет зависимости $Z \rightarrow x_1 | x_2 | \dots | x_m$ и группа переменных Z включает более одной переменной, является довольно распространенной, поскольку аргументы функции, как правило, не зависят друг от друга. Однако если предикат $P(X, Z)$ выражает частичную, а не определенную всюду функцию из $\text{Dom} X$ в $\text{Dom} Z$, то указанная зависимость может не выполняться.

Выполним дизъюнктивное разложение предиката P по переменным Z :

$$P(X, Z) = \bigvee_{(c_1, c_2, \dots, c_n) \in C} z_1^{c_1} z_2^{c_2} \dots z_n^{c_n} P(x_1, x_2, \dots, x_m, c_1, c_2, \dots, c_n) = \bigvee_{i=1}^l z_1^{c_{i1}} z_2^{c_{i2}} \dots z_n^{c_{in}} P(x_1, x_2, \dots, x_m, c_{i1}, c_{i2}, \dots, c_{in}), \quad (5)$$

где множество C определяется выражением (4), а l — число кортежей отношения C . Тогда функцию $p: A \rightarrow C$, описываемую уравнением $P(X, Z)=1$, в явном виде можно выразить системой вида

$$\begin{aligned} P(x_1, x_2, \dots, x_m, c_{11}, c_{12}, \dots, c_{1n}) &= z_1^{c_{11}} z_2^{c_{12}} \dots z_n^{c_{1n}}, \\ P(x_1, x_2, \dots, x_m, c_{21}, c_{22}, \dots, c_{2n}) &= z_1^{c_{21}} z_2^{c_{22}} \dots z_n^{c_{2n}}, \\ &\dots \\ P(x_1, x_2, \dots, x_m, c_{l1}, c_{l2}, \dots, c_{ln}) &= z_1^{c_{l1}} z_2^{c_{l2}} \dots z_n^{c_{ln}}, \end{aligned} \quad (6)$$

где $(c_{i1}, c_{i2}, \dots, c_{in}) \in C$, $i=1, 2, \dots, l$.

Выполним декомпозицию этих уравнений за счет введения вспомогательной переменной y со значениями $B = \{b_1, b_2, \dots, b_l\}$, которые будем понимать как уникальные имена кортежей отношения C . Каждому кортежу $(c_{i1}, c_{i2}, \dots, c_{in}) \in C$ произвольным образом ставим во взаимно однозначное соответствие его уникальное имя $b \in B$ с помощью следующих уравнений алгебры предикатов:

$$\begin{aligned} z_1^{c_{11}} z_2^{c_{12}} \dots z_n^{c_{1n}} &= y^{b_1}, \\ z_1^{c_{21}} z_2^{c_{22}} \dots z_n^{c_{2n}} &= y^{b_2}, \\ &\dots \\ z_1^{c_{l1}} z_2^{c_{l2}} \dots z_n^{c_{ln}} &= y^{b_l}. \end{aligned} \quad (7)$$

На основании (7) образуем вспомогательный предикат P_2 :

$$P_2(y, Z) = \bigvee_{i=1}^l y^{b_i} z_1^{c_{i1}} z_2^{c_{i2}} \dots z_n^{c_{in}}. \quad (8)$$

В силу такого построения предикат $P_2(y, Z)$ содержит пару функциональных зависимостей $y \leftrightarrow \{z_1, z_2, \dots, z_m\}$.

Преобразуя систему (6) на основе уравнений (7), получим

$$\begin{aligned} P(X, c_{11}, c_{12}, \dots, c_{1n}) &= y^{b_1}, \\ P(X, c_{21}, c_{22}, \dots, c_{2n}) &= y^{b_2}, \\ &\dots \\ P(X, c_{l1}, c_{l2}, \dots, c_{ln}) &= y^{b_l}. \end{aligned} \quad (9)$$

С помощью выражения (9) образуем вспомогательный предикат P_1 :

$$P_1(X, y) = \bigvee_{i=1}^l y^{b_i} P(X, c_{i1}, c_{i2}, \dots, c_{in}). \quad (10)$$

Т. к. $P_1(X, y)$ был образован из $P(X, Z)$ с зависимостью $\{z_1, z_2, \dots, z_m\} \rightarrow x_1 | x_2 | \dots | x_m$ путем взаимно однозначной замены каждого значения переменной Z значением переменной y , то предикат $P_1(X, y)$ будет содержать зависимость $y \rightarrow x_1 | x_2 | \dots | x_m$.

