

Э.Б.Мовшович М.Н.Кнепель М.С.Черкашин

# ФОРМАЛИЗАЦИЯ ГЕОЛОГИЧЕСКИХ ДАНЫХ ДЛЯ МАТЕМАТИЧЕСКОЙ ОБРАБОТКИ

---



Э.Б.Мовшович М.Н.Кнепель М.С.Черкашин

# ФОРМАЛИЗАЦИЯ ГЕОЛОГИЧЕСКИХ ДАННЫХ ДЛЯ МАТЕМАТИЧЕСКОЙ ОБРАБОТКИ

5144

Кольская комплексная геологическая  
экспедиция сверхглубокого бурения  
Н Г Б



МОСКВА «НЕДРА» 1987

Мовшович Э. Б., Кнпель М. Н., Черкашин М. С. Формализация геологических данных для математической обработки. — М.: Недра, 1987. — 190 с., ил.

На примере решения конкретных информационно-справочных и аналитических задач показано, как уровень формально-логической организации исходных геологических данных влияет на возможности их обработки и достоверность получаемых результатов. Рассмотрены актуальные проблемы геонформатики, систематизированы логические требования к организации данных и методические приемы их формализации.

Для геологов, связанных с использованием математических методов при решении геологических задач.

Табл. 13, ил. 40, список лит. — 48 назв.

Рецензент: А. А. Арбатов, д-р эконом. и канд. геол.-минер. наук (Всероссийский научно-исследовательский институт системных исследований)

## ПРЕДИСЛОВИЕ

Процесс внедрения математических методов (ММ) и ЭВМ в сфере геологической деятельности начался в действительно широком масштабе с начала 60-х годов. Все убыстряющееся развитие этого процесса, обусловленное объективными факторами, привело к оформлению во многих научных и производственных организациях подразделений разного масштаба, специализирующихся на внедрении ММ и разработке автоматизированных систем хранения, поиска и обработки геологической информации. Развитие указанных подразделений и соответствующих работ не может протекать и не протекает автономно, независимо от основной массы геологов, применяющих традиционные (неформализованные) средства геологической деятельности. Многочисленные коллективы геологов должны затрачивать большие усилия на подготовку исходных данных для автоматизированных систем хранения, поиска и обработки геологической информации. Значительное число геологов выступают в качестве создателей и заказчиков (соразработчиков) математических моделей решения разнообразных геологических задач, пользователей результатов решения этих задач, обращаются к автоматизированным системам со сложными селективными запросами на поиск и выдачу необходимых данных.

Насколько быстрым и эффективным будет процесс внедрения ММ и ЭВМ в геологическую деятельность, зависит не только и даже не столько от совершенства математических методов и разрабатываемых автоматизированных систем, сколько от многочисленных коллективов геологов, поставляющих исходные данные и использующих результаты их обработки. Сегодня, несмотря на накопленный в отрасли солидный положительный опыт применения ММ и автоматизированных средств решения разнообразных геологических задач информационно-справочного и информационно-аналитического характера, основная масса геологов еще не преодолела психологический барьер на пути обращения к указанным методам и средствам как к необходимому элементу геологической деятельности. И для этого имеются достаточно веские основания. Наряду с положительным опытом есть много примеров, когда при весьма ощутимых дополнительных затратах труда и времени геологов на подготовку и представление исходных данных в форме, приспособленной к конкретной автоматизированной системе или модели обработки, геологический результат оказывался малоудовлетворительным или, по крайней мере, не окупающим понесенных дополнительных усилий и технических средств. В таких случаях геолог, как правило, сначала слишком многого ожидает от применения ММ и ЭВМ, а затем слишком пессимистически оценивает их возможности. Отсюда вытекает

нередко высказываемое мнение о том, что применение ММ и ЭВМ оправдано лишь в крайне ограниченных областях (задачах) геологии, что геология это не та сфера, где внедрение ММ и вычислительной техники может ощутимо сказаться на результатах деятельности. Однако для столь пессимистических взглядов нет достаточно объективных оснований.

Во-первых, сейчас уже представляются очевидными и поддающимися систематизации истоки имеющегося негативного опыта, а это позволяет с помощью соответствующих методологических подходов и приемов воздействовать на качество результатов обработки.

Во-вторых, сейчас быстро развиваются так называемые человеко-машинные методы решения геологических задач. Суть их заключается в том, что в процессе решения задачи происходит как бы разделение труда между человеком и ЭВМ. На последнюю возложены отдельные фрагменты решения, которые поддаются математическому описанию (моделированию), а человек (геолог), работая в режиме диалога с ЭВМ, включает по ходу решения те данные, которые базируются на его опыте и интуиции и пока не поддаются формализации. Геолог, анализируя конечный результат, может многократно повторять решение, меняя значения вводимых им параметров, пока не получит удовлетворяющее его решение. Такое разделение труда между геологом и ЭВМ не только облегчает и убыстряет решение задачи, но, сохраняя за геологом возможность активно влиять на ход решения, существенно повышает достоверность результата (поскольку устраняется основная доля ошибок, обусловленных при полностью автоматизированной технологии неудачной формализацией тех элементов решения, которые при человеко-машинной технологии остаются за геологом).

В-третьих, следует отметить, что на результатах применения ММ и ЭВМ сказывается логическая организация не всей геологической информации<sup>1</sup> (ГИ), а лишь ее части. Информация, характеризующая определенные точки земных недр (координатно-привязанная информация), обычно агрегируется в виде таблиц или карт, профилей, разрезов, графиков (кривых), которые также могут быть переведены в табличную форму, т. е. она уже приведена или может быть приведена чисто механически к виду, подготовленному для применения ММ и ЭВМ. И поэтому совсем не случайно, что подавляющее большинство математических моделей решения геологических задач, в том числе с применением ЭВМ, ориентировано именно на координатно-привязанную ГИ и именно здесь достигнуты наиболее значительные положительные результаты.

<sup>1</sup> Геологическая информация здесь и далее понимается в широком смысле — как информация, характеризующая геологическое строение и освоение недр независимо от того, получена ли она чисто геологическими средствами или геофизическими, геохимическими, аэро- или космическими методами наблюдений (регистрации параметров недр).

Это не означает, конечно, что координатно-привязанная ГИ всегда совершенно безупречна с точки зрения возможностей применения ММ и ЭВМ и влияния на конечный результат. Здесь есть внутренние сложности, связанные с недостаточными представительностью и качеством данных, в частности с несовершенством или неунифицированностью методов измерения параметров и методов введения поправок на замеренные значения, неравномерностью сети наблюдений и др. Но тем не менее координатно-привязанная информация формально выражена на языке, легко воспринимаемом ММ и ЭВМ, и не требует от геолога какого-либо предварительного логического преобразования.

Гораздо сложнее обстоит дело с объектно-привязанной ГИ. Она выражена, как правило, на традиционном для геологии описательном языке, при котором логические связи между данными затушеваны. Самая главная проблема заключается в том, что единые общепринятые подходы к выделению большинства даже однотипных геологических объектов пока не выработаны. То же можно утверждать и в отношении понимания соотношения геообъектов разных уровней и описывающих их параметров. Это не может не сказаться на корректности логической организации объектно-привязанной ГИ, обуславливает ее неоднозначность, несопоставимость. В свою очередь, это затрудняет ее формализацию, т. е. запись способов получения, представления данных и связей между ними на языке зафиксированных правил. В то же время ММ и ЭВМ «понимают» и правильно «воспринимают» лишь формализованный язык.

Поскольку необходимость в специальной формально-логической подготовке (преобразовании) исходных для математической обработки данных возникает в подавляющем большинстве случаев при работе с объектно-привязанной ГИ, именно организации последней и уделено основное внимание.

Внедрение математических методов и ЭВМ в геологическую деятельность не столько породило, сколько обострило проблему логической организации (формализации) исходных геологических данных, подлежащих обработке. Эта проблема существовала и ранее, и именно она обуславливает недостаточно высокие прогностические возможности геологии, не позволяющей ей войти в число точных наук. Внедрение современной технологии сделало негативные аспекты состояния понятийно-терминологической базы геологии лишь более наглядными и очевидными, поскольку применение ММ и ЭВМ, основанное на формально зафиксированных логических связях и отношениях обрабатываемых данных, предъявляет достаточно жесткие логические требования к организации исходного фактического материала.

Проблема организации исходных данных для применения ММ и ЭВМ — целеобусловленная производная от более общих методологических проблем науки вообще и геологии и математики в частности, а также содержательных концепций геологии. Вполне естественно, что к настоящему времени насчитывается большое

число публикаций, имеющих прямое или опосредованное отношение к рассматриваемой проблеме. Это в первую очередь обобщающие разработки по общей методологии науки, отраженные в работах И. В. Блауберга, Е. К. Войшвилло, Д. П. Горского, Л. Друянова, Б. М. Кедрова, Н. И. Кондакова, Е. П. Никитина, К. Попа, А. И. Ракитова, С. С. Розовой, М. А. Розова, Г. И. Рузавина, В. Н. Садовского, А. В. Славина, А. И. Умова, Ю. А. Урманцева, В. А. Штоффа, Г. П. Щедровицкого, Э. Г. Юдина и др. Значительное число публикаций непосредственно затрагивает геологическую предметную область, включая ряд глубоких, практически важных (хотя и не всегда бесспорных) разработок по методологическим проблемам геологии и применения ММ в геологии. Среди них следует отметить известные работы И. И. Абрамовича, А. А. Арбатова, М. С. Арабаджи, А. Н. Бугайца, Ю. К. Буркова, В. П. Бухарцева, С. Ф. Васильченко, А. Б. Вистелиуса, Ю. А. Воронина, В. В. Грузы, Л. Ф. Дементьева, А. Н. Дмитриевского, Э. А. Еганова, Р. А. Жукова, Ю. Н. Каргодина, А. Э. Конторовича, Ю. А. Косыгина, Ф. П. Кренделева, И. В. Крутя, В. В. Марченко, В. И. Оноприенко, А. С. Поваренных, Д. А. Родионова, А. С. Смирновой, Б. И. Смирнова, В. А. Соловьева, Ю. А. Ткачева, А. А. Трофимука, И. П. Шаропова, А. М. Шурыгина и др.

Однако, во-первых, хотя многие работы имеют большую или меньшую область перекрытия с рассматриваемой проблемой фактографической информатики, последняя специально и целенаправленно, насколько известно, не освещалась. Во-вторых, за крайне редким исключением, имеющиеся работы ориентированы главным образом на читателя, имеющего солидную подготовку в области методологии науки и математических методов анализа. В-третьих, рассмотрение отдельных аспектов проблемы базировалось на материалах по различным аспектам геологии (главным образом общей геологии), но геологию полезных ископаемых и, в частности, геологию нефти и газа затронуло лишь в ничтожной степени. Это тем более разительно, что даже простой просмотр библиографии показывает, что не менее  $\frac{2}{3}$  всех опубликованных отечественных и зарубежных разработок по применению ММ и ЭВМ в геологии приходится именно на геологию нефти и газа.

В настоящей работе авторы поставили перед собой задачу рассмотреть причины, обуславливающие неадекватность исходного фактического материала применяемому математическому аппарату, и, опираясь на выполненные разными исследователями разработки по методологии геологических исследований и применению ММ и ЭВМ в геологии, вычленив, систематизировать и представить на традиционных для геологии языке и примерах только те элементы методологии и логики организации данных, которые необходимо учитывать геологу, если он собирается привлечь ММ и ЭВМ для повышения эффективности своей деятельности. Авторы твердо убеждены в том, что эффективное внедре-

ние ММ и ЭВМ в геологии не требует глубокого владения геологами математическим аппаратом и компьютерной техникой и технологией, а специалистами-математиками — геологической предметной областью. Расчет на это, учитывая требуемые усилия и время, представляется малореальным в широком масштабе. Поскольку математика — это прежде всего формальный стиль мышления [5], знание и соблюдение методологических (логических) требований к организации (формализации) данных — это тот легкодоступный общий язык, на котором может быть наиболее безболезненно обеспечено плодотворное взаимодействие геологов и математиков. Владея им, геолог обретает способность использовать ММ и ЭВМ «с открытыми глазами», т. е. с четким представлением о том, где, как и в какой мере он (через организацию своих данных) может активно, целенаправленно влиять на возможности и качество математической обработки (в том числе поиска) данных.

Материал в данной работе распределен по главам по принципу последовательного снижения уровня общности. Соответственно в главе 1 рассмотрены наиболее общие вопросы информатики, посвященные влиянию организации фактографических геологических данных на эффективность применения ММ и ЭВМ в геологии. Затем в главе 2 рассматриваются вопросы формализации исходных геологических данных, их соответствия математическому аппарату и языку (вне связи с применением ЭВМ). В завершающей части работы, в главе 3, освещены принципы логической организации ГИ при создании автоматизированных банков геологических данных и решении на их основе информационно-справочных и информационно-аналитических задач геологии.

Предисловие, заключение и главы 1—2 написаны Э. Б. Мовшовичем, глава 3 — всеми авторами книги.

# ИНФОРМАЦИЯ И ИНФОРМАТИКА В СТРУКТУРЕ ГЕОЛОГИЧЕСКОЙ ДЕЯТЕЛЬНОСТИ

## 1.1. ПРЕДПОСЫЛКИ МАТЕМАТИЗАЦИИ И КОМПЬЮТЕРИЗАЦИИ ГЕОЛОГИЧЕСКОЙ ДЕЯТЕЛЬНОСТИ

Одно из наиболее характерных проявлений научно-технического прогресса практически во всех отраслях знаний и деятельности — внедрение математических методов анализа и автоматизированных средств хранения, поиска и обработки информации. Широко известно замечание К. Маркса о том, что наука только тогда достигает совершенства, когда ей удается пользоваться математикой. Это положение в настоящее время в равной мере справедливо и по отношению к производственной практике, особенно в условиях геологоразведочной отрасли, где зачастую трудно провести формальную грань между прикладной наукой и производством.

Процесс внедрения математических методов и автоматизированных информационных и информационно-обрабатывающих систем в геологоразведочной отрасли развивается со все возрастающей активностью. Этому способствуют следующие факторы.

1. Перед всем народным хозяйством нашей страны поставлена задача интенсификации науки и производства. Применительно к геологоразведочной отрасли эта задача может быть реализована по двум основным направлениям. Первое заключается в повышении точности геологических прогнозов. Вероятностный характер геологоразведочного процесса настоятельно требует, чтобы любая геологическая оценка сопровождалась оценкой ее достоверности [35]. Второе направление состоит в динамической оптимизации геологоразведочного процесса в зависимости от степени подтверждаемости геологического прогноза, изменения наличных материально-технических ресурсов и т. п. Решение указанной задачи по обоим направлениям немислимо без применения достаточно сложного математического аппарата.

2. Основное назначение поисковой геологии — выявление скоплений минерального сырья, обеспечение требуемого прироста запасов полезных ископаемых. Эффективность достижения этого зависит от того, насколько глубоко познаны закономерности размещения полезных ископаемых в зависимости от тех или иных геологических условий. Закономерности в геологии почти исключительно статистические. В силу несоизмеримости длительности процессов формирования скоплений минерального сырья (например, нефти и газа) и человеческой жизни наблюдать этот процесс мы не можем и судим о нем на основе оставленных им следов, как правило затухающих влиянием различных сочетаний многочис-

ленных наложенных факторов. В результате познание геологических закономерностей в большинстве случаев требует системного анализа многоаспектной геологической информации по весьма представительному набору конкретных объектов<sup>1</sup>, что также невозможно без соответствующего математического аппарата.

