

УДК 621.142

© Д. В. Белоус, Ю. М. Шерстюк

ПОСТРОЕНИЕ МНОГОЦЕЛЕВЫХ АНАЛИТИЧЕСКИХ ИНФОРМАЦИОННЫХ СИСТЕМ С ЭЛЕМЕНТАМИ МЕТАУПРАВЛЕНИЯ

Основной проблемой создания многоцелевых аналитических информационных систем является обеспечение корпоративной деятельности множества экспертов с адаптацией используемых инструментов к классам решаемых задач и предметным областям. Описан подход к решению указанной проблемы, основанный на многоуровневой декомпозиции структуры и профилей многоцелевых аналитических информационных систем и базирующийся на применении метода метауправления, позволяющего внести элементы синтаксической вариантности.

ВВЕДЕНИЕ

Создание и внедрение многоцелевых аналитических информационных систем (МАИС), основанных на применении технологий интеллектуального анализа данных (ИАД), является одним из перспективных подходов к созданию высокоэффективных систем поддержки принятия решений на основе применения современных информационных технологий, математического аппарата и средств вычислительной техники [1].

Основными задачами, решение которых возлагается на МАИС, являются:

— оценка текущего и прогнозируемого состояния объекта управления и среды его функционирования;

— обнаружение и исследование скрытых закономерностей, факторов, тенденций и взаимосвязей;

— агрегация и интеграция информации, необходимой для обоснования и принятия решений;

— формирование альтернативных решений и поддержка выбора "оптимального" в соответствии с заданным критерием и результатами анализа сценариев развития ситуаций;

— моделирование процесса эволюции состояния объекта в нестационарной неоднородной среде и т.д.

Основными принципами построения подобных аналитических информационных систем, известными из мировой практики, являются:

— применение высокопроизводительных архитектур и современных платформ для создания вычислительной среды;

— ориентация на динамические данные, позволяющие строить временные ряды и осуществлять как ретроспективный анализ характеристик объектов, процессов и явлений, так и решение за-

дач оперативного и долгосрочного прогнозирования с выявлением неявных зависимостей и скрытых влияний (воздействий);

— использование множества современных, в том числе математических, методов различного направления, позволяющих эффективно решать задачи ИАД.

Особо следует отметить следующие обстоятельства.

Содержательное наполнение аналитических информационных систем в большой степени зависит от специфики предметной области и объекта управления, однако формальных методов его определения на сегодняшний день не существует, и выбор конкретных математических методов и обрабатываемых данных осуществляется исключительно эвристически с последующим их уточнением в ходе применения системы.

Направление исследований, связанное с созданием аналитических информационных систем, еще находится в начальной стадии своего развития. Оно определилось и бурно развивается с середины 90-х годов, поэтому в настоящее время имеются только лишь отдельные образцы соответствующих инструментальных средств и некоторый опыт их создания и применения.

Практически отсутствует отечественный опыт разработок в данной области.

Имеющиеся коммерческие образцы систем ИАД (изделия Microsoft, Oracle, SAS Institute и др.) при своей достаточно высокой стоимости имеют весьма существенный недостаток — сочетание фиксированного набора средств математического инструментария с частными программно реализованными методиками решения задач анализа. Это приводит как к существенному ограничению областей применения указанных систем и решаемых задач управления, так и к высокому риску неверного выбора потенциальными пользо-

вателями инструментария, адекватного решаемым ими задачам. В некоторой степени можно провести аналогию задачи выбора инструмента ИАД с задачей выбора языка программирования, свойства которого (уровень, надежность, функциональная полнота) оказывают существенное влияние на качество разрабатываемого программного продукта. Однако если для языков программирования широчайший опыт их использования позволяет формулировать некоторые практические рекомендации с позиций их применимости в проектах различного характера и назначения, для систем ИАД подобный подход весьма проблематичен вследствие слабой насыщенности рынка, отсутствия механизмов анализа опыта применения систем ИАД и уникальности прагматических аспектов их использования. Более того, в современных условиях факт, методики и результаты применения систем ИАД зачастую приобретают статус "ноу хау".