Нетрудно показать, что предикат P выражается через P_1 и P_2 в виде

$$P(X, Z) = \exists (P_1(X, y) \wedge P_2(y, Z)). \quad (11)$$

Применив тождество (2) к правой части равенства (11) и пользуясь законами для операций подстановки, получим

$$\begin{aligned} &\exists y (P_1(X, y) \wedge P_2(y, Z)) = \\ &= \bigvee_{b \in B} b (P_1(X, b) \wedge P_2(b, Z)) = \\ &= \bigvee_{b \in B} (y / b (P_1(X, y)) \wedge y / b (P_2(y, Z))) = \\ &= \bigvee_{i=1}^l (P_1(X, b_i) \wedge P_2(b_i, Z)) = \\ &= \bigvee_{i=1}^l (P(X, c_{i1}, c_{i2}, \dots, c_{in}) \wedge z_1^{c_{i1}} z_2^{c_{i2}} \dots z_n^{c_{in}}). \end{aligned}$$

Последнее выражение представляет собой дизъюнктивное разложение предиката P в виде (5), из чего следует справедливость (11).

После построения предикатов P_1 и P_2 можно выполнить их бинарную декомпозицию. Бинарная декомпозиция предиката $P_1(X, y)$ выполняется с помощью кванторов существования на основе зависимости конъюнкции $\{x_1, y\}, \{x_2, y\}, \dots, \{x_m, y\}$, бинарная декомпозиция $P_2(y, Z)$ — на основе зависимости $\{y, z_1\}, \{y, z_2\}, \dots, \{y, z_m\}$, которая имеет место в силу функциональной зависимости

$y \rightarrow \{z_1, z_2, \dots, z_m\}$ и следствия из теоремы Хеза. Таким образом, бинарная декомпозиция предиката $P(X, Z)$ выполнена.

Определим границы, в пределах которых должно находиться количество значений вспомогательной переменной, введенной с целью бинарной декомпозиции функционального предиката.

Пусть k — число значений вспомогательной переменной y . Когда предикат $P(X, Z)$ с зависимостью $X \rightarrow Z$ так же удовлетворяет зависимости $Z \rightarrow x_1 | x_2 | \dots | x_m$, то при его бинарной декомпозиции описанным выше способом имеем $k=l$, где l — количество элементов множества $C = P_1[Z]$. Это минимальное количество значений вспомогательной переменной, так как оно равно количеству элементов области значений функции $p: A \rightarrow C$. Поэтому данная ситуация устанавливает нижний предел для параметра k .

Как показано выше, верхним пределом параметра k является r — количество элементов множества $A = P_1[X]$ (когда вспомогательная переменная y вводится на основе декартовой декомпозиции предиката $P(X, Z)$ и количество ее значений равно количеству элементов области определения функции $p: A \rightarrow C$). Следовательно, при выполнении бинарной декомпозиции функционального предиката с помощью одной вспомогательной переменной параметр k должен лежать в пределах $l \leq k \leq r$.

Последняя и самая сложная ситуация с точки зрения бинарной декомпозиции функционального предиката $P(X, Z)$ имеет место тогда, когда зависимость $Z \rightarrow x_1 | x_2 | \dots | x_m$ не выполняется. В данном случае так же требуется вспомогательная переменная. Она будет введена после того, как будет сформирована система специальных подмножеств множества A , определяемого выражением (4). Значения вспомогательной переменной в данном случае будут выступать в качестве имен этих подмножеств. При таком подходе, с одной стороны, во вспомогательных предикатах будут выстроены зависимости, которые позволят выполнить их бинарную декомпозицию. С другой стороны, будет образовано минимально необходимое количество указанных подмножеств, что сведет количество значений вспомогательной переменной к минимуму.

Введем обозначения: $\bar{c}_i = (c_{i1}, c_{i2}, \dots, c_{in})$, $Z^{\bar{c}_i} = z_1^{c_{i1}} z_2^{c_{i2}} \dots z_n^{c_{in}}$, $i=1, 2, \dots, l$. Предикат $P(X, \bar{c}_i)$ описывает полный прообраз $\{X | (X, \bar{c}_i) = 1\}$ элемента $\bar{c}_i \in C$ при отображении $p: A \rightarrow C$. Обозначим множество $\{X | (X, \bar{c}_i) = 1\}$ через $s(\bar{c}_i)$. Множество всех полных прообразов $\{s(\bar{c}_1), s(\bar{c}_2), \dots, s(\bar{c}_l)\}$ является разбиением множества $A[M]$. Элементы разбиения называют слоями. Каждый элемент $\bar{c}_i = (c_{i1}, c_{i2}, \dots, c_{in}) \in C$ можно рассматривать как составное имя соответствующего ему слоя $s(\bar{c}_i)$, поскольку данное соответствие взаимно однозначно (это так называемое каноническое отображение [4]).

