

РОССИЙСКАЯ АКАДЕМИЯ НАУК
Пушкинский научный центр
Институт почвоведения и фотосинтеза

ПРЕПРИНТ

А.С.Комаров, Г.Л.Андриенко, Н.В.Андриенко,
М.Н.Брынских, Д.Б.Орлинский,
А.К.Благовидов

ИНФОРМАЦИОННО-АНАЛИТИЧЕСКАЯ
СИСТЕМА "ЗАПОВЕДНИК"

Пушкино-1994

РОССИЙСКАЯ АКАДЕМИЯ НАУК
Пушкинский научный центр
Институт почвоведения и фотосинтеза

ПРЕПРИНТ

А.С.Комаров, Г.Л.Андриенко, Н.В.Андриенко,
М.Н.Брынских, Д.Б.Орлинский,
А.К.Благовидов

ИНФОРМАЦИОННО-АНАЛИТИЧЕСКАЯ
СИСТЕМА "ЗАПОВЕДНИК"

Пушино-1994

А.С.Комаров, Г.Л.Андриенко, Н.В.Андриенко, М.Н.Брынских, Д.Б.Орлинский, А.К.Благовидов. Информационно-аналитическая система "Заповедник" / Препринт. Пушино: ПНЦ РАН, 1994, 51 с.

В работе предлагается проект информационно-аналитической системы для сбора, хранения, поиска и анализа данных экологического мониторинга в заповедниках. Система обеспечивает доступ к разнородным данным, включая тексты, таблицы различного формата, графические построения, карты и изображения. Обосновывается необходимость включения в систему базы знаний по экологическому мониторингу. Приводится архитектура системы. Описываются структура базы знаний и базы данных и методы отбора релевантной информации.

Использование базы знаний проблемной области обеспечивает пользователю возможность формулировки разнообразных запросов на его профессиональном языке и генерации отчетных документов в принятом виде. Включение в состав системы базы знаний по когнитивному представлению информации позволяет пользователю проводить визуальный анализ данных.

Изложение иллюстрируется примерами из Летописи природы Приокско-Тerrasного биосферного заповедника.

A.S.Komarov, G.L.Andrienko, N.V.Andrienko, M.N.Brynskich, D.B.Orlinsky, A.K.Blavogidov. Information and analytical system "Nature reserve" / Preprint: Pushchino: PRC, 1994, 51 p.

The communication suggests the project of the information and analytical system for the collection, storage, retrieval and analysis of data pertaining to the area of ecological monitoring in nature reserves. The system provides the access to heterogeneous data including texts, tables with different format, graphical constructions, maps and images. Substantiated is the necessity of the inclusion of knowledge base on ecological monitoring into the system. The architecture of the system is outlined.

The structure of the knowledge base and database and methods of relevant information retrieval are described. Exploitation of the domain knowledge base provides the possibility to the user to formulate diverse queries in his professional language and to generate report documents in the established form. Inclusion into the system of the knowledge base on cognitive information representation allows the user to carry out visual analysis of the data.

The communication is illustrated with the examples taken from the Chronicle of Nature of the Prioksko-Terrasny State Biosphere Reserve.

1. Мотивация

Основной научной деятельностью заповедников является экологический мониторинг - наблюдение за природными явлениями и состоянием природных объектов, фиксация, обработка и обобщение результатов наблюдений. Отчетным документом о проводимой в течение года научной работе является Летопись природы. В основном в Летописи природы отображаются обобщающие (суммарные, усредненные и т.д.) данные по отношению к текущим наблюдениям. Характерно, что различные разделы Летописи готовятся различными специалистами заповедника. В результате Летопись имеет вид разрозненных фрагментов информации, в то время как наблюдаемые объекты и явления тесно взаимосвязаны.

Информация, собираемая в заповедниках, представляет огромный научный и практический интерес. В частности, она должна учитываться при планировании развития экономики с точки зрения рационального природопользования и землепользования. Особое значение имеют данные о редких и исчезающих видах растений и животных. Данные экологического мониторинга в заповедниках также необходимо учитывать для прогнозирования развития экологической обстановки в стране и отдельных ее регионах.

Однако в настоящее время возможности использования информации, накопленной заповедниками, сильно ограничены. Во-первых, Летопись природы не тиражируется, и, следовательно, ее содержание практически недоступно широкому кругу специалистов. Во-вторых, Летопись неудобна для поиска в ней необходимой информации, выявления связей между объектами и явлениями природы, прослеживания динамики явлений, выявления закономерностей и т.д. В-третьих, представление данных в текстовом виде (в виде книги) значительно затрудняет манипулирование ими, т.е. переупорядочение данных, изменение способа представления, возможность различной числовой обработки. Такая деятельность, необходимая для реального использования Летописи природы, в настоящее время может осуществляться только вручную и, следовательно, с большими затратами времени, увеличением вероятности ошибок, снижением объема анализируемой информации.

Таким образом, существует необходимость в новом подходе к сбору и представлению данных по экологическому мониторингу в

заповедниках в целях обеспечения доступности информации, удобства доступа, обработки и анализа. Должен быть осуществлен переход к компьютерному представлению и хранению информации, к автоматизированным методам доступа и анализа информации на основе баз знаний и экспертных систем. Компьютерные технологии, в идеале, позволят организовать деятельность всех заповедников страны на единой основе. За счет этого станет возможным обмен информацией между заповедниками, проведение научной работы заповедника с учетом данных и результатов, полученных другими заповедниками.

Большую научную и практическую ценность имеет интеграция на этой основе результатов работы различных заповедников и заказников. С одной стороны, интеграция обеспечивает неоценимый материал для научных исследований и обобщений, с другой, позволяет получить целостную картину состояния природной среды в области, регионе, стране; является проводником единой государственной стратегии по организации и ведению экологического мониторинга, выраженной в проекте Единой государственной системы экологического мониторинга (ЕГСЭМ).

2. Цель работы

В данной работе предложен подход к решению указанных проблем, который заключается в создании компьютерной системы по сбору, хранению, поиску и обработке данных экологического мониторинга. Особенностью системы является ее гибкость, т.е. возможность адаптации к специфическим условиям каждого конкретного заповедника и заказчика. В качестве таких специфических условий выступают: 1) изучаемые в данном заповеднике объекты и явления природы, в том числе уникальные, не встречающиеся в других заповедниках; 2) принятые форматы хранения данных, в том числе форматы СУБД, электронных таблиц, геоинформационных систем и других систем работы с данными, традиционно используемых в заповеднике; 3) принятые или наиболее целесообразные для данного заповедника формы отчетности.

Предлагаемая система должна обеспечить выполнение следующего набора задач.

Задача 1. Ввод в компьютер информации по экологическому мониторингу сотрудниками заповедника. Система должна поддерживать как ввод в режиме диалога, так и подключение

заранее подготовленных данных в стандартных форматах (текстовом или принятом в той или иной СУБД или другой системе работы с данными). При вводе данных в режиме диалога система предоставляет экранную форму для заполнения, которая выбирается из соображений удобства или привычности для пользователя, но не из соображений соответствия внутренним форматам хранения данных.

Задача 2. Хранение информации всех видов, используемых в Летописи: таблиц, текстов, деловой графики, карт, рисунков, фотографий. Система должна быть универсальной и обеспечивать доступ к информации, хранящейся в разных форматах различных СУБД, графических пакетов, ГИС и других программных средств. Система обеспечивает доступ к информации любого вида, визуализацию информации на экране компьютера и печать на бумажных носителях.

Задача 3. Поиск и выборка информации по запросу пользователя. Система поддерживает удобный для пользователя интерфейс по составлению запросов. Пользователю предоставляются готовые тексты запросов с выделенными полями (пропусками) для подстановки необходимых терминов, называемые бланками запросов. В запросах используются только общеупотребительная лексика и профессиональный словарь пользователей. Таким образом, для составления запроса от пользователя не требуется знания формальных языков и форматов представления данных. При заполнении каждого поля бланка система предоставляет пользователю для выбора список подходящих терминов с учетом смысла запроса. Тем самым, во-первых, пользователь освобождается от необходимости ввода текста запроса, во-вторых, гарантируется правильная интерпретация запроса системой без подключения сложного аппарата анализа естественного языка.

Задача 4. Предоставление результатов поиска в удобном для восприятия и анализа виде. Эта задача имеет две стороны. Во-первых, это удобная организация просмотра информации в случае, когда по запросу системой найдено большое количество относящейся к делу (релевантной) информации, возможно, представленной в различном виде (тексты, графические изображения, карты и т.п.). Во-вторых, информация может храниться в форматах, неудобных для непосредственного восприятия пользователем (например, в формате какой-либо СУБД). В этом случае система должна преобразовать данные к читаемому виду.

В качестве решения первой стороны данной задачи предлагается оформлять результаты поиска в виде набора документов со ссылками между ними, отражающими содержательные связи (гипертекстовое представление). Активизация ссылки при просмотре одного из документов вызывает появление на экране другого документа, связанного с текущим.

Вторая часть задачи решается за счет табличных представлений, сопровождаемых графическими иллюстрациями (графиками, диаграммами и т.п.). Такие иллюстрации строятся системой автоматически исходя из характеристик данных и содержания запроса. Цель подобных графических построений - не только облегчить восприятие пользователем информации, но и стимулировать выявление особенностей и закономерностей в данных, связей между явлениями, активизировать процессы формирования гипотез и поиска обобщений. Класс изображений, удовлетворяющих перечисленным требованиям, носит название когнитивной графики.

Задача 5. Преобразование данных, найденных по запросу пользователя, в форматы, позволяющие проводить их статистическую и иную обработку применяемыми в заповеднике программными средствами.

Задача 6. Генерация отчетных документов. Система предоставляет возможность описать содержание и задать требуемые форматы отчетных документов, после чего обеспечивается автоматическая выборка нужных данных, преобразование их к указанным форматам и печать требуемых документов. Однажды заданные форматы по желанию пользователей сохраняются в системе и могут использоваться многократно. Генерируемые документы могут быть текстовыми, табличными, графическими, картографическими. Система, в частности, обеспечивает печать бумажного варианта Летописи природы в принятом в настоящее время виде.

Таким образом, предлагаемая система выполняет все функции информационно-поисковых систем, учитывая разнообразие форм представления информации. Кроме того, система обеспечивает поддержку аналитической деятельности пользователя. Эти свойства системы отражены в названии "информационно-аналитическая система". При создании системы в качестве основного принципа выдвигается организация взаимодействия со всеми категориями пользователей на их профессиональном языке, а не на формальном языке запросов. Принятый подход к

реализации этого принципа (подход на основе знаний) позволяет употребить в названии системы также термин "интеллектуальная".

Далее излагается проект интеллектуальной информационно-аналитической системы "Заповедник", решающей перечисленные выше задачи. Приложение содержит пример работы с системой различных категорий пользователей (эксперт, оператор по вводу данных, пользователь-аналитик). В связи с тем, что данная работа адресуется научному персоналу заповедников в качестве предложения по организации работы, авторы ожидают обсуждения как рассмотренных, так, возможно, и упущенных вопросов, замечаний и пожеланий от заинтересованных лиц.

Предлагается рассматривать публикацию описания проекта как приглашение к сотрудничеству всех лиц и организаций, заинтересованных в реализации данного проекта.

3. Структура и содержание Летописи природы заповедника

В Летописи природы отображаются прежде всего обобщенные результаты наблюдений за природными явлениями и объектами, проводившихся в течение года. В Летописи могут присутствовать также и сводные данные за несколько лет. Летопись представляет собой набор информационных единиц, которые будем называть документами. Отдельный документ содержит данные наблюдений за каким-либо одним объектом или явлением природы; за группой однородных или взаимосвязанных объектов. Под однородными понимаются объекты, характеризующиеся при помощи одних и тех же свойств. Как правило, данные по группам однородных объектов приводятся в виде таблиц. Взаимосвязи объектов и явлений описываются в виде текстов.

