УДК 004.4'22/.942

Т.В. СЕМЕНОВА, аспирант ХНУ имени В. Н. Каразина

ИСПОЛЬЗОВАНИЕ ШАБЛОНОВ ПОЛУСХЕМ В ПРОЦЕССЕ ПОСТРОЕНИЯ СТРУКТУРНО-ЛОГИЧЕСКОЙ МОДЕЛИ ПРЕДМЕТНОЙ ОБЛАСТИ ИНФОРМАЦИОННОЙ СИСТЕМЫ

У статті наведено аппарат шаблонних напівсхім, як розвиток теоріі напівсхім, що була запропонована у [1-2]. Розглядаються засоби використання шаблонних напівсхім у процесі побудови структурно-логічної моделі предметної області інформаційної системи.

1. Актуальность и постановка задачи.

Компьютеризованные информационные системы играют важную роль в современной промышленности, науке, технике и социальной сфере. Постоянно растущая сложность таких систем приводит к необходимости использования широкого спектра инструментальных программных средств для их разработки, что позволяет создателям системы построить комплекс согласованных моделей разрабатываемой системы и обеспечить таким образом повышение качества проектирования. Развитие программной индустрии в направлении автоматизации разработки информационных систем на базе комплексного моделирования существенных аспектов системы иллюстрируется созданием технологических решений, базирующихся на языке UML [3]. Интеграция моделей, регламентируемых языком UML, обеспечивает достаточно адекватное определение проектируемой системы.

Исследователи процессов разработки программных систем выделяют три основных аспекта, которые должны найти отражение в проектных документах на систему: данные, объекты или концепции и их структуры; архитектуру или невременные процессы; и динамику или поведение системы [4]. При этом для представления и анализа данных и их структуры используются структурно-логические модели. Общеизвестно, что процесс построения структурно-логической (концептуальной) модели предметной области является базовым и основополагающим при разработке информационной системы [5]. Ошибки и неточности, допущенные на этапе концептуального моделирования, являются наиболее дорогостоящими, поскольку требуют фактически полного повторения проектного процесса.

Сложность процесса концептуального моделирования объясняется тем, что он состоит в формализации и структуризации знаний, выделении базовых понятий предметной области и формулировке их определений, установление взаимосвязей между выделенными понятиями на основании представлений экспертов предметной области. Процесс концептуального моделирования требует доработок и уточнения модели, проверки корректности определенных в ней понятий предметной области и установленных взаимосвязей, и, следовательно, имеет итеративный характер. Следует отметить, что существующие CASE-средства [6-7] не поддерживают указанные выше особенности структурно-логического моделирования. Следует отметить, что в большинстве своем для этапа концептуального проектирования они предоставляют только графические редакторы и средства визуализации диаграммы.

В связи с этим работах [1, 2] авторами предложен алгебраический подход к построению концептуальных моделей предметных областей, базирующийся на введенном ими классе алгебраических объектов – полусхем. Этот подход позволил построить алгоритмы проверки корректности концептуального описания предметной области с точки зрения существования образцов понятий, определяемых полусхемой. Полусхема, корректная в том смысле, что все определяемые ее понятия имеют не менее одного образца, названа авторами схемой предметной области. Таким образом, схема может рассматриваться как концептуальная модель предметной области, поскольку фиксирует только структурные отношения между понятиями.

Настоящая работа посвящена дальнейшему исследованию схем, которые используются в качестве формализованной структурно-логической модели предметной области. В частности, рассматривается вопрос выделения шаблонных схем, которые позволят усовершенствовать процесс создания модели предметной области и облегчить ее анализ.

2. Основная часть.

2.1. Введение в теорию полусхем

Приведем основные понятия и определения теории полусхем [1].

N – конечное множество, элементы которого соответствуют именам понятий предметной области.

R – конечное множество, элементы которого соответствуют именам ролей, т.е. ссылкам внутри экземпляра понятия на его структурные части.

