

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ЯДРА СИСТЕМЫ МЕТАПРОГРАММИРОВАНИЯ

А. А. Седунов

Воронежский государственный университет

Поступила в редакцию 10.03.2011 г.

Аннотация. В данной работе представлена математическая модель структуры объектов, лежащей в основе разрабатываемой автором объектно-ориентированной инфраструктуры метапрограммирования TURBINE. Рассматриваются основные структуры данных, используемые для представления объектных моделей, отношения между объектами, а также основные разновидности и способы описания метаданных в объектных моделях.

Ключевые слова: метапрограммирование, ООП, теория, модель.

Annotation. The paper present mathematical model of object structure underlying object-oriented metaprogramming system TURBINE which is being developed by the author. It considers basic data structures which are used for presentation of object models, object relations and major cases/ways of description of metadata in object models.

Keywords: metaprogramming, OOP, theory, model.

ВВЕДЕНИЕ

На данный момент можно выделить несколько активно развивающихся подходов, направленных на преодоление ограничений, характерных для традиционной концепции объектно-ориентированного программирования, а также позволяющих улучшить модульность программных компонентов, разрабатываемых на основе методологии ООП. В контексте данной работы мы выделяем подход, основанный на декомпозиции классов, поскольку он может быть легко обобщен и приспособлен для модульного описания произвольных программных структур (помимо собственно классов). Метод основан на отказе от традиционного монолитного определения класса и замене его множественными определениями отдельных аспектов. Например, в языке Diesel (одна из реализаций метода декомпозиции классов) выделяются следующие аспекты:

- класс как символическое имя;
- класс как субъект отношения генерализации;
- класс как контракт, или набор операций;
- класс как структура данных, или набор определений слотов;

При наличии подходящего набора языковых средств для определения подобных аспектов становится возможной привязка этих опреде-

лений к конкретным модулям. Комбинируя модули, мы можем создавать контекстно-зависимые представления одного класса в зависимости от конкретного места, в котором употребляется его имя, например, изменяя набор методов, доступных его клиентам.

Далее мы представим математическую модель структуры объектов, лежащей в основе разрабатываемой автором объектно-ориентированной инфраструктуры метапрограммирования TURBINE и показано, каким образом метод декомпозиции позволяет достичь контекстно-зависимой интерпретации элементов программной структуры. Рассматриваются основные структуры данных, используемые для представления объектных моделей, отношения между объектами, а также основные разновидности и способы описания метаданных в объектных моделях.

ОСНОВНЫЕ ОБОЗНАЧЕНИЯ

Уточним ряд специфических обозначений, используемых в данной работе.

Для записи функций без присвоения им явного имени используется ограниченная форма лямбда-нотации вида $\lambda x.E$, или $\lambda x.E$, если множество S ясно из контекста. При этом $\text{Dom} \lambda x.E = S$.

Запись $Y \approx X$ означает, что функция является изоморфизмом, а множества X и Y равнозначны. Если функция φ ясна из контекста,

будем сокращать запись до $Y \approx X$. Аналогичное обозначение будем использовать для элементов X и Y , т. е.

$$Y \approx X, x \in X, y \in Y \Rightarrow y \approx x \leftrightarrow y = \varphi(x).$$

Если конструктор φ не обозначен явно, вместо $\varphi(x)$ будем писать \hat{x} , а вместо $\varphi^{-1}(y)$ – соответственно \hat{y} .

Пусть $F : A \rightarrow S$, где S – набор множеств, тогда

$$\prod F \stackrel{def}{=} \{f \mid \text{Dom } f = A \wedge \forall (x \in A) [f(x) \in F(x)]\}.$$

Если \mathbf{I} – некоторое фиксированное счетное множество идентификаторов, то элементы произведения $\prod X$, где $X \in S^{\mathbf{I}}$, S – набор множеств и $I \in \mathcal{P}\mathbf{I}$, назовем размеченными кортежами. Значение размеченного кортежа x в точке $i \in I$ обозначим через $x.i$. Размеченное декартово произведение при фиксированной мощности множества I будем записывать в виде $i_0 : X_0, \dots, i_{n-1} : X_{n-1}$, где $X_k = X(i_k)$, а $\text{Rng} \langle i_k \rangle_{k \in 0..n} = I$ (порядок идентификаторов i_k при этом роли не играет). Аналогичным образом для представления элементов такого произведения будем использовать обозначение $i_0 : x_0, \dots, i_{n-1} : x_{n-1}$, где $x_k \in X_k$, или, если порядок идентификаторов фиксирован, x_0, \dots, x_{n-1} .

Запись $Z = i_0 : X_0, \dots, i_{n-1} : X_{n-1}, \dots$ означает, что Z является объединением множеств вида $Z = \prod X$, причем $X(i_k) = X_k$ для всех $k \in 0..n$. Такую запись будем называть открытым произведением.

Нам также потребуются средства для представления размеченных объединений. Пусть $F : A \rightarrow S$, где S – набор множеств, тогда

$$\sum F \stackrel{def}{=} \{[x, s] \mid x \in A \wedge s \in F(x)\}.$$

Для элементов множества $\sum F$ введем операцию tag таким образом, что $[x, s] \in \sum F \Rightarrow \text{tag}[x, s] \stackrel{def}{=} x$, а сами элементы будем обозначать $x(s)$: $[x, s] \in \sum F \Rightarrow x(s) \stackrel{def}{=} [x, s]$. Если множество $F(x)$ состоит из одного элемента, то запись $x(s)$ сократим до x .

Пусть $X \in S^{\mathbf{I}}$, где S – набор множеств и $I \in \mathcal{P}\mathbf{I}$. Элементы множества $\sum X$ назовем размеченными вариантами (далее – варианты).

При фиксированной мощности множества I объединение $\sum X$ будем записывать в виде $i_0(X_0) + \dots + i_{n-1}(X_{n-1})$, где $X_k = X(i_k)$, а $\text{Rng} \langle i_k \rangle_{k \in 0..n} = I$. Если среди множеств X_k есть одинаковые, то соответствующие члены группируются в виде $(i_{k_1} + \dots + i_{k_p})(X_k)$. Скобки вокруг X_k могут быть опущены, если это не приводит к неоднозначности.