В целом при ознакомлении с образцами Летописи и методическими материалами по составлению Летописи выявляется стремление к представлению как можно большего числа данных в виде таблиц. Приводимые в Летописи графики, диаграммы и прочие графические построения носят иллюстративный характер и сопровождают, но не заменяют табличное представление данных. Тексты, как правило, содержат комментарии и обобщения, которые не могут быть представлены в табличном виде.

В качестве документов используются также картографические материалы: планы заповедника или отдельных его частей с показом ареалов обитания различных видов растений и

животных, маршрутов и площадок наблюдений с отметками результатов наблюдений и т.п.

Дополнительно к документам, содержащим фактические данные, в Летописи могут приводиться фотографии или рисунки явлений и объектов природы. Однако наибольший интерес с точки зрения научных и практических целей представляют фактические данные. В связи с этим рассмотрим в общем виде содержание документа Летописи, вводя при этом некоторую терминологию, которая будет использоваться в дальнейшем изложении.

Будем называть термином **"объект наблюдения"**, или просто **"объект"** любой объект или явление природы, за которым ведется наблюдение в заповеднике. Любой объект характеризуется рядом **свойств**, каждое из которых имеет множество отдельных проявлений, или **значений** этого свойства.

Рассмотрим пример. Одним из объектов наблюдения в заповеднике может быть определенный вид животных - лось. Свойства объекта: численность, количество самцов, количество самок, размещение особей по территории заповедника и т.д. Каждое свойство имеет свою область значений. У свойства "численность" - это множество неотрицательных целых чисел, у свойства "размещение" - набор номеров или названий кварталов заповедника либо других единиц его территориального деления.

Наблюдением за объектом назовем определение значений его свойств.

Значения свойств объекта зависят от условий наблюдения, среди которых прежде всего следует назвать время. Кроме того, для исследования природы представляет интерес определение свойств объекта в различных точках пространства, при различных состояниях среды, в которой находится объект, и т.п.

Факторы, в зависимости от которых свойства наблюдаемого объекта могут принимать различные значения, назовем **параметрами наблюдения**.

Параметры наблюдения чаще всего представляют собой свойства, при некоторых зафиксированных значениях которых происходит определение значений свойств изучаемого объекта. В качестве параметров могут выступать свойства наблюдаемого объекта или свойства среды, в которой существует объект. Так, при изучении свойств насекомых в различных видах леса параметром является вид леса - среды обитания насекомых. При наблюдении веса животных параметрами служат пол и возраст животных - свойства самих животных.

В качестве параметров наблюдения могут выступать также и объекты, если наблюдаются свойства, общие для нескольких объектов. Например, при изучении численности копытных одним из параметров наблюдения являются отдельные виды копытных.

Введенные понятия позволяют охарактеризовать в общем виде содержание документа, входящего в состав Летописи природы. В большинстве документов приводятся данные наблюдений за одним объектом или группой однородных объектов в зависимости по крайней мере от одного параметра. Другие документы приводят данные наблюдений за свойствами различных объектов в целях выявления или демонстрации взаимосвязей или влияния одних объектов на другие. Такие данные также зависят от параметров.

Параметры, при которых и в зависимости от которых велись наблюдения, явным образом указываются в документе. Часть параметров наблюдения принимает одно зафиксированное значение (например, время измерений - лето 1994 года, место измерений - квартал 36а), другие параметры могут принимать различные значения из заданного множества значений - **базового множества**. Базовое множество задается путем перечисления всех значений параметра, при которых проводилось определение (измерение, фиксация) рассматриваемого свойства объекта, или путем указания интервала и шага приращения значений параметра.

Как правило, если хотя бы один из параметров наблюдения варьирует (базовое множество содержит более одного значения), данные наблюдения представляются в виде таблицы.

В качестве примера рассмотрим таблицу 8.24 из методического пособия по ведению Летописи природы [1]:

**Сезонное размещение копытных животных по
основным местам обитания в 19 году.**

Места обитания	Сезоны года	Вид 1	Вид 2	...

В данном документе приводятся результаты наблюдения за свойством "численность" объекта "копытные животные" при следующих базовых множествах значений параметров:

- 1) Год наблюдения = {заданный год наблюдения}
- 2) Место обитания = {множество мест обитания}
- 3) Сезон года = {зима, весна, лето, осень}
- 4) Вид животного = {множество видов копытных, имеющих в заповеднике}.

Заметим, что значение первого параметра зафиксировано (базовое множество состоит из одного элемента), остальные параметры принимают различные значения из заданных базовых множеств. Значения наблюдаемых свойств объекта должны быть определены для всех возможных наборов значений параметров. Если записать данные наблюдений в виде последовательностей следующего вида (*кортежей*):

$VPar1, ..., VParN, VPI_1, ..., VPK,$

где $VPar1, ..., VParN$ - значения параметров, $VPI_1, ..., VPK$ - значения наблюдаемых свойств, то количество таких кортежей должно быть равно произведению числа элементов базовых множеств (мощности множеств) всех параметров.

Что касается документов комментирующего и аналитического характера, то они также посвящены рассмотрению тех или иных свойств объектов. Как правило, это не непосредственно наблюдаемые свойства, а некоторые оценки, интегрирующие в себе данные об одном или нескольких наблюдаемых свойствах для ряда значений параметров. Так, в подобных документах часто встречаются выводы типа: "Сухое и жаркое лето привело к низкой урожайности грибов".

В отличие от документов, содержащих фактические данные наблюдений, в документах комментирующего и аналитического характера может отсутствовать явная привязка к значениям параметров. Как правило, используется только один параметр с зафиксированным значением - год наблюдения. В связи с этим такие документы имеют вид текстов. Следует заметить, что

большое внимание в подобных текстах уделяется раскрытию взаимосвязей свойств различных объектов.

4. Подход. Общая схема информационной системы

Как видно из содержания предыдущего раздела, в Летописи природы рассматривается большое количество разнородных объектов, характеризующихся различными свойствами. Значения свойств варьируют в зависимости по крайней мере от одного, но чаще от ряда параметров. Всеобщим параметром является время, однако остальные параметры, влияющие на свойства объектов, различны для разных объектов. В связи с этим каждый документ Летописи имеет свою структуру. В случае создания на основе Летописи компьютерной базы данных схема такой БД окажется очень громоздкой из-за большого числа документов различной структуры, что потребует чрезмерных усилий для ее освоения и работы с ней пользователя. Это противоречит целям обеспечения доступности информации по экологическому мониторингу, задачам более широкого использования ее в научной и хозяйственной деятельности.

Таким образом, необходимо создать такую систему сбора и доступа к данным, с которой мог бы работать любой пользователь без предварительного изучения схемы БД и знания каких-либо других особенностей представления и организации информации. Система должна общаться с пользователем на его языке, т.е. при помощи понятий, обозначающих природные объекты, явления и их свойства, но не в терминах таблиц, записей, полей, ключей и т.п. Компьютерные средства с такими свойствами носят название *интеллектуальных информационных систем (ИИС)*.

Как указывалось в разделе 2, интеллектуальная информационная система по экологическому мониторингу в заповедниках должна выполнять следующие функции:

- облегчать, направлять и систематизировать ввод в компьютер данных о наблюдениях природы в заповеднике его сотрудниками;
- предоставлять конечным пользователям удобный интерфейс для составления запросов;
- обеспечивать поиск и выборку требуемой информации по запросам пользователей;
- осуществлять надлежащее оформление и представление информации с точки зрения удобства восприятия,

информативности и выразительности, возможности анализа информации и решения на ее основе различных задач;

- генерировать отчетные документы различного вида, в том числе автоматизировать создание бумажного варианта Летописи природы.

ИИС, выполняющую указанные функции, предполагается строить как *систему, основанную на знаниях*. Система должна содержать в себе базу знаний о проблемной области: наблюдаемых объектах и явлениях природы, их свойствах и отношениях между ними, принятых способах (методиках) определения значений свойств. Кроме того, в базе знаний должна содержаться информация о принятых в заповедниках формах отчетности.

За счет наличия такой базы знаний ИИС приобретает способность динамически генерировать и предоставлять специалистам заповедника подходящие формы для внесения данных в компьютер. Формы выбираются или генерируются системой из соображений удобства заполнения и могут на стадии ввода данных не совпадать с принятыми формами отчетности, генерация которых впоследствии будет проводиться системой.

Содержащиеся в базе знаний названия явлений природы, свойств, отношений составляют словарь (профессиональный язык), с помощью которого пользователь может формировать свои запросы. При этом достигается независимость пользовательского интерфейса от способа организации данных и от каких-либо формальных языков запросов, кроме языка проблемной области.

Для обеспечения доступа к информации, т.е. возможности выборки релевантной информации по запросу пользователя, производится *индексирование* (формальное описание содержания) каждой *единицы хранения информации* в терминах понятий *словаря проблемной области*. Единицами хранения считаются табличные документы, графики, карты, изображения, а также фрагменты текста, посвященные описанию одного объекта или одного вида взаимосвязи объектов. Индексирование документов осуществляется автоматически при вводе информации по предоставляемым системой формам либо путем опроса пользователя в случае подключения готовых документов или баз данных.

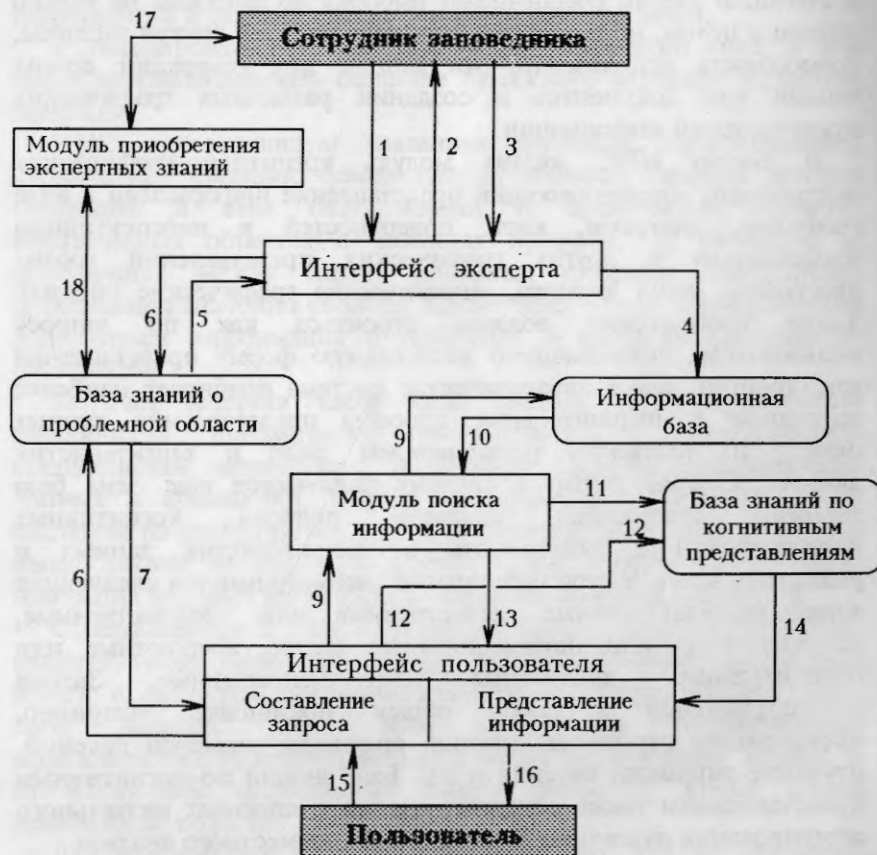
В *индексной записи*, соответствующей единице хранения информации, указываются: 1) описываемые объекты, 2) наблюдаемые свойства и 3) параметры, относительно которых

проводились наблюдения, и их значения. Индексирование строк и столбцов таблиц обеспечивает выборку по запросам не только таблиц в целом, но и доступ к каждому данному внутри таблицы, возможность использовать эти данные при генерации других таблиц или документов и создании различных графических представлений информации.