Для пары множеств X, Y обозначим $M_+(X,Y)$ – множество частичных, хотя бы где-то определенных, отображений из X в Y.

Определение 1.

Полусхемой предметной области называется тройка $S=(N,R,D)$, где N, R – конечные множества, $D \subset N \times M_+(R,N)$, для которой выполняется условие: для $n \in N$, $f,g \in M_+(R,N)$, $r \in R$ таких, что $(n,f) \in D$, $(n,g) \in D$ и $r \in dom(f) \cap dom(g)$, верно $f(r)=g(r)$.

Частичное отображение $\tau : N \times R \to N$ задается следующим образом: $\tau(n,r)$ определено в том и только том случае, если существует $f \in M(R,N)$, для которого $(n,f)\in D$, $r\in \mathrm{dom}(f)$ при этом $\tau(n,r)=f(r)$. Можно рассмотреть расширенный вариант частичного отображения $\tau : N\times R^{*}\to N$.

Определение 2.

Пусть тройка $S=(N,R,D)$ является полусхемой и для некоторого $n\in N$ существует $f\in M_{+}(R,N)$ такая, что $(n,f)\in D$. Тогда говорят, что для понятия n задан вариант определения f.

Определение 3.

Именующей нитью понятия n называется элемент (n,w) из множества $N\times R^{*}$, который удовлетворяет одному из следующих условий:

1. $w=e$ и $(n,\varepsilon)\in D$;
2. для $w=r_1r_2...r_k$ в последовательности $n_0=n$, $n_i=\tau(n_{i-1},r_i)$, где $i=1,\ldots,k$, все члены определены.

Множество именующих нитей понятия n обозначается через $T(n)$.

Можно рассмотреть расширенный вариант частичного отображения $\tau : T\to N$, где $T=\bigcup\limits_{n\in N} T(n)$.

Определение 4.

Пусть $n\in N,\ t\in T(n)$. Если $\tau(t)\in N_0$, то t называется терминальной именующей нитью, где $N_0=\{n\in N \mid (n,\varepsilon)\in D\}$. Элементы множества N_0 называются базовыми понятиями.

Через $T_0(n)$ обозначается множество терминальных именующих нитей понятия n, а через $F(T_0(n))$ множество всех конечных подмножеств множества $T_0(n)$.

Ключевым в предлагаемой модели является следующее

Определение 5.

Образцом понятия $n\in N$ называется конечное множество p терминальных именующих нитей этого понятия, удовлетворяющее следующему условию:

для всякой именующей нити $t=(n,w)\in p$ и для всякого представления w в виде w_1rw_2, где $w_1,w_2\in R^{*},\ r\in R$, найдется единственное отображение

$f \in M(R,N)$, для которого $(\tau(n,w_1),f) \in D$, $r \in dom(f)$, и для всех $r' \in dom(f)$ в p найдется нить вида $(n, w_1 r' v_{r'})$ для некоторого $v_{r'} \in R^*$.

Через $P(n)$ обозначается множество образцов понятия $n \in N$, а через $P = \bigcup_{n \in N} P(n)$ – множество всех образцов понятий из N.

Определение 6.

Полусхема $S = (N,R,D)$ называется схемой, если для любых $n \in N$ и $f \in M(R,N)$, для которых выполнено $(n,f) \in D$, можно построить хотя бы один образец $p \in P(n)$ в соответствии с вариантом определения f, т.е.

$$(n,f) \in D \text{ и } p = \bigcup_{r \in dom(f)} \left\{ t \in p \,\middle|\, \exists \left(w \in R^*\right) t = (n, rw) \right\}$$

Из сформулированного свойства схемы следует тот факт, что, если представленная в терминах полусхем предметная область является схемой, то все понятия определены корректно, так как имеют хотя бы один образец. Таким образом, процесс построения структурно-логической модели предметной области можно условно разбить на два этапа: первый - описание определений понятий предметной области и связей между ними в терминах полусхем; второй - проверка свойств схемы для построенной полусхемы.