По аналогии с открытыми произведениями можно ввести понятие открытого объединения $i_0(X_0) + \dots + i_{n-1}(X_{n-1}) + \dots$ как объединения (обычного \cup) множеств вида $\sum X$, где $X_k = X(i_k)$ при любом $k \in 0..n$.

Пусть $U = i_0(X_0) + \dots + i_{n-1}(X_{n-1})$ и $f_k : X_k \rightarrow Y$, тогда $f = i_0(f_0) \oplus \dots \oplus i_{n-1}(f_{n-1})$ обозначает такую функцию f , что

$$\begin{cases} f : U \rightarrow Y \\ z = i_k(x) \in U \Rightarrow f(z) = f_k(x) \end{cases}$$

ОСНОВНЫЕ ПОНЯТИЯ

Рассмотрим формальное определение структуры объекта. Для этого вначале введем два фиксированных счетных множества \mathbf{V} и \mathbf{L} , где \mathbf{L} – счетное множество ссылок с выделенным элементом $\text{null} \in \mathbf{L}$, а \mathbf{V} – множество простых значений, имеющее вид открытого размеченного объединения

$$\mathbf{V} = \text{int} : [\text{int}] + \text{bool} : [\text{bool}] + \text{void} : [\text{void}] + \dots, \quad (1)$$

где $[\text{int}] \approx \mathbb{N}_0$ – используется для представления неотрицательных целых чисел, $[\text{bool}] \approx \mathbf{B}$ – для представления логических значений и $[\text{void}] = \{\mathbf{void}\}$.

Определим множества $[\text{val}]$, $[\text{dic}]$ и $[\text{dat}]$ следующей системой уравнений:

$$\begin{cases} [\text{val}] \approx \text{pr } \mathbf{V} + \text{memb}[\text{dic}] \\ [\text{dic}] = \mathbf{I}^{[\text{dat}]} \\ [\text{dat}] \stackrel{def}{=} \text{val}[\text{val}] + \text{ref } \mathbf{L} \end{cases} \quad (2)$$

Используя аппарат теории доменов, можно показать существование минимального набора $\langle [\text{val}], [\text{dic}], [\text{dat}] \rangle$, удовлетворяющего. Далее мы будем использовать множества из этого набора. Элементы множества $[\text{val}]$ назовем значениями (или статическими структурами), элементы, элементы $[\text{dic}]$ – словарями, а элементы $[\text{dat}]$ – элементами данных.

Константа используется в качестве значения в словаре и выражает тот факт, что соответствующий идентификатор отсутствует в структуре (при этом формально словарь остается функцией, определенной на всем множестве \mathbf{I}).

Для произвольного множества Z через Z обозначим его расширение константой (при этом подразумевается, что $\notin Z$):

$$Z = Z \cup \{\}. \quad (3)$$

Для словаря введем функцию $\text{Dom} : [dic] \rightarrow \mathcal{P}\mathbf{I}$, которая описывает множество идентификаторов x , отображаемых на элементы данных (т. е. полный прообраз множества $[dat]$):

$$\text{Dom } d = d^{-1} [dat]. \quad (4)$$

Таким образом, $d(x) \neq \text{при } x \in \text{Dom } d$.

Множество кортежей, состоящих из элементов данных, обозначим $[seq]$:

$$[seq] = [dat]^*. \quad (5)$$

Словари и кортежи элементов данных объединим общим термином коллекции:

$$[col] = \text{dic} [dic] + \text{seq} [seq]. \quad (6)$$

Мы распространяем обозначение на произвольные коллекции. При этом в случае кортежей $\text{Dom } d = \text{Dom } d$. Кроме того, поскольку словари и кортежи имеют вид $X \rightarrow [dat]$, мы будем для упрощения записи интерпретировать элементы множества $[col]$ как функции вида $\text{dic } d \oplus \text{seq } s$.

Внутренняя структура объектов описывается следующим образом:

$$[obj] \approx \text{null} \{-\} + \text{bd}(\text{type} : \mathbf{T}, \text{self} : \mathbf{L}, \text{memb} : [dic]), \quad (7)$$

$$[list] \approx (\text{self} : \mathbf{L}, \text{itm} : [seq]). \quad (8)$$

Для элементов этих множеств введем наименования объект и список соответственно. Списки и объекты совместно будем называть динамическими структурами:

$$[dyn] = \text{obj} [obj] + \text{list} [list]. \quad (9)$$

Структурное различие между динамическими структурами и значениями состоит в наличии у первых компонента self , который содержит идентичность.

Множество \mathbf{T} используются для классификации объектов. Мы будем считать его константой, содержащей конечное число элементов – “типов” (термин “тип” взят в кавычки для того, чтобы его можно было отличить от типов данных, которые рассматриваются позже). Сами “типы” будут конкретизироваться по ходу дальнейшего построения теории.

Элементы $[val]$ и $[dyn]$ в совокупности назовем структурами:

$$[struct] = \text{val} [val] + \text{dyn} [dyn]. \quad (10)$$

Для значений определим селектор $\text{memb} : [val] \rightarrow [dic]$:

$$\begin{cases} \text{memb}(v | v \approx \text{pr } z) = \mathbf{dic}_0, \\ \text{memb}(v | v \approx \text{memb } z) = z, \end{cases} \quad (11)$$

где \mathbf{dic}_0 – это словарь, который отображает любой идентификатор на пустой кортеж:

$$\mathbf{dic}_0 = \gg (x \in \mathbf{I}). \quad (12)$$

Аналогично определим селекторы $\text{memb} : [dyn] \rightarrow [dic]$ и $\text{id} : [dyn] \rightarrow \mathbf{L}$ для динамических структур:

$$\text{id}(\text{obj } z | z \approx \text{null}) = \text{null} \quad (13)$$

$$\text{id}(\text{obj } z) = (\text{null}(\lambda _ . \text{null}) \oplus \oplus \text{bd}(\lambda v. (v. \text{self})))z^{\wedge} \quad (14)$$

$$\text{id}(\text{list } z) = z^{\wedge}. \text{self} \quad (15)$$

$$\text{memb}(\text{obj } z) = (\text{null}(\gg _ . \mathbf{dic}_0) \oplus \oplus \text{bd}(\gg v. (v. \text{memb})))z^{\wedge} \quad (16)$$

$$\begin{aligned} \text{memb}(\text{list } z) = \\ = \mathbf{dic}_0 ["length" \leftarrow [int] (|z^{\wedge}. \text{itm}|)]. \end{aligned} \quad (17)$$

Операцию type будем использовать для обозначения “типа” объекта:

$$\text{type}(\text{obj } z) = (\text{null}(\lambda _ . \text{Null}) \oplus \oplus \text{bd}(\lambda v. (v. \text{type})))z^{\wedge}, \quad (18)$$

где $\text{Null} \in \mathbf{T}$ – специальная константа.