В состав ИИС входит модуль когнитивно-графических построений, обеспечивающий представление информации в виде графиков, диаграмм, карт, поверхностей в перспективном изображении и других графических представлений (розы-диаграммы, лица Чернова, динамические графические образы). Такие изображения должны строиться как по запросу пользователя, указывающего желательную форму представления информации, так и автоматически: система подбирает наиболее наглядные и выразительные способы представления данных исходя из решаемых пользователем задач и характеристик данных. С этой целью в систему включается еще одна база знаний, содержащая правила подбора когнитивных представлений в зависимости от характеристик данных и решаемых задач. В этой базе знаний рассматриваются следующие характеристики данных: качественные или количественные, упорядоченная или неупорядоченная шкала, абсолютные или относительные, дискретные или непрерывные. Задачи рассматриваются в самой общей постановке, например, исследование структуры явления, сравнение размеров явлений, изучение динамики явления и т.п. База знаний по когнитивным представлениям также содержит знания о способах визуального агрегирования различных данных для их совместного анализа.

Следует учитывать, что каждый заповедник имеет свою специфику, выражающуюся в различных природно-климатических условиях, разных видах животных и растений и т.д. В связи с этим научным сотрудникам заповедников необходим инструмент для изменения базы знаний о проблемной области и учета в ней специфики конкретного заповедника. Поэтому в состав системы включается модуль автоматизированного приобретения экспертных знаний, который будет в диалоге с экспертом модифицировать базу знаний и обеспечивать ее корректность.

Общая структура системы и схема информационных потоков:



Числами на схеме обозначены:

- 1) указание описываемого явления и показателей
- 2) предоставление формы для заполнения
- 3) заполнение формы
- 4) внесение индексированных данных в базу данных
- 5) понятия проблемной области для описания явлений и генерации форм
- 6) поиск необходимых понятий проблемной области
- 7) понятия проблемной области для составления запроса
- 8) передача запроса на поиск информации
- 9) поиск информации
- 10) получение данных из информационной базы
- 11) описание характеристик отобранных данных
- 12) описание задачи
- 13) отобранные по запросу данные
- 14) рекомендуемые способы представления данных
- 15) составление запроса пользователем
- 16) получение пользователем требуемой информации
- 17) приобретение знаний о специфике заповедника
- 18) пополнение базы знаний проблемной области.

- 13) отобранные по запросу данные
- 14) рекомендуемые способы представления данных
- 15) составление запроса пользователем
- 16) получение пользователем требуемой информации
- 17) приобретение знаний о специфике заповедника
- 18) пополнение базы знаний проблемной области.

5. Организация базы знаний информационной системы

5.1. Понятия и их значения

База знаний (БЗ) интеллектуальной информационной системы содержит понятия проблемной области и отношения между ними. В области экологического мониторинга в заповедниках необходимыми являются понятия для обозначения наблюдаемых объектов и их свойств. Для краткости изложения будем называть такие понятия **объектами** и **свойствами**. В случае, когда необходимо будет явным образом различить понятие и реальный объект или реальное свойство, будем говорить "понятие-объект" и "понятие-свойство".

Любое понятие является общим именем некоторого множества сущностей, которые назовем **значениями** этого понятия. Поэтому каждое определенное в БЗ понятие считается **типом**. В общем случае полным определением понятия можно считать указание его имени и множества значений. Однако реально это не всегда возможно. Так, значениями понятия "заяц" является множество всех зайцев, когда-либо существовавших в мире (в прошлом, настоящем или будущем). То же обстоятельство имеет место и для любого другого понятия-объекта. В связи с этим понятия-объекты вводятся в базу знаний без задания множеств значений. Понятиями, для которых множество значений задается обязательно, являются свойства.

Существует несколько способов задания множества значений:

- Полное перечисление наименований возможных значений, например, свойство "цвет листа" имеет множество значений {"зеленый", "желтый", "желто-зеленый", "красный", "фиолетовый"}.
- Ссылка на тип. Кроме типов, определяемых как понятия, имеется ряд предопределенных типов: число, количество (неотрицательное целое число), строка, дата, логическое значение (истина или ложь). Например, у объекта "животный мир" есть

свойство "место обитания" со значениями типа "территория" и свойство "численность" со значениями типа "количество".

- Ограничение типа: значения свойства образуют подмножество множества всех значений указанного типа. Чаще всего ограничения применяются к числовым типам и к типу "дата" путем указания интервала значений, например, "17/05/1994".."25/06/1994".

- Значениями свойства могут быть наборы, состоящие из произвольного числа значений какого-либо типа. Например, значениями свойства "дни с осадками" являются наборы дат. При описании таких свойств будем использовать ключевое слово НАБОР: НАБОР("дата"), НАБОР("грибы") и т.д.

5.2. Отношения и их роль в базе знаний

С помощью *отношений* в БЗ задаются связи между понятиями. Понятия могут быть связаны следующими отношениями: родо-видовым, "имеет свойство", "имеет значение", "часть-целое", "объект-среда". В дальнейшем изложении будем использовать краткие наименования отношений, производные от слов английского языка: ISA, HAS-PROPErty, HAS-VALUE, PART-OF, HAS-ENVironment. Это не означает, что подобные идентификаторы являются частью языка взаимодействия системы с пользователями. Они применяются исключительно в целях выделения названий отношений в данном тексте.

Кроме перечисленных выше отношений, в системе используются теоретико-множественные отношения включения множеств (обозначается SUB) и принадлежности элемента множеству (IN).

5.2.1. Родо-видовое отношение. Иерархии понятий

При представлении знаний об объектах и свойствах существенную роль играет *родо-видовое отношение*, или ISA-отношение. Смысл этого отношения заключается в том, что все значения одного из понятий являются в то же время значениями другого понятия. Первое понятие называется *видовым* по отношению к второму, а второе - *родовым* по отношению к первому. Например, понятие "млекопитающее" является родовым по отношению к понятию "заяц" и видовым по отношению к понятию "животный мир". Формально это записывается следующим образом:

"заяц" ISA "млекопитающее",

"млекопитающее" ISA "животный мир".

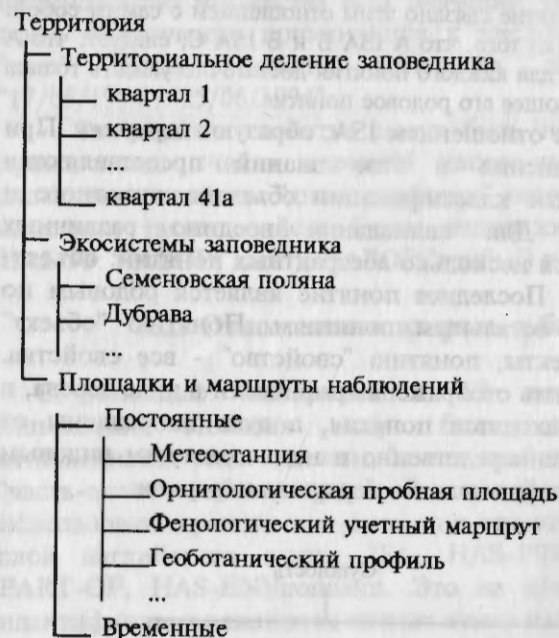
Отношение ISA бинарно (включает двух участников); рефлексивно (считается, что любое понятие связано этим отношением с самим собой: A ISA A) и транзитивно (из того, что A ISA B и B ISA C, следует, что A ISA C). Вследствие этого для каждого понятия достаточно указать только непосредственно включающее его родовое понятие.

Понятия, связанные отношением ISA, образуют *иерархии*. При помощи этого отношения в базе знаний представляются принятые иерархические классификации объектов животного и растительного мира. Для связывания воедино различных иерархий в БЗ вводится несколько абстрактных понятий: **объект**, **свойство** и **сущность**. Последнее понятие является родовым по отношению ко всем остальным понятиям. Понятию "объект" подчиняются все объекты, понятию "свойство" - все свойства. ISA-иерархия может быть отображена графически в виде дерева, в вершинах которого находятся понятия, а дуги направлены от родового понятия к непосредственно в него входящим видовым понятиям. Так, например, выглядит верхушка иерархии:



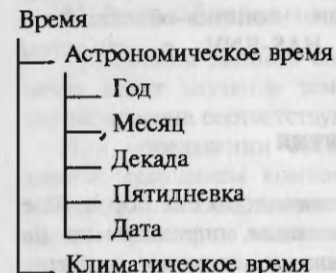
Понятия "Растительность" и "Животный мир" являются корневыми вершинами иерархий, в которые входят типы, классы, отряды, семейства и другие классификационные единицы растений и животных, вплоть до отдельных видов или, в случае редких видов, до отдельных популяций или даже особей. К другим узлам приведенной схемы также подключаются соответствующие им по тематике иерархии понятий.

Такой, например, вид имеет иерархия понятий, имеющих отношение к территории:



Отношением ISA могут быть связаны не только объекты, но и свойства. Во-первых, все свойства связаны этим отношением с понятием "свойство". Во-вторых, с помощью этого отношения можно указать, что некоторые понятия-свойства одинаковы по смыслу и представляют собой различные способы отображения одного и того же реального свойства, например, с различной детальностью, в абсолютных и относительных единицах и т.п. Так, существует как минимум 2 способа задания времени: астрономический (в виде даты) и климатический (сезоны). Астрономическое время может измеряться с разной детальностью. В подобных случаях в базу знаний вводится одно родовое свойство, которому подчиняются несколько видовых. Может быть несколько уровней подчинения. Считается, что множеством значений свойства более высокого уровня является объединение множеств значений подчиненных свойств.

Приведем в качестве примера иерархию свойств для обозначения времени, которые подчиняются родовому понятию "время":



5.2.2. Отношение "имеет свойство". Наследование свойств

Отношение "имеет свойство" (HAS-PROP) связывает объекты с их свойствами. Каждое свойство должно быть связано с объектом, к которому оно относится, так как не существует свойств без их носителей. В качестве носителя свойства "время" в системе принят выделенный в базе знаний объект "вселенная".

Система поддерживает принцип наследования свойств: свойство, определенное для некоторого объекта (связанное с этим объектом отношением HAS-PROP), считается определенным также для всех объектов, связанных с ним отношением ISA:

ЕСЛИ (A' ISA A) И (A HAS-PROP P) ТО (A' HAS-PROP P)

Однако для объектов, находящихся на более низких ступенях ISA-иерархии, свойства могут уточняться. Уточнением считается ограничение множества возможных значений свойства. Например, ранее было указано, что у объекта "животный мир" есть свойство "место обитания" с множеством значений "территория". Для объекта "бобр" множество значений этого свойства ограничивается до "реки и ручья".

Принимается по определению, что для каждого значения любого из свойств существует обозначающее его понятие, и что это понятие связано с понятием-свойством отношением "имеет значение", (HAS-VALUE). Например, "численность" HAS-VALUE "217".

Отношения "часть-целое" (PART-OF) и "объект-среда" (HAS-ENV) служат преимущественно для ориентировки при выборе понятий пользователем в ходе работы с системой. Кроме того, они позволяют ввести разумные ограничения при составлении запросов и индексировании документов, как будет показано далее. Отношением PART-OF связаны, например, понятия "возрастная группа" и "популяция растений", названия отдельных видов растений и понятие "растительная ассоциация". Отношение HAS-ENV имеет место между наименованиями обитающих в заповеднике птиц и насекомых и понятиями,

обозначающими различные экосистемы заповедника (сосняк сложный, березняк, дубрава, смешанный осиново-березовый лес с дубом и сосной и т.д.). Кроме того, средой для всех объектов является вселенная, поэтому при вводе в БЗ системы нового понятия-объекта оно автоматически связывается отношением HAS-ENV с объектом "вселенная".