Кроме того, для дальнейшего изложения материала нам потребуется понятие морфизма полусхем, которое более подробно обсуждается в [8].

Определение 7.

Пусть $S_1(N_1,R_1,D_1)$ и $S_2(N_2,R_2,D_2)$ - полусхемы. Морфизмом $\varphi : S_1 \to S_2$ называется пара отображений $\varphi = (\varphi_N, \varphi_R)$ таких, что $\varphi_N : N_1 \to N_2$, $\varphi_R : R_1 \to R_2$ и выполнено соотношение

$$(\forall n \in N_1)(\forall f \in nD_1)(\exists g \in \varphi_N(n)D_2) \quad \varphi_N \circ f = g \circ \varphi_R .$$

Для представления полусхем используется следующая графическая нотацию. Понятия (сущности) предметной области обозначаются прямоугольниками, а варианты определения – окружностями с закрашенным центром.

Приведем пример.

Рассмотрим предметную область, в которой определяется понятие курса (дисциплины). Курс характеризуется названием и количеством часов, а также перечнем тех курсов, которые необходимо изучить предварительно.

Графическое представление соответствующей полусхемы приведено на рис. 1.

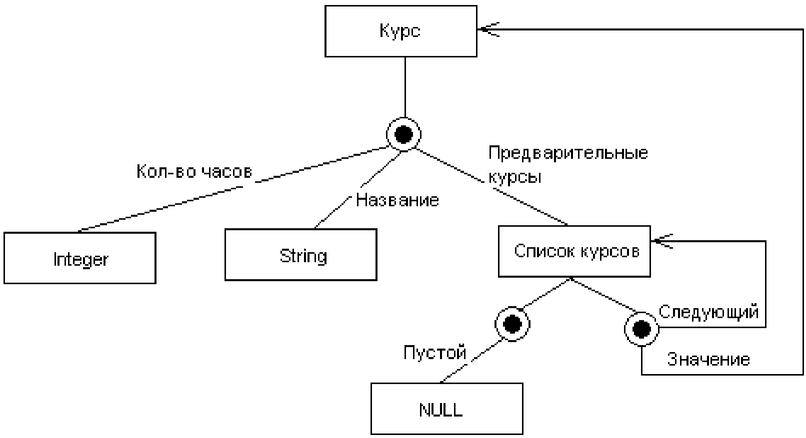

Рис. 1.

2.2. Понятие десигната

Перейдем к дальнейшему исследованию полусхем. Среди множества всех полусхем выделим подмножество таких, которые, упрощенно говоря, служат для определения некоторого единого понятия. Поскольку тройка $S=(N,R,D)$, полученная в результате определения одного понятия, также является полусхемой, это не противоречит общей теории. Такие полусхемы назовем десигнатами некоторого понятия.

Определение 8.

Пусть $S=(N,R,D)$ - полусхема и n,m - понятия из N. Будем говорить, что понятие m используется в определении понятия n, если существует такая нить $t=(n,w)$, что $\tau(t)=m$.

Введем ряд обозначений.

Пусть $N_k=\left\{n\in N \middle| (\exists t=(k,w))\tau(t)=n\right\}$ - множество понятий, использующихся в определении понятия k из N. Очевидно, что $k\in N_k$, так как $\tau(k,\varepsilon)=k$.

Пусть $R_k=\left\{r\in R \middle| (\exists n\in N_k)(\exists f\in M(R,N))(n,f)\in D\wedge r\in \mathrm{dom}(f)\right\}$ - множество ролей, использующихся в определении понятия k из N.

Пусть $D_k=\left\{(n,f)\in D \middle| n\in N_k\right\}$.

Утверждение 1.

Пусть $S=(N,R,D)$ - полусхема. Тройка $S_k=(N_k,R_k,D_k)$, где N_k,R_k,D_k определены выше, является полусхемой.

Определение 9.

Будем говорить, что полусхема $S=(N,R,D)$ является ***десигнатом*** понятия $n \in N$, если $S_n = S$.