Чтобы выполнить бинарную декомпозицию предиката $P(X, Z)$, используя минимум значений вспомогательной переменной, предлагается следующий способ действий.

Для каждого слоя $s(\bar{c}_i) = \{X | P(X(\bar{c}_i) = 1)\}$, $i=1,2,\dots,l$, строим такое семейство его подмножеств $\{D_1^i, D_2^i, \dots, D_{k_i}^i\}$, которое удовлетворяет условию

$$s(\bar{c}_i) = D_1^i \cup D_2^i \cup \dots \cup D_{k_i}^i, \quad (12)$$

где

$$D_1^i = A_1^{i1} \times A_2^{i2} \times \dots \times A_m^{i1},$$

$$D_2^i = A_1^{i2} \times A_2^{i2} \times \dots \times A_m^{i2},$$

...

$$D_{k_i}^i = A_1^{ik_i} \times A_2^{ik_i} \times \dots \times A_m^{ik_i}.$$

Каждое множество $A_1^s \subseteq \text{Dom} x_1, A_2^s \subseteq \text{Dom} x_2, \dots, A_m^s \subseteq \text{Dom} x_m$, $i=1,2,\dots,l$, $q=1,2,\dots,k_i$ выбирается произвольно в пределах своего домена. Множества $D_1^i, D_2^i, \dots, D_{k_i}^i$ назовем параллелепипедами слоя $s(\bar{c}_i)$ (по аналогии с термином из функционального анализа). В силу условия (12) семейство параллелепипедов $\{D_1^i, D_2^i, \dots, D_{k_i}^i\}$ является покрытием слоя $s(\bar{c}_i)$.

Таким образом, предлагается для каждого слоя $s(\bar{c}_i)$ подобрать минимальное количество m -мерных параллелепипедов пространства $\text{Dom } X$, которые совместно покрывают этот слой.

Для выполнения декомпозиции вводим вспомогательную переменную y со значениями $b_1, b_2, \dots, b_k$, которые будем понимать как имена всех параллелепипедов, построенных в рамках данной модели; k – общее число построенных параллелепипедов, $k=k_1+k_2+\dots+k_l$, $\{b_1, b_2, \dots, b_k\} = B$. Мы стремимся сделать множество значений вспомогательной переменной y минимальным, поэтому задача состоит в том, чтобы количество элементов k_i в каждом семействе $\{D_1^i, D_2^i, \dots, D_{k_i}^i\}$, удовлетворяющем условию (12), было минимальным.

Определим границы, в пределах которых будет колебаться количество значений k вспомогательной переменной y при указанном способе ее задания.

Пусть r_i – количество всех кортежей слоя $s(\bar{c}_i)$ разбиения множества A , $i=1,2,\dots,l$. Тогда $r_1+r_2+\dots+r_l=r$, r – число кортежей множества A .

Так как отдельно взятый кортеж $(a_1, a_2, \dots, a_m) \in A$ можно рассмотреть как параллелепипед $\{a_1\} \times \{a_2\} \times \dots \times \{a_m\} \subset A$, то для каждого слоя всегда можно подобрать покрытие, состоящее из таких одноэлементных параллелепипедов. Число таких параллелепипедов для слоя $s(\bar{c}_i)$ совпадает с числом всех кортежей этого слоя и равно r_i , $i=1,2,\dots,l$. Поэтому число r_i возьмем в качестве максимального значения для параметра k_i .

Минимально возможное значение параметра k_i равно 1, когда слой $s(\bar{c}_i)$ сам по себе является параллелепипедом пространства $\text{Dom } X$. Но этот случай возможен не для всех полных прообразов $s(\bar{c}_i)$, иначе, выполнялась бы зависимость $Z \rightarrow x_1 x_2 \dots x_m$. Следовательно, каждое число k_i , $i=1,2,\dots,l$ удовлетворяет неравенству $1 \leq k_i \leq r_i$. Так как $k_1+k_2+\dots+k_l=k$, то, выполнив суммирование по всем индексам $i=1,2,\dots,l$, получим предельные значения параметра k : $l \leq k \leq r$.