То же можно сказать и о предметах исследования геологии, в отношении которых формулируют геологические закономерности. По мере вынужденного и неизбежного усложнения геологических задач приходится переходить к гораздо более сложным предметам исследования. Наглядной иллюстрацией является намечаемая во многих регионах постепенная переориентация нефтегазопроисковых работ с традиционных антиклинальных ловушек на «неантиклинальные».

Усложнение предметов исследования переводит поисковый процесс на принципиально более сложный информационный уровень. Так, при выявлении неантиклинальных ловушек, весьма разнообразных как по морфологии, так и по генезису, наряду со структурным фактором на практике используют многочисленные литологические, палеогеографические, геофизические, геохимические параметры (в общей совокупности более 20) [37]. Столь большое число параметров означает чрезвычайно громоздкий комплекс регистрирующих эти параметры поисковых методов, что неприемлемо для промышленного производства. Поисковое значение этих параметров не оценивалось. Поэтому встают задачи минимизации поисковых признаков (и соответственно комплекса методов), отбора небольшого числа наиболее информативных. Подобного рода проблемы обуславливают широкое применение в геологии статистических методов оценки информативности признаков, классификации и ранжирования, распознавания образов и др.

Приведенные примеры отражают самый обобщенный и далеко не полный перечень факторов, обуславливающих внедрение методов математического анализа в геологическую деятельность. Специфика последней активно побуждает к сочетанию методов математического анализа с возможными, предоставляемыми современными вычислительными средствами, т. е. к автоматизации хранения, поиска и обработки геологической информации. Широкое привлечение ЭВМ неизбежно, и этому фактору суждено сыграть в геологоразведочной отрасли гораздо более революционизирующую роль, чем в большинстве других отраслей народного хозяйства. Для этого имеются следующие предпосылки.

1. Геологоразведочное производство, в отличие от многих отраслей экономики, не создает непосредственно материальных ценностей в традиционном их понимании, а лишь дает информацию об их величине, качестве, размещении в недрах. Сами

<sup>1</sup> Здесь и далее под термином «объект» понимается часть геологической реальности (конкретное месторождение, залежь и т. п.), а под термином «предмет» — понятийная модель объекта (понятие о месторождении вообще, залежи вообще и т. п.).

материальные ценности — скопления минерального сырья — были созданы природой задолго до появления геологической сферы деятельности человека, т. е. конечный продукт геологоразведочного производства — подготовленные запасы и ресурсы, — выражены в виде ГИ. И промежуточные геологические результаты также представлены в виде ГИ. Последняя рассматривается сейчас как специфическая форма очень важного и ценного материального ресурса.

2. Преобладающая часть фактографической ГИ имеет долговременное значение. Это связано в первую очередь с тем, что по мере накопления и теоретического осмысления полученного в процессе геологоразведочных работ эмпирического опыта поисковые концепции в геологии закономерно сменяются одна другой. Так, нефтегазовая геология в конце XIX в. ориентировалась на поиски нефти и газа вблизи их естественных выходов на поверхность. В первой половине XX в., когда практические возможности, предоставляемые этой концепцией, стали иссякать, возникла так называемая «антиклинальная теория», ориентировавшая поиск на антиклинальные (структурные) ловушки нефти и газа. Их выявление основывалось в основном на фиксации структуры (формы) поверхности слоев осадочных пород. Наличие антиклинального изгиба природных резервуаров нефти и газа дает основание предполагать здесь ловушку и проектировать поисковое бурение. Были разработаны весьма результативные методы выявления антиклинальных изгибов, в частности не выраженных на поверхности. В итоге к антиклинальным ловушкам приурочено подавляющее большинство открытых месторождений нефти и газа.

Однако сейчас фонд антиклинальных ловушек во многих районах в значительной степени исчерпан, что вызвало необходимость вновь изменять нефтегазопоисковую концепцию, которая все активнее переориентируется на так называемые «неантиклинальные» ловушки. Последние часто приурочены, например, к древним «погребенным» (захороненным под толщей более молодых слоев пород) рифам, дельтам, песчаным барам и пляжам, зонам выклинивания пород-коллекторов любого происхождения. Большое значение таких ловушек в нефтегазоносном отношении хорошо доказано геологоразведочной практикой как в Советском Союзе, так и в ряде зарубежных стран.

Для выявления неантиклинальных ловушек, весьма разнообразных по форме, условиям формирования и закономерностям размещения, совершенно недостаточно знания только структуры поверхности слоев, которая хорошо фиксируется геофизическими методами. Необходимо знать еще многие вещественные параметры перспективных пород и на этой основе установить их происхождение. А это означает, что переход от антиклинальных ловушек к неантиклинальным переводит нефтегазопоисковый процесс в информационном аспекте на принципиально иной уровень, требующий системного анализа широкого комплекса геологических факторов.

Такая смена поисковых концепций характерна не только для геологии нефти и газа, но и для геологии минерального сырья в целом. В условиях роста объема и совершенствования методов геологоразведочных работ развитие эмпирической основы нашего опыта, а соответственно и процесс смены поисковых концепций (парадигмы) все убыстряются. А с каждой такой перестройкой приходится под новым углом зрения пересматривать накопленный фактический материал (фактографическую ГИ). И это дает большой практический эффект, даже без привлечения дополнительной информации. Многочисленны примеры, когда исключительно за счет «разведки архивов» открыт целый ряд новых месторождений тех типов, о существовании которых вопрос ранее просто не ставился [20]. Это значит, что при каждом изменении поисковой концепции к фактографической ГИ приходится неоднократно обращаться, пересматривая ее под новым углом зрения.

К этому следует также добавить, что хотя конкретная ГИ характеризует определенный участок недр, она имеет не только местное значение. Геологическая информация по совокупности отдельных конкретных участков (районов), которые могут быть расположены в различных регионах мира, постоянно используется для уточнения общих закономерностей формирования и размещения скоплений минерального сырья, для реализации одного из весьма распространенных в геологии способов прогнозирования — метода сравнительной геологической аналогии.

3. Все более ощущается, что современный период развития геологии вообще и геологии нефти в частности характеризуется существенным усложнением геологических и технических задач. Для того чтобы обеспечить необходимые приросты разведанных запасов, приходится осваивать все большие глубины недр, нетрадиционные типы скоплений полезных ископаемых, вовлекать в разработку мелкие по запасам или содержанию полезных компонентов и сложно построенные месторождения, выходить в труднодоступные и суровые в физико-географическом отношении районы суши, а также на континентальный шельф, применять все более сложные методы как наземных, так и аэро- и космических наблюдений. Поэтому себестоимость, а соответственно и стоимость фактографической ГИ высока и все время растет. Существенно важно и то, что значительная часть этой информации невозобновляема (например, геологическая и геофизическая информация по скважинам).

Учитывая большие трудовые, временные, материальные и технические затраты на получение ГИ и долговременное значение последней, а также невозможность части информации, очень остро стоит вопрос об обеспечении ее сохранности и максимально полного использования. Становится все очевиднее, что на основе традиционных форм отображения и хранения ГИ (в виде сдаваемых в геологические фонды томов машинописных отчетов с картографическими иллюстрациями) решить эту проблему, в частности обеспечить оперативный доступ к информации, нельзя.

Весьма показательно, что преобладающая часть геологических данных, заключенных в отчетах и публикациях, не была востребована в фондах и технических библиотеках [20]. У потенциальных потребителей этой информации просто нет времени и физической возможности осуществлять выборку необходимой информации. Решение этой задачи, учитывая объемы и сложную логическую структуру геологической информации, требует использования современных автоматизированных средств хранения и оперативного поиска информации — автоматизированных банков данных (АБД).

4. Многие геологические задачи основываются на системном анализе многомерных связей большого числа параметров и принципиально не решаются «невооруженным мозгом» даже значительного коллектива исследователей. В то же время они не представляют трудностей при использовании современных автоматизированных систем обработки данных (АСОД), соединяющих возможности АБД по хранению и поиску исходных данных с возможностями их машинной обработки на основе математического моделирования отдельных геологических задач и геологоразведочного процесса.

Таким образом, перестройка всей геологической деятельности на основе системного подхода, применения ММ и ЭВМ объективно обусловлена и потому неизбежна. Эта перестройка уже началась, о ней свидетельствует большое число функционирующих и проектируемых автоматизированных систем хранения и аналитической обработки фактографической ГИ, которые предполагается интегрировать в рамках отраслевых АСУ. Переход к действительно широкому использованию ММ и автоматизированных систем обработки информации невозможен, пока это использование не будет восприниматься геологами как необходимый элемент технологии собственно геологической деятельности. А это подразумевает решение целого комплекса взаимосвязанных проблем, причем не только технических, но и методических, психологических, организационных и даже юридических.

В решении имеющихся технических проблем, связанных с совершенствованием ЭВМ и сопряженных с ними периферийных устройств (символьно-цифровые и графические дисплеи, графопостроители, устройства подготовки, ввода и вывода данных и др.), расширении технологических возможностей автоматизированных систем хранения и обработки ГИ, прогресс очевиден. Сменился целый ряд поколений ЭВМ и периферийных устройств, достигнуты почти фантастические объемы памяти и быстродействие. Но это не сказалось в той же мере на результативности автоматизированной обработки геологических данных.

Связано это в основном с двумя факторами. Во-первых, недостаточное внимание уделялось технологии общения геолога с ЭВМ. В подавляющем большинстве случаев между автоматизированной системой хранения и обработки геологических данных и геологом, выступающим в качестве внешнего пользователя такой

системы, стоят посредники в лице прикладных математиков (администраторов базы данных), которые переводят запрос на поиск нужных данных и способ их обработки с языка геологии на язык машинных инструкций (программ). Наличие таких посредников-переводчиков делает общение геолога с ЭВМ прежде всего менее производительным. Каждый, кто работал с материалом на незнакомом языке через переводчика, хорошо это знает. Но что гораздо важнее, при этом подразумевается, что геолог пока «не дорос» до общения со сложными компьютерными системами, и это создает психологический (и в определенной мере организационный) барьер на пути восприятия геологом работы с ЭВМ как элемента геологической деятельности. Преодоление этого является необходимой предпосылкой массового перехода к применению ММ и ЭВМ в геологическом анализе. Поэтому действительно широкое применение ММ и ЭВМ в геологии может быть обеспечено лишь упрощением процедуры общения геолога с ЭВМ, т. е. возможностью непосредственного контакта в терминах профессионального геологического языка.

При этом геологу вовсе не требуется обладать знаниями специалиста по ММ и автоматизированным системам. Необходимость в таком специалисте при этом не снимается, но функции его изменяются с последующих на сервисные или консультационные. Аналогом этой ситуации может служить современная организация автодела. Многие тысячи автолюбителей успешно пользуются автомашиной, освоив элементарные правила обращения с нею, но не обладая знаниями автомеханика, функции которого оказалось более эффективным передать централизованному автосервису. Требуется подобная организация технологии общения геолога с автоматизированной системой.

Второй фактор, определяющий не количественный, а качественный аспект внедрения ММ и ЭВМ в геологии, заключается в разрешении методологических (состояние геологического языка) и организационных проблем, в самом общем виде последовательно рассмотренных в разделах данной главы.

## 1.2. ТИПЫ И ПРОЦЕДУРЫ ПРЕОБРАЗОВАНИЯ ФАКТОГРАФИЧЕСКОЙ ИНФОРМАЦИИ В ПРОЦЕССЕ ГЕОЛОГИЧЕСКОЙ ДЕЯТЕЛЬНОСТИ

Геологическая информация существенно неоднородна. Системный анализ геологоразведочной деятельности показывает, что реализация геологоразведочного процесса, независимо от применяемых методов и технических средств, в функциональном плане сводится в конечном итоге к получению и (или) целенаправленному преобразованию фактографической ГИ. Учитывая это, а также информационную форму выражения результатов геологоразведочной деятельности, последнюю можно представить как систему, характеризующуюся определенными информационно-функциональными связями (рис. 1).

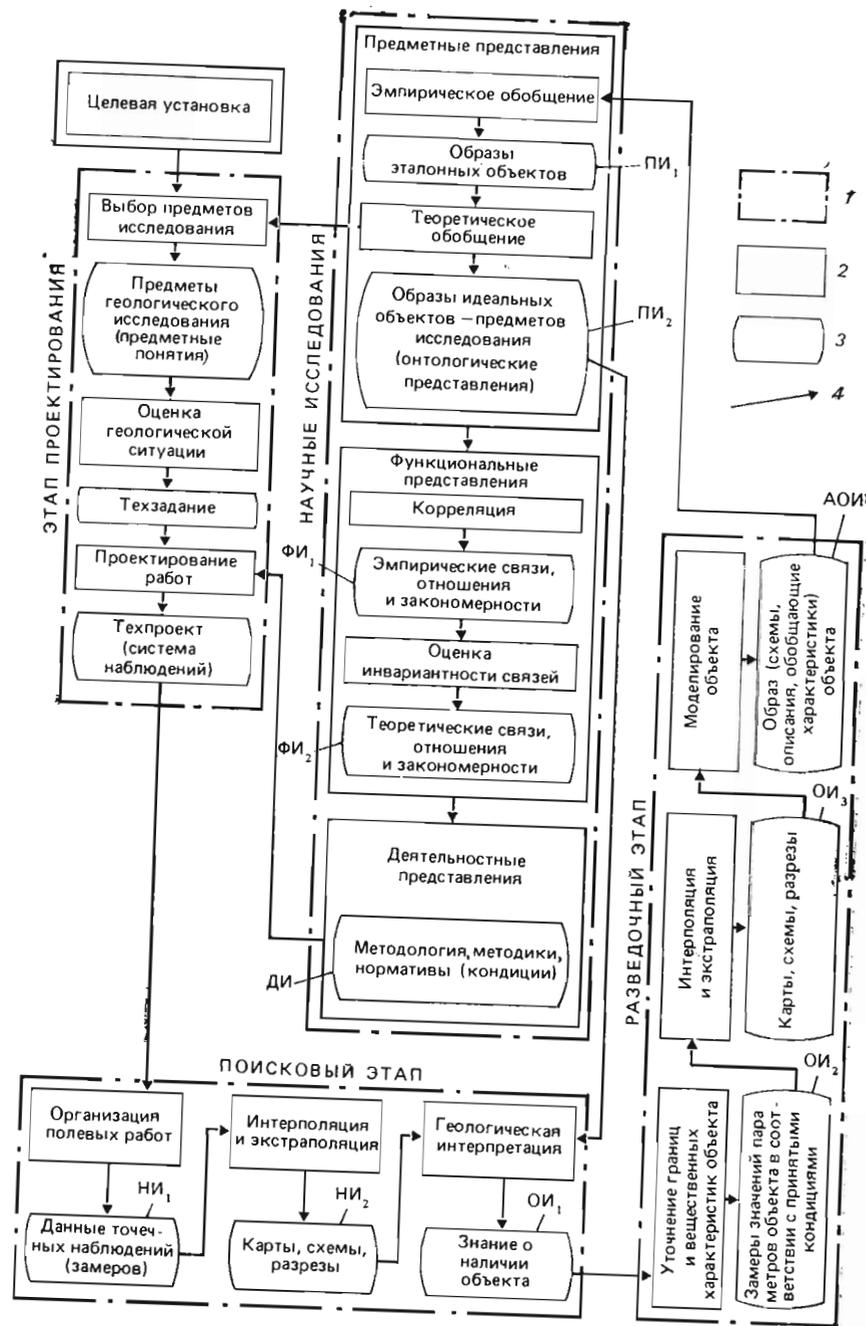


Рис. 1. Обобщенная информационная модель геологоразведочного процесса: 1 — основные структурно-функциональные элементы; 2 — основные процедуры; 3 — результаты реализации процедур; 4 — функциональные связи

Любая деятельность начинается не с нуля, она использует предшествующий опыт — предметный, функциональный и деятельностный. Первый из них зафиксирован в виде набора предметов исследования соответствующей (в данном случае — геологической) области знания. В арсенале геологии, где, как и в других областях знания, происходит интенсивный процесс дифференциации и специализации наук, имеется огромный, постоянно расширяющийся, набор предметов исследования. Формы их существования, во-первых, понятия, призванные логически оформить обособление отдельных предметов, зафиксировать их специфические признаки (на которых должно базироваться выделение объектов соответствующего типа) и, во-вторых, онтологические схемы. Последние представляют собой выраженные чаще всего графически или в виде формул образы идеальных предметов (репрезентирующих типы реальных объектов). Например, онтологическим представлением такого геологического предмета исследования, как антиклиналь, является в разрезе — выпуклая вверх линия, моделирующая пласт, а в плане — несколько замкнутых вложенных линий, моделирующих изгиб поверхности этого пласта.

