

Всё, что вообще может быть сказано, должно быть сказано ясно, а о чём невозможно говорить, о том следует молчать.

Л. Витгенштейн

Лекция 5. МЕТОДЫ СПЕЦИФИКАЦИИ СЕМАНТИКИ ФУНКЦИЙ

Основные подходы к спецификации семантики функций. Табличный подход, метод таблиц решений. Алгебраический подход: операционная, денотационная и аксиоматическая семантики. Языки спецификаций.

5.1. Основные подходы к спецификации семантики функций.

Для спецификации семантики функций используются следующие подходы: табличный, алгебраический и логический [5.1, стр.30-73], а также графический [5.2].

Табличный подход для определения функций хорошо известен ещё со средней школы. Он базируется на использовании таблиц. В программировании эти методы получили развитие в методе таблиц решений.

Алгебраический подход для определения функций базируется на использовании равенств. В этом случае для определения некоторого набора функций строится система равенств вида:

$$\begin{aligned} L_1 &= R_1, \\ \dots & \\ L_n &= R_n. \end{aligned} \tag{5.1}$$

где L_i и R_i , $i=1, \dots, n$, некоторые выражения, содержащие предопределенные операции, константы, переменные, от которых зависят определяемые функции (формальные параметры этих функций), и вхождения самих этих функций. Семантика определяемых функций извлекается в результате интерпретации этой системы равенств. Эта интерпретация может осуществляться по-разному (базироваться на разных системах правил), что порождает разные семантики. В настоящее время активно исследуются операционная, денотационная и аксиоматическая семантики.

Третий подход, логический, базируется на использовании предикатов - функций, у которых аргументами могут быть значения различных типов, а результатами являются логические значения (ИСТИНА и ЛОЖЬ). В этом случае набор функций может определяться с помощью системы предикатов. Заметим, что система равенств алгебраического подхода может быть задана с помощью следующей системы предикатов:

$$\text{РАВНО}(L_1, R_1),$$

(5.2)

...
РАВНО(Ln, Rn),

где предикат РАВНО истинен, если равны значения первого и второго его аргументов. Это говорит о том, что логический подход располагает не меньшими возможностями для определения функций, однако он требует от разработчиков ПС умения пользоваться методами математической логики, что, к сожалению, не для всех разработчиков оказывается приемлемым. Более подробно этот подход мы рассматривать не будем.

Графический подход также известен еще со средней школы. Но в данном случае речь идет не о задании функции с помощью графика, хотя при данном уровне развития компьютерной техники ввод в компьютер таких графиков возможен и они могли бы использоваться (с относительно небольшой точностью) для задания функций. В данном случае речь идет о графическом задании различных схем, выражающих сложную функцию через другие функции, связанными с какими-либо компонентами заданной схемы. Графическая схема может определять ситуации, когда для вычисления представляемой ею функции должны применяться связанные с этой схемой более простые функции. Графическая схема может достаточно точно определять часть семантики функции. Примером такой схемы может быть схема переходов состояний конечного автомата, такая, что в каждом из этих состояний должна выполняться некоторая дополнительная функция, указанная в схеме.

5.2. Метод таблиц решений.

Метод таблиц решений базируется на использовании таблиц следующего вида (см. табл. 5.1).

Переменные/ условия	Ситуации (комбинации значений)				
x1	a[1,1]	a[1,2]	...	a[1,m]	*
x2	a[2,1]	a[2,2]	...	a[2,m]	*
...			...		
xn	a[n,1]	a[n,2]	...	a[n,m]	*
s1	u[1,1]	u[1,2]	...	u[1,m]	u[1,m+1]
s2	u[2,1]	u[2,2]	...	u[2,m]	u[1,m+1]
...			...		
sk	u[k,1]	u[k,2]	...	u[k,m]	u[k,m+1]
Действия	Комбинации выполняемых действий				

Табл. 5.1. Общая схема таблиц решений.

Верхняя часть этой таблицы определяет различные ситуации, в которых требуется выполнять некоторые действия (операции). Каждая строка этой части задает ряд значений некоторой переменной или некоторого условия. Первый столбец этой части представляет собой список переменных или условий, от значений которых зависит выбор определяемых ситуаций. В каждом следующем столбце указывается комбинация значений этих переменных (условий), определяющая конкретную ситуацию. При этом последний столбец определяет ситуацию, отличную от предыдущих, т.е. для любых других комбинаций значений (будем обозначать их звездочкой *), отличных от первых, определяется одна и та же, (m+1)-ая, ситуация. Впрочем, в некоторых таблицах решений этот столбец может отсутствовать.