Таким образом, списки определяют специальный словарь, который связывает длину соответствующего кортежа элементов данных (представленную в виде значения $[int]$) с идентификатором “length”.

Операцию *memb* теперь можно поднять на уровень произвольной структуры, т. е. $\text{memb} : [\text{struct}] \rightarrow [\text{dic}]$:

$$\text{memb } z = \overset{\text{def}}{\left((\text{val} \oplus \text{dyn}) \text{memb} \right) z}. \quad (19)$$

Отметим, что *null* используется как для обозначения специальной ссылки, так и идентификатора в размеченных объединениях, связанных с объектами.

Для каждой структуры *s* определим собственную коллекцию $\text{Aol } s : [\text{struct}] \rightarrow [\text{col}]$ следующим образом:

$$\text{Aol}(\text{val } z) = \overset{\text{def}}{\text{memb } z} \quad (20)$$

$$\text{Aol}(\text{dyn } \text{obj } _) = \overset{\text{def}}{\text{memb } z} \quad (21)$$

$$\text{Aol}(\text{dyn list } v) = \overset{\text{def}}{v \hat{\cdot} \text{itm}} \quad (22)$$

Универсумом назовем размеченное объединение, включающее в себя элементы данных, структуры и коллекции:

$$\mathbf{U} = \overset{\text{def}}{\text{dat}} [\text{dat}] + \text{struct} [\text{struct}] + \text{col} [\text{col}]. \quad (23)$$

Для произвольного элемента универсума определим множество содержащихся в нем ссылок $\text{Ref} : \mathbf{U} \rightarrow \mathcal{PL}$:

$$\text{Ref} = \overset{\text{def}}{\emptyset} \quad (24)$$

$$\begin{aligned} \text{Ref}(\text{dat } d) &= \overset{\text{def}}{=} \\ &= \overset{\text{def}}{\left(\text{ref} (\gg x. \{x\}) \oplus \text{val} (\gg x. \text{Ref } x) \right) d} \end{aligned} \quad (25)$$

$$\text{Ref}(\text{struct } z) = \overset{\text{def}}{\text{Ref col } z} \quad (26)$$

$$\text{Ref}(\text{col } d) = \overset{\text{def}}{\bigcup_{x \in \text{Rng } d - \{\}} \text{Ref } x} \quad (27)$$

Для обозначения принадлежности объекта $z \in [\text{obj}]$ определенному “типу” *T* используем отношение $z : T$, заданное следующими правилами:

$$\frac{\text{type } z = T}{z : T} \quad (28)$$

$$\frac{z : T, T <: T'}{z : T'} \quad (29),$$

где $<:$ – отношение между “подтипами”, выражающее их классификацию. Общие свойства отношения $<:$ включают предпорядок и минимальность *Null*:

$$T <: T \quad (30)$$

$$\frac{T <: T', T' <: T''}{T <: T''} \quad (31)$$

$$\text{Null} <: T \quad (32)$$

“Тип” называется абстрактным, если он не может быть значением выражения *type z* для какой-либо динамической структуры. Абстрактные типы используются, главным образом, для классификации в отношениях $:$ и $<:$.

Имя “типа” мы также будем использовать для обозначения множества всех принадлежащих ему объектов:

$$\{e \in [\text{obj}] \mid e : T\} \quad (33)$$

Для связи ссылок с объектами введем понятие контекста. Ядром контекста называется множество **CCore** функций вида $f : \mathbf{L} \rightarrow [\text{dyn}]$, удовлетворяющих условию $f \text{ null} = \mathbf{null}$, где

$$\mathbf{null} = \overset{\text{def}}{\text{obj}(\hat{\cdot} \text{null})} \quad (34)$$

Множеством контекстов **C** называется множество всех кортежей, содержащих ядро контекста с идентификатором *core*:

$$\mathbf{!} = \overset{\text{def}}{\langle \text{core} : \mathbf{CCore}, \dots \rangle} \quad (35)$$

Контекст может отображать указанную метку на константу, которая, как и в случае со словарями, выражает неопределенность метки в этом контексте.

Для обозначения простейшего контекста введем константу **C₀**:

$$\mathbf{C}_0 = \overset{\text{def}}{\langle \text{core} : (\gg _) [\text{null} \leftarrow \mathbf{null}] \rangle} \quad (36)$$

Обозначение *Dom* естественным образом расширяется на ядро контекста:

$$\text{Dom } C.\text{core} = \overset{\text{def}}{\{l \in \mathbf{L} \mid C.\text{core } l \neq \}} \quad (37)$$

Если $C.\text{core } l = z$ и $z \in [\text{dyn}]$, то *l* называется внешней идентичностью динамической структуры *z*. При этом *id z* называется внутренней идентичностью.