В силу построения покрытий каждое значение k_i будет минимальным, $i=1,2,\dots,l$, поэтому и сумма $k_1+k_2+\dots+k_l=k$ будет минимальной.

Тогда в ситуации, когда зависимость $Z \rightarrow x_1 x_2 \dots x_m$ в функциональном предикате $P(X, Z)$ не выполняется, количество значений вспомогательной переменной, введенной для бинарной декомпозиции, будет находиться в допустимых пределах от l до r и при этом будет минимальным.

Перепишем выражение (13) в терминах алгебры предикатов:

$$P(X, \bar{c}_i) = D_1^i(X) \vee D_2^i(X) \vee \dots \vee D_{k_i}^i(X), \quad (13)$$

где

$$D_1^i(x_1, x_2, \dots, x_m) = A_1^{i1}(x_1) \wedge A_2^{i2}(x_2) \wedge \dots \wedge A_m^{i1}(x_m),$$

$$D_2^i(x_1, x_2, \dots, x_m) = A_1^{i2}(x_1) \wedge A_2^{i2}(x_2) \wedge \dots \wedge A_m^{i2}(x_m),$$

...

$$D_{k_i}^i(x_1, x_2, \dots, x_m) = A_1^{ik_i}(x_1) \wedge A_2^{ik_i}(x_2) \wedge \dots \wedge A_m^{ik_i}(x_m),$$

$i=1,2,\dots,l$.

После того как для каждого слоя $P(X, \bar{c}_i)$, $i=1,2,\dots,l$, построено минимальное покрытие (13) можно перейти к построению вспомогательных предикатов.

Каждому параллелепипеду произвольным образом взаимно однозначно сопоставим уникальное имя $b \in B$ (уникальное для всей модели, а не только в пределах данного слоя). Последовательно рассматривая каждый слой $P(X, \bar{c}_i)$, $i=1,2,\dots,l$, положим, что

$$D_1^i(X) = y^{b_{i1}}, D_2^i(X) = y^{b_{i2}}, \dots, D_{k_i}^i(X) = y^{b_{ik_i}}. \quad (14)$$

На основании (14) образуем вспомогательный предикат P_i :

$$P_i(X, y) =$$

$$= \bigvee_{j=1}^{k_i} (D_1^i(X) y^{b_{i1}} \vee D_2^i(X) y^{b_{i2}} \vee \dots \vee D_{k_i}^i(X) y^{b_{ik_i}}). \quad (15)$$

Так как $k_1+k_2+\dots+k_l=k$, то P_1 можно представить в более удобном виде:

$$P_1(X, y) = \bigvee_{j=1}^k D_j(X) y^{b_j}.$$

В соответствии с утверждением $X \rightarrow Y$ и $Y \cup Z \rightarrow W$, $X \cup Z \rightarrow W \wedge (Y \cup Z)$, предикат $P_1(X, y)$ удовлетворяет зависимости $y \rightarrow x_1 x_2 \dots x_m$, т. к. для каждого $b_j \in B$

$$P_1(X, b_j) = D_j(X) = A_1^j(x_1) \wedge A_2^j(x_2) \wedge \dots \wedge A_m^j(x_m).$$

Благодаря тому, что покрытие каждого слоя состоит только из параллелепипедов, вспомогательный предикат $P_1(X, y)$ удовлетворяет зависимости $y \rightarrow x_1 x_2 \dots x_m$ или, что то же самое, $\wedge \{x_1, y\}, \{x_2, y\}, \dots, \{x_m, y\}$. На основе этой зависимости можно выполнить бинарную декомпозицию P_1 .

Заметим, что система уравнений (14), $i=1,2,\dots,l$, равносильна одному уравнению $P_1(X, y)=1$.

Строим второй вспомогательный предикат.

Как говорилось выше, функцию $p: A \rightarrow C$, описываемую уравнением $P(X, Z)=1$, в явном виде

можно выразить системой равенств (6). С учетом (13) и (14) каждое уравнение из системы (6) можно представить в виде:

$$y^{b_{i1}} \vee y^{b_{i2}} \vee \dots \vee y^{b_{ik_i}} = Z^{\bar{c}_i}. \quad (16)$$

Уравнения (16), $i=1,2,\dots,l$, связывают имена параллелепипедов с именем слоя, покрытие которого они образуют.