Нижняя часть таблицы решений определяет действия, которые требуется выполнить в той или иной ситуации, определяемой в верхней части таблицы решений. Она также состоит из нескольких (k) строк, каждая из которых связана с каким-либо одним конкретным действием, указанным в первом поле (столбце) этой строки. В остальных полях (столбцах) этой строки (т.е. для $u[i, j]$, $i=1, \dots, m+1$, $j=1, \dots, k$) указывается, следует ли выполнять (при $u[i, j]= '+'$) это действие в данной ситуации или не следует (при $u[i, j]= '-'$). Таким образом, первый столбец нижней части этой таблицы представляет собой список обозначений действий, которые могут выполняться в той или иной ситуации, определяемой этой таблицей. В каждом следующем столбце этой части указывается комбинация действий, которые следует выполнить в ситуации, определяемой в том же столбце верхней части таблицы решений. Для ряда таблиц решений эти действия могут выполняться в произвольном порядке, но для некоторых таблиц решений этот порядок может быть предопределён, например, в порядке следования строк в нижней части этой таблицы.

Рассмотрим в качестве примера описание работы светофора у пешеходной дорожки. Переключение светофора в нормальных ситуациях должно производиться через фиксированное для каждого цвета число единиц времени (Ткр - для красного цвета, Тжёл - для жёлтого, Тзел - для зелёного). У светофора имеется счётчик таких единиц. При переключении светофора в счётчике устанавливается 0. Работа светофора усложняется необходимостью пропускать привилегированные машины (при их появлении на светофор поступает специальный сигнал) с минимальной задержкой, но при обеспечении безопасности пешеходов. Приведённая на рис. 5.2 таблица решений описывает работу такого светофора, управляющего порядком движения у него в каждую единицу времени. Звёздочка (*) в этой таблице означает произвольное значение соответствующего условия.

Условия	Ситуации								
Состояние светофора	Кр	Кр	Кр	Жёл	Жёл	Зел	Зел	Зел	
	Нет	Нет	Да	*	*	*	*	*	
	T=Tкр	T=Tжёл	T>Tзел	*	*	*	Нет	Да	Да
Появление привилегированной машины	Нет	Да	*	*	*	*	Нет	Да	
Включить красный	-	-	-	-	-	-	+	-	
Включить жёлтый	-	+	+	-	-	-	-	-	
Включить зелёный	-	-	-	-	+	-	-	-	
T:=0	-	+	+	-	+	-	+	-	
T:=T+1	+	-	-	+	-	+	-	+	
Освобождение пешеходной дорожки	-	-	-	+	-	-	-	-	
Пропуск пешеходов	+	+	+	-	-	-	-	-	
Пропуск машин	-	-	-	-	-	+	+	+	
Действия	Комбинации выполняемых действий								

Рис. 5.2. Таблица решений "Светофор у пешеходной дорожки".

5.3. Операционная семантика.

В операционной семантике алгебраического подхода к описанию семантики функций рассматривается следующий частный случай системы равенств (5.1):

$$\begin{aligned}
 f_1(x_1, x_2, \dots, x_k) &= E_1, \\
 &\dots\dots\dots \\
 f_n(x_1, x_2, \dots, x_k) &= E_n,
 \end{aligned}
 \tag{5.3}$$

где в левых частях этих равенств явно указаны определяемые функции, каждая из которых зависит (для простоты) от одних и тех же параметров

$$x_1, x_2, \dots, x_k,$$

а правые части этих равенств представляют собой выражения, содержащие, вообще говоря, вхождения этих функций, т.е. определяемые функции могут быть *рекурсивными*. Поэтому определяемая функция может иметь дополнительные вхождения в левые части этой системы равенств с постоянными значениями некоторых параметров. Таким образом, каждый параметр x_j вхождения определяемой функции f_i в левую часть равенств (5.3) будем понимать либо как переменную u_j , либо как константу s_{ij} , причём совокупность переменных

$$u_1, u_2, \dots, u_k$$

представляет входные данные, для которых требуется вычислять определяемые функции.

Операционная семантика интерпретирует эти равенства как систему подстановок. Под подстановкой

$$E \left| \begin{array}{l} s \\ T \end{array} \right.$$

выражения (терма) T в выражение E вместо символа s будем понимать *переписывание* выражения E с заменой каждого вхождения в него символа s на выражение T . Каждое равенство

$$f_i(x_1, x_2, \dots, x_k) = E_i$$

задает в параметрической форме множество правил подстановок вида

$$f_i(T_1, T_2, \dots, T_k) \rightarrow E_i \left| \begin{array}{l} x_1, x_2, \dots, x_k \\ T_1, T_2, \dots, T_k \end{array} \right.,$$

где T_1, T_2, \dots, T_k - конкретные аргументы (значения или определяющие их выражения) данной функции. Каждое такое правило допускает замену в каком-либо выражении вхождения его левой части на её правую часть.