Динамическая структура *z* называется валидной в контексте *C*, (запись: $\square_C z$), если все содержащиеся в ней ссылки определены в этом контексте и внутренняя идентичность совпадает с внешней:

$$\square_C z \Leftrightarrow \overset{\text{def}}{\begin{cases} \text{Ref } z \subseteq \text{Dom } C.\text{core} \\ \text{id } z \in \text{Dom } C.\text{core} \\ C.\text{core}(\text{id } z) = z \end{cases}} \quad (38)$$

Поднимем предикат валидности на уровень множества \mathbf{U} :

$$\square_C v \stackrel{def}{\Leftrightarrow} \begin{cases} \square_C z, v = \text{struct}(\text{dyn } z) \\ \text{Ref } z \subseteq \text{Dom } C.\text{core} \end{cases} \quad (39)$$

Ядро контекста $!$ называется валидным (запись: $\square_C.\text{core}$), если множество $\text{Dom } C.\text{core}$ конечно и любая динамическая структура, содержащаяся в его множестве значений, валидна в C :

$$\square_C.\text{core} \stackrel{def}{\Leftrightarrow} \text{Dom } C.\text{core} \in \mathcal{FL} \wedge \bigwedge (z \in \text{Rng } C.\text{core})[\square_C z]. \quad (40)$$

Контекст называется валидным, если валидны все его компоненты и выполняется условие верификации:

$$\square C \stackrel{def}{\Leftrightarrow} \text{ver } C \wedge \bigvee (z \in \text{Rng } C)[\square_C z]. \quad (41)$$

Далее мы уточним определение верификации. При изложении настоящего параграфа мы будем считать условие верификации тождественно-истинным.

Понятие контекста позволяет нам распространить некоторые свойства динамических структур на соответствующие им ссылки. Пусть \circ – некоторое бинарное отношение, левый операнд которого является динамической структурой. Введем обозначение $l \square_C x$ для ссылки l , которой в контексте C соответствует такая структура y , что $y \circ x$:

$$l \square_C x \stackrel{def}{\Leftrightarrow} l \in \text{Dom } C.\text{core} \wedge C.\text{core}(l) \circ x. \quad (42)$$

Обозначение контекста будем опускать (и писать, например, $l \in x$), если это не приводит к неоднозначности.

Рассмотрим теперь операции доступа к компонентам структур. Для начала введем механизм адресации. Значение $x \in \mathbf{X}$, где $\mathbf{X} = \mathbf{I} \cup \mathbb{N}$, будем называть элементом адреса, а кортеж $X \in \mathbf{X}^*$ – адресом. На множестве адресов введем частичный порядок \leq так, чтобы неравенство $X_1 \leq X_2$ имело место тогда и только тогда, когда X_1 является префиксом X_2 :

$$X_1 \leq X_2 \stackrel{def}{\Leftrightarrow} \exists (X \in \mathbf{X}^*)[X_2 = X_1 X]. \quad (43)$$

$$X_1 < X_2 \stackrel{def}{\Leftrightarrow} X_1 \leq X_2 \wedge X_1 \neq X_2. \quad (44)$$

Там, где это не приводит к противоречиям, адрес, состоящий из одного элемента $\langle x \rangle$, будем обозначать x .

Пусть $!$ – контекст, X – адрес и $u \in \mathbf{U}$. Выражение вида $u[X]_C \in [dat]$ назовем извлечением. Определим для него следующие правила редукции:

$$[-]_C \rightarrow \quad (45)$$

$$\text{dat } d[\langle \rangle]_C \rightarrow d \quad (46)$$

$$\frac{u \in \mathbf{U}, \text{tag } u \neq \text{dat}}{u[\langle \rangle]_C \rightarrow} \quad (47)$$

$$\text{dat}(\text{val } v)[X]_C \rightarrow \text{struct}(\text{val } v)[X]_C \quad (48)$$

$$\frac{l \notin \text{Dom } C}{\text{dat}(\text{ref } l)[X]_C \rightarrow} \quad (49)$$

$$\frac{l \in \text{Dom } C}{\text{dat}(\text{ref } l)[X]_C \rightarrow \text{struct}(\text{val } C.\text{core } l)[X]_C} \quad (50)$$

$$\text{struct } z[X]_C \rightarrow \text{col}(\text{col } z)[X]_C \quad (51)$$

$$\frac{u = \text{col } z, x \notin \text{Dom } z}{u[\langle x \rangle]_C \rightarrow} \quad (52)$$

$$\frac{u = \text{col } z, x \in \text{Dom } z}{u[\langle x \rangle]_C \rightarrow \text{dat } z(x)[X]_C} \quad (53)$$

Редукцию, состоящую из произвольного (возможно, пустого) конечного числа шагов, обозначим символом \rightarrow^* :

$$\alpha \rightarrow^* \alpha \quad (54)$$

$$\frac{\alpha \rightarrow \alpha', \alpha' \rightarrow^* \alpha''}{\alpha \rightarrow^* \alpha''}, \quad (55)$$

где α обозначает произвольное извлечение.

Можно показать, что редукция данного вида всегда приводит к единственному элементу данных:

Теорема 1.

$$u \in \mathbf{U}, X \in \mathbf{X}^* \Rightarrow \exists! (d \in [dat]) \times \left[u[X]_C \rightarrow^* d[\langle \rangle]_C \right]. \quad (56)$$

Таким образом, извлечение $[-]_C$ можно интерпретировать как функцию вида $\mathbf{C} \rightarrow \mathbf{U} \times \mathbf{X}^* \rightarrow [dat]$. Далее под $\text{eval } u[X]_C$ будем понимать значение этой функции.

В некоторых случаях необходимо наложить на элементы, соответствующие префиксам адреса X , определенные ограничения. В этом

случае мы будем использовать запись $d = \text{eval } u [X]_C$, где $\varphi : [dat] \rightarrow B$ - предикат:

$$\begin{aligned} d &= u [X]_C \stackrel{def}{\Leftrightarrow} \\ \stackrel{def}{\Leftrightarrow} &\left\{ \begin{array}{l} d = \text{eval } u [X]_C \\ \forall (\langle \rangle \leq X' < X) [\varphi (\text{eval } u [X']_C)] \end{array} \right. \end{aligned} \quad (57)$$

Если $\varphi(d)$ имеет вид $d \in Y$, то запись $d = \text{eval } u [X]_C$ будем сокращать до $d = \text{eval } u [X]_C$.

Теорема 2.

Результат любого извлечения валиден при условии валидности контекста:

$$\square C \Rightarrow \square_C (\text{eval } u [X]_C) \quad (58)$$

Таким образом, извлечение сохраняет свойство валидности при переходе от контекста к своему результату.

