

РАЗРАБОТКА КОМПОНЕНТОВ АИС НА КОНЦЕПТУАЛЬНОМ УРОВНЕ

В.В. Бойко

Любая информационная система (ИС) характеризуется тремя компонентами: носителем, в состав которого входят технические средства системы; информационным фондом, представляющим собой совокупность информационных объектов и отношений; процедурами, реализующими процессы накопления, поиска, обновления и обработки информации. Носитель называют физическим компонентом, информационный фонд - информационным компонентом, процедуры - функциональным компонентом системы [1].

Одним из подклассов ИС являются автоматизированные информационные системы (АИС), построенные на основе персональных компьютеров, базы данных и программного обеспечения управления данными. Между компонентами таких АИС существует тесная взаимосвязь и взаимозависимость. Предполагаемые информационные процессы определяют структуру базы данных. Вся совокупность процедур манипулирования данными, а также технические возможности ЭВМ влияют на организацию пользовательского интерфейса и эффективность реализации различных способов отображения меню, форм, отчетов.

Зависимость компонентов АИС создает проблему реструктуризации системы при изменении информационных потребностей пользователей.

В рамках традиционных технологий, ориентированных на, так называемые, логические модели данных и коммерческие системы управления базами данных (СУБД), внесение изменений в процессы поиска и модификации данных влечет за собой повторное программирование, компиляцию и тестирование компонентов АИС. Это приводит к длительным срокам перепроектирования системы.

Сократить затраты можно, используя технологии и инструментальные средства семантического моделирования данных.

Семантические модели обладают большими выразительными возможностями для представления знаний о данных и, вместе с тем, достаточно просты для понимания как разработчиками, так и пользователями АИС [3]. Они позволяют моделировать данные на концептуальном уровне, независимом от уровня СУБД.

Математически хорошо определенные правила описания семантических структур способствуют созданию универсальных процедур манипулирования объектами базы данных. Включение этих процедур в состав АИС повышает гибкость и надежность системы, сокращает сроки разработки.

Задача моделирования данных на концептуальном уровне применительно к проектированию АИС может быть сформулирована следующим образом: при условии, что известна информация для хранения и требования к ее обработке, определить статические и динамические свойства информационного компонента системы, обеспечивающие выполнение ожидаемых информационных процессов

Под статическими свойствами имеются ввиду закономерности, которым подчиняются состояния базы данных АИС. Под динамическими свойствами - закономерности, описывающие переходы из состояния в состояние. Статические свойства находят свое выражение в схеме базы данных. Динамические свойства определяются в спецификациях запросов и действий. Поиск по запросу и выполнение действий - неотъемлемые части большинства информационных процессов АИС.

Анализ семантических моделей показывает, что большинство из них хорошо приспособлено для моделирования статических свойств АИС. Для моделирования динамических свойств на концептуальном уровне, на наш взгляд, предпочтительней расширенная семантическая иерархическая модель (SHM+) [2] и модель Смита [3]. Заложенные в них идеи определения схем и действий были применены и развиты нами при разработке семантической модели и технологии АСПЕКТ.

С целью преодоления сложности проекта в данной технологии используются два принципа: принцип абстракции и принцип локализации. Принцип абстракции состоит в подавлении несущественных деталей для того, чтобы выделить наиболее уместные, которые могут быть рассмотрены и именованы как целое. Принцип локализации заставляет проектировщика моделировать каждое свойство независимо (локализовано) и затем объединять свойства для выполнения полного проекта.

Семантическая модель технологии АСПЕКТ предоставляет для описания структуры базы данных два ключевых понятия: объект и тип объектов, а также

четыре формы абстракции данных: агрегация, обобщение, специализация и унификация.

Объект - первичное, неопределяемое понятие, синоним понятий "вещь", "реалия", "сущность". Термин "объект" всегда ассоциируется с различимостью вещей. Поэтому предполагается, что каждый объект имеет имя. Совокупность объектов, обладающих общими свойствами, обозначается понятием "тип".