Интерпретация системы равенств (5.3) для получения значений определяемых функций в рамках операционной семантики производится следующим образом. Пусть задан набор входных данных (аргументов)

$$d_1, d_2, \dots, d_k.$$

На первом шаге осуществляется подстановка этих данных в левые и правые части равенств с выполнением там, где это возможно, предопределённых операций и с *переписыванием* получаемых в результате этого равенств. В результате этого будет сформирована исходная *преобразуемая* система равенств.

На каждом следующем шаге по *текущей преобразуемой* системы равенств производится формирование новой *преобразуемой* системы равенств (которая становится *текущей*) путем *переписывания* этих равенств со следующими преобразованиями.

(1) Если правая часть очередного равенства является каким-либо значением, то это равенство не изменяется (это значение и является значением функции, указанной в левой части этого равенства).

(2) В противном случае правая часть является выражением, содержащим вхождение каких-либо определяемых функций с теми или иными наборами аргументов. В этом случае для каждого вхождения функции в эту правую часть (с конкретным набором аргументов) просматриваются левые части преобразуемых равенств и выполняются следующие действия.

(2.1) Если для этого вхождения в *текущей преобразуемой* системе равенств находится совпадающая с ним левая часть некоторого равенства, то проверяется правая часть этого равенства

(2.1.1) и в случае, если она является уже вычисленным значением, производится подстановка этого значения вместо указанного вхождения определяемой функции,

(2.1.2) если же эта правая часть не является вычисленным значением, то указанное вхождение *переписывается* в неизменном виде.

(2.2) В том же случае, если для указанного вхождения в *текущей преобразуемой* системе равенств не находится совпадающей с ним левой части никакого равенства, то к новой *преобразуемой* системе равенств дописывается *новое* равенство. Это равенство получается из исходного равенства для определяемой функции

$$f_i(y_1, y_2, \dots, y_k),$$

вхождение которой в данный момент исследуется, путем подстановки аргументов этой функции из исследуемого вхождения вместо параметров y_1, y_2, \dots, y_k этой функции (с выполнением предопределённых операций там, где это возможно).

Эти шаги будут осуществляться до тех пор, пока все определяемые функции не будут иметь вычисленные значения.

В качестве примера операционной семантики рассмотрим определение функции $F(n)=n!$. Она определяется следующей системой равенств:

$$F(0)=1,$$

$$F(n)=F(n-1)*n.$$

Для вычисления значения $F(3)$ осуществляются следующие шаги.

1-й шаг:

$$F(0)=1, F(3)=F(2) * 3.$$

2-й шаг:

$$F(0)=1, F(3)=F(2) * 3, F(2)=F(1) * 2.$$

3-й шаг:

$$F(0)=1, F(3)=F(2) * 3, F(2)=F(1) * 2, F(1)=F(0) * 1.$$

4-й шаг:

обозначает *цепление* множеств Z_1, Z_2, \dots, Z_k , причём вхождение в эту цепочку символа a_j представляет множество из одного элемента $\{a_j\}$. Это означает, что phi определяет множество цепочек

$$\{p_1 p_2 \dots p_k \mid p_j \in Z_j, j=1, \dots, k\},$$

причём цепочка $p_1 p_2 \dots p_k$ представляет собой последовательность записанных друг за другом цепочек p_1, p_2, \dots, p_k (результат выполнения операции *конкатенации* цепочек). Таким образом, каждая правая часть уравнений системы (5.4) представляет собой объединение множеств цепочек.

Решением системы (5.4) является набор языков

$$(L_1, L_2, \dots, L_n),$$

если все уравнения системы (5.4) превращаются в тождество при $X_1 = L_1, X_2 = L_2, \dots, X_n = L_n$.