СУЩНОСТИ И МЕТАИНФОРМАЦИЯ

Перейдем теперь к рассмотрению более специфичной части ядра, предназначенной для описания метамodelей. В основе этой структуры лежат специальные объекты, называемые сущностями.

Для начала разделим все объекты на 2 категории. Объекты, имеющие “тип” Pure называются чистыми объектами. Объекты, “тип” которых отличен от Pure, называются сущностями. Абстрактный “тип” сущностей обозначается Ent.

Константа **null**, являясь объектом с “типом” Null, относится как к чистым объектам, так и к сущностям.

Выделяются 4 основных “подтипа” Ent: Attr (атрибуты), Link (связи), Binder (связывающие сущности) и Plent (простые сущности). Для этих “типов” вводятся соответствующие правила классификации:

$$\left\{ \begin{array}{l} \text{Attr} <: \text{Ent} \\ \text{Link} <: \text{Ent} \\ \text{Binder} <: \text{Ent} \\ \text{Plent} <: \text{Ent} \end{array} \right. \quad (59)$$

Кроме того, вводится свойство замкнутости “типа” Ent, запрещающее вводить для него новые непосредственные подтипы.

$$\begin{aligned} z : \text{Ent} &\Rightarrow z : \text{Attr} \vee z : \text{Link} \vee \\ &\vee z : \text{Binder} \vee z : \text{Plent} \end{aligned} \quad (60)$$

Атрибуты и связи объединяются “типом” Wrapper (промежуточные сущности) со следующими свойствами:

$$\left\{ \begin{array}{l} \text{Attr} <: \text{Wrapper} \\ \text{Link} <: \text{Wrapper} \\ \text{Wrapper} <: \text{Ent} \end{array} \right. \quad (61)$$

Правила классификации “типов” можно представить в виде графа, где вершинами являются сами типы, а наличие дуги соответствует отношению $<:$.

Отличие объектов-сущностей от чистых объектов состоит в том, что с ними ассоциируется метаянформация, позволяющая отслеживать их связи с другими объектами. Для хранения этой метаянформации мы расширим структуру контекста, добавив к ней ряд функций, заданных для ссылок на определенные виды сущностей. Таким образом, мы переопределяем множество $!$, введенное в предыдущем параграфе следующим образом:

$$\begin{aligned} &\stackrel{def}{C} \leftarrow ! \times \\ &\times \left(\begin{array}{l} \text{owner} : \mathbf{L} \rightarrow \mathbf{L}, \text{target} : \mathbf{L} \rightarrow \mathbf{L}, \\ \text{attrs} : \mathbf{L} \rightarrow (\mathbf{L}^*), \text{deps} : \mathbf{L} \rightarrow (\mathbf{L}^*), \\ \text{scope} : \mathbf{L} \rightarrow \mathbf{L}, \text{bnds} : \mathbf{L} \rightarrow (\mathbf{L}^*), \dots \end{array} \right) \end{aligned} \quad (62)$$

Для корректного использования переопределенного понятия нам потребуется также расширить константу $!_0$:

$$\begin{aligned} &\stackrel{def}{C}_0 \leftarrow !_0 \times \\ &\times \left[\left\langle \begin{array}{l} \text{owner} : \gg x., \text{target} : \gg x., \\ \text{attrs} : \gg x., \text{deps} : \gg x., \\ \text{scope} : \gg x., \text{bnds} : \gg x. \end{array} \right\rangle \right] \end{aligned} \quad (63)$$

Кроме того, необходимо определить валидность для новых компонентов контекста.

Функция owner формирует граф композиции, в котором ссылка l предшествует ссылке l' , если $C. \text{owner } l = l'$. Для обозначения ссылок, которые могут быть связаны функцией owner, вводится предикат cntd и наименование узел. В рассматриваемой модели к таким ссылкам относятся ссылки на сущности и списки:

$$\text{cntd}_C l \stackrel{def}{\Leftrightarrow} l \boxed{C} \text{Ent} \vee \text{tag}(C. \text{core } l) = \text{arr} \quad (64)$$

Для ссылок, не являющихся узлами, значением owner считается. Таким образом, полное определение валидности функции owner имеет вид:

$$f = C.\text{owner} \Rightarrow_{\square_C} f \stackrel{\text{def}}{\leftrightarrow} \begin{cases} l = \text{null} \vee \neg \text{cntd}_C l \rightarrow f(l) = \\ \text{cntd}_C l \wedge e = f(l) \rightarrow \text{cntd}_C e \wedge \square_{\hat{f}.\text{core}} e. \end{cases} \quad (65)$$

Функция `target` определена для ссылок “типа” `Wrapper`, которые предназначены для отслеживания узлов и формирует граф ассоциации. Для всех прочих ссылок значение `target` полагается равным:

$$f = C.\text{target} \Rightarrow_{\square_C} f \stackrel{\text{def}}{\leftrightarrow} \begin{cases} l = \text{null} \vee \neg l \stackrel{\square_C}{\vdash} \text{Wrapper} \rightarrow f(l) = \\ l \stackrel{\square_C}{\vdash} \text{Attr} \rightarrow f(l) \stackrel{\square_C}{\vdash} \text{Ent} \wedge \square_C f(l) \\ l \stackrel{\square_C}{\vdash} \text{Link} \rightarrow \text{cntd } f(l) \wedge \square_C f(l). \end{cases} \quad (66)$$

При этом для ссылки “типа” `Attr` значение функции `target` может быть только ссылка на сущность.