Таким образом, при реализации МАИС основной проблемой становится обеспечение корпоративной деятельности множества экспертов с динамической адаптацией (т.е. непосредственно в процессе эксплуатации) используемых инструментов к классам решаемых задач и предметным областям. При этом каких-либо ограничений на используемые частные методики ИАД, состав математического инструментария, а также задачи и области применения МАИС на момент ее разработки принципиально не накладываются. Последнее порождает потенциальную неоднозначность в проектировании программных приложений МАИС, неразрешимую в рамках традиционных архитектур и технологий создания корпоративных информационных систем.

ОСНОВНЫЕ ПРИНЦИПЫ ФОРМИРОВАНИЯ МАИС

С учетом вышеизложенного, а также результатов анализа подходов к созданию адаптивных программных систем и опыта применения метода метауправления можно сформулировать два базовых принципа системной интеграции, обеспечивающих формирование МАИС нового поколения:

1) многоуровневый подход к синтезу архитектуры и профилей МАИС и формализации процессов ее функционирования;

2) создание МАИС как синтаксически вариантной информационной системы, вариантность которой реализуется на основе метауправления в виде совокупности процессов формирования и обработки ряда компонентов метаинформации.

Результатами применения многоуровневого подхода к синтезу архитектуры и профилей МАИС и формализации процессов ее функционирования являются (рис. 1):

— декомпозиция процессов функционирования МАИС;

— потенциальный состав пользователей МАИС и распределение их возможностей и обязанностей с позиций участия в функционировании МАИС;

— двухступенчатый характер хранения и обработки предметных сведений;

— функциональная, организационная и техническая структуры, архитектура и профили МАИС и соответствующие отношения отображения (структуры — на профили, а профили — на архитектурные компоненты).

ДЕКОМПОЗИЦИЯ ПРОЦЕССОВ ФУНКЦИОНИРОВАНИЯ МАИС

Процессы функционирования МАИС, являющиеся основой для синтеза ее функциональной структуры, определяются в рамках классического выделения управляющего, основного (содержательного) и обеспечивающего процессов, и формализуются как три взаимосвязанных уровня функционирования МАИС:

— стратегический (управляющий): уровень управления, охватываемый предметной областью (Про) и классом потенциально решаемых задач в рамках динамически формулируемых конкретных проблем анализа и прогностики, представляющий собой процесс управления структурой и содержанием хранимых предметных сведений (включая процессы их сбора), а также множеством программно реализованных математических методов, обеспечивающих решение задач ИАД;

— оперативный (основной): уровень отдельных задач в рамках общей решаемой проблемы, на котором по отношению к каждой задаче выполняются этапы ее постановки, анализа и решения;

— технологический (обеспечивающий): уровень функционирования служб МАИС, составляющих ее технологические профили (телекоммуникации, разграничение доступа, защита информации, архивация результатов решения задач и т.д.).

Для оперативного уровня дальнейшая декомпозиция предполагает:

— разбиение процесса решения задачи на ряд стадий, регламентируемых методологией решения задач с использованием МАИС;

— разбиение стадий решения на этапы, в той или иной мере имеющие эвристический характер и определяемые частными методиками их выполнения;

— разбиение этапов на отдельные операции, выполняемые на основе использования соответствующих формальных методов.