На основании (16) образуем вспомогательный предикат P_2 :

$$P_2(y, Z) = \bigvee_{i=1}^l (y^{b_{i1}} y^{b_{i2}} \vee \dots \vee y^{b_{ik_i}}) Z^{\bar{c}_i}. \quad (17)$$

Система уравнений (16), $i=1,2,\dots,l$, равносильна одному уравнению $P_2(y, Z)=1$.

Заметим, что предикат $P_1(X, y)$ описывает всюду определенное многозначное отображение из множества A во множество B , т. к. параллелепипеды, покрывающие один и тот же слой, могут пересекаться, а значит, кортеж $\bar{a} \in A$ может входить в разные параллелепипеды. Но параллелепипеды, покрывающие разные слои разбиения пересекаться не могут, поэтому каждый параллелепипед является подмножеством только одного слоя. Следовательно, предикат $P_2(y, Z)$ описывает функцию из множества B во множество C , т. е. содержит функциональную зависимость $y \rightarrow \{z_1, z_2, \dots, z_m\}$. Тогда в силу следствия из теоремы Хеза предикат $P_2(y, Z)$ удовлетворяет зависимости конъюнкции $\wedge \{y, z_1\}, \{y, z_2\}, \dots, \{y, z_m\}$, которая позволяет выполнить его бинарную декомпозицию.

Предикаты P_1 и P_2 образованы так, что представляют собой результат декомпозиции исходного предиката P . Покажем, что P выражается через P_1 и P_2 в виде (11).

Действительно, пользуясь тождеством

$$\exists x_i(P) = \bigvee_{a \in \text{Dom} x_i} x_i / a(P)$$

и законами для операций подстановки, имеем

$$\begin{aligned} \exists y(P_1(X, y) \wedge P_2(y, Z)) &= \\ &= \bigvee_{b \in B} y / b(P_1(X, y) \wedge P_2(y, Z)) = \\ &= \bigvee_{b \in B} (y / b(P_1(X, y)) \wedge y / b(P_2(y, Z))) = \\ &= \bigvee_{i=1}^{\kappa} (P_1(X, b_{ij}) \wedge (P_2(b_{ij}, Z))) = \\ &= \bigvee_{i=1}^l (P_1(X, b_{i1}) \wedge P_2(b_{i1}, Z) \vee P_1(X, b_{i2}) \wedge P_2(b_{i2}, Z) \vee \dots \vee P_1(X, b_{ik_i}) \wedge P_2(b_{ik_i}, Z)) = \\ &= \bigvee_{i=1}^l (D_{i1}(X) \wedge Z^{\bar{c}_i} \vee D_{i2}(X) \wedge Z^{\bar{c}_i} \vee \dots \vee \\ &\quad \vee D_{ik_i}(X) \wedge Z^{\bar{c}_i} = \\ &= \bigvee_{i=1}^l (D_{i1}(X) \vee D_{i2}(X) \vee \dots \vee D_{ik_i}(X) \wedge Z^{\bar{c}_i}). \end{aligned}$$

Далее, используя (13) и (5), получим

$$\exists y(P_1(X, y) \wedge P_2(y, Z)) = \bigvee_{i=1}^l (P(X, \bar{c}_i) Z^{\bar{c}_i} = P(X, Z),$$

что и требовалось.

После того, как вспомогательные предикаты построены, остается выполнить их бинарную декомпозицию. Бинарная декомпозиция предиката $P_1(X, y)$ выполняется с помощью кванторов существования на основе зависимости конъюнкции $\wedge \{x_1, y\}, \{x_2, y\}, \dots, \{x_m, y\}$, $P_2(y, Z)$ — на основе зависимости $\wedge \{y, z_1\}, \{y, z_2\}, \dots, \{y, z_m\}$. Тем самым бинарная декомпозиция исходного предиката $P(X, Z)$ выполнена; при этом k — число значений переменной y в полученных предикатах — является минимальным и лежит в пределах $l \leq k \leq r$.

На основе описанного метода разработан алгоритм выполнения бинарной декомпозиции функциональных предикатов.

Пусть задан конечный предикат $P(X, Z)$ с функциональной зависимостью $X \rightarrow Z$. Для выполнения его бинарной декомпозиции с помощью одной вспомогательной переменной, принимающей минимальное возможное количество значений, надо выполнить следующие шаги.