Поясним введенное определение на примере.

Пример 1.

Пусть полусхема $S=(N,R,D)$ определена следующим образом:

$$N=\{n,n_1,n_2,n_3,n_4\}$$

$$R=\{r_1,r_2,r_3,r_4\}$$

$$D=\{(n,f),(n_2,g)\},$$

где функции $f,g \in M(R,N)$ определены следующим образом:

$$f(r_1)=n_1,\ f(r_2)=n_2,\ \ g(r_3)=n_3,\ g(r_4)=n_4.$$

Графическое представление полусхемы $S=(N,R,D)$ приведено на рис. 2.

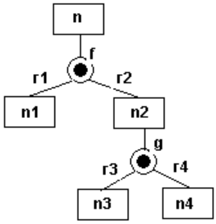

Рис. 2.

Приведенная полусхема S является десигнатом понятия n, так как $S_n = S$. С другой стороны, полусхема S не является десигнатом понятия n_2, поскольку $N_{n_2} = \{n_2,n_3,n_4\}$, а следовательно $N_{n_2} \neq N$.

Отметим, что полусхема может являться десигнатом сразу относительно нескольких понятий.

2.3. Операции на полусхемах

Для дальнейшего исследования необходимо ввести ряд операций.

Определение 10.

Пусть $S_1=(N_1,R_1,D_1)$ и $S_2=(N_2,R_2,D_2)$ - полусхемы. Будем говорить, что полусхема S_2 слабо включает полусхему S_1 (обозначение $S_1 \subset S_2$), если выполнены следующие условия:

1) $N_1 \subset N_2$, $R_1 \subset R_2$

2) существует морфизм $\varphi : S_1 \to S_2$ такой, что

$\varphi = \varphi(i_N, i_R)$, $i_N : N_1 \to N_2$, $i_R : R_1 \to R_2$

Поясним приведенное определение на примере.

Пример 2.

Пусть полусхема $S_1 = (N_1, R_1, D_1)$ определена следующим образом:

$N_1 = \{n, n_3, n_4\}$,

$R_1 = \{r_3, r_4\}$,

$D = \{(n, g)\}$,

где функция $g \in M(R, N)$ определены следующим образом: $g(r_3) = n_3$, $g(r_4) = n_4$.

Графическое представление полусхемы $S_1 = (N_1, R_1, D_1)$ приведено на рис. 3.

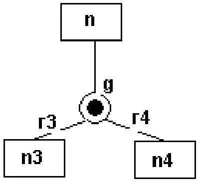

Рис. 3.

Пусть полусхема $S_2 = (N_2, R_2, D_2)$ определена следующим образом:

$N_1 = \{n, n_1, n_2, n_3, n_4\}$

$R_1 = \{r_1, r_2, r_3, r_4\}$

$D = \{(n, f), (n, g)\}$,

где функции $f, g \in M(R, N)$ определены следующим образом: $f(r_1) = n_1$, $f(r_2) = n_2$, $g(r_3) = n_3$, $g(r_4) = n_4$. Графическое представление полусхемы $S_2 = (N_2, R_2, D_2)$ приведено на рис. 4.

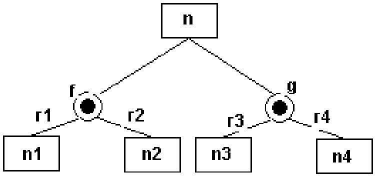

Рис. 4.

Полусхема S_2 слабо включает полусхему S_1.

Как следует из примера, слабое включение в одну сторону не влечет за собой слабого включения в обратную сторону.

Замечание

Пусть $S_1 = (N_1, R_1, D_1)$ - полусхема и $S_2 = (N_2, R_2, D_2)$ - схема. При этом S_2 слабо включает S_1. Тогда S_1 не обязательно является схемой.

Определение 11.