Понятия и онтологические схемы, представляющие собой предельно абстрактную геологическую информацию (ПИ<sub>2</sub>), тесно взаимосвязаны: и те и другие есть продукт теоретического обобщения. Ему обычно предшествует эмпирическое геологическое обобщение, в результате которого формируются образы уже идеализированных (не реальных), но пока еще ориентированных на конкретную региональную среду предметов (например, образы типовых для данного района, региона залежи, месторождения, дислокации и т. д.). Такая первичная (ПИ<sub>1</sub>) и вторичная (ПИ<sub>2</sub>) предметная, т. е. идеализированная, геологическая информация определяет стратегию всей геологической деятельности и в большой мере достоверность геологических выводов.

В зависимости от целевой установки из предшествующего предметного опыта выбираются предметы исследования, отвечающие поставленной цели. Набор предметов исследования в разных случаях может быть различным. Если цель заключается, например, в оценке ресурсов нефти и газа, то в качестве предметов будут выступать в первую очередь нефтегазоносные (потенциально нефтегазоносные) бассейны или провинции, их структурно-тектонические элементы, литолого-стратиграфические (оценочные) комплексы, зоны нефтегазонакопления (выявленные и возможные). Если цель — реализация сформированной поисковой концепции в конкретном районе, например поиски скоплений нефти и газа, приуроченных к погребенным рифам (барам, дельтам), то в качестве предметов исследования будут выступать рифы, фации, ловушки, породы-коллекторы, залежи и др.

Выбор предмета исследования предопределяет набор характеризующих его и подлежащих регистрации параметров (основные из них должны вытекать из определения понятия). Это первый шаг на пути определения оптимальной системы наблюдений —

одного из основных результатов проектирования геологоразведочных работ. Однако на пути к достижению этого результата требуется выполнить еще ряд проектных разработок. Прежде всего необходимо провести оценку геологической ситуации, поскольку регистрация значений одних и тех же геологических параметров может осуществляться разными методами и с разной эффективностью в зависимости от глубины залегания, литологии, дислоцированности пород и т. п. Данные, полученные в результате реализации процедур отбора предметов исследования и оценки геологической ситуации, представляют собой основу для формулирования технического задания на геологоразведочные работы и последующего проектирования последних.

На этой стадии приходится привлекать также предшествующие функциональные и деятельностные представления, зафиксированные в виде используемых закономерностей, связей, а также в методологических регулятивах, методиках, нормативах. Проектирование любых геологоразведочных работ чаще всего в неявной форме обязательно опирается на эти представления. Поэтому сформулированная в техническом задании геологическая задача уже подразумевает наличие определенных представлений о деятельности, необходимой для ее выполнения. Эти деятельностные представления в процессе технического проектирования фиксируются в виде технического проекта, на основе которого и осуществляется организация полевых работ.

Каждая операция в описанном выше крупном звене технологии геологической деятельности, названном этапом проектирования, несет на себе прямое или опосредованное влияние предметной ГИ. Это влияние в не меньшей степени продолжает сказываться и на последующих звеньях технологической цепи, где в процессе полевых работ получается ГИ принципиально иных типов. В результате непосредственной регистрации параметров недр получаем прежде всего дискретные значения геологических параметров, отвечающие определенной (условной) координатно описываемой точке геологического пространства. В качестве примеров можно указать значения силы тяжести или геологического возраста в определенной точке земной поверхности, концентрации того или иного элемента в локализованном по месту образце (в том числе керне скважин), глубины кровли (подошвы) отдельного горизонта или его вещественный состав в определенной точке недр и т. п. Это *первичная необъективированная* фактографическая информация (НИ<sub>1</sub>). Она характеризуется четкой координатной привязкой, но еще не соотнесена с каким-либо геологическим объектом, который может быть выделен позднее в данном месте, и потому выражена в нейтральной (необъективированной) форме, т. е. выступает как знание как бы независимое от представлений потребителя информации о тех или иных принятых предметах геологического исследования (ожидаемых типах объектов). Оговорка «как бы» употреблена здесь не случайно. Косвенное влияние существующих предметных представлений тем не менее

ощутимо сказывается, проявляясь уже в самом выборе регистрируемых параметров. Суть одного из основных положений методологии науки заключается в том, что «система понятий определяет систему наблюдений» (ведь мы не станем наблюдать, измерять, регистрировать то, о чем мы не имеем понятия). В понятии (если его определение логически корректно оформлено) должны быть отражены признаки, выделяющие этот предмет среди других и подлежащие регистрации при выделении соответствующих объектов. Это положение зафиксировано, в частности, в сформулированном в теории понятий правиле реализуемости [46], согласно которому определение понятия (предмета) должно строиться только на таких его признаках, констатация которых технически или в крайнем случае принципиально возможна. К сожалению, в определениях многих понятий геологии правило реализуемости не соблюдено. Приведем примеры определения такого исходного для геологии нефти и газа понятия, как «ловушка»: 1) объем породы, могущий вместить нефть или газ вне зависимости от ее формы и условий возникновения, но при наличии способности к аккумуляции и консервации нефти и газа в ней [6, т. 1, с. 397]; 2) минимальный по объему участок земной коры, который при этом в силу структурно-генетической характеристики способен заключать единичную залежь нефти и (или) газа [33, с. 38].

Из приведенных определений видно, что в качестве признаков ловушки приводятся такие, как «наличие способности к аккумуляции и консервации нефти и газа», «в силу своей структурно-генетической характеристики способен заключать единичную залежь нефти и (или) газа». Подобные признаки характерны и для других определений ловушки, приведенных в учебниках, словарях, справочниках и монографиях. Подробно этот вопрос рассмотрен в работе [43]. Такие «признаки» не дают никаких указаний на то, посредством каких процедур (измерений, наблюдений) можно выделить ловушку.

Тем не менее даже при таких определениях понятий геологи выделяют свои объекты (правда, как правило, далеко не однозначно). И это возможно только потому, что понятие существует не само по себе, а за ним стоят предшествующие знания и опыт: а) выделения какого-то числа обобщаемых понятием отдельных объектов, каждый из которых уникален; б) изображения (образов) каждого из этих объектов в виде карт, разрезов, схем, описаний; в) вычленения из образов этих объектов того общего, что свойственно совокупности последних, т. е. формирования образов идеальных предметов.

Понятие стоит над этим опытом, оно связывает, согласовывает между собой его элементы и может изменяться, если по мере накопления опыта возникает необходимость в пересогласовании указанных элементов. Именно в силу стоящего за понятием опыта мы знаем (с той или иной степенью уверенности), какие параметры нам необходимо фиксировать в соответствующих случаях,

даже если из определения понятия этого не следует. Но определение понятия в явном или неявном виде должно отражать представление об исходной фактографической ГИ, на которой строится система наблюдений. В результате реализации последней мы получаем первичную необъективированную ГИ (НИ<sub>1</sub>).

Геолог, обычно оперирующий графическими образами, воспринять НИ<sub>1</sub> непосредственно, как правило, не в состоянии. Для этого она должна быть преобразована и представлена не в дискретной, а связной форме — в виде графических документов (схем, карт, разрезов, графиков). Это вторичная необъективированная геологическая информация (НИ<sub>2</sub>), представляющая основу для последующей геологической интерпретации (только ее геолог может зрительно воспринять, чтобы анализировать с геологической точки зрения). НИ<sub>2</sub> так же, как и НИ<sub>1</sub>, координатно привязана, т. е. каждой точке отраженного на графическом документе геологического пространства отвечает конкретное значение соответствующего параметра. При этом для одной части таких точек значение параметра получено в результате непосредственного измерения, а для другой — на основе интерполяции и экстраполяции. Соответственно переход от НИ<sub>1</sub> к НИ<sub>2</sub> требует применения процедур расчета значений между точками измерения. Эти процедуры основаны в основном на методах и концепциях математики, физики, графики (картографии) и в меньшей степени — на геологических знаниях, выступающих, как правило, в неявном виде.

В силу того что указанные методы и концепции в принципе формализуемы, процедуры преобразования НИ<sub>1</sub> в НИ<sub>2</sub> легко поддаются формализации. Поэтому зачастую процедуры получения НИ<sub>1</sub> и преобразования ее в НИ<sub>2</sub> объединяются в один технологический цикл, на выходе которого представляется сразу НИ<sub>2</sub>. Таков, например, процесс автоматизированного получения разнообразных каротажных диаграмм, сейсмодновременных или сейсмогеологических разрезов и др.

С геологической точки зрения НИ<sub>2</sub> также необъективированная, поскольку собственно геологические предметные представления в явном виде в ее получении не задействованы. Для того чтобы эту информацию воспринять геологически, требуется переосмыслить ее с помощью процедуры геологической интерпретации. Последняя заключается в том, что вторичная необъективированная ГИ рассматривается через призму предметных представлений геологии и на этой основе проводится выделение соответствующих конкретных геологических объектов. Например, для того чтобы по данным съемки (геологической, геохимической, геофизической) и бурения выделить залежь полезного ископаемого (или антиклиналь, или разлом и т. п.), необходимо прежде всего иметь представление о том, что такое залежь (антиклиналь, разлом и т. п.) и как она выражается графически (онтологические схемы), спроецировать соответствующее онтологическое представление на графически выраженный фактический материал (НИ<sub>2</sub>), установить, есть ли в анализируемом геологическом простран-

стве места, где наблюдается соответствие между предметным (онтологическим) представлением и графической выраженностью фактического материала, и, наконец, идентифицировать такие места, как соответствующие конкретные геологические объекты (залежи, антиклинали, разломы и т. п.).

Процедура проецирования абстрактных предметных (онтологических) представлений на НИ<sub>2</sub>, в результате которой вычленяется геологический объект, составляет суть геологической интерпретации данных наблюдения. Итогом является знание о наличии в данном месте геологического пространства ранее неизвестного геологического объекта определенного типа. Это знание фиксируется на карте или схеме контуром выделенного объекта, что представляет собой *первичную объективированную* геологическую информацию — ОИ<sub>1</sub>. На ее основе, т. е. посредством сопоставления контуров геологических объектов, решают многочисленные задачи по определению наличия и тесноты связей разнотипных объектов (например, оценка приуроченности органогенных построек к поднятиям палеорельефа в подстилающих слоях). Естественно, контуры объектов могут уточняться в процессе последующих работ по изучению выделенного объекта. Этим завершается еще одно крупное звено технологии геологической деятельности, названное нами поисковым этапом.

Знание о наличии объекта дает представление в лучшем случае о его наиболее общих типовых характеристиках, но не индивидуализирует этот объект среди других объектов того же типа. Для этого нужна дополнительная ГИ, которая получается на этапе разведки. В процессе наблюдений (измерений), имеющих целью уточнить границы, морфологические и вещественные характеристики выделенного геологического объекта, проводят замеры значений интересующих параметров изучаемого объекта в отдельных его точках. Это *вторичная* ОИ<sub>2</sub>. Как и в случае НИ<sub>1</sub>, отдельные замеры параметров не приспособлены к характерному для геологов образному восприятию информации. Поэтому ОИ<sub>2</sub> посредством процедуры интерполяционных расчетов преобразуется в связную форму (карты, схемы, разрезы). Последние представляют собой *третичную объективированную* геологическую информацию ОИ<sub>3</sub>. Она отличается от НИ<sub>2</sub> (так же, как и ОИ<sub>2</sub> от НИ<sub>1</sub>) тем, что имеет и координатную и объектную привязку, т. е. она выступает одновременно в качестве характеристики как координатно-зафиксированной точки геологического пространства, так и расположенного в пределах этого пространства геологического объекта, во имя которого и организуется получение этой информации. На этой и только этой ГИ базируется решение многих важнейших типовых задач геологии, например задачи оптимизации разработки и подсчета запасов.

После того как завершается процедура обособления и целенаправленного изучения свойств объекта, в соответствии с условиями, принятыми на основе предшествующего деятельностного опыта, ОИ<sub>3</sub> обобщается в пределах выделенного объекта и «изыма-

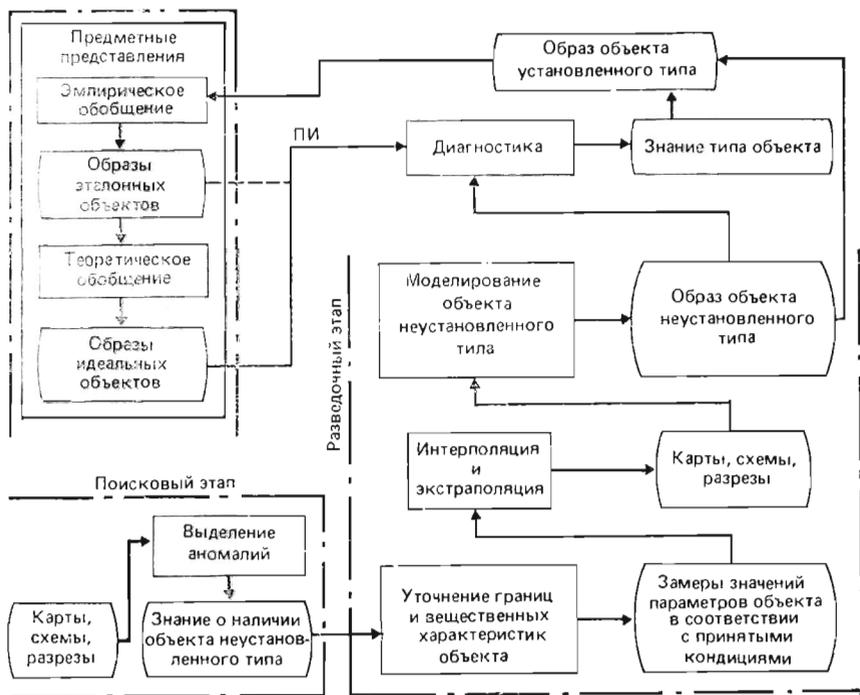


Рис. 2. Возможная модификация информационной модели геологоразведочного процесса.

Условные обозначения те же, что на рис. 1.

ется» из координатного каркаса. Так получается *агрегированная объективированная* геологическая информация — АОИ. Эта информация уже только объектно-привязанная. Она характеризует не какую-либо точку объекта, а весь объект в целом. Такие, например, параметры залежи нефти, как площадь, запасы, состав флюидов, высота, положение водонефтяного контакта, газовый фактор и др., представляют собой типичную АОИ. Она существует в форме описаний, разрезов, карт, схем, таблиц, создающих воспринимаемый нами образ данного геологического объекта. Этот образ получается в результате реализации геологом процедуры знакового (языкового, графического) моделирования конкретных геологических объектов. В общем объеме разнообразной ГИ, хранящейся в настоящее время в автоматизированных банках данных, преобладающая часть приходится именно на АОИ (кадастры месторождений, залежей, различных геологических тел и т. п.).