1. Выполнить дизъюнктивное разложение предиката $P(X, Z)$ по переменным Z (5). С помощью утверждения: $X \rightarrow Y$ и $Y \cup Z \rightarrow W$, $X \cup Z \rightarrow W \setminus (Y \cup Z)$, по виду разложения проверить, выполняется зависимость $Z \rightarrow x_1 | x_2 | \dots | x_m$ в предикате $P(X, Z)$ или нет.

Если указанная зависимость выполняется, то дальнейшие шаги зависят от количества переменных $Z = \{z_1, z_2, \dots, z_m\}$: если $n=1$, то перейти к пункту 2; если $n>1$, то перейти к п. 3. Если зависимость $Z \rightarrow x_1 | x_2 | \dots | x_m$ не выполняется, то перейти к п. 5.

2. В данном случае предикат $P(X, z)$ удовлетворяет зависимости конъюнкции $\wedge \{z, x_1\}, \{z, x_2\}, \dots, \{z, x_m\}$. С помощью кванторов существования на основе данной зависимости выполнить бинарную декомпозицию $P(X, z)$. Бинарная декомпозиция предиката $P(X, Z)$ выполнена.

3. По формулам (7) ввести вспомогательную переменную.

4. По формулам (8) образовать вспомогательный предикат $P_2(y, Z)$. По формулам (9), (10) образовать вспомогательный предикат $P_1(X, y)$. Перейти к пункту 7.

5. Каждый предикат $P(X, \bar{c}_i)$, $i=1,2,\dots,l$ представить в виде выражения (13) так, чтобы число слагаемых было минимальным. Ввести вспомогательную переменную по формулам (14).

6. По формулам (15) образовать вспомогательный предикат $P_1(X, y)$. По формулам (16), (17) образовать вспомогательный предикат $P_2(y, Z)$.

7. Выполнить бинарную декомпозицию вспомогательных предикатов на основе содержащихся в них зависимостей конъюнкции с помощью кванторов существования: декомпозицию предиката $P_1(X, y)$ выполнить на основе зависимости $\wedge \{x_1, y\}, \{x_2, y\}, \dots, \{x_m, y\}$ декомпозицию предиката $P_2(y, Z)$ выполнить на основе зависимости $\wedge \{y, z_1\}, \{y, z_2\}, \dots, \{y, z_m\}$. Так как равенство

(11) выражает неэквивалентную декомпозицию предиката $P(X, Z)$ на предикаты $P_1(X, y)$ и $P_2(y, Z)$, то бинарная декомпозиция предиката $P(X, Z)$, выполнена.

Отметим некоторые достоинства и недостатки предложенного метода для декомпозиции функциональных предикатов.

Приведенный выше алгоритм является скорее инструкцией, чем алгоритмом в строгом смысле, поскольку пункт 5 является далеко не тривиальной задачей. Возможность полной алгоритмизации этой инструкции зависит от того, является ли задача, которая определяется пятым пунктом, алгоритмически разрешимой.

Достоинством данного метода является возможность обрабатывать формулу или отношение не целиком, а по частям: для каждого полного прообраза его минимальное покрытие параллелепипедами подбирается отдельно.

СПИСОК ЛИТЕРАТУРЫ

1. Павловский Ю.Н. Геометрическая теория декомпозиции и некоторые ее приложения. — М.: Вычислительный центр им. А.А. Дородницына РАН, 2011. — 93 с.
2. Бондаренко М.Ф. Модели языка // Бионика интеллекта. — 2004. — № 1. — С. 27–37.

В дальнейшем, на основе предложенного метода предполагается построить модель логической сети для морфологической модели путем бинарной декомпозиции последней. Полученные при этом результаты исследований будут опубликованы дополнительно.

Выводы

1. Проведен анализ единственного обоснованного метода бинарной декомпозиции конечных предикатов — метода декартовой декомпозиции.
2. Впервые разработан и обоснован частный метод бинарной декомпозиции для функциональных предикатов, который отличается от общего метода декартовой декомпозиции тем, что он основан на анализе зависимостей и количество значений вспомогательной переменной сведено к минимуму.
3. На основе предложенного метода разработан алгоритм бинарной декомпозиции функциональных предикатов.

3. Шабанов-Кушнаренко Ю.П. Теория интеллекта. Проблемы и перспективы. — Харьков: Высшая школа, 1987. — 159 с.
4. Мальцев А.И. Алгебраические системы. — М.: Наука, 1970. — 392 с.

Поступила 10.01.2012 г.