5.3. Составные понятия

В отличие от *простых* понятий, рассматриваемых до сих пор, в базе знаний присутствуют также и *составные понятия*, определяемые на основе других понятий. Например, составным является понятие результата наблюдения (факта), так как для его определения необходимо указать, какой объект и какое из его свойств наблюдалось и какое значение свойства было зафиксировано. Таким образом, понятие факта состоит из трех компонент: имени наблюдаемого объекта, наименования наблюдаемого свойства и значения свойства. Заметим, что такое определение понятия "факт" является обобщенным и позволяет ввести множество конкретных фактов, различающихся по наблюдаемым объектам, свойствам и значениям: "максимальная масса плода земляники равна 0.3 г", "процент повреждения листовой поверхности березы равен 8.3%", "дата образования устойчивого снежного покрова - 3 декабря" и т.д. Удобно принять, что составное понятие вводится как некоторая формула с переменными, вместо которых могут подставляться те или иные понятия базы знаний. Результатом подстановки является новое понятие, называемое *экземпляром* составного понятия.

В общем случае составные понятия определяются в базе знаний путем указания имени понятия и списка компонент. Каждая компонента задается в виде:

<переменная>:<название компоненты>:<ограничения>,

где: <переменная> - произвольный идентификатор, не встречающийся в других компонентах данного составного понятия; <название компоненты> - произвольный текст, разъясняющий смысл компоненты; <ограничения> указывают, какие понятия могут подставляться вместо данной переменной. Так, при определении понятия "факт" необходимо указать, что одна из его компонент обязательно является объектом, другая - свойством этого объекта, а третья - значением свойства. Ограничения задаются при помощи отношений над понятиями БЗ и компонентами данного составного понятия.

Составные понятия включаются в общую иерархию понятий, т.е. являются объектами либо свойствами (подчинены соответствующим понятиям по отношению ISA). Значениями составного свойства считаются кортежи, составленные из значений компонент этого свойства в порядке, заданном при определении составного свойства.

Пример составного свойства:

Имя: "абсолютный максимум температур"

Компоненты:

#1: "достигнутое значение температуры": #1 ISA "число"

#2: "дата наблюдения максимума": #2 ISA "дата"

Значениями данного свойства являются пары, в которых на первом месте стоит значение температуры, а на втором - дата, когда была зафиксирована соответствующая температура.

При определении составных свойств принята упрощенная форма записи: описанием компоненты может считаться указание множества значений или типа соответствующего понятия. Компонента, описанная таким образом, не содержит переменной. Так, например, определяются свойства, связанные с измерением времени (компоненты составных свойств указаны в фигурных скобках):

Время

- Астрономическое время
 - Год, количество: от 1900 до 2100
 - Месяц, {январь, февраль, ..., декабрь; Год}
 - Декада, {I, II, III; Месяц}
 - Пятидневка, {число: от 1 до 6; Месяц}
 - Дата, {число: от 1 до 31; Месяц}
- Климатическое время {Сезон; Год}

Будем в дальнейшем называть составные понятия, в которых имеется хотя бы одна переменная, *понятиями-формулами*. Для задания экземпляра понятия-формулы указывается имя этого понятия, за которым в скобках следует список понятий, подставляемых вместо переменных в порядке их перечисления в определении составного понятия.

Приведем примеры понятий-формул и их экземпляров:

Имя: "факт"

Родовое понятие: "объект"

Компоненты:

#1: наблюдаемый объект: #1 ISA "объект"

#2: свойство объекта: #2 ISA "свойство", #1 HAS-PROP #2

#3: наблюдаемое значение свойства: #2 HAS-VALUE #3

Примеры порожденных составных понятий - экземпляров формулы:

1) "факт"("копытные", "численность", "217")

2) "факт"("ландыш", "начало цветения", "12.05.1994")

3) "факт"("травяное растение", "место обитания", "поляна Семеновская")

Имя: "причинно-следственная связь"

Родовое понятие: "объект"

Компоненты:

#1: причина: #1 ISA "факт"

#2: следствие: #2 ISA "факт"

Примеры экземпляров понятия-формулы:

1) "причинно-следственная связь"

("факт"("лето", "опенка", "жаркое"),

"факт"("грибы", "урожайность", "низкая"))

2) "причинно-следственная связь"

("факт"("насекомые", "численность", "меньше нормы"),

"факт"("птицы", "численность", "меньше нормы"))

Весьма важным понятием для описания данных экологического мониторинга является понятие параметра, которое обязательно должно быть определено в базе знаний. Как было указано в разделе 3, в качестве параметра наблюдения может выступать свойство объекта или среды, в котором он обитает, либо группа однородных объектов. В соответствии с этим приведем определения понятий "параметр-свойство" и "параметр-группа", принятые в БЗ системы.

Имя: "параметр-свойство"

Родовое понятие: "свойство"

Компоненты:

#1: носитель свойства: #1 ISA "объект"

#2: название параметра: #2 ISA "свойство", #1 HAS-PROP #2

#3: базовое множество: #3 ISA НАБОР(x), #2 HAS-VALUE x

#4: значение: #4 IN #3

Примеры порожденных составных понятий:

1) "параметр-свойство"("объект", "сезон", {"зима", "весна", "лето"}, "весна")

2) "параметр-свойство"("погода", "максимальная температура июля", {"20"..40"}, "37")

3) "параметр-свойство"("объект", "местоположение", {"квартал 1"..квартал 40а"}, "квартал 36")

Имя: "параметр-группа"

Родовое понятие: "объект"

Компоненты:

#1: общее имя группы: #1 ISA "объект"

#2: базовое множество: #2 ISA НАБОР(x), x ISA #1

#3: значение: #3 IN #2

Примеры порожденных составных понятий-экземпляров:

1) "параметр-группа"("грибы", {"мухомор", "бледная поганка", "ложный опенок"}, "мухомор");

2) "параметр-группа"("метеорологические явления", {"образование устойчивого снежного покрова", "санный путь", "разрушение устойчивого снежного покрова", "начало безморозных ночей", "первый сход снега", ...}, "санный путь").

Составные понятия используются в основном при описании документов (индексировании). Полагается, что каждый документ посвящен по крайней мере одному понятию, простому или составному.

6. Организация данных

Данные наблюдений объектов хранятся в *базе данных*. Способ представления данных зависит от характера данных (таблицы, графические построения, карты, тексты, изображения). Для каждой единицы данных в системе хранится описание, или индексная запись.

Индексные записи хранятся отдельно от данных и образуют индекс, с помощью которого осуществляется поиск информации. Индексная запись имеет следующее содержание:

- наименование документа;
- тип документа (таблица, графическое построение, карта, текст или изображение);
- указатель на документ (имя файла, идентификатор и т.п.);
- предмет наблюдения. Задается в виде набора понятий, в том числе экземпляров понятий-формул;
- параметры наблюдения - множество экземпляров составных понятий типа "параметр-свойство" или "параметр-группа". Параметры могут быть фиксированными (наблюдения проводились при одном и том же значении параметра) или переменными (наблюдения проводились при всех значениях параметра из заданного множества). Базовое множество фиксированного параметра состоит ровно из одного элемента.

Чаще всего предмет наблюдения задается при помощи составных понятий типа "факт". Во многих документах приводится ряд однородных фактов, для которых совпадают наблюдаемые объекты (переменная #1) и свойства объекта (переменная #2). Такие факты различаются только значениями свойства, найденными при различных значениях параметров. Выносить все значения свойств и все наборы значений параметров в индексную запись нецелесообразно: индексная запись должна служить для поиска информации, но не дублировать ее.

Большинство документов, в которых приводятся однородные наборы фактов, представляются в системе в виде наборов кортежей. Каждый кортеж содержит значения параметров и результаты наблюдений при этих значениях. Во всех кортежах, представляющих один документ, сохраняется порядок следования элементов, т.е. значения каждого параметра и каждого наблюдаемого свойства находятся во всех кортежах на одном и том же месте.

Индексная запись, соответствующая документам рассматриваемого вида, вместо конкретных значений фактов и параметров содержит ссылки на номера элементов в кортежах в виде &n, где n - номер элемента. Для документов других видов в качестве указателя наличия нескольких значений свойства или параметра используется специальный символ *.

6.1. Табличные данные

Табличные данные хранятся в виде кортежей, состоящих из значений наблюдаемых свойств и значений параметров, при которых были установлены эти значения свойств. Такой способ представления является универсальным, применимым независимо от сложности исходного табличного документа и числа параметров наблюдения.

Значения параметров берутся из базовых множеств, указанных в индексной записи соответствующего документа. Отображать в кортежах значения фиксированных параметров не имеет смысла, поскольку они одни и те же для всех кортежей. Поэтому при использовании кортеж пополняется неявно заданными (т.е. взятыми из индексной записи) значениями фиксированных параметров.

В качестве примера рассмотрим одну из таблиц Летописи природы Приокско-Террасного биосферного заповедника за 1992 год [2] (табл.25, стр.40):

Индексная запись:

Наименование: Динамика текущего отпада 160-летнего насаждения за период 1983-1992 гг.

Тип: Таблица

Предмет наблюдения:

"факт" ("Сосновое насаждение", "Число стволов", &2);

"факт" ("Сосновое насаждение", "Площадь поперечного сечения", &3);

"факт" ("Сосновое насаждение", "Запас", &4);

"факт" ("Сосновое насаждение", "Площадь боковой поверхности", &5);

Параметры:

"параметр-свойство" ("вселенная", "Год", {"1983"..1992"}, &1);

"параметр-группа" ("Сосновое насаждение",

{"160-летний сосняк"}, "160-летний сосняк");

"параметр-свойство" ("Сосновое насаждение", "Местоположение", {"квартал 36а"}, "квартал 36а");

Данные:

1983, 5, 0.41, 4.9, 77.4

1984, 5, 0.33, 3.8, 66.3

...

1992, 6, 0.48, 5.9, 100.0

6.2. Графические построения

Под графическими построениями понимаются графики, диаграммы, гистограммы и прочие способы наглядного показа количественных данных. Хранение подобных документов в виде готовых изображений нецелесообразно по следующим причинам:

1. Такое изображение легко построить автоматически при наличии исходных данных. Хранение исходных данных вместо изображения требует меньшего расхода памяти.

2. Любые изменения данных потребуют изменения изображения. При хранении изображений в готовом виде внесение в них изменений весьма затруднительно.

3. При хранении исходных данных сохраняется их точность.

4. При автоматическом построении изображения по данным пользователь может варьировать параметры (масштабы шкал, размеры изображения, количество одновременно отображаемых показателей, цветовое решение и т.д.), добиваясь максимальной выразительности и преследуя различные цели, например, вывод рисунка на печать для помещения его в отчет. Такие манипуляции невозможны при рассмотрении готового графического изображения.

В связи с этим в системе принято во всех случаях, когда возможно автоматическое воспроизведение графического изображения, хранить не само изображение, а исходные данные. Исходные данные хранятся так же, как и для табличных документов. Дополнительно указывается тип изображения (график, гистограмма и т.п.) и параметры визуализации, номенклатура которых зависит от типа изображения. Например, для графика должны сохраняться наименования показателей,

соответствующих осям, масштабы осей, цветовое решение для всех отображаемых показателей.