АОИ — исходная основа для эмпирического и теоретического обобщений, в результате которых из образов конкретных объектов отбрасывается все то, что является уникальным, и вычленяется то общее, что присуще всей совокупности обобщаемых объектов. Это общее представляется в предметной, т. е. абстракт-

ной идеализированной, форме, существенно расширяющей познавательные возможности. Таков системообразующий нормальный кругооборот геологической информации, при котором из ГИ одного типа путем соответствующих процедур преобразования обеспечивается получение ГИ другого типа.

В геологоразведочной практике возможны некоторые модификации отдельных элементов отраженной на рис. 1 информационной модели геологоразведочного процесса. Например, возникает ситуация, когда при получении НИ<sub>2</sub> на картах, схемах, разрезах визуально четко проявляются те или иные неожиданные и не сразу объяснимые аномалии свойств (мощности, литологии, силы тяжести и др.). В таком случае вместо процедуры геологической интерпретации осуществляется процедура выделения (проведения контуров) аномалии, т. е. выделения объекта неустановленного типа (рис. 2). Далее, основываясь на этой информации (ОИ<sub>1</sub>), проводится изучение этого объекта по стандартной схеме и после получения образа объекта посредством процедуры диагностики устанавливается его тип. Суть этой процедуры заключается в последовательном сопоставлении образа объекта с тем или иным набором образов (онтологических представлений) разнотипных предметов исследования. По достаточному соответствию одному из них и определяется тип объекта.

Таким образом, процедура диагностики существенно противоположна процедуре геологической интерпретации. В последнем случае проводится сканирование карты или набора карт образом одного предмета. Конечный результат этого — локализация соответствующего объекта. При диагностике локализация объекта является исходным пунктом, а на заключительной стадии осуществляется «сканирование» совокупности предметов исследования образом этого объекта.

Поскольку процедуры преобразования ГИ четко соответствуют определенным звеньям технологии (этапам) геологоразведочного процесса, именно с этими звеньями связаны и типы получаемой в результате соответствующих процедур ГИ. Отсюда вытекает и следующее исключительно важное положение: поскольку на каждом звене технологии геологоразведочной деятельности решаются свойственные только ему типы геологических задач, существует четкая связь между этими задачами и типами получаемой в этом звене ГИ.

В отмеченном кругообороте геологической информации исключительно важная роль принадлежит геологическим фондам, призванным осуществлять сбор, хранение и поиск ГИ всех указанных на рис. 1 типов. Обобщенная структура геологических фондов, основанных на современных автоматизированных средствах, приведена на рис. 3.

Следует отметить, что термины, обозначающие на рис. 1 отдельные звенья технологии геологоразведочной деятельности, в частности «поисковый этап», «разведочный этап», имеют в данном контексте значительно более широкое смысловое содержание,

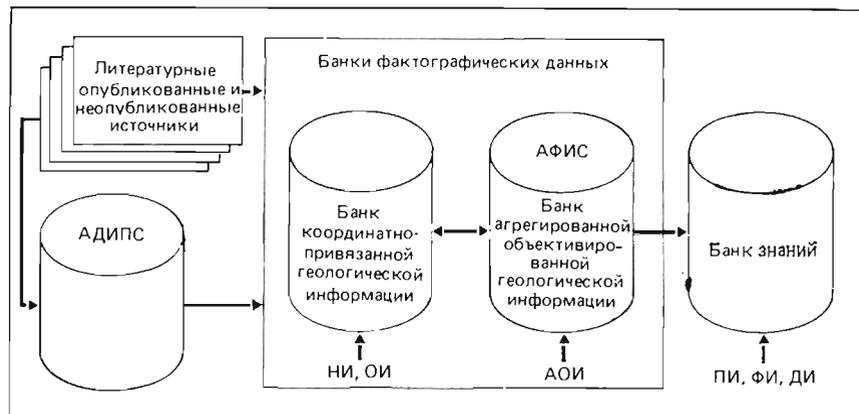


Рис. 3. Обобщенная структура геологических фондов, базирующихся на средствах банков данных. АДИПС и АФИС см. в гл. 3

чем принято. Дело в том, что в геологоразведочном производстве говорят о поисковом и разведочном этапах, как правило, применительно лишь к конечному продукту — скоплениям минерального сырья (залежи, месторождения). Но показанная на рис. 1 технология универсальна: она реализуется для любого геологического объекта, даже если он изучается вне связи со скоплениями минерального сырья или на самых начальных подходах к таким скоплениям. Например, на пути выявления скоплений нефти и газа, приуроченных к ловушкам неантиклинального типа, приходится (по характерному для геологии принципу последовательного приближения) изучать формации, фации, палеогеоморфологические объекты (погребенные рифы, бары, палеоделты и др.), приуроченные к ним природные резервуары и ловушки и, наконец, залежи. Рассмотренная технология, в том числе все процедуры, представленные под терминами «проектирование», «поисковый этап» и «разведочный этап», реализуются для каждого из этих типов объектов. В этом смысле, во-первых, можно говорить об этапах проектирования, поиска и разведки применительно и к формации, и к фации и т. п., а во-вторых, нет принципиальной разницы между работами региональными и детальными. Региональность или детальность работ проявляется лишь в масштабности (иерархических отношениях) объектов, последовательно изучаемых по универсальной технологии.

Время от времени на пути развития геологических наук, в частности под влиянием тех или иных целевых установок, возникает потребность в переоценке предшествующего предметного, а соответственно и деятельностного опыта, задающих стратегию геологоразведочной деятельности. Основной фактор, обуславливающий такую потребность, — осознание (весьма небезболезнен-

ное) существования противоречий между: а) логической организацией ГИ, обусловленной существующими предметными представлениями, с одной стороны, и общенаучными методологическими подходами (например, системным) и междисциплинарными (например, математическими) методами анализа и обработки данных — с другой; б) получаемым на практике эмпирическим материалом и предметными теоретическими представлениями («сюрпризы» природы). В таких случаях происходит определенное временное разрегулирование системы геологического знания. Но всякая корректно организованная система обладает устойчивостью, что обеспечивается наличием механизмов регуляции в условиях внешних возмущающих воздействий. Такие механизмы регуляции системы геологического знания приведены на рис. 4 и 5.

Установка на интенсификацию геологоразведочных работ и повышение их эффективности может быть реализована по двум основным направлениям — совершенствования способов и средств получения фактического материала и обработки ГИ (рис. 4). Третье возможное направление, связанное с совершенствованием организационной структуры и методов управления, в значительной степени зависит от достижений на двух первых. Разработки, ведущиеся в рамках первого направления (в частности, совершенствование дистанционных методов наблюдения технологии бурения на суше и на шельфе, повышение разрешающей способности геофизических и геохимических методов и др.), безусловно, очень важны. Они обеспечивают более эффективное получение ГИ, но при этом исходят из существующей логической организации ГИ, несовершенство которой существенно ограничивает возможности повышения эффективности геологоразведочных работ.

На втором из указанных направлений основная и реальная отдача возможны на пути внедрения современных методологических подходов (в первую очередь системного), математических методов и автоматизированных средств обработки данных, призванных оптимизировать (ускорить и удешевить) процесс получения геологических выводов и, главное, повысить их достоверность. Необходимость в этом остро ощущается все большим числом геологов, и это обусловило в первую очередь быстро расширяющееся внедрение математических методов и ЭВМ при решении геологических задач.

Однако опыт этого внедрения убедительно показал, что логическая организация предметного знания в геологии (а соответственно и фактографической ГИ) не всегда согласуется с формальным математическим аппаратом обработки. Это часто ведет к существенному искажению и невысокой надежности получаемых при этом геологических заключений. Вытекающая отсюда проблематика достаточно четко сформулирована в работах Ю. А. Воронина, Э. А. Еганова и других представителей новосибирской школы математической геологии. В настоящее время уже очевидно, что одностороннее приспособление математических методов к традиционной для геологии логической организации ГИ — путь,

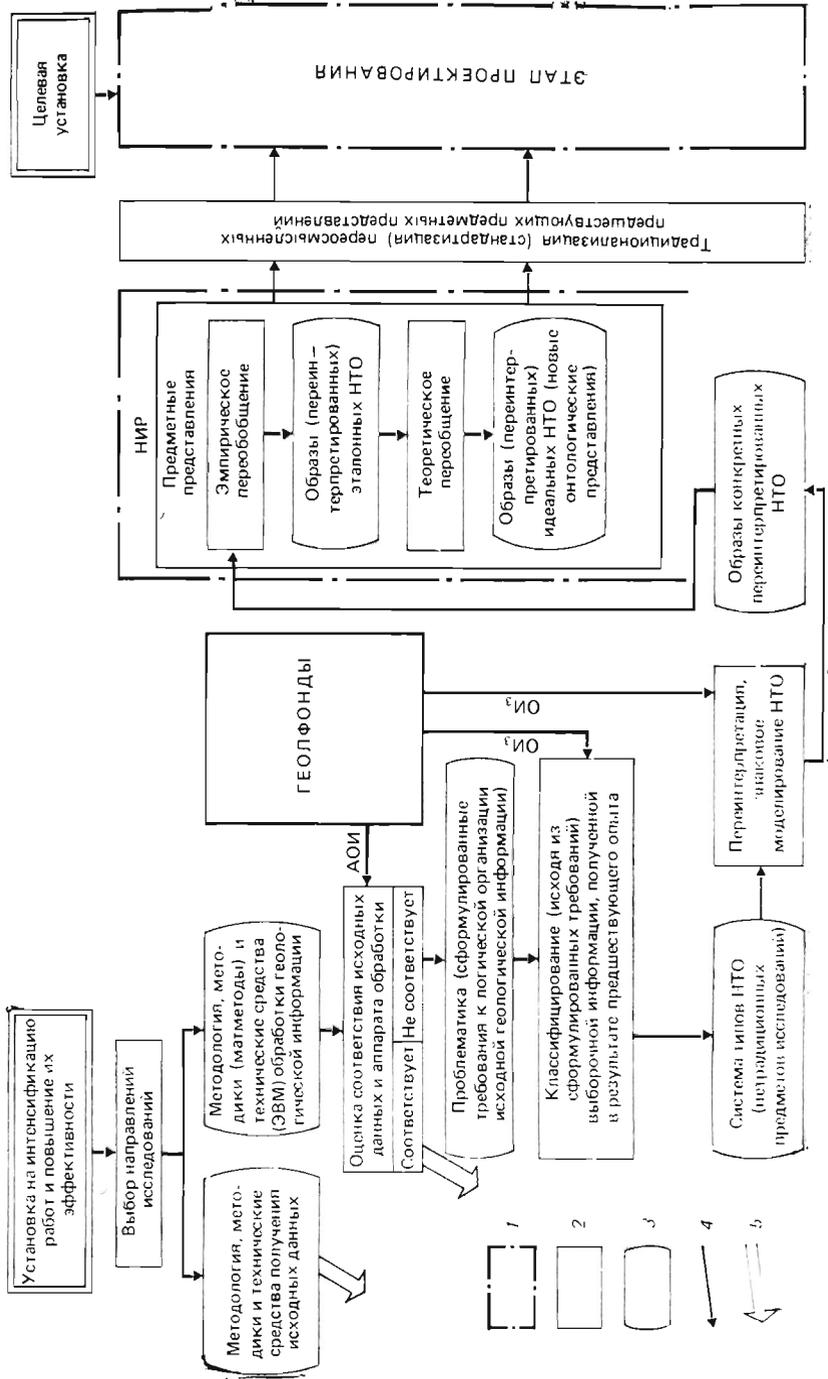


Рис. 4. Схема регуляции предметного знания в геологии минерального сырья (в условиях несоответствия требованиям методологии и методов обработки информации):

1 — основные структурно-функциональные элементы (детальнее их структуру и связи см. на рис. 1); 2 — процедуры регуляции в условиях нарушения равновесия системы предметного и деятельностного знания; 3 — результаты реализации процедур регуляции; 4 — функциональные связи; 5 — направления исследований, не вызывающие возмущающих воздействий на существующее предметное знание

в значительной степени себя исчерпавший. Настало время и геологии несколько перестроить структуру знания с учетом требований, предъявляемых междисциплинарным математическим аппаратом обработки данных, как это происходит сейчас в других областях знания [9].

Оптимальное средство переорганизации предметного геологического знания — системный подход, являющийся характерной чертой современного этапа НТР и открывающий перспективы более глубокого уровня познания окружающего нас мира. Если конкретизировать наиболее важные направления воздействия системного подхода в геологии, то прежде всего следует отметить, что представить геологический предмет в виде системы — значит формально зафиксировать ее элементы, связи (отношения), процедуры преобразования, т. е. характеризующую данный предмет и соответствующие объекты ГИ. А именно такое формализованное представление информации и есть тот язык, который корректно воспринимается математическими методами и ЭВМ.

Сейчас среди геологов быстро растет понимание этой важнейшей миссии системного подхода в геологии, стремление использовать предоставляемые этим подходом возможности [18, 28, 39, 40].

В результате такой направленной переорганизации данных на той или иной методологической, в частности системной, основе возникают нетрадиционные для данного момента типы объектов (предметы исследований), из которых наиболее адекватные реальности со временем будут восприняты и признаны, войдут в число традиционных. Процесс этот объективно неизбежен и потому ускорение его и управление им (включение в структуру научно-исследовательских работ отрасли) существенно повысили бы эффективность не только применения математических методов и ЭВМ в геологии, но и геологоразведочной деятельности в целом.