Агрегацией называется форма абстракции, в которой отношение между объектами рассматривается как объект более высокого уровня абстракции. Агрегация, таким образом, устанавливает связь "компонент агрегированного объекта" между объектами разных уровней абстракции. Объекты, не относящиеся к агрегированным (агрегатам), называются простыми. В результате агрегации устанавливается функциональная зависимость кортежа объектов - компонентов от агрегированного объекта и агрегированного объекта от совокупности объектов - компонентов. Эта совокупность называется идентификатором.

Специализация - форма абстракции, позволяющая рассматривать каждый объект подтипа как объект супертипа. Специализация определяет отношение "является подтипом" между двумя типами объектов.

Обобщение - форма абстракции, которая также устанавливает отношение "является подтипом", но в отличие от специализации образует тип как результат объединения подтипов.

И последняя форма абстракции - унификация. Унификация означает отвлечение от индивидуальных различий и выявление идентичных свойств у объектов нескольких типов. Унификация устанавливает отношение эквивалентности на множестве типов.

Формируемое агрегацией отношение "компонент агрегированного объекта" обладает всеми свойствами отношения строго частичного порядка, а значит, определяет на множестве объектов ациклический ориентированный граф. В теории баз данных такой граф называют иерархией [4]. Отношение "является подтипом", которое создается в результате обобщения и специализации, также представляет собой отношение строго частичного порядка. Это отношение определяет иерархию типов. Унификация образует отношение эквивалентности, а значит, и классы эквивалентности на множестве типов.

Структура, состоящая из множества объектов, множества типов объектов и отношения "компонент агрегированного объекта", называется в технологии АСПЕКТ структурой семантической базы данных.

Каждая форма абстракции определяет ограничения целостности, которым должна удовлетворять структура базы данных. Агрегация устанавливает следующие ограничения:

- для каждого агрегированного объекта в структуре семантической базы данных должны существовать все его компоненты;
- агрегированному объекту должен соответствовать единственный набор компонентов;
- компонентам, входящим в идентификатор, должен соответствовать единственный агрегированный объект.

Специализация вносит ограничение на формирование подтипов и супертипов: каждый объект подтипа должен быть объектом супертипа.

Сохранение семантики обобщения требует кроме того, чтобы каждый объект супертипа был объектом одного из подтипов (категорий обобщения).

Унификация создает ограничение на вхождение объектов в некоторую совокупность типов: каждый объект одного типа должен быть объектом другого типа.

Для описания структуры базы данных на концептуальном уровне в технологии АСПЕКТ используются три вида схем.

Множество упорядоченных пар, состоящих из имени типа агрегированного объекта и имени типа компонента, названо схемой иерархии объектов. Локальной схемой иерархии объектов названо подмножество упорядоченных пар, определяющее дерево - частный случай иерархии. Аналогично, множество упорядоченных пар имен супертипов и подтипов, связанных обобщением, получило название "схема иерархии типов", а их подмножество, определяющее дерево, - "локальная схема иерархии типов". Множество пар имен типов, связанных унификацией, названо схемой класса эквивалентных типов.

Локальные схемы иерархии объектов, локальные схемы иерархии типов и схемы классов эквивалентных типов иначе называются фрагментами схемы семантической базы данных. Локальная схема иерархии объектов, характеризующая связь агрегированного объекта с его компонентами, получила название

"аспект". Каждому аспекту всегда соответствует иерархия объектов одной агрегации.

Интеграция фрагментов в общую схему базы данных производится по правилам алгоритма, основанного на принципе проверки согласованности семантики аспектов. Подробно алгоритм изложен в [5]. Правила построения фрагментов рассматривались в [6]. Процедуры редактирования и согласования фрагментов позволяют определить все структурные (статические) свойства информационного компонента АИС. Примеры фрагментов схемы показаны на рис. 1.

Для описания динамических свойств модель предоставляет примитивные операции над объектами, две формы абстракции управления и две формы ограничений на объекты. Считается, что каждая операция базы данных выполняется над одним объектом. Операции разделены на две категории: операции преобразования и операции поиска. К операциям преобразования отнесены: INSERT (ввести), DELETE (удалить), UPDATE (обновить). К операциям поиска отнесены: FIND (найти объект в базе данных), CREATE (создать объект, используя некоторую процедуру), REQUEST (запросить объект у пользователя).