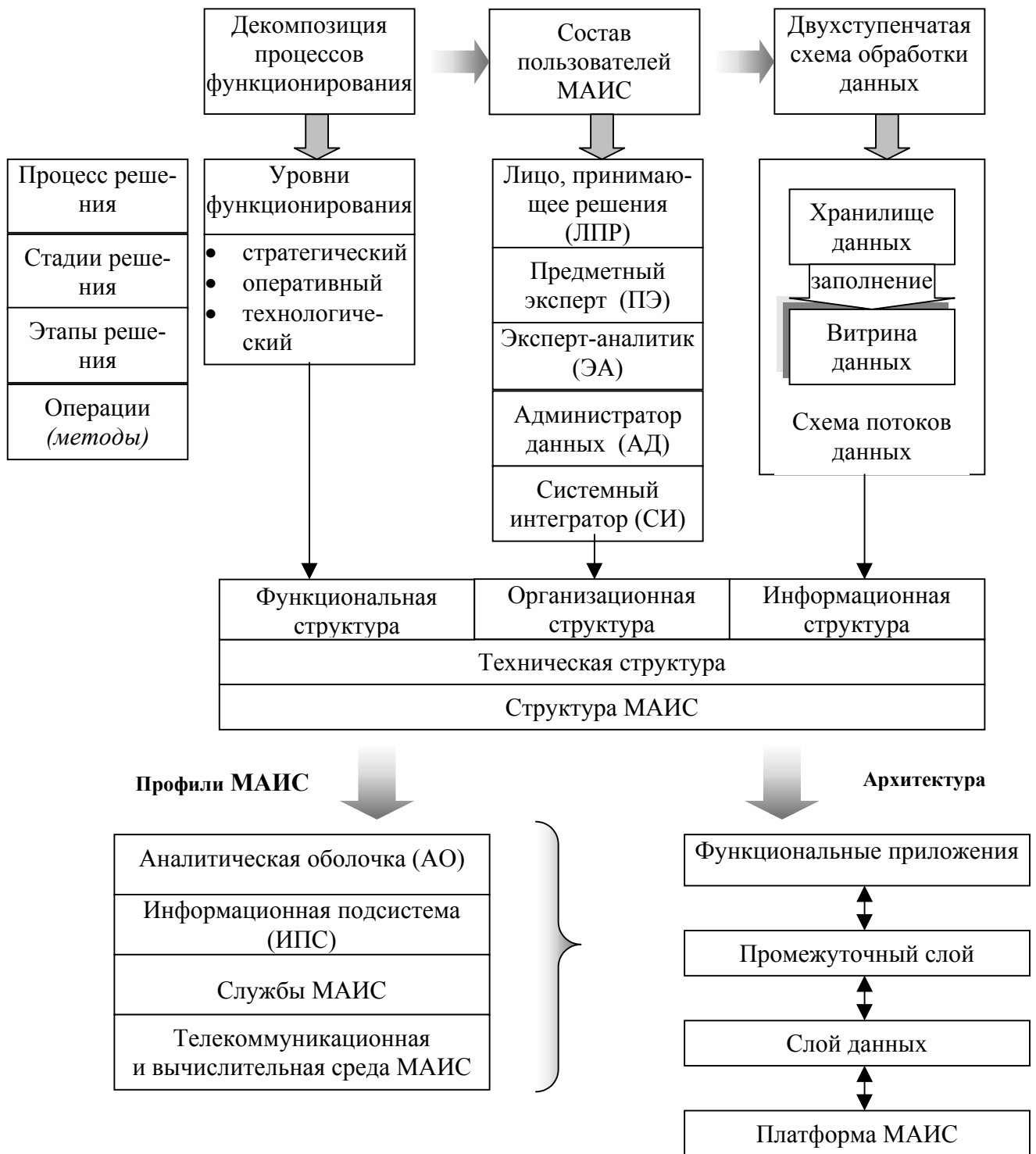


Рис. 1. Объекты и результаты применения многоуровневого подхода

ПОТЕНЦИАЛЬНЫЙ СОСТАВ ПОЛЬЗОВАТЕЛЕЙ

Потенциальный состав пользователей МАИС, являющийся основой для синтеза организационной структуры, определяется как совокупность следующих должностных лиц:

— лицо, принимающее решения (ЛПР): руководитель, осуществляющий постановку проблем и задач и являющийся для МАИС конечным потребителем получаемых результатов;

— предметный эксперт (ПЭ): специалист предметной области, несущий основную нагрузку по определению состава и содержательному анализу предметных сведений а также результатов их обработки;

— эксперт-аналитик (ЭА): специалист, осуществляющий выбор методик решения задач и выполняющий их решение на основе формальных методов обработки;

— администратор данных (АД): лицо, выполняющее действия по сбору необходимых предметных сведений и помещению их в информационное хранилище;

— системный интегратор (СИ): лицо, осуществляющее управление службами МАИС, ее конфигурированием и содержательным наполнением.

Приведенный перечень пользователей достаточно условен, а распределение обязанностей описано весьма поверхностно. Так, например, ПЭ может одновременно выполнять и функции ЛПР, а эксперта-аналитика может и не быть в случае автоматической реализации соответствующих процессов.