Функция `attrs` используется для привязки набора атрибутов к ссылке на сущность. Для ссылок на другие структуры привязка атрибутов не разрешена, поэтому значением функции `attrs` для таких ссылок мы считаем:

$$f = C.\text{attrs} \Rightarrow_{\square_C} f \stackrel{\text{def}}{\leftrightarrow} \begin{cases} \neg l \stackrel{\square_C}{\vdash} \text{Ent} \rightarrow f(l) = \\ f \text{ null} = \langle \rangle \\ L = f(l) \rightarrow |L| = |\text{Rng } L| \\ l \stackrel{\square_C}{\vdash} \text{Ent} \wedge l' \in \text{Rng } f(l) \rightarrow \\ \rightarrow l' \stackrel{\square_C}{\vdash} \text{Attr} \wedge \square_C l' \end{cases} \quad (67)$$

Таким образом, аргумент функции `attrs` представляет собой ссылку на сущность, а возвращаемое значение – кортеж ссылок на атрибуты, связанных с этой сущностью (заметим, что второе условие в гарантирует отсутствие двух идентичных атрибутов в этом кортеже).

Функции `scope` и `bnds` связывают метаинформацию с сущностями “типа” `Binder`, назначение которых состоит в ограничении доступности ссылок на другие объекты. Функция `scope` отображает ссылку `Binder` на ссылку, которой соответствует некоторая сущность (она может относиться к любому “типу”), называемая областью видимости. Валидность функции `scope` определяется аналогично валидности `target`:

$$f = C.\text{scope} \Rightarrow_{\square_C} f \stackrel{\text{def}}{\leftrightarrow} \begin{cases} l = \text{null} \vee \neg l \stackrel{\square_C}{\vdash} \text{Binder} \rightarrow f(l) = \\ l \stackrel{\square_C}{\vdash} \text{Binder} \rightarrow f(l) \stackrel{\square_C}{\vdash} \text{Ent} \wedge \square_C f(l) \end{cases} \quad (68)$$

Функция `bnds` используется для присоединения к сущности “типа” `Binder` набора связанных сущностей:

$$f = C.\text{bnds} \Rightarrow \begin{cases} l = \text{null} \vee \neg l \stackrel{\square_C}{\vdash} \text{Binder} \rightarrow \\ \square_C f \stackrel{\text{def}}{\leftrightarrow} \begin{cases} \rightarrow f(l) = L = f(l) \rightarrow |L| = |\text{Rng } L| \\ l \stackrel{\square_C}{\vdash} \text{Binder} \wedge l' \in \text{Rng } f(l) \rightarrow \\ \rightarrow l' \stackrel{\square_C}{\vdash} \text{Ent} \wedge \square_C l' \end{cases} \end{cases} \quad (69)$$

Неформально смысл связывающей сущности b можно охарактеризовать следующим образом: ссылка на связанную сущность (т. е. из кортежа `bnds b`) может использоваться только в области видимости `scope b`, или в объекте, который является прямым или косвенным компонентом этой области видимости. Для уточнения этого свойства необходимо прежде всего формализовать отношения композиции и ассоциации.

Будем говорить, что узел l является компонентом узла l' (а l' соответственно контейнером l), если выполняется одно из следующих условий:

- l' – ссылка на список, и l – один из элементов этого списка;
- l' – ссылка на сущность, и l содержится в наборе атрибутов $C.\text{attrs } l'$;
- l' – ссылка на сущность, и l можно извлечь из l' при условии, что элементы, соответствующие собственным префиксам адреса, являются значениями;
- l' – ссылка на сущность “типа” `Binder`, и l либо совпадает с $C.\text{owner } l'$, либо содержится в наборе $C.\text{bnds } l'$.

Для формализации отношения композиции потребуется, кроме того, ввести расширенное множество адресов \mathbf{X} :

$$\begin{cases} \hat{\mathbf{X}} \stackrel{\text{def}}{=} \mathbf{X} \cup \mathbf{X}', \\ \mathbf{X}' \stackrel{\text{def}}{=} \{\text{scope}\} \cup \text{attr } \mathbf{L} \cup \text{bnd } \mathbf{L}, \end{cases} \quad (70)$$

где $\mathbf{X} \cap \mathbf{X}' = \emptyset$.

Адрес вида `attr n` обозначает атрибут, который является компонентом данной сущности и имеет идентичность n , адрес `bnd n` – связанную

сущность с идентичностью n , score – область видимости (последние два варианта применимы только в том случае, если данная сущность имеет “тип” Binder).

Таким образом, определение отношения композиции (обозначение \prec_C) можно представить в виде следующих правил вывода (в них подразумевается указанное выше условие $\text{cntd}_C l \wedge \text{cntd}_C l'$):

$$\frac{\text{tag}(C.\text{core } l') = \text{list}, l = \text{eval}(l'[\langle n \rangle]_C)}{\langle n \rangle} \quad (71)$$

$$\frac{l'[\cdot]_C \text{Ent}, l \in \text{Rng}(C.\text{attrs } l')}{\langle \text{attr } l \rangle} \quad (72)$$

$$\frac{l'[\cdot]_C \text{Ent}, l = \text{eval}(l'[X]_C), X \neq \langle \rangle}{\langle X \rangle} \quad (73)$$

$$\frac{b[\cdot]_C \text{Binder}, l = C.\text{scope } b}{\langle \text{scope} \rangle} \quad (74)$$

$$\frac{b[\cdot]_C \text{Binder}, l \in \text{Rng}(C.\text{bnds } b)}{\langle \text{bnd } l \rangle} \quad (75)$$

Если конкретное значение адреса X не важно, запись $l \prec_C l'$ будем сокращать до $l \prec_! l'$:

$$l \prec_! l' \stackrel{\text{def}}{\Leftrightarrow} \exists (X \in \mathbf{X}^*) \left[l \overset{X}{\prec}_! l' \right] \quad (76)$$

Очевидно, что null не может быть контейнером для какой-либо ссылки.

Узел l называется корнем, если никакой другой узел не является его контейнером.

Расширим отношение \prec_C на кортежи адресов (будем говорить о нем, как о косвенной композиции):

$$\langle \rangle \prec_C^* l \quad (77)$$

$$\frac{l \overset{X}{\prec}_C l', l' \overset{X}{\prec}_C^* l''}{\langle X \rangle} \quad (78)$$

Отношение композиции в данном контексте должно быть согласовано с функцией owner этого же контекста. Сформулируем это ограничение в виде предиката ver_C , составленного

из конъюнкции следующих утверждений (здесь подразумевается, что $\text{cntd}_C l \wedge \text{cntd}_C l'$):

$$\forall (l, l' | l \prec_C l') [C.\text{owner } l = l'] \quad (79)$$

$$\forall (l, e | e = C.\text{owner } l \wedge e \neq \text{null}) \times [l \prec_C e] \quad (80)$$

$$\neg l \prec_C^* l \quad (81)$$

Из этих свойств, в частности, следует, что ссылка l является корнем тогда и только тогда, когда $\text{cntd}_C l$ и $C.\text{owner } l = \text{null}$.