В технологии АСПЕКТ построение схемы базы данных всегда предшествует разработке спецификаций. Спецификация запроса составляется из схемы запроса и ограничений на объекты. Схема запроса, в общем случае, представляет собой несколько трансформированных локальных схем иерархии объектов - фрагментов запроса. Для получения необходимого фрагмента запроса используются операторы "удалить тип" и "выделить тип":

- Удалить тип(имя фрагмента, имя типа);
- Выделить тип(имя фрагмента, имя типа).

Первый оператор удаляет из фрагмента имя, указанное пользователем. Если оно является корневой вершиной, то удаляется и ассоциированная с ним локальная схема. Второй оператор удаляет все имена, не входящие в схему с корневой вершиной, указанной в операторе. Для каждого фрагмента запроса могут быть определены ограничения. Для ввода ограничений на простые объекты используется оператор

- Добавить ограничение(имя фрагмента, ограничение простых объектов).

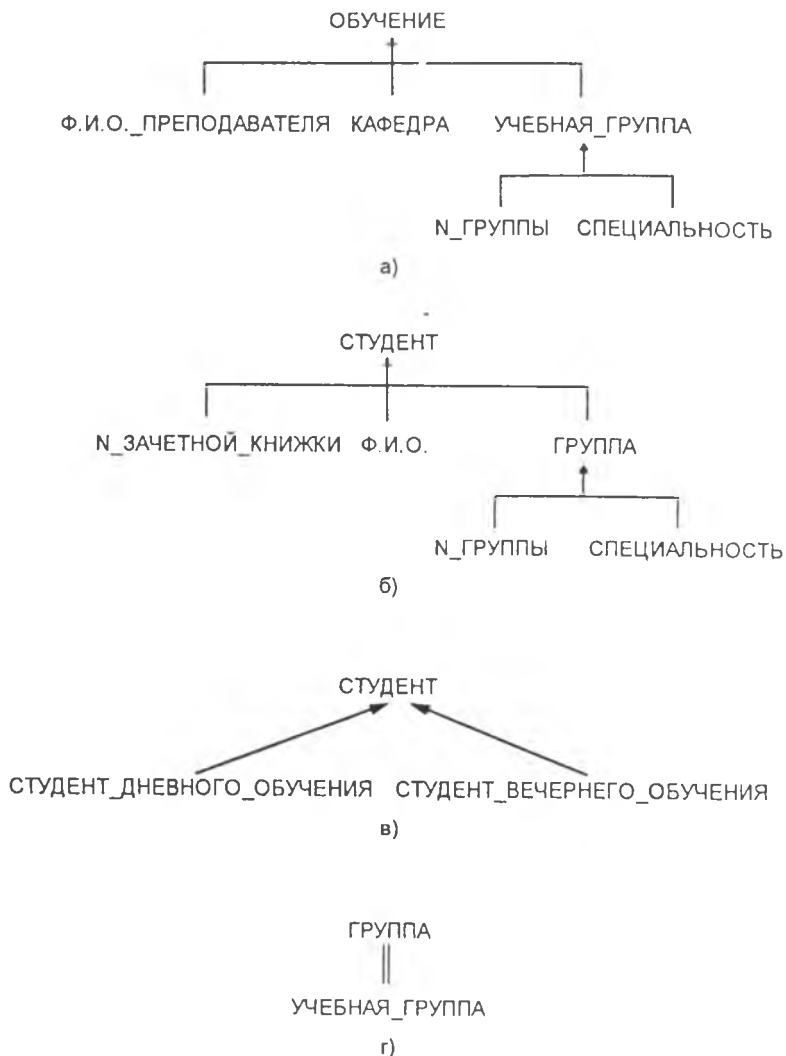


Рис. 1. Фрагменты схемы базы данных "ВУЗ"

а, б - локальные схемы иерархии объектов; в - локальная схема иерархии типов
 г - схема класса эквивалентных типов

Второй параметр оператора является булевским выражением, состоящим из имени типа, знака отношения (=, <, >, <=, >=) и границы множества объектов. Граница множества объектов может быть обозначена простым объектом или именем типа простых объектов. Ограничения на агрегаты задаются с помощью оператора

Добавить ограничение(имя целевого фрагмента, ограничение агрегатов, имя контекстного фрагмента). В этом операторе в булевском выражении используется в качестве знака отношения символ "=", а границей является имя типа агрегата.