ДВЕ СТУПЕНИ ХРАНЕНИЯ И ОБРАБОТКИ

Двухступенчатый характер хранения и обработки предметных сведений следует из специфики функционирования МАИС и предполагает разделение информационного хранилища на две части — хранилище данных и витрины данных. В первом приближении хранилище данных (ХД) предназначено для хранения всех предметных сведений, собранных АД и совместно используемых другими пользователями путем выбора необходимых сведений и копирования их в витрины данных (ВД). Последние представляют собой локальные информационные хранилища, цикл жизни которых ограничен рамками процесса решения одной задачи. В ВД размещаются как исходные предметные сведения, необходимые для решения задачи, так и все результаты, получаемые в процессе ее решения. Совокупность ХД и ВД, а также информационных связей (потоков данных) является основой для синтеза информационной структуры МАИС.

Перечисленные выше результаты проектирования МАИС позволяют синтезировать ее функциональную, организационную, информационную и техническую структуры, основные элементы которых схематично показаны на рис.2. Следует отметить, что имеется положительный опыт формализации описания большинства перечисленных выше проектных решений по созданию МАИС с использованием методологий IDEF0 и IDEF1x.

ПРОФИЛИ И АРХИТЕКТУРА МАИС

Профили МАИС определяются как совокупность:

— функциональных профилей решения задач, интегрируемых аналитической оболочки (АО) МАИС;

— информационных профилей, обеспечивающих адаптацию структуры ХД к конкретным Про, ведение (заполнение) ХД и выборку из него сведений, заносимых в витрины данных — информационная подсистема (ИПС) МАИС;

— профилей служб МАИС: конфигурирования, защиты информации, управления составом разного рода библиотек;

— профилей телекоммуникационной и вычислительной среды МАИС, включая удаленный доступ пользователей как к ресурсам самой МАИС, так и внешним источникам информации с использованием протоколов LAN и WAN.

Архитектура МАИС имеет многоуровневую организацию, основанную на применении архитектуры "клиент—сервер", СОМ-технологии и механизма RPC. В качестве основных архитектурных компонентов МАИС выступают:

- слой приложений (application layer);
- промежуточный слой (middleware);
- слой данных (data layer);
- платформа МАИС.

Слой приложений составляют функциональные приложения специального программного обеспечения МАИС, которые:

— реализуют модели и методы (алгоритмы) аналитической и экспертной обработок;

— образуют содержательное наполнение автоматизированных рабочих мест (АРМ) пользователей.

Промежуточный слой составляют средства доступа к данным и манипулирования ими, образующие интерфейс слоя приложений и слоя данных.

Слой данных (слой хранения) образуют информационные массивы различного содержания и назначения, основу которых составляют общее разделяемое хранилище данных и витрины данных приложений.