Рассмотрим график динамики температур по месяцам. Горизонтальная ось представляет параметр

"параметр-свойство" ("вселенная", "месяц", {"январь,1982".. "декабрь,1982"}, &1)

при этом 1 см оси соответствует 1 месяцу. Вертикальная ось задает шкалу для значений температуры от -40 до +40 градусов, 1 см соответствует 10 градусам. На графике отображаются факты:

"факт" ("погода", "минимальная температура", &2)

"факт" ("погода", "максимальная температура", &3)

"факт" ("погода", "средняя температура", &4)

причем первые два факта показываются тонкими линиями синего и красного цветов, третий факт отображается толстой линией черного цвета.

6.3. Карты

С точки зрения системы, карта состоит из следующих элементов:

- Картографическая основа, служащая в основном для ориентировки при рассмотрении тематической информации. Основа состоит из отдельных слоев, объединяющих однородные объекты (реки, леса, строения и т.п.). Каждый слой есть набор контуров этих объектов с указанием параметров отображения (общих для всех объектов слоя): цвет, штриховка, ширина линий и т.п..

- Набор территориально распределенных объектов, к которым привязана тематическая информация. Объекты могут быть точечными (размерность не выражается в масштабе карты), линейными (протяженные объекты, длина которых выражается в масштабе карты) и площадными (размерность выражается в масштабе карты). Местоположение точечных объектов указывается парой координат, для других видов объектов должны быть заданы контуры.

- Собственно данные с указанием, к какому объекту относится тот или иной элемент данных. На карте может одновременно быть показано несколько тематически различных наборов данных.

- Способ отображения каждого набора данных на карте и его параметры. В тематической картографии [3] приняты следующие способы картографического изображения тематических данных: значковый способ, точечный, знаки движения, локализованные диаграммы, изолинии, картограмма, картодиаграмма, количественный фон, качественный фон, ареалы. Параметры способов изображения - это

градации, на которые разбивается множество возможных значений отображаемого показателя, размеры значков, цветовое решение.

При наличии описания всех этих элементов карты сама карта может быть сгенерирована в тот момент, когда она понадобилась. С точки зрения реализации, ИИС работает во взаимодействии с некоторой геоинформационной системой (ГИС), которая отвечает за построение в нужный момент требуемой карты. При этом должны быть заранее подготовлены картографические слои, заданы координаты и контуры объектов. Данные, отображаемые на карте, хранятся так же, как табличные данные.

Вообще говоря, в качестве ГИС, ответственной за работу с картами, может быть выбрана любая из существующих геоинформационных систем, удовлетворяющая двум требованиям:

- 1) обеспечение достаточного для пользователей набора способов картографического изображения;

- 2) возможность автоматического построения карты по описанию, заданному извне.

Приведем пример карты, которая может быть предоставлена системой по запросу пользователя. На карте отображаются данные о плотности птичьих гнезд по кварталам заповедника. Кварталы являются площадными объектами, к ним относятся corteжи данных. Данные отображаются в виде картограммы, плотность гнезд выражается при помощи интенсивности окраски в желто-коричневых тонах, шкала возможных значений разбивается на 6 равномерных градаций.

6.4. Тексты

Тексты хранятся в текстовых файлах, в которых может быть задано разбиение на фрагменты. Фрагменты отделяются друг от друга служебными символами и помечаются идентификаторами. Индексная запись может задаваться как для файла в целом, так и для каждого фрагмента. Имеет смысл разбивать на фрагменты текст, в котором последовательно рассматривается несколько объектов или приводится ряд выводов относительно связей между объектами. Это позволит более точно находить информацию, релевантную запросу пользователя.

Приведем в качестве примера индексное описание текста из работы [2, стр.60], описывающего причины первого появления нового вида птиц в заповеднике:

Наименование: Дополнение к фауне Приокско-Террасного заповедника

Тип: текст

Предмет наблюдения:

"факт"("мухоловка-белошейка", "первое появление в заповеднике", "истина")

"факт"("мухоловка-белошейка", "местонахождение", *)

"причинно-следственная связь" ("факт" ("лето", "характеристика", "жаркое"), "факт"("мухоловка-белошейка", "первое появление в заповеднике", "истина"))

Параметры:

"параметр-свойство"("вселенная", "год", {1992}, 1992)

6.5. Изображения

Под изображениями понимаются неизменяемые графические документы (фотографии, рисунки, схемы и т.д.). Изображения хранятся в отдельных файлах в стандартных растровых форматах (PCX, BMP и т.д.). Необходимо заметить, что в разряд изображений не входят графики, диаграммы и прочие иллюстративные графические построения. Как и всем остальным документам, каждому изображению соответствует индексная запись, идентифицирующая отображаемый объект и параметры (место, время, способ съемки и т.д.)

7. Информационная система: взаимодействие с пользователем

Пользователь взаимодействует с информационной системой с целью получения и анализа необходимых ему данных. В связи с этим ИИС должна:

- 1) позволять пользователю сформулировать запрос, т.е. указать, какая информация его интересует;
- 2) обеспечивать выборку из информационной базы интересующих пользователя сведений;
- 3) представлять отобранные данные в удобном для анализа виде.

Взаимодействие пользователя с системой должно быть организовано таким образом, чтобы от пользователя не требовалось знание способа организации данных, структуры и содержания БЗ и других элементов внутреннего устройства. Диалог с системой должен вестись с использованием только терминов проблемной области (профессионального словаря пользователя) и общепринятой лексики.

7.1. Запросы и интерфейс по их составлению

В своем запросе пользователь должен указать системе, документы с какими понятиями и параметрами следует искать и какую информацию нужно из них извлечь и представить для анализа. Однако представление запроса в формальном виде, воспринимаемом системой, неудобно для пользователя и противоречит принятым принципам разработки ИИС. Поэтому при составлении запроса пользователь оперирует с *бланком запроса*. Бланк запроса отображается на экране компьютера, там же помещаются списки понятий проблемной области, определенных в базе знаний системы. Задача пользователя состоит в том, чтобы выбрать из списков интересующие его понятия и подставить в *поля* бланка запроса. Заполненный бланк автоматически переводится в формальную запись запроса, которая уже может непосредственно сопоставляться с формальными описаниями документов (индексными записями).

Для того, чтобы избежать построения бессмысленных запросов, для каждого поля бланка запроса задаются ограничения: какого типа понятие может быть подставлено в это поле и как оно должно быть связано с остальными понятиями, используемыми в запросе. Ограничения скрыты от пользователя, их выполнение достигается за счет того, что система предоставляет пользователю для выбора только те понятия, которые удовлетворяют ограничениям. Например, если в какое-то поле бланка должно подставляться понятие-объект, то система предоставит пользователю для выбора список понятий-объектов. Если в следующее поле бланка необходимо подставить свойство объекта, выбранного ранее, то система формирует список только из свойств данного объекта.

Таким образом, заполнение бланка запроса происходит последовательно, и уже выбранные понятия влияют на выбор остальных понятий.

7.1.1. Выбор понятий

Несмотря на наличие ограничений, список понятий для подстановки в бланк запроса может быть очень велик. В первую очередь это относится к выбору объекта, так как набор свойств ограничивается выбранным объектом-носителем. Для облегчения поиска пользователем нужного понятия создается не один

линейный список, включающий все понятия, а серия иерархически связанных списков в соответствии с иерархией объектов. При этом выбор объекта для подстановки в поле бланка в общем случае разбивается на ряд шагов. На каждом шаге производится выбор из относительно небольшого списка понятий, относящихся к одному уровню иерархии. В результате выбора одного из понятий списка текущий список заменяется списком понятий, непосредственно подчиненных выбранному в ISA-иерархии. Пользователь может продолжить процедуру выбора, перейдя к рассмотрению нового списка. Если пользователь отказывается от дальнейшего выбора, система подставляет в текущее поле бланка последнее выбранное понятие. Для удобства пользователя последнее выбранное понятие постоянно показывается на экране в выделенном месте.

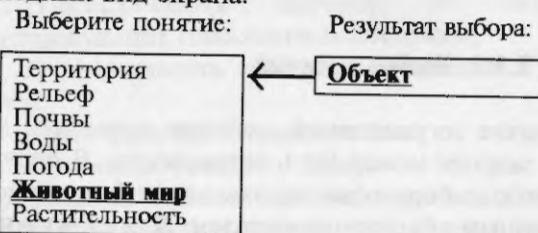
Поскольку существует возможность ошибочного выбора на каком-либо уровне иерархии, желательно предоставить пользователю возможность возврата к предыдущим выборам для их изменения. С этой целью на экране отображается весь путь, по которому пользователь пришел к текущему понятию, т.е. все родовые понятия в порядке их подчиненности. При выборе пользователем какого-либо из этих понятий система обеспечивает возврат к рассмотрению списка непосредственно подчиненных ему понятий.

В качестве дополнительного сервиса система позволяет пользователю локализовать положение определенного объекта в иерархии за счет контекстного поиска по полному или частичному его наименованию.

Рассмотрим пример выбора понятия "копытные". В примере схематически изображается часть экрана компьютера, в которой происходит процесс выбора; примеры бланков запросов будут приведены позднее.

Шаг 1.

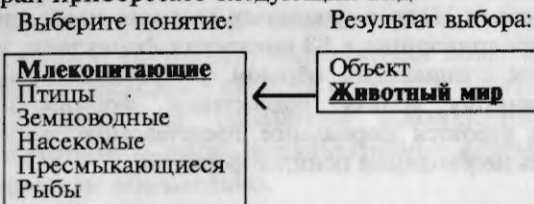
Исходный вид экрана:



Пользователь выбирает объект "животный мир".

Шаг 2.

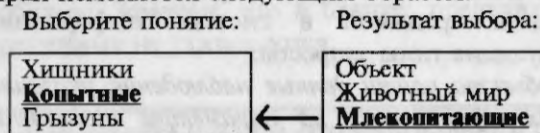
Экран приобретает следующий вид:



Пользователь выбирает "млекопитающие".

Шаг 3.

Экран изменяется следующим образом:



Выбрав объект "копытные", пользователь указывает системе, что выбор завершен.

При выборе понятий пользователю иногда удобно не только проходить по родо-видовой иерархии объектов, но перейти от объекта-системы к составным частям этой системы, от понятия, обозначающего некоторую среду обитания животных и растений, до объектов, населяющих эту среду. В связи с этим всегда, когда есть возможность, система предоставляет дополнительно к списку объектов, связанных с последним выбранным объектом отношениями ISA, списки объектов, связанных с этим объектом отношениями PART-OF и HAS-ENV.

Выбор понятий для полей, требующих подстановки свойства, производится после выбора объекта-носителя свойства. В связи с этим в список свойств для выбора включаются только свойства этого объекта. При этом учитывается наследование свойств: свойствами объекта являются также свойства родовых по отношению к нему объектов. Тем не менее список свойств остается относительно небольшим. Никаких специальных мер по его структуризации в системе не предпринимается, за исключением случаев наличия ISA-иерархий свойств (например, свойств группы "время") и составных свойств. Выбор родового свойства влечет появление списка свойств, подчиненных ему в ISA-иерархии. Если пользователь отказывается от дальнейшего выбора, это значит, что его устраивает любое из свойств данной группы. Выбор составного свойства приводит к появлению списка составляющих его свойств. При отказе от выбора в этом списке система считает, что пользователя интересуют все компоненты составного свойства.

Необходимо заметить, что пользователю не предоставляются для выбора понятия-формулы, так как их использование связано с определенными сложностями. Если в некоторых запросах необходимо использовать понятия, представленные в БЗ понятиями-формулами, то в этих целях используются специальным образом построенные бланки запросов. По заполненному бланку подсистемой формирования запросов автоматически строится формальное представление запроса, которое и будет включать необходимые понятия-формулы.