Нетрадиционные геологические объекты (НТО) возникают и из-за «сюрпризов» природы, т. е. тогда, когда на основе получаемого эмпирического материала вырисовывается геологическая ситуация, не согласующаяся с теоретическими предметными представлениями (рис. 5). Например, в процессе бурения обнаруживается залежь нефти. Она не может быть вне ловушки, но при определении типа последней выясняется, что ее нельзя соотносить ни с одним из традиционных типов ловушек. В качестве другого примера можно указать многочисленные случаи несоответствия эмпирических данных и существующих теоретических представлений в процессе изучения геологии Мирового океана. Это впослед-

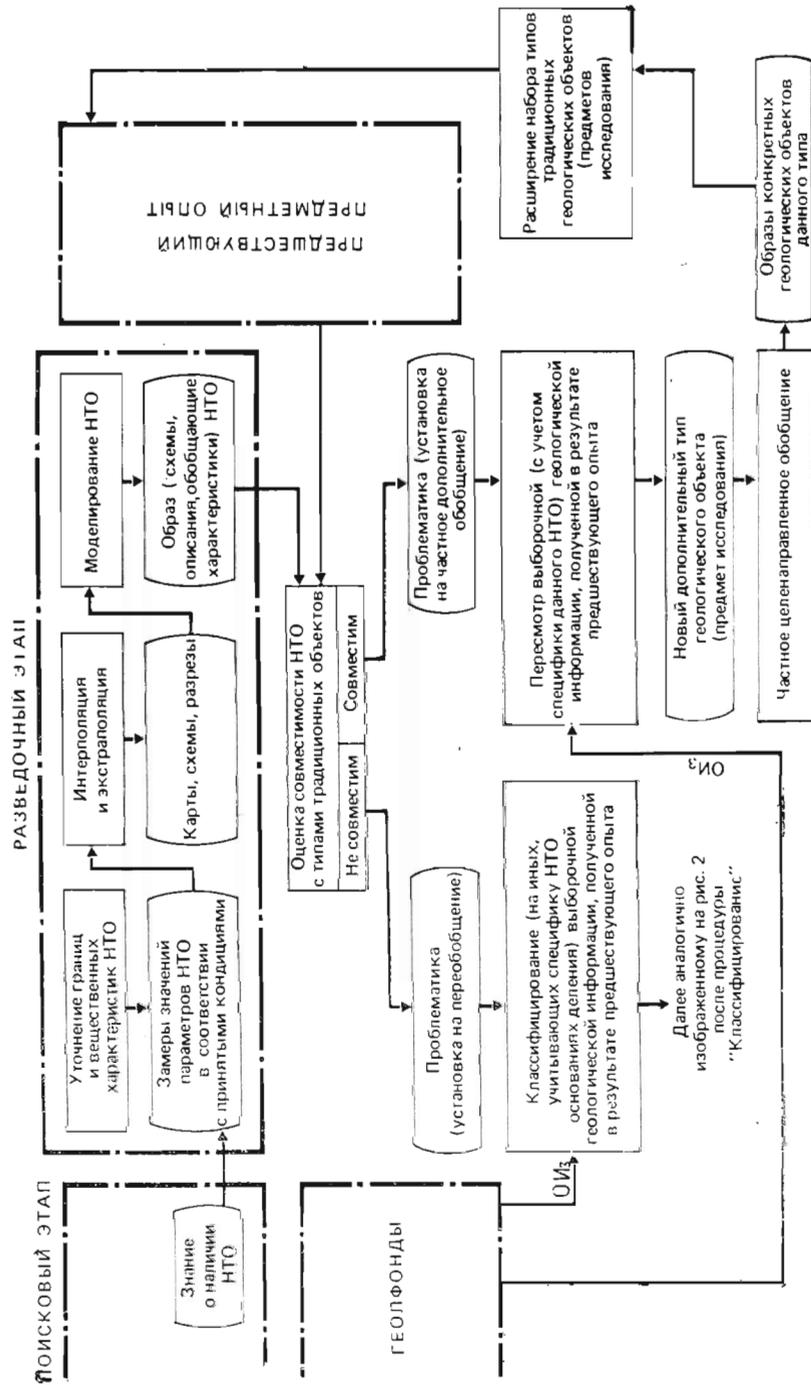


Рис. 5. Схема регуляции предметного знания в геологии минерального сырья при появлении нетрадиционного типа геологических объектов.

Условные обозначения те же, что на рис. 4

ствии привело к возникновению (в рамках теории литосферных плит) целого набора новых, нетрадиционных с позиции прежнего опыта предметов исследования. Так, в результате трудностей с реализацией процедуры геологической интерпретации возникает знание о существовании НТО. Затем идет изучение его по стандартной схеме («этап разведки») с итогом в виде образа НТО, а далее решается вопрос о совместимости НТО с типами традиционных объектов (предметами исследования), т. е. о том, можно ли безболезненно, без явных противоречий встроить этот новый тип объекта в существующую структуру предметного знания. Если нельзя, то неизбежна кардинальная перестройка структуры предметного знания в рамках соответствующей проблематики (так произошло в проблематике глобальной тектоники), если можно, то дальнейшая деятельность ограничивается обобщением данных по новому типу объектов и добавлением нового предмета исследования к уже существующим традиционным предметам.

Выбор между «нельзя» и «можно» очень часто определяется не только и не столько научными, сколько психологическими соображениями. Геолог, столкнувшийся с появлением НТО, даже если последний не согласуется с существующим предметным знанием, как правило, психологически и методологически не готов «замахнуться» на существенную перестройку всей совокупности предметов исследования. К этому могут побудить лишь действительно фундаментальные и всем очевидные открытия, которые весьма редки. В обычных случаях представляется гораздо более безболезненным «закрывать глаза» на наметившуюся несовместимость и заняться целенаправленным изучением данного НТО независимо от других типов объектов. В этом одна из основных причин обилия узких специализаций в геологии. Так, в геологии нефти и газа это привело к появлению групп геологов и целых направлений, специализирующихся на выявлении предельно узких, специфических типов ловушек, связанных, например, только с погребенными рифами, или только с барями, или только с зонами выклинивания, или только с так называемыми «запечатанными» залежами. Под эти узкие специализации разрабатываются многочисленные методики работ, что в конечном итоге приводит к распылению сил и средств.

Сегодня геология уже явственно стоит перед выбором — либо продолжение безудержного увеличения числа в значительной мере независимо вводимых предметов исследования, либо методологически контролируемое ограничение числа системно-увязываемых предметов геологии. Невозможность долго идти по первому пути уже ощущается большинством геологов. Переход же на второй путь требует значительных и, главное, организованных интеллек-

туальных затрат. Периодическая системная реорганизация предметов геологии — дело нетривиальное, равносильное получению принципиально нового научного результата, требующее очень глубокого проникновения в области как проблемно-содержательного, так и методологического знания. Но это единственное средство, способное остановить неконтролируемое разрастание совокупности предметов исследования геологии, противопоставить ему тенденцию к системности, характерную, как отмечалось, для современного этапа НТР. В этом заключается еще одна важная миссия системного подхода в геологии.

Процедура получения объективированной информации (геологической интерпретации) для подавляющего большинства геологических предметов и объектов в настоящее время логически не оформлена (не формализована). Это связано с тем, что из определений этих предметов исследования и из стоящих за этими определениями, как правило, неоднозначных онтологических представлений зачастую нельзя вывести и сформулировать в явном виде правило (алгоритм) выделения соответствующих объектов в конкретных геологических условиях. Это обстоятельство, более чем какое-либо другое, негативно отражается на состоянии информационного фонда и, следовательно, на эффективности геологоразведочной деятельности в целом.

Отсутствие фиксированных правил выделения геологических объектов на практике приводит к тому, что в одних и тех же геологических условиях границы однотипных объектов проводятся произвольно. Ни одна из многочисленных (зачастую противоречивых) дефиниций даже таких фундаментальных понятий, как месторождение, нефтегазоносные бассейны, металлогенические провинции, не регламентирует строго способ выделения соответствующих объектов на практике. Это достаточно наглядно показано в гл. 2. В результате неоднозначно выделяемые на практике геологические объекты оказываются несопоставимыми. Искать же объективные закономерности размещения произвольно выделенных объектов — занятие малоперспективное. В этом — один из барьеров на пути к более глубокому познанию закономерностей размещения скоплений минерального сырья, а соответственно и повышения эффективности поисков этих скоплений.

Такая ситуация порождает огромные скрытые потери дорогостоящей геологической информации. Как отмечалось, объективированная информация, составляющая значительную, а часто и основную часть информационного фонда, жестко привязана к соответствующему геологическому объекту, способу его выделения. Например, такие параметры месторождения, как общие запасы, размеры, число залежей, тип, средняя глубина, стратиграфическая приуроченность и многие другие, непосредственно зависят от принципа объединения залежей в данное месторождение. Использование такой информации под углом зрения новой поисковой концепции, которая обычно подразумевает иной принцип выделения объектов (например, группирования залежей) или вообще иные

объекты, как правило, невозможно. Так, имеющаяся в кадастрах месторождений нефти и газа обширнейшая объективированная информация, организованная исходя из «антиклинальной» теории поисков, оказалась неподдающейся целенаправленному анализу с позиций выявления неантиклинальных ловушек нефти и газа. Здесь может быть условно проведена аналогия с нефтью, которая при современной технологии извлекается из залежи на поверхность в среднем лишь на 30—40%. Преобладающая менее активная часть залежи остается в недрах из-за сцепленности нефти с вмещающей породой. Так и в данном случае, значительная часть информации из-за жесткой «сцепленности» с несопоставимыми геологическими объектами также оказывается несопоставимой, пассивной, не поддающейся переработке под новым углом зрения.

Для преодоления этого приходится, сохраняя накопленную ранее объективированную информацию, вновь обращаться к координатно-привязанной информации (НИ<sub>2</sub>) и на ее основе формировать новую объективированную информацию, привязанную к другим, иначе выделенным типам геологических объектов. Если бы процедура геологической интерпретации (т. е. выделения объектов) была формализована, то такой проблемы не существовало бы. Вполне достаточно было бы хранить НИ<sub>1</sub> и (или) НИ<sub>2</sub>, а также библиотеку алгоритмов выделения объектов в геологической среде, т. е. получения соответствующей им объективированной информации. При этом открываются широкие возможности более целенаправленного получения последней по запросам. Разработать алгоритмы выделения всех геологических объектов трудно, и потому массивы объективированной информации безусловно сохраняются, но надо стремиться свести их к минимуму.

Одно из следствий хранения ОИ вне связи с НИ — неизбежность многократного дублирования значительной части хранимой информации. Так, многие геологические параметры, общие для сотен и даже тысяч месторождений различных видов минерального сырья (например, тектоническое положение — древняя платформа), повторяются при описании каждого расположенного в пределах этой платформы месторождения, вместо того чтобы храниться в одном месте в виде модели геологической среды, выраженной первичной или вторичной необъективированной информацией.

Таким образом, неформализованность процедуры геологической интерпретации, во-первых, обуславливает несопоставимость объективированной информации, что, независимо от способа и технических средств ее обработки, негативно сказывается на качестве получаемых выводов и соответственно на эффективности геологоразведочных работ. Во-вторых, тот же фактор обуславливает огромные непроизводительные финансовые, материальные, трудовые, технические затраты на хранение колоссального и все возрастающего объема объективированной информации, поддающейся лишь ограниченной переработке. Со временем на это будет расходоваться все большая часть усилий по созданию, развитию

и функционированию отраслевой сети автоматизированных банков геологических данных. Такова все возрастающая цена, которую приходится платить за отсутствие формализованной технологии обработки геологической информации. Совершенствование процедуры геологической интерпретации (получения ОИ) — одна из стержневых методологических проблем геологии, от решения которой в значительной мере зависят темпы и глубина НТП в геологической отрасли. Эта проблема одновременно относится и к сфере геологической науки, и к сфере геоинформатики. В связи с этим важно рассмотреть соотношение между различными сферами геологоразведочной деятельности.

### 1.3. ИНФОРМАТИКА В ГЕОЛОГИЧЕСКОЙ ДЕЯТЕЛЬНОСТИ: ФУНКЦИОНАЛЬНЫЙ И ОРГАНИЗАЦИОННЫЙ АСПЕКТЫ

Эффективность геологической деятельности определяется не только развитием технических средств и методов регистрации параметров недр, но и в не меньшей степени возможностями обработки получаемой при этом фактографической ГИ. Исторически обусловленная система получения и преобразования информации в геологии довлеет над методами ее обработки, вынуждая последние занимать пассивную позицию приспособления. Но широкое внедрение ММ и ЭВМ потребует принципиального изменения ситуации: ММ должны влиять на способ сбора данных, вся система геологических наблюдений и опробования должна определяться требованиями последующей их обработки математическими методами [5].

Весь комплекс рассмотренных выше методологических проблем, связанных с получением и преобразованием информации, является предметом исследования информатики — дисциплины, изучающей закономерности получения, отбора, хранения передачи, преобразования и применения информации в производственной и научной деятельности [19]. В последнее время именно в связи с широким использованием в народном хозяйстве автоматизированных средств обработки информации понимание функций информатики в результате введения принципиальной поправки на научно-технический прогресс существенно расширилось. Сейчас информатика — это комплексная научная и технологическая дисциплина, изучающая прежде всего важнейшие аспекты разработки, проектирования, создания, «встраивания» машинных систем обработки данных, а также их воздействие на жизнь общества, механизированная технология сбора, передачи, переработки и использования информации. Таким образом, информатика в современном понимании синтезирует трансформированные под соответствующим углом зрения разработки, полученные в рамках ряда других дисциплин, в первую очередь методологии и логики науки, прикладной математики и кибернетики, компьютерной техники и технологии, психологии и социологии, а также той или иной предметной области.