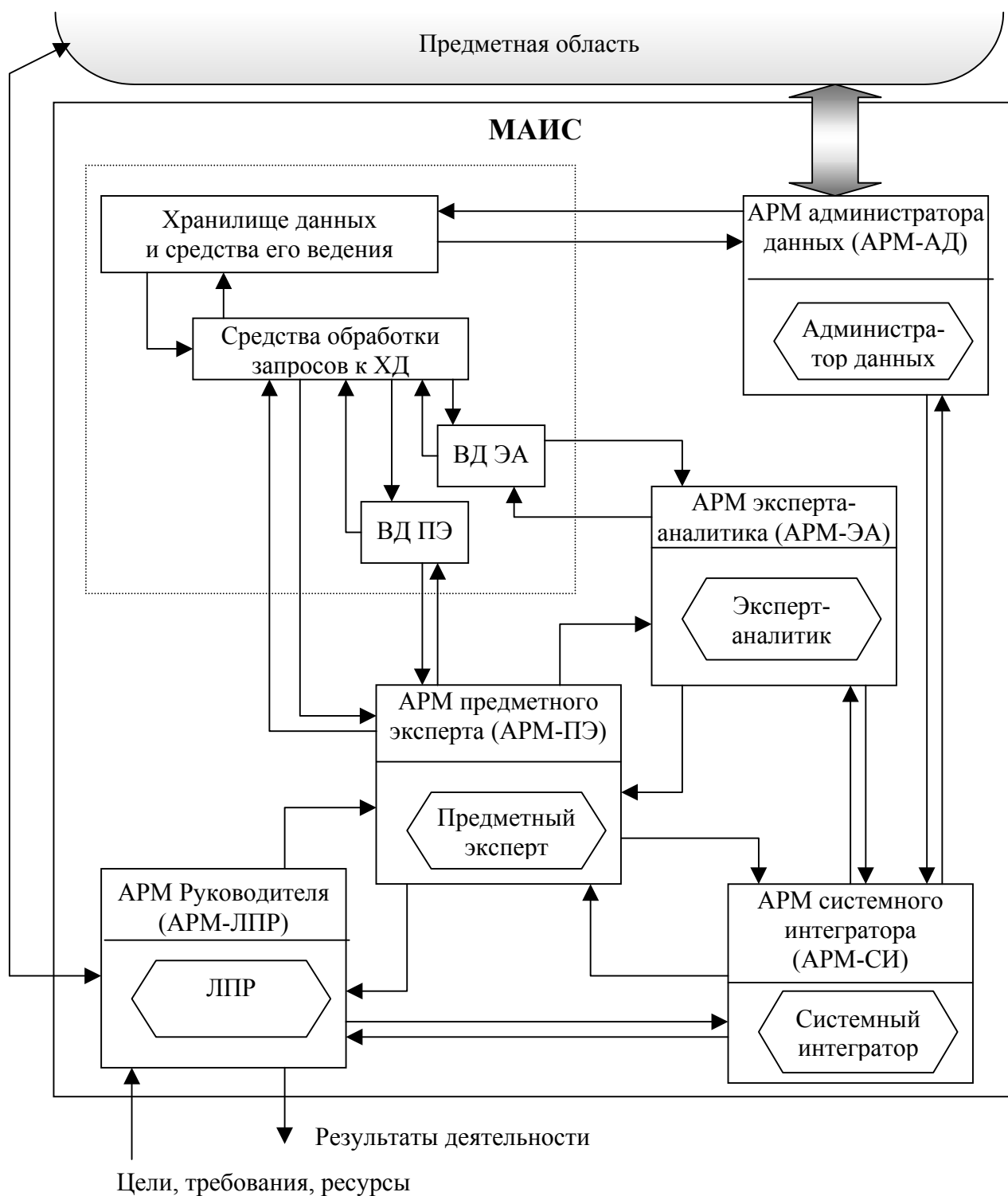


Рис. 2. Основные структурные компоненты МАИС

Платформу МАИС составляют используемые средства вычислительной техники, сетевое оборудование и средства общесистемного программного обеспечения.

Отмеченные в начале статьи недостатки существующих программных реализаций средств ИАД (равно как и абсолютного большинства современных информационных систем) прежде всего связаны с тем, что они создаются как синтаксически неизменяемые. Последнее означает, что структуры хранения данных и алгоритмы их обработки в таких системах жестко фиксированы. Существующие технологии программирования обеспечивают быстрое создание подобных программных систем, автоматизируя фиксацию синтаксиса систем в виде совокупностей моделей данных и логики приложений с последующей автоматической кодогенерацией.

Однако изменение требований к программной реализации (как следствие динамики условий и среды ее эксплуатации) приводит к необходимости модернизации программного кода с привлечением программистов и целым рядом возможных негативных последствий (риск внесения ошибок, изменение привычного интерфейса и т.д.). Кроме того, уникальность применения систем ИАД при использовании синтаксически неизменных реализаций по сути ведет к росту количества их адаптированных версий, внося путаницу в администрирование систем и существенно затрудняя процессы их сопровождения и развития.

МЕТАУПРАВЛЕНИЕ В МАИС

Вышеизложенное, а также ориентация на создание многофункциональных систем ИАД в сочетании с потенциальной неформализуемостью структур хранения предметных сведений и содержания процессов решения задач на этапе проектирования и реализации таких систем приводят к необходимости создания новых технологий реализации и организации функционирования МАИС. Следствием работ в данном направлении явился второй базовый принцип системной интеграции МАИС (создание МАИС как синтаксически вариативной информационной системы) и разработка соответствующей технологии.