7.1.2. Бланки запросов

Бланки запросов составляются в соответствии с типами запросов, с которыми пользователь может обратиться к системе. Исходя из содержания хранимой в системе информации, рассматриваются следующие типы запросов:

1) *Для заданного объекта найти данные наблюдения указанных свойств при заданных ограничениях на параметры наблюдения.* Другими словами, пользователь задает системе вопрос: Какими были значения указанных свойств объекта при заданных условиях?

2) *Для заданного объекта найти, при каких значениях параметров достигались заданные значения свойств.*

3) *Сопоставить данные наблюдений свойств по крайней мере двух различных объектов при одних и тех же значениях параметров.* Такие запросы направлены на выявление взаимосвязей между явлениями природы.

4) *Найти указанный объект.* Данный запрос может применяться для поиска изображений или карт, на которых отображен интересующий пользователя объект.

5) *Запросы на поиск экземпляров понятий-формул.* Например, пользователя может интересовать, в каком документе приводится тот или иной факт, что является причиной низкого урожая грибов в 1992 году и т.п. По сути, с каждым понятием-формулой связан свой тип запроса и, следовательно, должен быть задан отдельный бланк запроса. Следует сказать, что способ определения понятий-формул в БЗ системы позволяет последней генерировать нужные бланки автоматически, по мере необходимости.

Выбор типа запроса производится по меню, включающем текстовые описания смысла запросов первых 4-х типов и наименования всех понятий-формул, присутствующих в базе знаний системы. После определения типа запроса пользователю предлагается заполнить бланк запроса.

Бланк запроса представляет собой предложение или набор предложений на естественном языке с пропусками - полями для

заполнения. Некоторые фрагменты бланка являются повторяемыми, т.е. предоставляют возможность включения в запрос нескольких однородных частей, заполненных различными понятиями. Например, пользователь может выбрать и подставить в запрос несколько свойств объекта и задать произвольное число ограничений на параметры наблюдений. В бланке могут присутствовать также факультативные фрагменты, заполнять поля которых не обязательно.

Рассмотрим бланки, связанные с каждым типом запроса. Поля бланка будем отмечать переменной вида &i, где i - номер поля в бланке запроса. Повторяемые фрагменты будем выделять фигурными скобками, факультативные фрагменты - квадратными скобками. Заметим, что в бланке, предъявляемом пользователю, переменные не указываются.

1)

Для объекта (группы объектов с общим именем) &1
найти данные наблюдения {свойства &2}
при условии, что [{параметр &3 принимает значение &4}]
в зависимости от
[{параметра &5, принимающего значения из области &6}]
[по следующим объектам группы: {&7}]

С этим бланком запроса связаны следующие ограничения:

&1 ISA "объект" - это значит, что в первое поле бланка должно подставляться понятие-объект.

&1 HAS-PROP &2 - во второе поле бланка должно подставляться свойство объекта, выбранного для первого поля. Таких свойств может быть выбрано несколько.

(&1 HAS-PROP &3) OR

((&1 PART-OF &x) AND (&x HAS-PROP &3)) OR

((&1 HAS-ENV &x) AND (&x HAS-PROP &3)) - третье поле бланка предназначено для задания одного или нескольких фиксированных параметров. В качестве фиксированного параметра наблюдения может выступать свойство выбранного объекта, свойство среды данного объекта либо свойство другого объекта, частью которого является выбранный объект.

&3 HAS-VALUE &4 - в четвертое поле должно подставляться значение фиксированного параметра.

(&1 HAS-PROP &5) OR

((&1 PART-OF &x) AND (&x HAS-PROP &5)) OR

((&1 HAS-ENV &x) AND (&x HAS-PROP &5)) - пятое поле предназначено для задания изменяемых параметров. Ограничения аналогичны тем, что указаны для фиксированных параметров.

(&x IN &6) AND (&5 HAS-VALUE &x) - в шестом поле указывается множество значений изменяемого параметра.

(&x IN &7) AND (&x ISA &1) - седьмое поле предназначено для задания параметра-группы. На этом месте должно быть указано множество видовых объектов по отношению к ранее выбранному объекту.

Пример заполненного бланка запроса первого типа:

Для объекта (группы объектов с общим именем) "погода"
найти данные наблюдения
свойства "минимальная температура воздуха"
и свойства "максимальная температура воздуха"
в зависимости от
параметра "месяц", принимающего значения
из области (ноябрь, 1991)..(октябрь, 1992)
и параметра "место наблюдений", принимающего значения
из области {Долы, полигон, кв. 36а}

В данном запросе не заданы фиксированные параметры и параметр-группа, в связи с этим соответствующие факультативные фрагменты исключены из запроса.

2)

Для объекта &1 установить,
при каких значениях {параметра &2}
{свойство &3 принимало значение &4}

Ограничения:

&1 ISA "объект" - в первое поле бланка подставляется понятие-объект.

(&1 HAS-PROP &2) OR

((&1 PART-OF &x) AND (&x HAS-PROP &2)) OR

((&1 HAS-ENV &x) AND (&x HAS-PROP &2)) - параметр определяется так же, как и в предыдущем типе запроса.

&1 HAS-PROP &3 - в третьем поле должно быть свойство выбранного объекта.

(&3 HAS-VALUE &4) OR (&4="минимум") OR (&4="максимум") - в четвертом поле должно быть значение указанного свойства либо одно из ключевых слов "минимум" или "максимум". Для свойств с нечисловыми значениями использование этих ключевых слов не имеет смысла.

Пример запроса второго типа:

Для объекта "погода" установить,
при каких значениях параметра "дата"
и параметра "место наблюдения"
свойство "температура воздуха" принимало значение
"минимум"

3)

Для объектов &1 и &2
выявить взаимосвязь между значениями
{свойства &3} первого объекта
и {свойства &4} второго объекта
при условии, что [{параметр &5 принимает значение &6}]
в зависимости от
{параметра &7, принимающего значения из области &8}]

Для краткости, не будем приводить ограничения для данного и последующих типов запросов, так как они задаются аналогично ограничениям для первых двух типов.

Необходимо заметить, что набор понятий, подставляемых в качестве параметров в данный бланк запроса, ограничивается как первым, так и вторым из выбранных для сопоставления объектов: параметр должен быть либо общим свойством обоих объектов, либо свойством среды, общей для объектов, либо свойством другого объекта, частями которого являются оба выбранных объекта.

Пример запроса:

Для объектов "лето" и "клубника"
выявить взаимосвязь между значениями
свойства "характеристика сезона" первого объекта
и свойства "урожайность" второго объекта
в зависимости от
параметра "год", принимающего значения из области
1983..1992

4)

Найти объект &1
при условии, что {параметр &2 принимает значение &3}

Например:

Найти объект "Дубрава"
при условии, что параметр "сезон" принимает значение
("лето", 1992)

Данный запрос поручает системе найти фотоснимок или рисунок дубравы, сделанный летом 1992 года, либо карту, отображающую территорию заповедника по состоянию на лето 1992 года, на которой показана дубрава.

5) Бланк запроса на поиск экземпляра понятия-формулы составляется системой автоматически по определению понятия-формулы, заданному в базе знаний, по следующей схеме:

Найти <имя понятия-формулы>:
<имя-компоненты>: &1
<имя-компоненты>: &2
...
<имя-компоненты>: &N
при условии, что [{параметр &N+1 принимает значение
&N+2}]

где N - число компонент понятия-формулы. В качестве ограничений на поля бланка принимаются ограничения, заданные для переменных понятия-формулы в его определении. Параметры подчиняются тем же ограничениям, что и в запросах других типов.

В связи с формальным способом построения запросов данного типа, в них не учитывается то обстоятельство, что для одной и той же формулы пользователь мог бы построить несколько различных запросов, в зависимости от того, какая из компонент считается известной, а какую необходимо найти. Например, для понятия-формулы "причинно-следственная связь" запросы могут быть следующего вида:

а). Найти документ, в котором говорится о причинно-следственной связи заданных фактов.

б). Найти следствие заданного факта.

в). Найти причину указанного факта.

С целью предоставления пользователю возможности формулировать подобные запросы принимается, что пользователь может подставлять на место некоторых компонент символ "?" (вопросительный знак). Это означает, что система должна:

1) найти документ, в котором рассматривается указанное понятие-формула при заданных значениях параметров, причем компоненты, заданные в запросе пользователя, совпадают с компонентами, указанными в документе;

2) найти понятия, подставляемые на место символов "?" в сформированный пользователем запрос.

Пример запроса подобного вида:

Найти факт:
наблюдаемый объект: "яблоня"
свойство объекта: "начало цветения"
наблюдаемое значение свойства: ?
при условии, что параметр "год" принимает значение "1992"

7.2. Поиск и выборка информации

7.2.1. Интерпретация запросов

Заполненный пользователем бланк запроса должен быть переведен системой во внутренний формат, пригодный для сравнения с описанием документа. Для удобства и быстроты сравнения внутренний формат запроса должен совпадать с форматом описания документа, точнее, той его части, в которой описывается содержание документа: предмет наблюдения и параметры. Следовательно, в формальной записи запроса должны присутствовать 2 раздела: предмет наблюдения и параметры. Предмет наблюдения, как и в описании документа, представляет собой набор понятий, в том числе экземпляров понятий-формул. Раздел "параметры" состоит из экземпляров понятий-формул типа "параметр-свойство" или "параметр-группа". Раздел "параметры" формируется на основе фрагментов запроса, начинающихся словами "при условии, что ...", "в зависимости от ...", "по следующим объектам группы...". Способ перевода этих фрагментов запроса в понятия-формулы очевиден, необходимо лишь указать, что на место значения изменяемого параметра подставляется спецсимвол, обозначающий, что значение может быть произвольным в пределах базового множества. Например:

1) ...: при условии, что параметр "сезон" принимает значение "весна" переводится в: параметр-свойство (<объект, указанный в запросе>, "сезон", {"весна"}, "весна");

2) ... в зависимости от параметра "местоположение", принимающего значения из области {"квартал 36а", "квартал 40", "квартал 41"} переводится в: параметр-свойство(<объект, указанный в запросе>, "местоположение", {"квартал 36а", "квартал 40", "квартал 41"}, ?), где "?" - спецсимвол.

Раздел "предмет наблюдения" для запросов типа 1,2,3 состоит из экземпляров понятия "факт" с подстановкой указанных в запросе объектов, свойств и их значений, либо спецсимвола вместо значений свойств. Для запроса типа 4 этот раздел содержит имя объекта, интересующего пользователя. Перевод в понятия-формулы запросов 5-го типа очевиден, так как они представляют собой не что иное, как записи экземпляров понятий-формул.

Пример:

Запрос: <Для объекта "яблоня" найти данные наблюдения свойства "начало цветения" ...> переводится в: факт("яблоня", "начало цветения", ?)

7.2.2. Выборка информации

Поиск релевантных запросу документов происходит следующим образом. Система сопоставляет раздел "предмет наблюдения" формальной записи запроса (в дальнейшем изложении - запроса) с соответствующим разделом описания документа. При этом каждое понятие, включенное в запрос, сопоставляется со всеми понятиями, содержащимися в описании документа. Будем считать, что понятие А запроса соответствует понятию В документа в следующих случаях:

а) А и В - простые понятия и выполняется: $V \text{ ISA } A$;
б) А и В - составные понятия, в том числе экземпляры понятий-формул, и выполняется:

(б.1) $V \text{ ISA } A$;

(б.2) для каждой компоненты СА понятия А и соответствующей компоненты СВ понятия В:

(б.2.1) СА - спецсимвол $\langle ? \rangle$;

(б.2.2) СВ - ссылка на номер элемента в кортеже;

(б.2.3) СВ ISA СА, если СВ и СА - понятия;

(б.2.4) СВ IN СА, если СВ - понятие, СА - множество понятий;

(б.2.5) $\text{СВ} * \text{СА} \langle \rangle = 0$, если СВ и СА - множества понятий.