Сфера приложения средств, методов и требований информа-

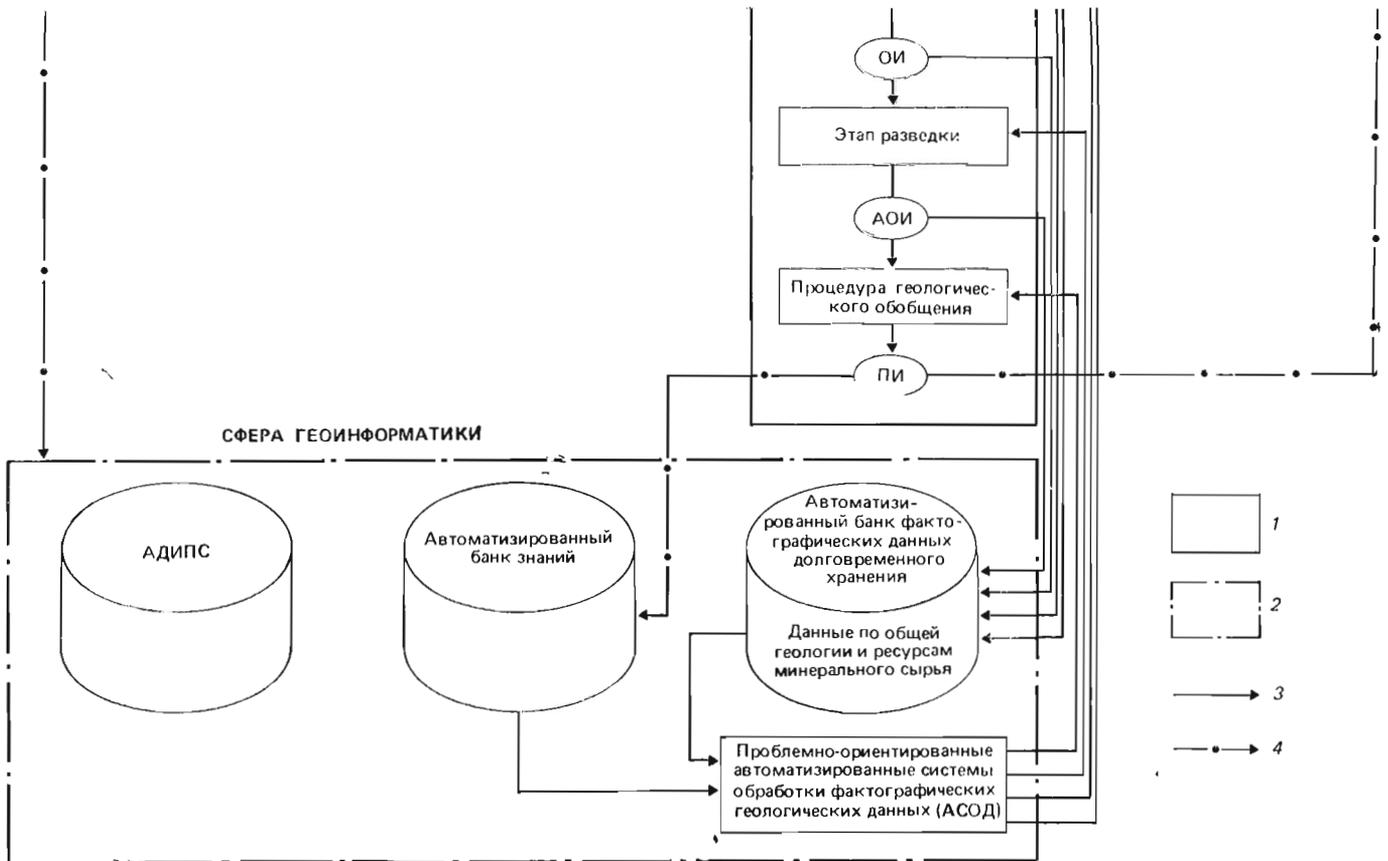
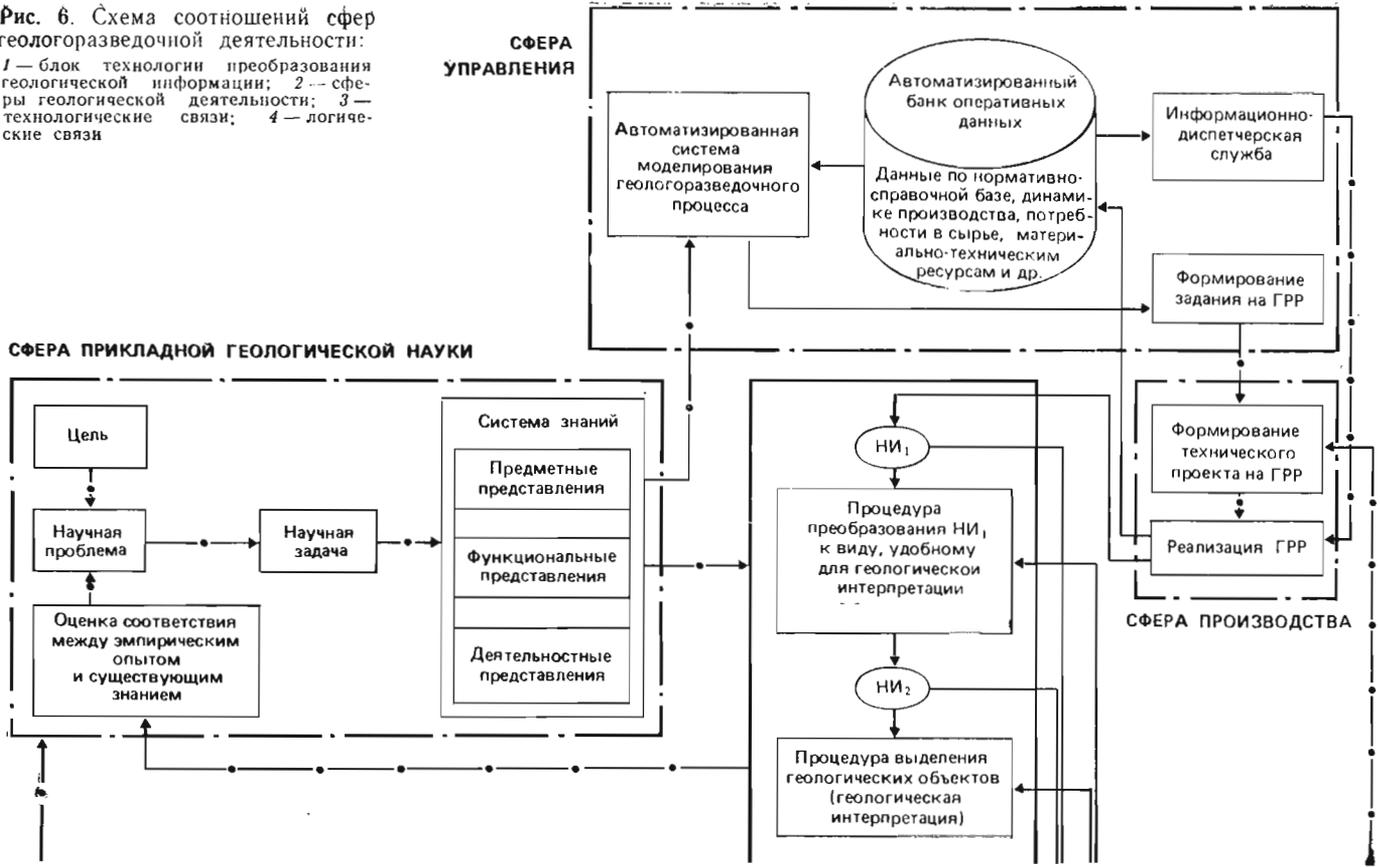
тики к геологической информации в дальнейшем для краткости обозначается термином «*геоинформатика*».

В связи с информационным выражением результатов геологоразведочного производства любая сфера деятельности в рамках последнего достаточно четко соотносится с определенными типами информации и соответствующими процедурами ее получения и преобразования. Поэтому для очерчивания сфер деятельности и анализа их соотношения удобно воспользоваться приведенной выше типологией геологической информации. Намечающиеся здесь связи в сильно генерализованном виде отражены на рис. 6. На этой схеме структура выделенных сфер деятельности детализирована лишь в той мере, в какой это нужно для дальнейшего изложения, хотя фактически она много сложнее.

Широкое внедрение ММ и автоматизированных средств обработки геологической и оперативной геологоразведочной информации породило существенное изменение отраслевой функциональной структуры, обусловленное резким повышением роли и обособлением сфер управления и геоинформатики. До этого вся отраслевая организационная структура основывалась на принципе двух основных сфер деятельности — прикладной науки и производства, в рамках которых реализовывались функции управления и геоинформатики. При этом функции сфер науки и производства не были четко (формально) разграничены. Для этого имелась достаточно солидная теоретическая основа в виде представления о том, что «геологические производственные организации являются научно-исследовательскими в обычном понимании этого слова» [4, с. 93]. На практике это привело к тому, что объективная установка на интеграцию науки и производства начала реализовываться в направлении не только внедрения в производство достижений науки (как элементов производственной технологии), но и дублирования производством функций науки. В производственных организациях появились многочисленные тематические партии, группы обобщения и т. п.

На стадии, когда обработка информации в подавляющей части осуществлялась «вручную», это было в некоторой степени оправдано, поскольку мощностей НИИ было недостаточно для научного переосмысления всего объема информации, получаемого производством. Но в условиях широкого внедрения ММ и ЭВМ при обработке данных издержки такого подхода становятся гораздо более ощутимыми. Вопрос здесь не только и даже не столько в нерациональном использовании сил и средств, сколько в усугублении состояния фактографической геологической информации. В условиях упоминавшейся неформализованности процедур преобразования информации и при независимой геологической интерпретации многочисленными разобшенными мелкими коллективами невозможно обеспечить приемлемый уровень унификации геологических объектов, к которым привязывается объективированная информация. Кроме того, снижается мера ответственности специализированных научных организаций за результативность

Рис. 6. Схема соотношений сфер геологоразведочной деятельности: 1 — блок технологии преобразования геологической информации; 2 — сферы геологической деятельности; 3 — технологические связи; 4 — логические связи



производства, поскольку программа последнего часто полностью или в значительной степени определяется собственными тематическими группами, а рекомендации научных организаций оказываются на уровне советов, которые учитывать не обязательно. Поэтому четкое разграничение функций сферы науки и сферы производства становится насущной потребностью.

Граница между ними — динамическая, она смещается по мере того, как та или иная научная разработка становится элементом производственной технологии. Последнее возможно лишь при условии, что соответствующая процедура формализована, представлена в виде инструкции, методики, алгоритма, обеспечивающих однозначность действий в однотипной ситуации. Сегодня положение таково, что в общем случае раздел между сферой науки и производством проходит на уровне между необъективированной и объективированной геологической информацией. Если для получения последней нет четкого алгоритма, то это уже не технология и не производство, а область научно-исследовательских изысканий. Но в частных, пока редких случаях, когда для отдельных предметов геологического исследования процедура геологической интерпретации оформлена логически корректно, граница между научной и производственной деятельностью может смещаться, и тогда получение объективированной информации становится уже функцией производства. Поскольку автоматизации поддаются лишь формализованные, т. е. подготовленные наукой для внедрения на производстве, процедуры, то научный уровень производства на каждый период определяется в конечном итоге тем, каковы возможности (пределы) автоматизации получения и преобразования геологической информации.

Ситуация неразграниченности (смещения) функций, подобная той, что отмечалась выше для сфер науки и производства, создавалась сейчас также для сфер управления и геоинформатики. Причина этого заключается, видимо, в том, что в этих сферах широко используются средства автоматизированного банка данных. Поэтому зачастую считается, что реализация задач управления и геоинформатики может быть осуществлена одним коллективом и едиными технико-технологическими средствами. Однако по существу информационные и, как следствие, эксплуатационно-технологические параметры автоматизированных систем в сфере управления и сфере геоинформатики принципиально различны. В первой используется в основном оперативная (быстро теряющая актуальность, быстро обновляемая, не подлежащая длительному хранению) информация. В ее составе собственно геологическая информация составляет незначительную часть и выступает преимущественно в качестве специфического дополнения к нормативно-справочной базе. Используется законченная, увязанная сверху донизу система классификации и кодирования информации при однозначности (единственности) данных и относительно жестком задании связей. Необходимая предпосылка функционирования таких систем — наличие не только прямой, но и обратной связи

с управляемым объектом и соответствующие требования ко времени реакции системы. Это и есть АСУ в ее подлинном смысле.

В сфере же геоинформатики оперируют только с собственно геологической информацией, имеющей непреходящее значение и подлежащей долговременному хранению. В силу специфики геологического языка и отмечавшейся недоразработанности процедур преобразования геологической информации, логические связи между отдельными элементами последней здесь несравненно сложнее (в частности, многовариантность связей, множественность значений параметров), чем в информации, используемой в сфере управления, а машинное представление этих связей требует углубленных геологических знаний. Актуализация геологической информации сводится не столько к динамическому обновлению, сколько к дополнению и корректировке. Из-за отсутствия единой системы классификации и кодирования средства банка данных часто могут быть использованы лишь в весьма ограниченном виде. Отсутствует такой специфический признак АСУ как обратная связь.

Для того чтобы отличить от действительных АСУ, системы, работающие с геологическими и геолого-экономическими данными, часто обозначают термином АСОД (автоматизированная система обработки данных). Попытки совмещения АСУ и АСОД в рамках одной системы приводят к универсализации, достигаемой неизбежно за счет потери качественных свойств отдельных элементов (усложнение реализации запросов, снижение уровня специализации и др.). Самое главное заключается в том, что при этом затушевываются информационные связи между различными сферами геологоразведочной деятельности, обуславливая нечеткость информационно-функциональной структуры и несоответствие ей структуры организационной. Одно из проистекающих отсюда следствий — охарактеризованное положение с формализацией процедур преобразования и унификацией системы классификации и кодирования собственно геологической информации. Не в последнюю очередь это связано с тем, что разработка соответствующих проблем, относящихся к сферам геологической науки и информатики, возложена на сферу управления (АСУ). Совмещение под флагом АСУ самостоятельных сфер деятельности представляется искусственным и нерациональным.

Четкое формализованное разграничение функций сфер науки и производства, сфер управления и геоинформатики нельзя осуществить изолированно. Это возможно только в соответствующих организационных рамках специализации и кооперирования всех выделенных сфер в получении и преобразовании определенных типов информации (см. рис. 6). В этой структуре, где сфера управления выступает в качестве постановщика целей и задач, сфера науки разрабатывает общую стратегию достижения целей и частные поисковые концепции, а сфера производства реализует последние, сфера геоинформатики выступает как информацион-

но-технологический сервис сфер геологической науки и производства, как ОТК продуктов геологоразведочного процесса, представленных в виде геологической информации.

Среди важнейших задач, реализация которых требует специализированных средств и методов информатики, отметим следующие: формирование требований по формализации процедуры выделения и способов задания геологических объектов; типизация последних и геологических задач; унификация и агрегирование наборов регистрируемых геологических параметров; совершенствование системы сбора ГИ с учетом применения ММ и ЭВМ; формирование отраслевой системы классификации и кодирования данных; разработка критериев оценки достоверности исходных геологических данных и влияния этого на результат обработки последних; стандартизация машиноориентированной отчетной документации по результатам работ; разработка принципов моделирования геологической среды на основе системной увязки совокупности фиксируемых параметров; стандартизация информационных запросов пользователей; оптимизация структуры информационного фонда; разработка заданий на системное проектирование банков данных и АСОД геологии.

Решение указанных задач, направленных в совокупности на разработку принципов формально-логической организации и иерархизации (агрегирования) геологических данных применительно к ММ и машинной технологии обработки, возможно только на основе централизованного объединения и координации усилий ведущих научных геологических центров. Реализуемые ныне программы формирования интегрированных банков геологических данных в рамках отраслевых АСУ министерств, ведущих геологические изыскания, создают необходимую основу объединения усилий. Для обеспечения такой координации возникает потребность в создании методологической службы геологической отрасли (организационного оформления сферы геоинформатики). Назначение этой службы не только и даже не столько в том, чтобы выполнять «санитарные» функции по расчистке образовавшихся «информационных завалов», сколько в создании новой структуры геологического знания на основе системного подхода — методологии современного этапа научно-технического прогресса. Любая методологическая разработка в геологии — задача труднейшая и нетривиальная, поскольку она в конечном итоге направлена на перестройку геологического мышления. Но это задача огромнейшей важности и затраты на ее реализацию многократно окупаемы. Здесь опять можно провести условную аналогию с нефтью. Увеличение коэффициента нефтеотдачи, составляющего в среднем 30—32 %, всего на 1 % обеспечило бы прирост извлекаемых запасов нефти только для развитых капиталистических и развивающихся стран, примерно на 2,5 млрд. т, что равносильно открытию новых месторождений, равных всем нефтяным месторождениям Северного моря. Подобно этому увеличение отдачи от обработки имеющейся геологической информации было бы равно-

сильно получению новой информации в результате огромных дополнительных объемов геологоразведочных работ.

Повышение результативности обработки ГИ требует также, чтобы понимание целей и средств геоинформатики было достоянием широкого круга геологов, входило бы в число элементов профессиональной подготовки. Это может быть достигнуто лишь введением соответствующих кафедр (курсов лекций) в отраслевых ИПК и вузах геологического профиля. Начало этому в нашей стране положено созданием на факультете газонефтяной геологии, геофизики и геохимии Московского института нефтехимической и газовой промышленности им. И. И. Губкина первой в стране кафедры системно-геологических исследований литосферы.

Вместе с тем в условиях интенсификации науки и производства, вызывающих быстрое увеличение объемов хранимой и обрабатываемой ГИ, а также возрастание «цены» достоверности результата обработки ГИ, повышается значение фактора времени. Геологу уже в ближайшее время придется практически использовать в своей работе ММ и ЭВМ и потому необходимо уже сейчас иметь представление о том, что вызывает необходимость «состыковки» языков геологии и математики, в чем выражается и как сказывается отсутствие такой «состыковки», каковы доступные средства и методы реализации этой задачи. Рассмотрению этих вопросов на конкретных примерах геологии и посвящены последующие главы этой книги.

## МЕТОДОЛОГИЧЕСКИЕ ОСНОВЫ ФОРМАЛИЗАЦИИ ГЕОЛОГИЧЕСКОЙ ИНФОРМАЦИИ

### 2.1. ИДЕАЛИЗАЦИЯ И ФОРМАЛИЗАЦИЯ. ИХ ЗНАЧЕНИЕ И ФУНКЦИИ В ГЕОИНФОРМАТИКЕ И ГЕОЛОГИЧЕСКОМ ПОЗНАНИИ

Идеализованная информация (в частности, ПИ<sub>1</sub> и ФИ<sub>2</sub>) возникает как результат идеализации — мысленного конструирования понятий об объектах, не существующих и неосуществимых в действительности, но таких, для которых имеются прообразы в реальном мире [44]. В свою очередь, идеализация представляет собой частный случай абстрагирования — мысленного отвлечения тех или иных определенных свойств от множества свойств исследуемого предмета [19]. В реальном мире нельзя увидеть «ловушку вообще», «залежь вообще», а только конкретную ловушку, конкретную залежь. Подобно этому выбор того или иного признака в качестве основания деления при построении классификации означает отвлечение от всех других признаков, свойств классифицируемых предметов и эта операция абстрагирования повторяется на каждом таксономическом уровне (класс, подкласс, вид и т. д.).