Пусть $S_1 = (N_1, R_1, D_1)$ и $S_2 = (N_2, R_2, D_2)$ - полусхемы. Будем говорить, что полусхема S_2 сильно включает полусхему S_1 (обозначение $S_1 \prec S_2$), если выполнены следующие условия:

1) $S_1 \subset S_2$

2) $(\forall n \in N_1)(\forall f : (n, f) \in D_2)(n, f) \in D_1$

Для полусхем S_1 и S_2 из Примера 2 сильное включение не выполнено, поскольку для понятия $n \in N_1$ (а значит $n \in N_2$) в полусхеме S_2 имеется вариант определения f, т.е. $(n,f) \in D_2$, но при этом $(n,f) \notin D_1$.

<u>Утверждение 2.</u>

Пусть $S_1 = (N_1, R_1, D_1)$ - полусхема и $S_2 = (N_2, R_2, D_2)$ - схема. Если $S_1 \prec S_2$, то тогда S_1 также является схемой.

<u>Определение 12.</u>

Пусть $S_1 = (N_1, R_1, D_1)$ и $S_2 = (N_2, R_2, D_2)$ - полусхемы. Полусхема $S = (N, R, D)$ является пересечением полусхем S_1 и S_2 (обозначение $S = S_1 \bigcap S_2$), если $S \prec S_1$ и $S \prec S_2$ и S является максимальной в том смысле, что не существует такой полусхемы $\tilde{S}$, что $S \prec \tilde{S}$ и каждая из полусхем S_1 и S_2 сильно включает $\tilde{S}$.

<u>Определение 13.</u>

Пусть $S_1 = (N_1, R_1, D_1)$ и $S_2 = (N_2, R_2, D_2)$ - полусхемы и $\varphi(\varphi_N, \varphi_R)$ - морфизм из S_1 в S_2. Образом морфизма φ будем называть тройку $\tilde{S} = (\tilde{N}, \tilde{R}, \tilde{D})$ такую, что:

1) $\tilde{N} = \varphi_N(N_1)$, $\tilde{R} = \varphi_R(R_1)$

2) для каждого $n \in \tilde{N}$ выполнено

$$n\tilde{D} = \left\{ g \in M(\tilde{R}, \tilde{N}) \middle| (\exists n \in N_1)(\exists f \in nD_1)\, \varphi_N \circ f = g \circ \varphi_R \right\}$$

2.4. Шаблонная полусхема.

<u>Определение 14.</u>

Пусть $S_1 = (N_1, R_1, D_1)$ и $S_2 = (N_2, R_2, D_2)$ - полусхемы. Будем говорить, что S_2 является ***встроенным шаблоном*** для S_1, если выполнены следующие условия:

1) существует морфизм $\varphi : S_2 \to S_1$

2) S_2 является концептом

3) $\varphi(S_2) \prec S_1$

<u>Пример 3.</u>

Пусть полусхема S_1 задана графически, как показано на Рис. 5.

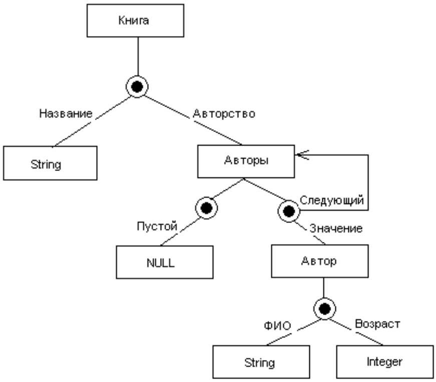

Рис.5.

Полусхема S_2, представленная на Рис. 6, является встроенным шаблоном для S_1.

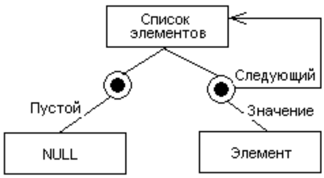

Рис. 6.