Если ни одно из понятий раздела "предмет наблюдения" описания документа не соответствует какому-либо понятию того же раздела запроса, процесс сопоставления запроса с этим документом прекращается и система переходит к анализу следующей записи индекса, соответствующей следующему документу. Если же какие-либо понятия соответствуют, то система переходит к проверке соответствия параметров. Документ считается релевантным запросу, если для всех параметров, заданных в запросе, имеются соответствующие параметры в описании документа. При этом соответствие параметров понимается так же, как и соответствие остальных понятий.

Независимо от того, признан ли текущий документ релевантным, система просматривает индекс до конца и выбирает ВСЕ релевантные документы. Тексты и изображения, отобранные по запросу, предъявляются пользователю без изменений. В то же время над таблицами, картами (точнее, описаниями карт) и описаниями графических построений выполняются следующие действия:

- исключаются элементы кортежей, соответствующие объектам или свойствам, отсутствующим в запросе; для карт исключаются группы отображаемых объектов, отсутствующие в запросе, и соответствующие им таблицы;

- исключаются кортежи, значения параметров в которых не удовлетворяют заданным пользователем ограничениям;

- если отобрано несколько таблиц с различными показателями, но одними и теми же параметрами и с совпадающими или пересекающимися базовыми множествами этих параметров, то эти таблицы объединяются путем добавления в кортежи одной из них новых элементов, соответствующих показателям из других таблиц. При этом в случае частичного несовпадения базовых множеств нескольких параметров в результирующей таблице имеются пропуски;

- если отобрано несколько таблиц с одними и теми же показателями и с совпадающим набором параметров, но различными (т.е. непересекающимися) базовыми множествами значений параметров, то эти таблицы объединяются путем добавления к одной из них кортежей из остальных таблиц.

Таким образом, за счет описанных действий над табличными документами системой могут быть сгенерированы новые документы, ранее отсутствовавшие в базе данных. Благодаря исключению нерелевантной информации предъявляемые пользователю документы содержат только ту информацию, которая его интересует. В целом вследствие способности системы к преобразованию документов достигаются следующие цели:

1). Возможность сведения в один документ информации из нескольких исходных документов в интересах сопоставления свойств объектов и выявления взаимосвязей.

2). Возможность автоматического построения документа, отражающего зависимость показателей от параметров.

Рассмотрим пример. Пусть в базе данных присутствует информация по ряду свойств объекта за несколько лет, но данные за каждый год хранятся в разных табличных документах различной структуры. По запросу пользователя система может создать новую таблицу, отражающую динамику свойств за указанный пользователем период.

7.3. Представление отобранной информации

В случае, если по запросу пользователя было отобрано несколько документов, их необходимо представить для просмотра в связанном виде. Подходящим способом представления связанных разнородных документов является *гипертекст* [4,5]. Каждому документу соответствует *узел* гипертекста, его представляющий. *Ссылки* между документами (и узлами гипертекста) устанавливаются таким образом, чтобы обеспечивать переходы от документа к документу для полного просмотра всей релевантной информации. Между парой документов устанавливается ссылка в том случае, если в них рассматривается общее понятие. В целом документы связываются между собой таким образом, чтобы

вначале рассматривались текстовые документы, затем таблицы, графические построения, карты и изображения. Дополнительно устанавливаются ссылки от текста к изображению или карте, если в тексте идет речь об объекте, показанном на этом изображении или карте. От карт и изображений устанавливаются ссылки к описаниям объектов, отображенных на них.

Кроме документов, отобранных или созданных системой при интерпретации запроса пользователя, система может автоматически создавать новые документы - графические построения, иллюстрирующие таблицы. Приведем несколько примеров:

1). По таблице, показывающей динамику некоторого числового показателя по годам (месяцам и т.п.), может быть построен график, наглядно отображающий эту динамику.

2). Зависимость свойств объекта от его местоположения удобно анализировать при помощи карты. Подходящий способ картографического изображения и его параметры могут быть автоматически подобраны системой.

3). По данным о численности животных по видам может быть построена структурная диаграмма, показывающая соотношение животных различных видов.

В ряде случаев возможно построение нескольких изображений, дополняющих друг друга с точки зрения возможностей визуального анализа информации. Изображения могут быть как двумерными, так и трехмерными. Иногда удобно рассматривать не только статичные изображения, но и динамические: мультфильм, показывающий динамику графического построения в зависимости от времени или какого-либо другого показателя.

Способность системы автоматически строить иллюстративные графические изображения достигается за счет включения в нее проблемно-независимой базы знаний по графическим построениям, задающей правила выбора подходящего способа изображения. Выбор зависит от 1) характеристик отображаемых свойств; 2) характеристик параметров; 3) количества одновременно рассматриваемых свойств; 4) количества параметров, от которых зависят рассматриваемые свойства. Принимаются во внимание следующие характеристики шкалы значений свойств и параметров:

- упорядоченная или неупорядоченная шкала;
- дискретная или непрерывная шкала;

- значения абсолютные или относительные;
- являются ли значения пространственно-распределенными.

В последнем случае система может автоматически сгенерировать компьютерную карту, используя базу знаний по тематической картографии [6].

На основе правил выбора способов графической иллюстрации данных система может самостоятельно сгенерировать следующие виды документов: графики; гистограммы; структурные диаграммы различной формы; интегрированные изображения, состоящие из набора наглядных значков; поверхности в перспективном изображении; карты с различными способами картографического изображения. Документы всех перечисленных видов могут быть как статическими, так и динамическими.

Параметры графических изображений подбираются системой автоматически. Однако система предоставляет пользователю возможность варьирования параметров в интересах достижения наиболее выразительного изображения.

8. Информационная система: взаимодействие с экспертом

Под экспертом понимается сотрудник заповедника, ведущий работу по пополнению базы данных системы и использующий систему для генерации отчетных документов. Кроме того, время от времени может возникать необходимость в пополнении или корректировке базы знаний в связи с изменениями, происходящими в заповеднике (например, появление новых видов), или новыми требованиями, предъявляемыми к деятельности заповедника (например, необходимость учета новых показателей). Таким образом, выделяется 3 основных вида взаимодействия эксперта с ИИС: внесение данных, генерация отчетных документов и модификация базы знаний.

8.1. Внесение данных

По возможности все данные хранятся в системе в виде файлов стандартных форматов из соображений их доступности для других программ и систем. Для табличных данных таким форматом является DBF, для растровых изображений - BMP и PCX. Для картографической информации в состав системы включается ряд

конверторов для преобразования внутреннего формата векторных карт системы в форматы систем AutoCAD и ProDesign и из них.

Возможны 2 режима ввода информации в базу данных: (1) диалоговый, применяемый для таблиц, графических построений, карт и текстов; (2) пакетный: подключение к базе данных готовых текстов, растровых изображений; импорт данных из баз данных в формат DBF. В обоих случаях собственно вводу данных предшествуют следующие шаги:

1. Эксперт формулирует название вводимого документа.
2. Эксперт указывает тип документа (текст, таблица и т.д.).
3. Эксперт описывает содержание документа: какие объекты в нем рассматриваются, какие свойства от каких параметров зависят, каковы базовые множества параметров.

Описание содержания документа происходит аналогично процедуре построения запроса пользователем, т.е. путем заполнения бланка, который затем автоматически переводится во внутреннее представление, используемое в индексной записи. Применяются бланки 2-х типов: бланк для документов, представляющих результаты наблюдений свойств объектов и бланк для документов, содержащих экземпляры понятий-формул.

Бланк первого типа имеет следующий вид:

В документе рассматривается
{ объект ... [со следующими свойствами: {...}] }
[при следующих фиксированных параметрах:
{ параметр ... принимает значение ... }]
[в зависимости от
{ параметра ... принимающего значения из области ... }]
[по группе объектов {...}]

Как и в случае запросов, бланки описаний по понятиям-формулам строятся автоматически. Для выбора подходящего бланка описания эксперту предоставляется список, состоящий из наименования "Данные наблюдений свойств объектов" и наименований имеющихся в базе знаний понятий-формул. Подстановка необходимых понятий в поля бланков производится так же, как и при формировании запроса пользователем.

После заполнения экспертом бланка описания документа система формирует индексную запись о новом документе. Затем, если ввод осуществляется в пакетном режиме, система запрашивает имя файла с данными и устанавливает соответствующую ссылку в индексной записи. При импорте табличных данных эксперт дополнительно указывает системе, что

содержится в каждом поле данных, т.е. устанавливает соответствие между полями данных и заданными в описании документа свойствами и параметрами.

При вводе документов в диалоговом режиме эксперту предоставляются средства заполнения документа в зависимости от его типа.

1. Для ввода текстов используется текстовый редактор.

2. При вводе табличных данных система создает форму для ввода в виде таблицы, столбцы которой соответствуют параметрам и свойствам. Все возможные значения параметров заранее подставляются системой в соответствующие поля таблицы. Кроме того, в случае заполнения сводной таблицы, часть данных которой уже есть в базе данных, эти данные также автоматически заносятся в формируемую таблицу. Эксперту остается только заполнить пропуски значениями свойств для уже указанных системой значений параметров.

3. Для ввода графических построений вводится соответствующая таблица и указываются вид и параметры графика.

4. При формировании картографических документов вначале нужно подготовить картографическую основу. Для этого производится импорт векторных изображений известных форматов. Данная процедура выполняется только один раз при вводе первого картографического документа. Данные в документ вносятся в виде таблиц, которые привязываются к территориально-распределенным объектам карты. По описанию таблиц система сама определяет территорию, которая отображается на карте, после чего эксперту предъявляется карта с отмеченными на ней объектами. Выбор способа картографического изображения и его параметров производится экспертом либо самостоятельно, либо на основе базы знаний по тематической картографии.

8.2. Генерация отчетных документов

Как указывалось в разделе 2, в состав Летописи природы включаются таблицы, тексты, графические построения, карты и изображения. Система обеспечивает печать любого имеющегося в базе данных или сгенерированного по запросу документа с учетом возможностей используемого принтера (размер страницы, разрешение, ориентация бумаги и пр.). Перед выводом на печать система отображает документ на экране в том виде, в каком он

будет отображен на бумаге, и дает возможность эксперту изменить внешний вид документа: разбивку на страницы; шрифты, размер символов, выравнивание текста для текстовых документов; масштаб, цветовое решение и набор отображаемых объектов для карт и графических построений и т.п.

Табличные документы требуют дополнительной предобработки, поскольку их внешний вид, расположение и наименование строк и столбцов регламентируются методическими рекомендациями [1]. Система хранит все регламентируемые форматы таблиц и автоматически представляет табличные данные в требуемом виде при выводе на печать. В системе принята следующая дисциплина подготовки табличных документов к печати.

После выбора или создания документа система ищет готовый формат таблицы для данных такого вида. При этом принимаются во внимание описываемые объекты, свойства и параметры, но не значения свойств и параметров. Если подходящий формат найден, то в соответствии с ним строится образ табличного документа. В противном случае образ строится автоматически, при этом столбцы соответствуют параметрам и свойствам объектов и помечаются их наименованиями.

На следующем шаге эксперту предоставляются средства преобразования таблицы. Он может: изменить название столбца; изменить порядок столбцов таблицы; изменить ориентацию таблицы (превратить строки в столбцы и наоборот); добавить новый столбец; удалить столбец или строку и т.д. Путем указанных преобразований эксперт приводит таблицу к требуемому виду. Созданный или измененный экспертом формат таблицы может быть сохранен для дальнейшего использования.

Таким образом, формат таблицы задается экспертом неявно, путем преобразования заготовки, подобранной или созданной системой.

При формировании отчетных документов имеет значение вопрос подбора печатаемых документов и их последовательности. Для удобства работы эксперта система создает оглавления имеющихся в базе данных документов за каждый год. Структура оглавления соответствует принятой структуре летописи природы и последовательности ее разделов.