Неформально синтаксическая вариантность в первом приближении означает независимость программного кода функциональных приложений от содержания обрабатываемых данных с одновременной возможностью изменения состава хранимых данных, множества доступных элементарных операций и динамического задания на этом множестве необходимых алгоритмов обработки.

В качестве методологической основы создания синтаксически вариантных информационных сис-

тем может быть использован метод метауправления, который предусматривает деление обрабатываемой информации на два уровня [2].

Первый является уровнем, на котором задаются определенные и используемые в информационной системе понятия, семантические ограничения и формы представления обрабатываемой информации. Совокупность метаинформации и определяет синтаксическую настройку ИС, а семантическая настройка реализуется (в основном) через механизм динамически загружаемых процедур.

На втором уровне пользователи получают возможность (через соответствующие приложения) осуществлять манипуляции с обрабатываемой предметной информацией в терминах предметных данных и операций, заданных на метауровне.

Таким образом, метауправление дает возможность выполнения существенных изменений в ИС без ее перепрограммирования или при минимальных изменениях программного кода — изменения описываются пользователем на метауровне в процессе эксплуатации системы и приводят к соответствующей самонастройке программных средств и корректировке структур хранимых данных.

Рассматривая МАИС как формальную лингвистическую систему, позволяющую создавать и интерпретировать дескриптивные задания процессов преобразования информации, введем следующие обозначения [3]:

γ — алфавит системы как множество структурных единиц хранимых данных γ_d и множество имеющихся алгоритмов обработки γ_a ;

v — множество правил продукции, т.е. множество правил соединения символов алфавита в выражения (цепочки), допустимые (правильные) с позиций возможности их интерпретации в МАИС;

$\lambda(\gamma, v)$ — множество правильно построенных (истинно интерпретируемых) выражений, описывающих как этапы решения задач в МАИС (λ_e), так и связи (отношения) хранимых в ней данных (λ_d);

ω — множество правил интерпретации правильно построенных выражений с целью получения результата этапа решения задачи (т.е. $\omega(\varepsilon, I) \rightarrow \langle r, I^* \rangle$, $\varepsilon \in \lambda_e(\gamma, v)$) — реализуемый процесс вычислений, I и I^* — состояния информационной базы МАИС соответственно до и после выполнения ε , r — результат выполнения ε ($r \subset I^*$);

$\langle \gamma, v \rangle$ — синтаксис системы;

$\langle \lambda, \omega \rangle$ — семантика системы.

Следует отметить, что множество γ_a представляет собой совокупность имен программно реализованных алгоритмов выполнения операций, γ_d соответствует именам сущностей ПрО и их атрибутов (но никак не экземплярам этих сущностей Y_d), и имеющихся между ними связей, а содержимое информационной базы составляют λ_d и Y_d . При

этом содержимому хранилища данных соответствуют λ_{ds} и Y_{ds} , а содержимому витрины данных — λ_{dt} и Y_{dt} ($\lambda_{ds} \cup \lambda_{dt} = \lambda_d$, $Y_{ds} \cup Y_{dt} = Y_d$, $\lambda_{ds} \cup Y_{ds} = I_s$ (содержимое ХД), $\lambda_{dt} \cup Y_{dt} = I_t$ (содержимое ВД), $I_s \cup I_t = I$).

Для того чтобы МАИС была синтаксически вариантной системой, необходимо ввести правила изменения синтаксиса $\alpha = \alpha_\gamma \cup \alpha_v$. Очевидно, что применение α ведет к изменению λ в рамках остающегося неизменным ω . Последнее и обеспечивает инвариантность интерпретирующей части лингвистической системы. Изменения α_γ в конечном счете сводятся к модификации логической структуры ХД и ВД (т.е. λ_d) и производной модификации Y_d , а изменения α_v — к модификации множества алгоритмов обработки.

Таким образом, МАИС как синтаксически вариантная система определяется совокупностью $\langle \gamma, v, \alpha, \lambda, \omega, I \rangle$. Эта совокупность величин — текущее состояние системы, определяющее ее потенциальные возможности по решению задач $((\lambda_d, Y_d) \times \gamma_a \rightarrow \lambda_e)$, а α — правила адаптации МАИС, реализуемые исключительно лингвистическими средствами.