Свойство исключает появление циклов в графе композиции при выполнении условия ver_C (хотя в общем случае отношение косвенной композиции является рефлексивным – см.). В частности, ссылка на сущность в таком контексте не может быть своим собственным компонентом.

Путем к узлу l в контексте $!$ называется такая последовательность $\text{Path}_C l \in (\mathbf{L})^*$, что

$$\text{Path}_C l = \{l_n\} \stackrel{\text{def}}{\Leftrightarrow} \begin{cases} l_0 = l \\ l_{n+1} = \begin{cases} C.\text{owner } l, l_n \neq \end{cases} \end{cases} \quad (82)$$

Из определения непосредственно следует, что если $l_i, l_j \in \mathbf{L}$ и $i < j$, то первая ссылка является косвенным компонентом второй:

$$i < j \Rightarrow l_i \prec_C^* l_j \quad (83)$$

Заметим, что в валидном контексте в силу конечности $\text{Dom } C.\text{core}$ любой путь, начиная с некоторого номера, состоит только из. Действительно, если бы это было не так, то путь не содержал бы (иначе все последующие элементы также были бы равны), т. е. состоял бы из корректных ссылок. Пусть общее число ссылок в контексте $N = |\text{Dom } C.\text{core}|$, тогда по крайней мере l_N совпадает с одной из предшествующих меток – обозначим ее l_k . Но поскольку $l_k \prec_C^* l_N$, то равенство $l_k = l_N$ противоречит условию валидности контекста. Таким образом, любой путь, начиная с некоторого номера, состоит только из.

В силу сказанного мы можем заменить бесконечную последовательность кортежем, содержащим только элементы пути, отличные от null – такой кортеж мы также будем называть путем:

$$\begin{cases} \text{path}_C : \mathbf{L} \rightarrow \mathbf{L}^* \\ \text{path}_C l = \langle l_i \rangle_{i=0}^n \stackrel{\text{def}}{\Leftrightarrow} \text{Path}_C l = \lambda n. \left[\left\langle l_i \right\rangle_{i=0}^n \right] \end{cases} \quad (84)$$

Обратным путем $\text{htap}_C l$ называется кортеж, составленный из элементов пути, расположенных в обратном порядке.

$$\text{htap}_C l = \langle l_{n-i-1} \rangle_{i=0}^n \stackrel{\text{def}}{\Leftrightarrow} \text{path}_C l = \langle l_i \rangle_{i=0}^n \quad (85)$$

Перейдем теперь к отношению ассоциации. Назначение сущностей “типа” *Wrapper* состоит в хранении (посредством функции *target* расширенного контекста) узла, который не является компонентом данного контейнера (прямое хранение такой ссылки в списке, словаре сущности или соответствующей ей функции *attrs* при условии их валидности невозможно в силу ограничений, наложенных на отношение композиции). При этом сущности “типа” *Wrapper*, как и любые другие, могут иметь собственные компоненты.

Если узел l является контейнером для ссылки l' , указывающей на сущность “типа” *Wrapper* и $C.\text{target } l' = l''$, где l'' – узел, то l ассоциирован с (или использует) l'' :

$$\frac{\text{cntd}_C l, \text{cntd}_C l'', l' \stackrel{x}{\sqsubseteq}_C \text{Wrapper}, l' \prec_C l, C.\text{target } l' = l''}{l \rightsquigarrow_{\hat{C}} l''} \quad (86)$$

Заметим, что узел, доступная через ссылку на чистый объект, не рассматривается в качестве компонента сущности, содержащей эту ссылку (например, в своем словаре). В этом смысле чистые объекты для структуры, состоящей из узлов, играют роль “барьеров”, разрывающих композиционные связи между ее элементами.

Ограничение на применение ассоциаций представим в виде предиката $\text{obscured}(l, l')$, выражающего тот факт, что ссылка l не может использоваться ссылкой l' . Критерий валидности (обозначим его $\text{ver}_{\rightsquigarrow} C$) при этом имеет вид:

$$\text{ver}_{\rightsquigarrow} C \stackrel{\text{def}}{\Leftrightarrow} \neg \exists (l, l' \mid l' \rightsquigarrow_1 l) \text{obscured}(l, l') \quad (87)$$

В рассматриваемой модели мы вводим единственное правило для предиката obscured . Это правило формализует рассмотренную выше семантику сущностей “типа” *Binder*.

$$\frac{b \stackrel{\square}{\sqsubseteq}_C \text{Binder}, l \in \text{Rng}(C.\text{bnds } b), \neg(l' \prec_C^* C.\text{scope } b)}{\text{obscured}(l, l')} \quad (88)$$

Помимо композиции и ассоциации мы вводим еще одно вспомогательное отношение зависимости представленное компонентом кон-

текста deps . Ссылка на сущность e называется зависящей от ссылки на атрибут a в контексте C , если $a \in \text{Rng } C.\text{deps } e$. Валидность функции deps определяется следующим образом:

$$f = C.\text{deps} \Rightarrow \square_C f \Leftrightarrow \begin{cases} \neg l \stackrel{\square}{\sqsubseteq}_C \text{Ent} \rightarrow f(l) = \\ f \text{ null} = \langle \rangle \\ L = f(l) \rightarrow |L| = |\text{Rng } L| \\ l \stackrel{\square}{\sqsubseteq}_C \text{Ent} \wedge l' \in \text{Rng } f(l) \rightarrow \\ \rightarrow l' \stackrel{\square}{\sqsubseteq}_C \text{Attr} \wedge \square_C l' \wedge \neg \text{obscured}(l', l) \end{cases} \quad (89)$$

Как и отношение ассоциативности, зависимость требует достижимости атрибута (выражаемой предикатом obscured).