Познавательная сила абстракций (идеализаций) заключается в том, что они, отвлекаясь от частных, несущественных признаков, фиксируют то общее и существенное, что присуще всему множеству охватываемых абстракций однородных в каком-то отношении предметов. А знание этого общего и существенного — это инструмент более глубокого познания как отдельных предметов данного множества, так и всего множества в целом. Вскрыть закономерности развития какого-либо предмета, явления можно только в форме абстракции. Именно на это обстоятельство обратил внимание В. И. Ленин, который отмечал, что все научные (правильные, серьезные, не вздорные) абстракции отражают природу глубже, вернее, полнее<sup>1</sup>. При этом всякая научная идеализация в конечном итоге обязана своим возникновением реальному миру и отображает (моделирует) его. Именно этим научная идеализация принципиально отличается от всякого идеализма. Вскрыв существенное с помощью абстракции (идеализации), исследователь возвращается к конкретному, но он уже обогащен знанием закономерностей возникновения и развития конкретного, что расширяет предсказательные возможности.

Познавательные возможности абстракций были оценены на начальных этапах развития науки. Наглядным примером может служить геометрия, которая строилась как абстрактная наука,

изучавшая общие свойства идеальных фигур. Такие фигуры в чистом виде в природе не встречаются (линии, имеющие длину, но не имеющие ширины, точки без длины и ширины и т. п.). Отвлеченные понятия были восприняты не сразу и не без борьбы. Так, Протагор — видный философ времени становления геометрии, выступал против этой науки на том основании, что никто не видел линий без ширины или чтобы линия касалась круга только в одной точке, а не по маленькому отрезку. Однако вскоре оказалось, что такой подход позволяет решать практические задачи, которые невозможно решить иными способами. Например, на практике из-за погрешности измерения (даже при современных точных приборах, не говоря уже о тех, которыми пользовались в Древней Греции), нельзя доказать, что сумма углов треугольника равна 180°. Таким образом, уже древние греки понимали, что установить правильность какого-либо свойства для всех объектов некоторого класса можно только путем абстракции и логических операций.

Тем не менее в геологии и в наши дни любая теоретическая разработка считается тем более верной и приемлемой, чем ближе она к эмпирическому уровню. Более или менее глубокая абстракция часто рассматривается как отрыв от реальности, появляются обвинения в надуманности, искусственности, нежизненности и т. п. Такое положение, возможно, является следствием того, что геологические науки в целом еще никак не переросли эмпирический индуктивный уровень, и поэтому теоретические разработки рассматриваются с позиции именно этого уровня. И это при том условии, что на примере многих развитых наук уже давно и убедительно доказано, что ни одна гипотеза и теория не была выведена непосредственно из индуктивного обобщения эмпирических (чувственно воспринимаемых) фактов.

В основе всех теоретических построений лежат абстракции, тем более глубокие, чем фундаментальней и универсальней теория или гипотеза. В самом деле, в результате индуктивного обобщения каких эмпирических наблюдений данных могли возникнуть статистическая физика, электродинамика, квантовая механика, частная и общая теория относительности? Методологический анализ этих теорий, изучение истории их возникновения, наконец, свидетельства самих их создателей приводили к единодушному выводу о невозможности на пути индуктивного обобщения эмпирических данных придти к открытию основных теоретических принципов. Этот вывод поддерживают в настоящее время по существу все специалисты по методологии науки.

Абстрактность свойств и моделей объектов исследования, оснований деления в классификациях и т. п. никоим образом не может рассматриваться как слабость, недостаток соответствующих теоретических построений. При этом происходит огрубление геологической реальности, но, как совершенно справедливо отмечают И. И. Абрамович и др. [27], упрощение и идеализация выступают как плата за увеличивающиеся познавательные возможности.

<sup>1</sup> Ленин В. И. Полн. собр. соч. т. 29, с. 152.

Именно недооценка этого и является основной причиной того, что для многих объектов геологии нет соответствующих формально зафиксированных идеальных предметов. Последнее же в не меньшей степени, чем недостаток информации, обуславливает неоднородность геологической интерпретации строения недр (см. гл. 1).

Одной из важнейших форм абстрагирования (идеализации), особенно на современном этапе НТР, является формализация, определяемая как «отображение результатов мышления в точных понятиях или утверждениях. В этом смысле формализация противопоставляется интуитивному мышлению. В математике и формальной логике, где формализация наиболее развита, под формализацией обычно понимается отображение содержательного знания в знаковом формализме» [44, с. 743]. Последнее связано с тем, что язык можно рассматривать как знаковую систему, а его элементы как знаки особого рода. Поэтому любая языковая деятельность есть не что иное, как знаковое моделирование отображаемых знаками объектов в процессах коммуникации и познания. Знаки языка науки — это термины и символы, которые всегда имеют содержательную интерпретацию, т. е. им всегда поставлены в соответствие определенные предметы. Если же отвлечься от содержательного значения знаков (т. е. смысла терминов), а рассматривать только форму знаков и правила оперирования ими, базирующиеся исключительно на графическом представлении знаков, то мы получим законченную формальную систему. Наглядный результат конечной формализации являет любая алгебраическая формула, например  $(a-b)^2 = a^2 - 2ab + b^2$ . Здесь мы отвлекаемся от конкретных значений знаков  $a$  и  $b$  и оперируем лишь формой знака и правилами взаимоотношения знаков.

Правда, это отвлечение не абсолютно. Фактически и практически формализованные системы строят всегда применительно к некоторому содержанию, лишь потом абстрагируются от него и в целях наибольшей точности создают формальную систему, чтобы снова дать ей содержательную интерпретацию [47]. Формализованная система выполняет свою задачу лишь в том случае, когда ее элементы и отношения могут быть содержательно интерпретированы.

Формальные системы могут быть построены только при определенном уровне подготовленности (точности и адекватности) предметного языка соответствующей науки. Язык считается точным, если все его термины однозначно определены и каждое предложение, содержащее такие термины, построено по заранее определенным правилам; язык считается адекватным, если получаемые в нем предложения могут описывать все существующие ситуации в области объектов, информацию о которых выражает, хранит и передает данный язык. Таким образом, формализация языка начинается по существу с достижения однозначности и недвусмысленности терминов, т. е. со строгого в логическом отношении определения вновь вводимых понятий и уточнения в логическом плане уже используемых понятий (их экспликации).

И уже на этой основе строится настоящий формализованный язык, отличающийся от уточненного предметного языка еще большей точностью и адекватностью. Это достигается [47]: введением вместо слов обычного языка знаков, образующих алфавит формализованного языка (т. е. список исходных знаков — имен, терминов); семантическими (смысловыми) правилами, определяющими значения исходных знаков — имен, терминов; точной и явной формулировкой правил построения из исходных знаков (терминов, имен, выражений) сложных знаков и знаковых систем; точной и явной формулировкой правил перехода от одних сложных знаков и знаковых выражений к другим.

Напомним, что современные кибернетические средства обработки информации «понимают» только формализованный язык, и поэтому возможности и эффективность использования автоматизированных систем для обработки геологической информации определяются состоянием понятийно-терминологического аппарата геологических наук. Пока еще низкая научная отдача от применения ЭВМ и математических методов в геологии обусловлена прежде всего неподготовленностью (логическим несовершенством) предметного языка. Поэтому формализация представляет необходимую предпосылку внедрения математических методов и кибернетических средств обработки данных в отдельные отрасли естествознания, в том числе и геологию.

Однако значение формализации не исчерпывается созданием условий для математизации. Есть и другие функции формализации в науке и практике вообще и в геологии в частности. Они подробно освещены в специальных работах [34, 47]. Укажем лишь некоторые из функций формализации:

формализация приводит к получению нового знания. Знаковая система, в которой каждое понятие находится в единстве со способом его построения и оперирования, дает возможность вскрыть глубокие внутренние закономерности объектов исследования;

относительная самостоятельность знаковых систем позволяет расширить границы их применения и выявить существенную общность разнородных явлений. Следовательно, формализация может приводить к получению данных о предметах и процессах, еще не познанных. В этом случае можно говорить о предсказательной функции формализации;

формализация дает возможность конкретно фиксировать и легко обозревать (свертывать) знания. Как средство фиксации знания формализация обеспечивает краткость, четкость, определенность, логическую строгость. Так как формализация осуществляется в рамках явно и четко сформулированных правил, устраняются правила и соображения, которые подразумеваются, т. е. не сформулированы явно. Последние весьма опасны, ибо их истинность или ложность не установлены и могут вообще не поддаваться практической или теоретической проверке. Метод формализации помогает, таким образом, исключить из теоретических построений сомнительные, молчаливо предполагаемые посылки;

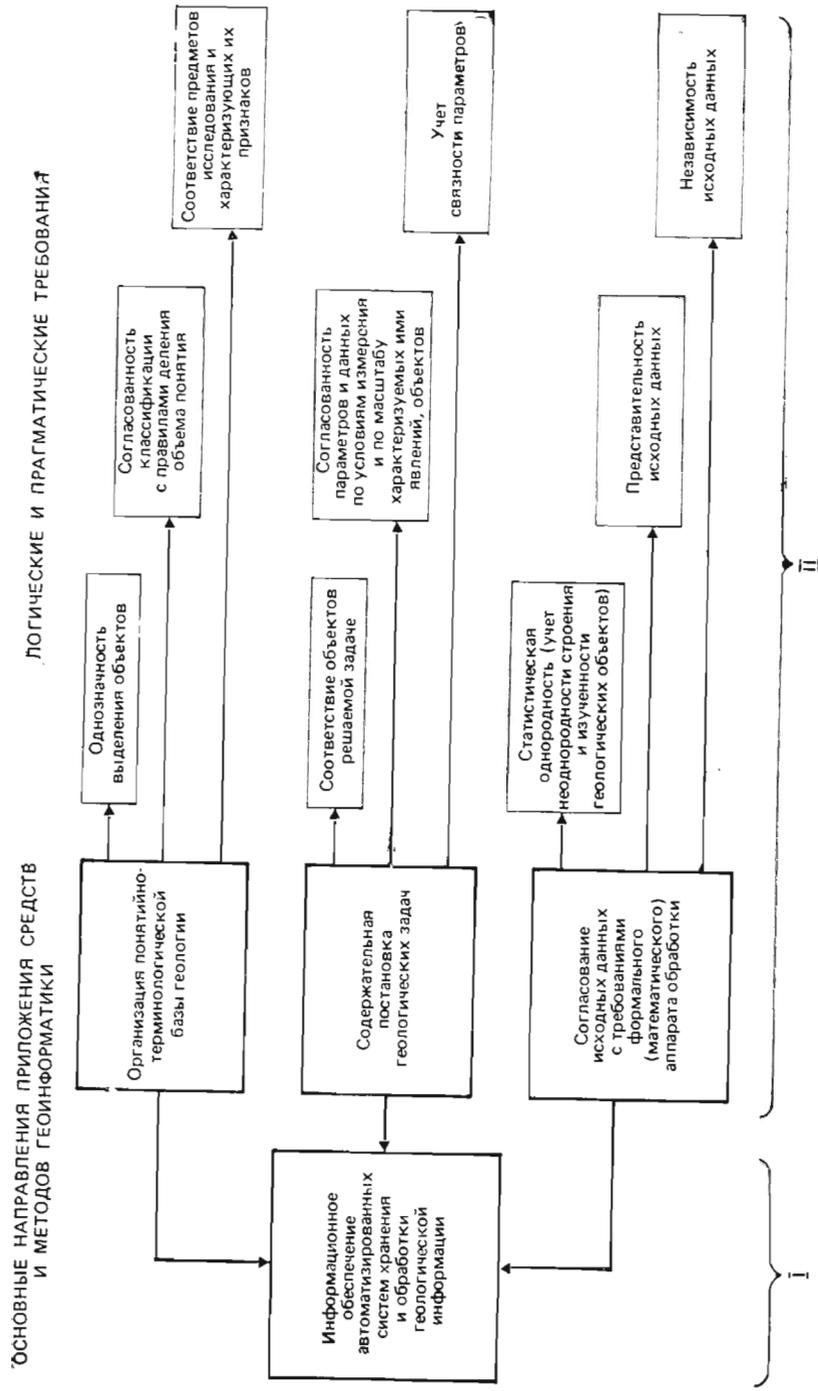


Рис. 7. Обобщенная схема приложения средств и методов геoinформатики в геологии.

I — описание см. в гл. 3, структуру — на рис. 25; II — описание см. в гл. 2

формализация дает дополнительные (формальные) возможности проверки истинности той или иной теоретической схемы.

Знаковые системы — конечный результат формализации. Общий современный уровень развития геологического языка таков, что о построении формальных знаковых систем в геологии можно говорить лишь как об отдаленной перспективе (достижение этого уровня выдвинуло бы геологию в число точных наук). Но для того чтобы достигнуть этой стадии формализации, необходимо преодолеть ее предшествующие стадии, в частности, осуществить логическое совершенствование системы понятий и терминов, задание связей между предметами исследований. Уже одно это позволило бы резко повысить надежность (достоверность) геологических построений и, как следствие, эффективность геологоразведочных работ. Поскольку в различных науках и отдельных проблемах достигнута и достижима различные стадии формализации, в термин «формализация» разные исследователи вкладывают различный смысл. Но независимо от стадии формализации ее результатом является исключение неоднозначности, неоднозначности, не сформулированных явно посылок и соображений. А поскольку развитие науки и практики, минуя формализацию, невозможно (степень формализованности характеризует уровень теоретической развитости науки), то перспективными представляются лишь те научные, методические, технологические разработки, которые поддаются формализации. Формализация — один из важнейших инструментов информатики вообще и геoinформатики в частности.

В последующих разделах данной главы рассмотрены некоторые важнейшие методологические принципы и требования, на которых основывается формализация информации. Выбор указанных принципов и требований базируется на обобщении авторами опыта (в том числе и собственного) методологического анализа и математизации геологической деятельности. На основе этого обобщения были вычленены наиболее часто встречающиеся типы логических нарушений в организации геологических данных. Каждому из этих типов было поставлено в соответствие то или иное методологическое требование (рис. 7), несоблюдение которого и обусловило данный тип логического нарушения. Поэтому методологические требования, представленные на рис. 7, отражают в основном наиболее распространенные случаи логических нарушений и не претендуют на исчерпывающую полноту охвата. Они расположены на рисунке в соответствии со значимостью, которая уменьшается слева направо и сверху вниз. Эта значимость учтена и в последовательности изложения. Подробнее основные направления применения средств и методов геoinформатики в геологии

рассмотрены ниже в главе 2, а применительно к информационному обеспечению АФИС — в главе 3.

Форма изложения материала в большей части главы (см. разд. 2.2—2.4) в значительной степени базируется на подходе «от обратного», т. е. описание приемов и установок «как надо» организовывать данные проводится на фоне систематизированных примеров того, чего следует избегать. Такая форма изложения избрана сознательно, поскольку она обеспечивает читателю большую наглядность и психологически облегчает восприятие излагаемого материала.

Следует отметить также, что рассматриваемые в данной главе вопросы касаются внутренней логики организации данных, не затрагивая вопросы внешнего представления данных для вычислительного процесса (например, вида матрицы данных, стыкующейся с конкретным вычислительным методом) и внешней формализации данных.