В спецификации запроса только в одном фрагменте могут находиться имена типов простых объектов, отмеченные меткой операции FIND. Эта метка указывает на объекты поиска. Поскольку поиск объектов составляет цель запроса, соответствующий фрагмент назван целевым. Наряду с целевым в спецификацию, содержащую ограничения на агрегаты, входят, так называемые, контекстные фрагменты. Они включаются в спецификацию для определения границы множества агрегатов, введенного в целевом фрагменте. Запрос, заданный спецификацией, интерпретируется, в конечном счете, как логическое выражение, определенное на объектах схемы, ограничениях на объекты и отношении" компонент агрегированного объекта". Для примера рассмотрим следующий запрос: получить фамилии преподавателей, работающих с группами студентов дневного обучения специальности "АИС". Исходная схема базы данных приведена на рис. 1.

На первой стадии разработки спецификации выберем фрагмент ОБУЧЕНИЕ с объектом поиска Ф.И.О._ПРЕПОДАВАТЕЛЯ. Затем удалим имена КАФЕДРА, N_ГРУППЫ с помощью операторов:

Удалить тип(ОБУЧЕНИЕ, КАФЕДРА);

Удалить тип(ОБУЧЕНИЕ, N_ГРУППЫ);

Введем ограничение для объектов УЧЕБНАЯ_ГРУППА:

Добавить ограничение(ОБУЧЕНИЕ, УЧЕБНАЯ_ГРУППА = ГРУППА,
СТУДЕНТ_ДНЕВНОГО_ОБУЧЕНИЯ).

Введем ограничение для объектов контекстного фрагмента:

Добавить ограничение(СТУДЕНТ_ДНЕВНОГО_ОБУЧЕНИЯ,
СПЕЦИАЛЬНОСТЬ = 'АИС').

Из контекстного фрагмента удаляем имена, не относящиеся к запросу:

Удалить тип(СТУДЕНТ_ДНЕВНОГО_ОБУЧЕНИЯ. N_ЗАЧЕТНОЙ_КНИЖКИ);

Удалить тип(СТУДЕНТ_ДНЕВНОГО_ОБУЧЕНИЯ. Ф.И.О.);

Удалить тип(СТУДЕНТ_ДНЕВНОГО_ОБУЧЕНИЯ. N_ГРУППЫ).

Из целевого фрагмента удаляем имя СПЕЦИАЛЬНОСТЬ, поскольку ограничение на объекты этого типа определено в контекстном фрагменте:

Удалить тип(ОБУЧЕНИЕ. СПЕЦИАЛЬНОСТЬ).

В итоге получаем схему запроса, изображенную на рис.2.

Важная особенность модели АСПЕКТ, а также моделей Смита [4] и SHM+ [3] заключается в поддержке семантического релятивизма. В соответствии с принципом семантического релятивизма любой объект может рассматриваться как независимо, так и в терминах связи, в которой он участвует (компонент или категория). Это свойство используется в технологии АСПЕКТ для спецификации действий над объектами базы данных.

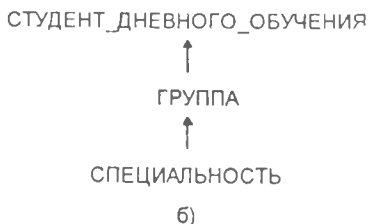
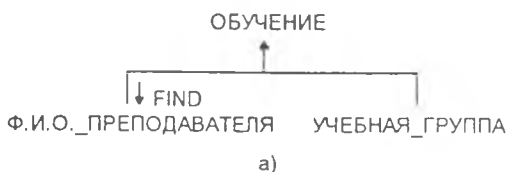


Рис.2. Схема запроса

а - целевой фрагмент; б - контекстный фрагмент.