Теперь мы можем конкретизировать предикат верификации контекста, определив его в виде:

$$\text{ver } C \stackrel{\text{def}}{\Leftrightarrow} \text{ver}_{\prec} C \wedge \text{ver}_{\rightsquigarrow} C \quad (90)$$

Основное назначение атрибутов состоит в связывании с сущностями контекстно-зависимых наборов метаданных. Далее мы рассмотрим правила построения таких наборов, опираясь на введенные отношения композиции и ассоциации.

Пусть e – ссылка на сущность в контексте C . Локальным дескриптором e называется кортеж, образованный слиянием кортежей атрибутов, принадлежащих косвенным контейнерам e и самой сущности e , указывающих на e и взятых в порядке обратного пути e :

$$\begin{aligned} \text{filter}(\text{htap}_C e) \left(\gg l \stackrel{\square}{\sqsubseteq}_C \text{Ent} \setminus \text{Null} \right) &= \\ &= \langle l_i \rangle_{i=1}^n \stackrel{\text{def}}{\Rightarrow} \text{locdsc}_C e = \\ &= \text{concat} \left(\text{filter}(C.\text{attrs } l_i C.\text{deps } l_i) \times \right. \\ &\quad \left. \times \left(\gg a.C.\text{target } a = e \right) \right)_{i=1}^n \end{aligned} \quad (91)$$

Таким образом, локальный дескриптор описывает метаданные сущности, выводимые из ее расположения в композиционном графе контекста. Использование операции слияния означает, что один и тот же атрибут не встречается в дескрипторе более одного раза. Кроме того, поскольку порядок построения пути определяется обратным путем, атрибуты, введенные сущностями, находящимися ближе к e , имеют больший индекс в дескрипторе. Эти условия обобщают принятое во многих языках правило

приоритета определений, находящихся внутри некоторой конструкции по отношению к внешним определениям.

Гибкая модификация метаданных сущности достигается за счет возможности расширения локального дескриптора, атрибутами, доступными через композиционные связи узлов “типа” Wrapper, которые ссылаются на исходную сущность. Тем самым один и тот же объект может обладать различным набором метаданных в зависимости от того, какая ссылка используется для обращения к нему.

Для учета ассоциативных связей мы введем понятие полного дескриптора. В отличие от локального, полный дескриптор определен для сущностей-обертки, имеющих “тип” Wrapper (доступ к исходной сущности можно получить с помощью функции target), поскольку его структура зависит не только от самой сущности, но и от ее обертки.

$$\begin{aligned} \text{fulldsc}_C w & \stackrel{\text{def}}{=} \\ & \stackrel{\text{def}}{=} \text{concat} \left(\text{locdsc}_C C.\text{target } w, \text{extdsc}_C w \right), \end{aligned} \quad (92)$$

где $\text{extdsc}_C w$ – вспомогательное обозначение внешнего дескриптора:

$$\begin{aligned} \text{filter} \left(\text{htap}_C w \right) \left(\gg ll \sqsubseteq_C \text{Ent} \setminus \text{Null} \right) & = \\ & = \left(l_i \right)_{i=1}^n \stackrel{\text{def}}{\Rightarrow} \text{extdsc}_C w = \\ & = \text{concat} \left(\text{filter} \left(C.\text{attrs } l_i C.\text{deps } l_i \right) \times \right. \\ & \quad \left. \times \left(\gg a.C.\text{target } a = e \right) \right)_{i=1}^n \end{aligned} \quad (93)$$

Заметим, что поскольку обертки сами являются сущностями, для них определен также и собственный локальный дескриптор (однако он, вообще говоря, не имеет отношения к дескриптору сущности, на которую указывает обертка).

ЗАКЛЮЧЕНИЕ

Представленная модель объектной структуры может быть использована в качестве основы

Седунов Алексей Александрович – ассистент кафедры программирования и информационных технологий Воронежского государственного университета, тел. 8-961-182-31-28, e-mail: alexey.sedunov@gmail.com

для реализации инфраструктуры метапрограммирования. Описанные в статье “типы” определяют набор структурных элементов, который представляется оптимальным с точки зрения основных шаблонов, возникающих при построении объектных моделей, в частности, отношений между объектами (композиция, ассоциация, зависимость). Кроме того, разработанный здесь механизм атрибутов, основанный на принципах контекстно-ориентированного программирования, позволяет обеспечить высокий потенциал расширяемости и модульности разрабатываемых моделей на уровне ядра соответствующей инфраструктуры.

В рамках проводимого автором исследования также разработаны теоретические основания вычислительных процессов, основанных на механизме объектных действий, поддерживающего модификацию структуры объектов и списков, манипуляции с метаданными контекста, а также обработку событий и, как следствие, отслеживание изменений в структуре моделей (с возможностью повторов и откатов). Однако подробное описание данного механизма выходит за рамки настоящей статьи.

СПИСОК ЛИТЕРАТУРЫ

1. Clark A., Sammut P., Willans J. Superlanguages. Developing Languages and Applications with XMF, 2008. <http://itcentre.tvu.ac.uk/~clark/Papers/Superlanguages.pdf>.
2. Clark A., Sammut P., Willans J. Applied Metamodeling. A Foundation for Language Driven Development (Second Edition), 2008. <http://www.ceteva.com/home/Papers/Applied%20Metamodelling%20%28Second%20Edition%29.pdf>.
3. Gilad Bracha. Newspeak Programming Language Draft Specification, Version 0.6, 2010. <http://bracha.org/newspeak-spec.pdf>.
4. Chambers C. The Diesel Language Specification and Rationale, 2006. <http://www.cs.washington.edu/research/projects/cecil/www/Release/doc-diesel-lang/diesel-spec.pdf>.
5. Bruce K. B. Foundations of Object-Oriented Languages. Types and Semantics. The MIT Press, 2002.

Sedunov A.A. – Assistant Department of Programming and Information Technologies, Voronezh State University, Tel. 8-961-182-31-28, e-mail: alexey.sedunov@gmail.com