Рассмотрим в качестве примера частный случай системы (5.4), состоящий из одного уравнения

$$X = a X \cup b X \cup X \cup c$$

с алфавитом $A = \{a, b, c\}$. Решением этого уравнения является язык

$$L = \{\text{phi } c \mid \text{phi} \in \{a, b\}^*\}.$$

Система (5.4) может иметь несколько решений. Так в рассмотренном примере помимо L решениями являются также

$$L_1 = L \cup \{\text{phi } a \mid \text{phi} \in \{a, b\}^*\}$$

и

$$L_2 = L \cup \{\text{phi } b \mid \text{phi} \in \{a, b\}^*\}.$$

В соответствии с денотационной семантикой в качестве *определяемого* решения системы (5.4) принимается наименьшее решение. Решение

$$(L_1, L_2, \dots, L_n)$$

системы (5.4) называется *наименьшим*, если для любого другого решения

$$(L_1', L_2', \dots, L_n')$$

выполняются соотношения

$$L_1 \subseteq L_1', L_2 \subseteq L_2', \dots, L_n \subseteq L_n'.$$

Так в рассмотренном примере наименьшим (а значит, определяемым денотационной семантикой) является решение L .

В качестве метода решения систем уравнений (5.3) и (5.4) можно использовать метод последовательных приближений. Сущность этого метода для системы (5.4) заключается в следующем. Обозначим правые части уравнений системы (5.4) операторами

$$T_i(X_1, X_2, \dots, X_n).$$

Тогда система (5.4) примет вид

$$X_1 = T_1(X_1, X_2, \dots, X_n),$$

$$X_2 = T_2(X_1, X_2, \dots, X_n),$$

$$\dots \dots \dots \quad (5.5)$$

$$X_n = T_n(X_1, X_2, \dots, X_n).$$

$$t_{i1} * t_{i2} * \dots * t_{ik} \rightarrow t_{i0}.$$

Такое аксиоматическое описание получит конкретную интерпретацию, если будут заданы конкретные типы данных $t_i = t_i'$, $i=1, \dots, r$, и конкретные значения констант $c_i = c_i'$, $i=1, \dots, l$. В таком случае говорят, что задана одна конкретная интерпретация A символов сигнатуры S , называемая *алгебраической системой*

$$A = (t_1', \dots, t_r', f_1', \dots, f_m', c_1', \dots, c_l'),$$

где f_i' , $i=1, \dots, m$, конкретная функция, представляющая символ f_i . Таким образом, аксиоматическое описание (S, E) определяет класс алгебраических систем (частный случай: одну алгебраическую систему), удовлетворяющих системе аксиом E , т.е. превращающих в тождества равенства системы E после подстановки в них f_i' , $i=1, \dots, m$, и c_i' , $i=1, \dots, l$, вместо f_i и c_i соответственно.

В программировании в качестве алгебраической системы можно рассматривать, например, тип данных, при этом определяемые функции представляют операции, применимые к данным этого типа. Так К. Хоор построил аксиоматическое определение набора типов данных [5.4], которые потом Н. Вирт использовал при создании языка Паскаль.

В качестве примера рассмотрим систему равенств

$$\text{УДАЛИТЬ}(\text{ДОБАВИТЬ}(m, d)) = m,$$

$$\text{ВЕРХ}(\text{ДОБАВИТЬ}(m, d)) = d,$$

$$\text{УДАЛИТЬ}(\text{ПУСТ}) = \text{ПУСТ},$$

$$\text{ВЕРХ}(\text{ПУСТ}) = \text{ДНО},$$

где УДАЛИТЬ, ДОБАВИТЬ, ВЕРХ - символы функций, а ПУСТ и ДНО - символы констант, образующие сигнатуру этой системы. Пусть D , D_1 и M - некоторые типы данных, такие, что $m \in M$, $d \in D$, $\text{ПУСТ} \in M$, $\text{ДНО} \in D_1$, а функциональные символы представляют функции следующих типов:

$$\text{УДАЛИТЬ}: M \rightarrow M,$$

$$\text{ДОБАВИТЬ}: M * D \rightarrow M,$$

$$\text{ВЕРХ}: M \rightarrow D_1.$$

Данная сигнатура вместе с указанной системой равенств, рассматриваемой как набор аксиом, образует некоторое аксиоматическое описание.

С помощью этого аксиоматического описания определим абстрактный тип данных, называемый *магазином*. Для этого зададим следующую интерпретацию символов её сигнатуры: пусть D - множество значений, которые могут быть элементами магазина, $D_1 = D \cup \{\text{ДНО}\}$, а M - множество состояний магазина, $M = \{d_1, d_2, \dots, d_n \mid d_i \in D, i=1, \dots, n, n \geq 0\}$, $\text{ПУСТ} = \{\}$, ДНО - особое значение (зависящее от реализации магазина), не принадлежащее D . Тогда указанный набор аксиом определяет свойства магазина.