Полная спецификация действия образуется из описания схемы действия и спецификаций пред- и постусловий. Действие - есть проблемно-ориентированная операция создания или удаления одного агрегированного объекта с соблюдением всех семантических ограничений. Для связывания действий с другими операциями используются абстракции управления. Две формы абстракции управления, - последовательность и выбор, являются поведенческими аналогами двух форм абстракции данных - агрегации и обобщения. Агрегации соответствует либо последовательное, либо параллельное связывание операций. Поэтому действие над агрегатом вызывает последовательное или параллельное выполнение операций (по одной операции для каждого компонента). Действие над объектом обобщения вызывает выполнение операции над объектом одной или нескольких категорий в зависимости от принадлежности объекта обобщения той или иной категории (конструкция if - then).

Синтаксис имен операций, вызываемых при выполнении действий, задается следующим грамматическим правилом:

op ::= INSERT | UPDATE | DELETE | FIND | REQUEST | CREATE |
 or V or | or & or,

где "&" обозначает связку "и" (последовательность или параллельность), "V" обозначает "или" (выбор).

Графической формой описания поведенческих свойств проблемно-ориентированной операции в технологии АСПЕКТ является схема действия. Схема действия представляет собой аспект с нанесенными на каждой дуге метками операций (по одной метке для каждой частичной операции). Метка операции - это имя вызываемой операции плюс символ "->" (стрелка), задающий контекст вызова. Стрелка указывает на объект, для которого операция вызывается, и направлена от объекта, вызывающего его. Двойная стрелка "=>" используется, чтобы отличить действие от составляющих операций. При этом действия и операции поиска всегда адресованы объектам, которые в соответствии с принципом семантического релятивизма рассматриваются независимо, а операции преобразования всегда обращены к объектам, участвующим в связи "компонент агрегированного объекта". Состояние базы данных до и после выполнения действия характеризуют спецификации пред- и постусловий. Предусловие есть логическое выра-

жение, которое должно выполняться в структуре базы данных до вызова действия. Постусловие есть логическое выражение, которое выполняется в структуре базы данных после успешного завершения действия. Поэтому спецификация предусловия (постусловия), подобно спецификации запроса, включает схему предусловия (постусловия) и ограничения.

Рассмотрим этапы определения спецификации на примере проблемно-ориентированной операции зачисления студента в конкретную группу (N 631). Это действие моделируется с помощью операции INSERT (ввести). Схема действия показана на рис.3,а.

Спецификация предусловия включает в себя схему предусловия (рис.3,б) и ограничение, заданное оператором

Добавить ограничение(ГРУППА, N_ГРУППЫ = 631).

Спецификация постусловия полностью определяется по завершению действия. Схема постусловия для данной операции приведена на рис.3,в. Ограничения, которым должны удовлетворять объекты схемы постусловия, относятся к N_ГРУППЫ, N_ЗАЧЕТНОЙ_КНИЖКИ и Ф.И.О. Ограничение на объекты N_ГРУППЫ определено ранее. Остальные ограничения определяются по результатам запроса объектов у пользователя.

Нетрудно заметить, что множество упорядоченных пар <предусловие, постусловие>, характеризующих действия над некоторой иерархией объектов, образует отношение строго частичного порядка и, следовательно, определяет распространение действий (основанных на операциях INSERT или DELETE) по иерархии объектов. Что касается операции UPDATE, то она применима только к простым объектам. Агрегаты можно вводить в базу данных и удалять. Изменять допускается только компоненты агрегатов.

Выполнение действия над агрегатом гарантирует соблюдение всех семантических ограничений целостности. В нашем примере, понятие СТУДЕНТ - обобщение категорий: СТУДЕНТ_ДНЕВНОГО_ОБУЧЕНИЯ и СТУДЕНТ_ВЕЧЕРНЕГО_ОБУЧЕНИЯ. Поэтому, ввод объекта для обобщения СТУДЕНТ вызывает ввод такого же объекта для одной из категорий. В данном случае, - для категории СТУДЕНТ_ДНЕВНОГО_ОБУЧЕНИЯ.