Приведем определение соответствующего морфизма $\varphi : S_2 \to S_1$:

$\varphi_N(\text{Список\_элементов}) = \text{Авторы}$, $\varphi_N(\text{Null}) = \text{Null}$,

$\varphi_N(\text{Элемент}) = \text{Автор}$, $\varphi_R(\text{Пустой}) = \text{Пустой}$,

$\varphi_R(\text{Следующий}) = \text{Следующий}$, $\varphi_R(\text{Значение}) = \text{Значение}$

Следует отметить, что список элементов часто встречается в определении понятий, а потому выделение его как шаблона позволит оптимизировать процесс проектирования.

Утверждение 4.

Если S_2 - шаблонная полусхема для S_1, а S_3 - шаблонная полусхема для S_2, то S_3 является шаблонной полусхемой и для S_1.

Обозначим через $T(S)$ множество шаблонных полусхем для S.

Очевидно, что в силу конечности множеств N, R, D множество $T(S)$ является конечным.

Можно провести факторизацию множества $T(S)$ относительно отношения изоморфности полусхем Q. Тогда $T(S)/Q$ - фактор множество, а $[\tilde{S}] \in T(S)/Q$ - представитель класса шаблонных полусхем.

3. Выводы.

Выделение набора шаблонных полусхем для заданной предметной области дает возможность выделить типичные для нее определяющие структуры. Это позволяет объективировать шаблонные конструкции, свойственные данной предметной области, что может стать полезным при дальнейшем проектировании и разработке информационной системы.

Шаблонные полусхемы могут применяться также на этапе проектирования исходной полусхемы. Опыт показывает, что существует набор стандартных конструкций (шаблонов), которые используются в большинстве предметных областей. К таким шаблонным конструкциям относится, например, список понятий. Определение списка в терминах полусхем является несложной, но трудоемкой задачей, если в определении понятий список будет участвовать достаточно часто. Решением указанной проблемы может стать выделение списка в качестве шаблонной полусхемы.

Еще одним важным применением шаблонных полусхем может стать определение шаблонных конструкций самим проектировщиком. Это особенно актуально, если в предметной области используются типичные конструкции для определения большого числа понятий. Следует отметить, что выделение шаблонных полусхем еще на этапе проектирования, позволяет разработчику структурировать свои знания о предметной области, что является несомненным достоинством и дает преимущества на последующих этапах проектирования и разработки информационной системы.

Список литературы: 1. Жолткевич Г.Н., Семенова Т.В. К проблеме формализации концептуального моделирования информационных систем // Вісник Харківського національного університету. Серія «Математичне моделювання. Інформаційні технології. Автоматизовані системи управління», №605, 2003, с. 33-42. 2. Жолткевич Г.Н., Семенова Т.В., Федорченко К.А. Представление полусхем предметных областей информационных систем средствами реляционных баз данных // Вісник Харківського національного університету. Серія «Математичне моделювання. Інформаційні технології. Автоматизовані системи управління», №629, 2004, с. 11-24. 3. Буч. Г., Якобсон А., Рамбо Дж. UML. Классика CS. 2-е изд. / Пер. с англ.; Под общей редакцией проф. С. Орлова – СПб.: Питер, 2006. – 736 с. 4. Сомервилл, Ифн. Инженерия программного обеспечения, 6-у издание. : Пер. с англ. – М.: Издательский дом «Вильямс», 2002. – 624с.: ил. 5. Советов Б.Я., Яковлев С.А. Моделирование систем. – М.: Высш. шк., 1985. – 271 с., ил. 6. Боггс У., Боггс М. UML и Rational Rose. - М.: "ЛОРИ", 2000. - 582 с. 7. Кватрани, Терри Rational Rose 2000 и UML. Визуальное моделирование : перевод с : англ. / Терри Кватрани ; Предисл. Грейди Буч ; перевод А.Б. Литвин . - М. : ДМК Пресс, 2001 . - 176 с. 8. Семенова Т.В. Морфизмы полусхем и их приложения. // Вісник Харківського національного університету. Серія «Математичне моделювання. Інформаційні технології. Автоматизовані системи управління», №703, 2005, с. 198-206.

Поступила в редколлегию 03.04.06