С аксиоматической семантикой связана логика равенств (эквациональная логика), изучаемая в курсе "Математическая логика". Эта логика содержит правила вывода из заданного набора аксиом других формул.

5.6. Языки спецификаций.

Как уже отмечалось, функциональная спецификация представляет собой математически точное, но, как правило, не формальное описание поведения ПС. Однако, формализованное представление функциональной спецификации имеет ряд достоинств, главным из которых является возможность применять некоторые виды автоматизированного контроля функциональной спецификации.

Под *языком спецификаций* понимается формальный язык, предназначенный для спецификации функций. В нём используется ряд средств, позволяющих фиксировать синтаксис и выражать семантику описываемых функций. Различие между языками программирования и языками спецификации может быть весьма условным: если язык спецификаций имеет реализацию на компьютере, позволяющую как-то выполнять представленные на нём спецификации (например, с помощью интерпретатора), то такой язык является и языком программирования, может быть, и не позволяющий создавать эффективные программы. Однако, для языка спецификаций важно не эффективность выполнения спецификации на компьютере, а её выразительность. Язык спецификации, не являющийся языком программирования, также может быть полезен в процессе разработки ПС (для автоматизации контроля, тестирования и т.п.).

Язык спецификации может базироваться на каком-либо из рассмотренных методов описания семантики функций, а также поддерживать спецификацию функций для какой-либо конкретной предметной области.

Упражнения к Лекции 5.

5.1. Функции

function $F(x, y: integer): integer;$

function $G(x, y: integer): integer;$

function $R(x, y: integer): integer;$

определены с помощью операционной семантики равенствами:

$$R(x, y) = x * (y - 1),$$

$$F(x, y) = R(x + 1, y) - R(x, y - 1),$$

$$G(x, y) = F(x, R(x, y)).$$

Найти значения $G(3, 3)$.

5.2. Функции

function $F(n: integer): integer;$

function $G(n: \text{integer}): \text{integer};$
определены с помощью операционной семантики равенствами:

$$\begin{aligned} F(0) &= 1, \\ G(0) &= 2, \\ F(n) &= G(n-1), \\ G(n) &= F(n-1) + G(n-1). \end{aligned}$$

Найти значения $F(3)$ и $G(3)$.

5.3. Формальные языки E и T определены над алфавитом $\{ 'a', '*', '&', '<', '>' \}$

с помощью денотационной семантики равенствами

$$\begin{aligned} E &= T \cup '*' T \cup E '&' T, \\ T &= 'a' \cup 'a*' \cup '<' E '>' \end{aligned}$$

Какие из следующих строк

$$\begin{aligned} '*a&*a*&a*', \\ '*a&<a&a*>', \\ '*<*a*&a>&<*a*>*' \end{aligned}$$

принадлежат языку E и какие из них не принадлежат языку E .

5.4. Тип R определён с помощью следующей аксиоматической семантики.

Описания:

```
type R = record P1, P2, P3: CHAR end;
function READ(S: R): CHAR; {READ: R → CHAR}
function SHIFT(S: R): R; {SHIFT: R → R}
function ADD(S: R, C: CHAR): R; {ADD: R * CHAR → R}
function REMOVE(S: R): R; {REMOVE: R → R}
var X, Y, Z: CHAR;
U: R;
```

Аксиомы:

$$\begin{aligned} \text{SHIFT}(\text{ADD}(\text{ADD}(\text{ADD}(U, X), Y), Z)) &= \\ &\text{ADD}(\text{ADD}(\text{ADD}(U, Y), Z), X); \\ \text{REMOVE}(U) &= \text{SHIFT}(\text{ADD}(U, '#')); \\ \text{READ}(\text{SHIFT}(\text{ADD}(U, X))) &= X; \end{aligned}$$

Найти значение:

$$\text{READ}(\text{SHIFT}(\text{SHIFT}(\text{REMOVE}(\text{ADD}(\text{ADD}(U, 'a'), 'b'))))) =$$

Литература к Лекции 5.

- 5.1. В.Н. Агафонов. Спецификация программ: понятийные средства и их организация. - Новосибирск: Наука (Сибирское отделение), 1987.
- 5.2. Ian Sommerville. Software Engineering. - Addison-Wesley Publishing Company, 1992.
- 5.3. Д. Скотт. Теория решеток, типы данных и семантика / Данные в языках программирования. - М.: Мир, 1982. - С. 25-53.
- 5.4. К. Хоор. О структурной организации данных / У. Дал, Э. Дейкстра, К. Хоор. Структурное программирование. - М.: Мир, 1975. -

C. 98-197.