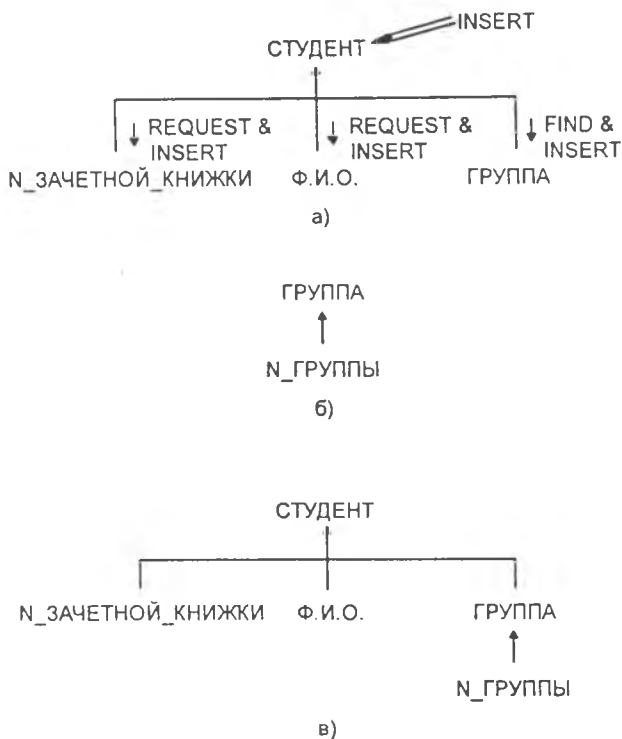


Рис. 3. Схемы, входящие в спецификацию проблемно-ориентированной операции
 а - схема действия; б - схема предусловия; в - схема постусловия.

На основе модели и технологии АСПЕКТ созданы инструментальные средства разработки компонентов АИС на концептуальном уровне. Инструментальные средства состоят из двух комплексов. Первый назван комплексом проектирования АИС, второй - комплексом реализации. В комплекс проектирования входят программы-дизайнеры схем, запросов и действий. В комплекс реализации включены процессоры схем, запросов и действий, а также генераторы стандартных форм и отчетов. Разработка компонентов АИС с их помощью заключается в проектирова-

нии спецификаций и связывании спецификаций с соответствующим процессором. Инструментальные средства реализованы на ПЭВМ IBM/PC.

ЛИТЕРАТУРА

1. Автоматизированные информационные системы. Криницкий Н.А., Мионов Г.А., Фролов Г.Д. / Под ред. А.А. Дородницына. - М.: Наука. Главная редакция физико-математической литературы, 1982.-384с.
2. Hull R., King R. Symantic database modelling: survey, application and reseach issues // ACM Computer Survey. - 1987. - Vol.19, No.3. - pp. 201 - 260.
3. Brodie M.L., Ridjanovic D. On the Design and Specification of Database Transactions // On Conceptual Modelling / Brodie, Mylopoulos and Schmidt, Eds. Springer Verlag, New York. - 1984. - pp. 277 - 311.
4. Смит Дж., Смит Д. Принципы концептуального проектирования баз данных // Требования и спецификации в разработке программ. Сборник статей. Пер. с англ. - М.: Мир. - 1984. - с. 165 -198.
5. Бойко В.В. Интеграция представлений пользователей о семантической структуре базы данных в технологии проектирования "АСПЕКТ" // Тезисы докладов научно-технической конференции "Перспективные информационные технологии в Высшей школе", Самара, сентябрь 1993г. - Самара: НПЦ "Авиатор". -1993. - с. 38.
6. Бойко В.В. Оптимизация концептуальной схемы в процессе проектирования реляционной базы данных // Тезисы докладов научно-практической конференции "Проблемы информатики", Самара, 11 - 18 мая 1991г. - Тольятти: ПП "Современник". 1991. - с. 24 -25.