8.3. Модификация базы знаний

База знаний системы отображает классификацию природных объектов и задает свойства этих объектов. Эти знания независимы от структуры и содержания данных, регламентированных форматов документов и т.п. Поэтому разумно предположить, что в процессе эксплуатации системы не возникнет необходимости коренного изменения ее базы знаний даже в случае выработки новых стандартов ведения заповедного дела и новых форматов отчетных документов.

Однако БЗ не может оставаться и полностью неизменной, поскольку возможно изменение видового состава животных и растений заповедника, появление необходимости в учете ранее не учитываемых свойств объектов, появление интереса к анализу ранее не рассматриваемых закономерностей и связей. Таким образом, в системе должна обеспечиваться возможность пополнения БЗ новыми объектами и свойствами, частичного изменения ее структуры, дополнения системы рассматриваемых понятий и отношений между ними. Для этих целей применяются как редактор базы знаний, так и специализированные модули автоматизированного приобретения экспертных знаний, стимулирующие работу эксперта. Эти методы используют ряд психологических приемов, игровые процедуры и т.д. [7-10], помогающие эксперту вспомнить и сформулировать соответствующие знания. Подробное рассмотрение этих вопросов выходит за рамки настоящей работы.

9. Замечания о реализации

Из описания проекта системы следует обязательность следующих требований к программной реализации ИИС:

1. Организация многооконного графического интерфейса системы с пользователем.
2. Стандартизация способов взаимодействия всех категорий пользователей с системой.
3. Обеспечение работы системы на различных конфигурациях технических средств, имеющихся в наличии в заповедниках; возможность работы с различными типами периферийных устройств (мониторы, принтеры).
4. Обеспечение возможности одновременной работы ИИС с другими прикладными программами для организации межпрограммного обмена информацией.

Исходя из перечисленных требований, в качестве среды разработки принята MS Windows 3.1, язык программирования системы - BORLAND C++. Вследствие этого для функционирования системы нужен IBM-совместимый персональный компьютер, обеспечивающий работу MS Windows в стандартном или расширенном режиме. Это соответствует следующим минимальной и рекомендуемой конфигурациям аппаратуры:

	Минимальная конфигурация	Рекомендуемая конфигурация
Процессор	80286	80386/80486
Арифметический сопроцессор	не обязателен	желателен
Оперативная память	1 Mb	4 Mb и более
Видеомонитор	произвольный	VGA/SVGA
Свободная память на диске	1 Mb	1 Mb и более
Дополнительные устройства	мышь, принтер	

Порядок реализации проекта основывается на общих принципах разработки систем, основанных на знаниях [11]. Вначале реализуется прототип системы, демонстрирующий принципиальную возможность создания системы и ее основные функции. На данном этапе будут реализованы следующие модули системы:

- интерфейс составления запроса пользователем;
- интерфейс диалогового ввода и индексации документов;
- модуль поиска информации.

Кроме того, будут созданы прототипы базы знаний о проблемной области и БЗ по когнитивному представлению информации.

Эксперименты и апробирование прототипа позволят выработать окончательные требования к ИМС, после чего будет осуществлена разработка рабочей версии системы в объеме, заданном функциональной схемой, которая приведена в 4-м разделе.

Работа выполнена в рамках Государственной научно-технической программы "Экологическая безопасность России" Министерства экологии и природных ресурсов (задания N 5.4.4.9, 11.3.6, 5.4.3.10).

Приложение. Пример ввода и использования данных по метеорологическим наблюдениям

Рассмотрим работу различных категорий пользователей системы с одним из разделов Летописи природы - данными метеорологических наблюдений. В базе знаний проблемной области описывается следующая система понятий:

Объект "погода" имеет свойства "средняя температура дня", "максимальная температура дня", "минимальная температура дня", "относительная влажность на 15 час", "среднесуточное давление", "сумма осадков", "скорость ветра (средняя)", "скорость ветра (максимальная)", "преобладающее направление ветра", "высота снежного покрова" и "атмосферные явления". Все свойства, кроме последнего, имеют тип "число", свойство "атмосферные явления" имеет тип "строка". Перечисленные свойства объединены в составное свойство "метеорологические наблюдения".

Эксперт задает форму для ввода данных метеорологических наблюдений за день и описание отчетной формы "Таблица метеорологических наблюдений за месяц".

Данные наблюдений за каждый день хранятся в таблице со следующей индексной записью:

Наименование: Метеорологические наблюдения

Тип: Таблица

Предмет наблюдения:

"факт" ("Погода", "средняя температура дня", &2);

"факт" ("Погода", "минимальная температура дня", &3);

...

"факт" ("Погода", "высота снежного покрова", &11);

"факт" ("Погода", "атмосферные явления", &12);

Параметры:

"параметр-свойство" ("вселенная", "Дата", {1.1.1980..31.12.2000}, &1);

"параметр-свойство" ("Погода", "Место наблюдения", {"станция фонового мониторинга"}, "станция фонового мониторинга");

Для ввода данных оператор использует форму:

Таблица метеорологических наблюдений

Дата: _____
Температура:
средняя: _____ макс: _____ мин: _____
Относительная влажность на 15 ч (%): _____
Среднесуточное давление (мб): _____
Сумма осадков (мм): _____
Скорость ветра (м/с), средняя: _____ максимальная: _____
Преобладающее направление ветра: _____
Атмосферные явления: _____
Высота снежного покрова (см): _____

Пример хранимых данных:

01.01.1992 -13.4 -4.2 ... 16 *uv
02.01.1992 -8.6 -4.6 ... 16 *u
...
31.01.1992 -3.9 -1.0 ... 23 *Lu

Примеры запросов различных типов, с которыми пользователь может обратиться к системе:

Для объекта "ПОГОДА"
найти данные наблюдения "СУММА ОСАДКОВ",
"ВЫСОТА СНЕЖНОГО ПОКРОВА"
в зависимости от
параметра "ДАТА",
принимającego значения из области 1/1/1992..13/1/1992

Для объекта "ПОГОДА" установить,
при каких значениях параметра
"СРЕДНЕСУТОЧНОЕ ДАВЛЕНИЕ"
свойство "СУММА ОСАДКОВ"
принимало значение МАКСИМУМ

Для объекта "ПОГОДА"
выявить взаимосвязь между значениями
свойства "КОЛИЧЕСТВО ОСАДКОВ"
и свойства "ТЕМПЕРАТУРА ВОЗДУХА (СРЕДНЯЯ)"
при условии,
что параметр "ПРЕОБЛАДАЮЩЕЕ НАПРАВЛЕНИЕ
ВЕТРА"
принимает значение 90..180 градусов
в зависимости от
параметра "ДАТА", принимающего значения из области
1/1/1992..31/12/1992

Найти ФАКТ:
НАБЛЮДАЕМЫЙ ОБЪЕКТ: ПОГОДА
НАБЛЮДАЕМОЕ СВОЙСТВО: АТМОСФЕРНЫЕ
ЯВЛЕНИЯ
ЗНАЧЕНИЕ: **uv

Отметим, что, помимо отбора интересующих пользователя данных, система автоматически генерирует ряд графических построений, наглядно демонстрирующих особенности данных. Характеристики данных и содержание запросов 1) и 3) позволяют системе автоматически построить графики динамики соответствующих показателей.

На основании заданного экспертом описания отчетных форм могут быть сгенерированы и напечатаны отчетные документы.

Литература

1. К.П.Филонов, Ю.Д.Нухимовская. Летопись природы в заповедниках СССР. М.: Наука, 1985. 143 с.
2. Летопись природы. Приокско-Террасный государственный биосферный заповедник. Книга 46. - Данки, 1992. 202 с.
3. К.А.Салишев. Картография. М.: Высшая школа, 1971. 248 с.
4. J.Conklin. A survey of hypertext / IEEE Computer (1987), V.20, N.9. P.17-41.
5. И.П.Белецкая. Гипертекстовые системы: современное состояние. "Знания-диалог-решение". Отчет РГ-26 КНВВТ. Киев, 1990 Ч.1. С.269-282.

6. A.V.Simonov, G.L.Andrienko, N.V.Andrienko, L.P.Simonova. Cartographic foundations of the intellectual GIS. Intercarto'94 Int.Conf. GIS World, 1994.

7. Г.Л.Андриенко. Методы сопоставления для автоматизированного приобретения экспертных знаний. Автореф. дис. ... канд.физ.-мат.наук. М.: МГУ, 1992. 15 с.

8. Н.В.Андриенко. Инструментальные средства формирования баз процедурных знаний. Автореф. дис. ... канд.физ.-мат.наук. М.: МГУ, 1993. 16 с.

9. Г.Л.Андриенко, Н.В.Андриенко. Игровые процедуры сопоставления в инженерии знаний. Доклады III конф. по ИИ. Тверь, 1992. Т.1, С.93-96.

10. G.L.Andrienko, N.V.Andrienko. AFORIZM approach: creating situations to facilitate expertise transfer. Lecture Notes in Artificial Intelligence, Springer Verlag, 1994. (in press)

11. Э.В.Попов. Экспертные системы: решение неформализованных задач в диалоге с ЭВМ. М.: Наука, 1987. - 288 с.

СОДЕРЖАНИЕ

1. Мотивация.....	3
2. Цель работы.....	4
3. Структура и содержание Летописи природы заповедника	7
4. Подход. Общая схема информационной системы.....	11
5. Организация базы знаний информационной системы.....	15
5.1. Понятия и их значения.....	15
5.2. Отношения и их роль в базе знаний.....	16
5.2.1. Родо-видовое отношение. Иерархии понятий	16
5.2.2. Отношение "имеет свойство". Наследование свойств	19
5.3. Составные понятия	20
6. Организация данных.....	23
6.1. Табличные данные	24
6.2. Графические построения.....	25
6.3. Карты	26
6.4. Тексты.....	27
6.5. Изображения	28
7. Информационная система: взаимодействие с пользователем.....	28
7.1. Запросы и интерфейс по их составлению.....	29
7.1.1. Выбор понятий	29
7.1.2. Бланки запросов.....	32
7.2. Поиск и выборка информации	37
7.2.1. Интерпретация запросов.....	37
7.2.2. Выборка информации	38
7.3. Представление отобранной информации.....	39
8. Информационная система: взаимодействие с экспертом	41
8.1. Внесение данных.....	41
8.2. Генерация отчетных документов.....	43
8.3. Модификация базы знаний.....	45
9. Замечания о реализации.....	45
Приложение.	47
Литература	49

6. А.В.Сидоров, Г.И.Сидорова, Н.А.Сидорова	1
Содержание	1
1. Матрица	1
2. Матрица	1
3. Матрица	1
4. Матрица	1
5. Матрица	1
6. Матрица	1
7. Матрица	1
8. Матрица	1
9. Матрица	1
10. Матрица	1
11. Матрица	1
12. Матрица	1
13. Матрица	1
14. Матрица	1
15. Матрица	1
16. Матрица	1
17. Матрица	1
18. Матрица	1
19. Матрица	1
20. Матрица	1
21. Матрица	1
22. Матрица	1
23. Матрица	1
24. Матрица	1
25. Матрица	1
26. Матрица	1
27. Матрица	1
28. Матрица	1
29. Матрица	1
30. Матрица	1
31. Матрица	1
32. Матрица	1
33. Матрица	1
34. Матрица	1
35. Матрица	1
36. Матрица	1
37. Матрица	1
38. Матрица	1
39. Матрица	1
40. Матрица	1
41. Матрица	1
42. Матрица	1
43. Матрица	1
44. Матрица	1
45. Матрица	1
46. Матрица	1
47. Матрица	1
48. Матрица	1
49. Матрица	1
50. Матрица	1