Очевидно, что лингвистические средства реализации правил α представляют собой метаязык L , который применительно к МАИС является объединением следующих пяти языков.

L1 — язык описания алфавита γ_d и структуры данных λ_{ds} , а также возможных правил их выборки $\lambda_s \subset \lambda_e$ (как частный случай операции, входящей в γ_e и задаваемой на γ_d).

L2 — язык описания запросов на выборку сведений из хранилища данных и помещение их в витрины данных как множества вида $\langle y, P \rangle$, где $y \subset \gamma_{ds}$ — выбираемые сведения, $P = \{ \lambda_s(\gamma_{ds}) \}$ — правила отбора экземпляров данных в формируемую выборку.

L3 — язык описания алфавита γ_{dt} и структуры данных λ_{dt} .

L4 — язык описания алфавита операций $\{ \gamma_e / \gamma_s \}$ и правил выполнения этих операций на γ_{dt} .

L5 — язык описания запросов на выполнение операций обработки сведений, хранящихся в ВД, как множества вида $\langle z, U \rangle$, где z — операция обработки ($z \in \{ \gamma_e / \gamma_s \}$), $U = \{ u | u \in \gamma_{dt} \}$ — обрабатываемые сведения из ВД.

Любое предложение каждого из перечисленных метаязыков представляет собой компонент метаинформации: номенклатор понятий предметной области, запрос на заполнение витрины данных, спецификации структуры витрины данных, номенклатор операций обработки и запрос на выполнение операций для L1–L5 соответственно.

Также в МАИС вводятся описания синтаксиса всех типов компонентов метаинформации — т.е. грамматики $G(L1)–G(L5)$, а семантика этих мета-

языков реализуется интерпретирующей системой через механизм ассоциированных процедур.

Для компонентов метаинформации в МАИС имеются три существенные особенности:

— возможность наличия взаимосвязей между номенклаторами понятий различных доменов ПрО, за счет чего реализуется последовательное наращивание непротиворечивой понятийной модели выделяемого фрагмента реального мира;

— выполнение запроса на заполнение витрины данных приводит к изменениям в I_t (в общем случае как в λ_{dt} , так и в Y_{dt});

— выполнение запросов на выполнение операций приводит к вызовам ассоциированных процедур обработки (т.е. программно реализованных алгоритмов аналитической оболочки), которые помещают результаты своей работы в ВД, также изменяя тем самым любые части I_t .

О РЕАЛИЗАЦИИ МАИС

При наличии соответствующих интерактивных средств технология работы пользователей МАИС строится на основе формирования компонентов МАИС и анализа результатов их обработки, а обрабатываемые программы интегрируются со средствами интерпретации этих компонентов через механизм ассоциированных процедур, составляя основное содержание промежуточного слоя архитектуры МАИС.

Для реализации метода метауправления в МАИС на сегодняшний день наиболее целесообразно использовать язык XML (eXtensible Markup Language, "Расширяемый язык разметки"), разработанный Консорциумом W3C, осуществляя доступ к содержимому компонентов метаинформации через стандартизованный интерфейс DOM (Document Object Model).

СПИСОК ЛИТЕРАТУРЫ

1. Львов В. // СУБД. 1997. №3. С. 30–40.
2. Демитрович Я., Кнут Е., Радо П. Автоматизированные методы спецификации. М.: Мир, 1989. 115 с.
3. Губинский А.И. Надежность и качество функционирования эргатических систем. Л.: Наука, 1982. 270 с.

Военный университет связи, Санкт-Петербург

Материал поступил в редакцию 20.10.2000.

CONSTRUCTION OF MULTIOBJECT ANALYTICAL INFORMATION SYSTEMS WITH THE ELEMENTS OF METAMANAGEMENT

D. V. Belous, Yu. M. Sherstyuk

Military University of Communication, Saint-Petersburg

The main problem of creating multiobject analytical information systems is to provide corporate (united) work of many experts with the adaptation of the instruments being used to the classes of tasks being solved and fields of subject. The paper describes an approach to solving this problem. It is based on the multilevel decomposition of the multiobject analytical information systems structure and profile, as well as using metamanagement, which allows one to introduce the elements of syntactic plurality.