В связи с последним хотелось бы сделать следующее пояснение. Формализация направлена в конечном итоге на то, чтобы обеспечить однозначность понимания и действий в однотипных ситуациях. При корректно осуществленной формализации в формализованном виде должны быть представлены как форма, так и стоящее за ними содержание. Однако в геологии часто проявляется тенденция ограничить формализацию формой представления данных, не затрагивая характеризующее ими содержание. Это подробно проиллюстрировано ниже (см. разд. 2.1) на примере отсутствия правила объединения (закона композиции) залежей при выделении месторождения, в результате чего в однотипных геологических ситуациях число выделяемых месторождений нефти и газа различно. Но при этом ряд данных, характеризующих неоднозначно выделяемые месторождения (например, величина разведанных запасов, число залежей и др.), выражен в предельно формализованной — числовой форме. Вот такие случаи, когда данные представлены в формализованной форме, но выделение и представление характеризующего ими объекта не формализовано, мы и обозначаем термином «внешняя формализация». Если в модели решения геологической задачи задействована информация, базирующаяся на внешней формализации, то корректность результата решения не может быть обеспечена.

## 2.2. ОРГАНИЗАЦИЯ ПОНЯТИЙНО-ТЕРМИНОЛОГИЧЕСКОЙ БАЗЫ

Совокупность используемых в геологии понятий и соответствующих им терминов образует основу геологического языка, на котором зафиксирована геологическая информация. Основные проблемы, связанные с применением ММ и ЭВМ в геологии, лежат главным образом в сфере логической организации геологического языка и обусловлены резкими различиями в степени формализации (логической строгости) языков геологии и математики.

Эффективность, а зачастую и сама возможность применения ММ и ЭВМ для обработки геологической информации зависят прежде всего от «стыкуемости» языков геологии и математики, от возможности трансляции геологических представлений, выраженных на описательном языке геологии, на строго формализованный язык математики.

Здесь как и при любом взаимодействии разных языков требуется переводчик, и по самой сути данной проблемы эту роль в состоянии выполнить только логика и методология науки, которые располагают достаточным арсеналом средств для решения подобных проблем. Эти средства в совокупности оформлены в теорию понятий, определений и терминов. Современное обобщение приемов и правил определения понятий, принципов образования терминов приведено в работах Е. К. Войшвилло, Д. П. Горского, В. П. Даниленко, И. С. Квитко, Д. С. Лотте, К. Попа, А. А. Реформатского и др. Имеется значительное число работ, посвященных творческому приложению указанных логических средств в геологии. Среди них следует отметить работы М. Г. Бергера, А. М. Боровикова, С. Ф. Васильченко, Н. Б. Вассоевича, Ю. Н. Карогодина, Ю. А. Косыгина, В. А. Соловьева, А. А. Трофимука, И. П. Шаропова и др. В некоторых из этих работ [18, 46] рассмотрение проблем понятийно-терминологической базы отдельных проблем геологии сопровождается кратким суммированием основных положений теории понятий, определений и терминов. Поэтому во избежание дублирования этих работ мы ограничимся лишь самыми общими вводными замечаниями, необходимыми для понимания излагаемого материала.

Основным в теории понятий, определений и терминов является положение о триединстве категорий — реалии (явления, предмета), понятия и термина. Термин имеет *смысл* и *значение*, каковыми являются соответственно понятие (т. е. способ указания предмета, модель объекта) и сам предмет. Понятие характеризуется *содержанием* и *объемом*. Под содержанием понятия понимается совокупность отличительных признаков, ядром которых являются существенные признаки<sup>1</sup>, а *под объемом понятия* — множество предметов, каждый из которых имеет признаки, зафиксированные в данном понятии, т. е. совокупность предметов, охватываемых этим понятием [19].

Форма существования понятия — его научная дефиниция, представляющая собой выраженный в форме предложения результат логической операции *определения понятия*. Основной логический прием определения понятия — *определение через ближайший род и видовое отличие*. Этот прием заключается в том, что отыскиваются ближайший род для определяемого понятия и отличительные признаки, имеющиеся только у данного вида предметов и

<sup>1</sup> Существенные характеристики должны быть отражены в определении предмета исследования и на них должно базироваться выделение соответствующих объектов. Устранение таких характеристик (хотя бы одной из них) равносильно уничтожению самого предмета.

отсутствующие у всех других видов предметов, входящих в этот ближайший род. Определение начинается с указания рода. «Что значит...— по мысли В. И. Ленина,— дать «определение»? Это значит, прежде всего, подвести данное понятие под другое, более широкое»<sup>1</sup>. Но в каждый род входит много видов. Для того чтобы обособить данный вид, надо найти и указать тот специфический существенный признак, который отличает этот вид от всех остальных видов, входящих в указанный род.

Однако определения понятия через ближайший род и видовое отличие недостаточно для того, чтобы обеспечить корректность определения понятия. Необходимо выполнение ряда дополнительных требований, входящих в число логических *правил определения понятия*. В работах различных специалистов-логиков и исследователей, занимающихся «логизацией» геологии, число и набор таких правил не совпадает. Так, Н. И. Кондаков [19] приводит семь правил, Д. П. Горский [8] — восемь, И. П. Шарاپов [46] — десять, Ю. Н. Карагодин [18] — двенадцать, и эти правила в значительной мере не перекрываются. Такое положение отражает неодинаковую оценку относительной значимости отдельных правил. В разделе 1.2 уже охарактеризовано одно из них — правило реализуемости, некоторые правила будут рассмотрены ниже в связи с конкретными примерами геологических определений.

В силу тесной связи понятия и определяющего его термина, часто на первый план выдвигают терминологический аспект и, рассматривая указанные правила, говорят об определении терминов (а не понятий). Это отражает то обстоятельство, что учет указанных правил, затрагивающих значение и смысл терминов, в равной мере обязателен при формировании системы терминов.

Вместе с тем имеется ряд дополнительных требований к форме терминов, применительно к геологии, обобщенных в работе [2]. Однако следует сразу же отметить, что требования к форме терминов носят «косметический» характер, в том смысле, что несоблюдение их (в отличие от правил определения понятия) не сказывается на геологических результатах. Это вызвано тем, что связь знака<sup>2</sup> и обозначаемого им предмета условна и произвольна и никак не детерминируется свойствами самого предмета. Предметы определяют не форму знака (термина), а его значение. Разумеется стихийное возникновение или сознательный выбор тех или иных знаков всегда так или иначе обусловлен и мотивирован. Но эта обусловленность и мотивация имеют очень мало общего с необходимостью и относятся к области прагматики [47]. Сказанным мы ни в коей мере не хотели принизить значение требований, корректирующих форму терминов и способствующих удобству практического пользования ими, а стремились лишь

<sup>1</sup> Ленин В. И. Полн. собр. соч., т. 18, с. 149.

<sup>2</sup> В настоящее время подавляющее большинство лингвистов рассматривают язык как знаковую систему, а его элементы (слова, в частности термины) — как знаки особого рода.

определить место этих правил в общей системе логической организации понятийно-терминологической базы науки.

Несоблюдение принципов организации понятийно-терминологического аппарата привело к такому положению, что, как отмечают Ю. А. Косыгин и В. А. Соловьев, определений большинства геологических понятий или нет, или же они определены в логическом отношении не строго, причем это касается не только сложных, но и относительно простых и часто используемых терминов.

Издержки логического оформления предметов геологического исследования непосредственно сказываются (через геологические классификации) и на процедуре задания соотношения между отдельными типами объектов и параметрами последних. Большинство геологических классификаций не отвечает формально-логическим требованиям, что усугубляет несопоставимость объектно-привязанной информации.

Альтернативы перестройке понятийно-терминологической базы геологии на более корректной логической основе не существует. В этом, как отмечалось, состоит одна из важнейших задач геoinформатики и системного подхода.

### 2.2.1. Логические основы обеспечения сопоставимости геологических объектов и характеризующей их информации

Решение практически любой геологической задачи, любая геологическая закономерность базируются на использовании связей различной степени сложности. В свою очередь, безусловной предпосылкой корректности вывода является сопоставимость объектов, в отношении которых формулируются эти связи. Сопоставимость же геологических объектов может быть обеспечена лишь при обязательном соблюдении двух условий.

1. В рамках обобщения один термин должен иметь единственное значение, обозначать один и тот же предмет, явление. В теории понятий это условие известно как «правило однозначности». Вряд ли будет корректным обобщение, например, по так называемым литологическим ловушкам, выделенным различными исследователями, если каждый из них под этим термином понимал геологические тела, существенно различные по морфологии, генезису и вещественному составу.

2. Должно быть сформулировано в явном виде правило (алгоритм) выделения объекта из геологической реальности. Это правило должно вытекать из дефиниции соответствующего предмета геологического исследования (разумеется, если дефиниция согласуется с упоминавшимися выше правилами определения понятий). Но поскольку данное условие часто не соблюдается, то на практике это приводит к тому, что в одних и тех же геологических условиях границы однотипных объектов проводятся по-разному. Такие границы и объекты в методологической литературе принято называть произвольными [10]. Естественно, что произвольно выделяемые объекты часто несопоставимы.

В геологии нарушение указанных условий — весьма распространенное явление, и это иллюстрируется на конкретных примерах популярных в геологии предметов исследования. Однако, прежде чем перейти к этим примерам, важно уяснить истоки существующего положения. Эти истоки лежат в сфере философского мировоззрения и обусловлены тем, что формирование основных понятий геологии происходило в условиях, когда подавляющее большинство геологов (как правило, подсознательно) придерживались концепции «естественных» или «природных» тел.

В основе этой концепции, которую для краткости будем называть концепцией естественности, лежит постулат об априорной изначальной расчлененности мира на составляющие его «естественные» или «природные» объекты. Последние, в том числе геологические объекты, уже как бы однозначно «выделены» и обособлены самой природой и предлагаются нам для изучения в готовом, фиксированном виде.

Задача исследователя заключается в том, чтобы выбрать из объектов разного ранга те, которые более всего согласуются с целями исследования, изучить их сущность, условия формирования и закономерности пространственного размещения. Наблюдаемые исследователем однозначно выделенные природой объекты оформляются логически и включаются в арсенал основных понятий. Таким образом, исторически сложившиеся предметы и объекты изучения геологических наук принимаются некритически, как сами собой разумеющиеся, как преподнесенные самой природой. Применительно к геологии, наиболее четкое теоретическое обоснование концепции естественности, хотя и не в таком отпрепарированном виде, а с оговорками и частичным признанием роли субъекта в формировании предметов и объектов геологии, дано в [14, 21].

Однако в последние годы концепция естественности была подвергнута критике, как не отвечающая современному уровню теории познания [10, 15, 16, 27].

Основные возражения связаны с недооценкой роли познающего субъекта в выделении предметов исследования. Как подчеркивал В. И. Ленин, сознание человека не только отражает объективный мир, но и творит его<sup>1</sup>.

Наличие инвариантного в вещах, явлениях и процессах не детерминирует жестко однозначности производимого людьми членения действительности на дискретные единицы. Об этом убедительно свидетельствуют данные этнографии и языкознания. Так, у одних народов используются такие членения действительности, которые отсутствуют у других. Например, как показал М. О. Косвен в своих очерках истории первобытной культуры, в языке эскимосов существует до 20 слов для обозначения льда в разных состояниях его образования и таяния. Вычленение многих видов состояния льда у эскимосов произошло потому, что лед в их жизни и деятельности играл важнейшую роль. Этот пример наглядно

<sup>1</sup> Ленин В. И. Полн. собр. соч., т. 29, с. 194.

подтверждает то сформулированное в теории познания положение, что членение действительности детерминируется не только свойствами самих предметов, но и потребностями общественной практики [8]. Практика, как указывал В. И. Ленин, должна войти в полное «определение» предмета<sup>1</sup> в том смысле, что на ее основе очерчиваются границы предметов, а тем самым и их специфика.

Именно практика, цель, задача определяют выбор того или иного варианта расчленения действительности, выделения объектов исследования в соответствующих границах.

Любой предмет исследования и соответствующий геологический объект отражают реальную действительность не полностью, а в каком-то «срезе», определяемом ограниченным набором отобранных геологом свойств, признаков. Выбранный набор свойств определяет и границы выделяемых по этим свойствам объектов. Так как число свойств, признаков, «могущее быть установленным на материальных объектах, не ограничено, то очевидно, что количество разнообразных границ внутри любой заданной материальной области тоже не может быть выражено определенным числом, оно бесконечно. Выходит, субъекту всегда представляется весьма богатый выбор возможностей вычленения материальных объектов по разнообразным наборам признаков. Весь вопрос в том, какую из этих возможностей и из каких соображений он реализует» [15, с. 265].

Изложенные выше возражения против концепции естественности образуют систему взглядов на конструктивизацию природы, получившую название *модельно-целевой концепции*. Последняя имеет прочное обоснование в рамках теории познания диалектического материализма, признающей активную роль субъекта в познании. Основной практический вывод, вытекающий из модельно-целевой концепции, заключается в том, что единственно возможного и единственно правильного (научного) расчленения геологической реальности не существует, что в каждом выделенном предмете исследования и объекте присутствуют две тесно переплетенные составляющие — объективная и субъективная, отражающая цели, опыт и методы познания. Из этого основного вывода вытекают два очень важных (с точки зрения возможностей представления и использования геологической информации) следствия.

1. Геолог-исследователь не обязан ограничиваться исторически сложившимися предметами исследования. Если они не позволяют решить поставленную задачу, он вправе выделить более соответствующие задаче предметы на основе выбранного им для этого набора признаков. Это вполне естественное и закономерное положение. Новые предметы, оформленные в виде новых понятий, появлялись в геологии постоянно.

В последнее время привлечение современных технических средств, предоставляемых НТР, и выход на океан дали резкий приток новой геологической информации и новых геологических

<sup>1</sup> Ленин В. И. Полн. собр. соч., т. 42, с. 290.

идей. Оформление последних сопровождается введением новых предметов исследований (понятий). Новые идеи базируются не только на новом материале, но и на пересмотре старого под новым углом зрения. При этом, чем менее жестко старый фактический материал привязан к предшествующим предметам и объектам геологии, тем легче он поддается анализу с новых позиций, тем шире круг задач, в которых этот материал может быть корректно использован. Приемы хранения геологического материала, расширяющие возможности его включения в иную геологическую концепцию, а также соответствующие примеры (см. 2.5).

2. Поскольку выделение предметов и объектов геологии производится не природой, а человеком, встает вопрос о корректности процедуры выделения, которая должна обеспечивать однозначность проведения границ объектов по выбранным свойствам. «Если принимаемые правила проведения границ по каким-либо свойствам однозначны и используются корректно, то любые получаемые результаты в равной мере «естественны» в том смысле, что им соответствует нечто в объективном мире, и в равной мере «искусственны», поскольку они получаются на основе правил, вводимых исследователем, при изучении свойств, им отбираемых и изучаемых, на основе методов, которые он применяет. Абсолютно искусственными, или произвольными, не имеющими объективной основы в той области, которая используется, можно назвать лишь такие границы, для проведения которых нельзя подобрать алгоритм, основанный на свойствах, устанавливаемых в этой области» [10, с. 61].

В геологии исторически обусловленное выделение произвольных предметов и объектов — не редкость. Ниже мы проиллюстрируем существующее положение на примерах нескольких из наиболее популярных среди геологов предметов и объектов. Представления о предметах геологического изучения логически оформляются в виде понятия, формой существования которого является дефиниция (определение). Поэтому при рассмотрении примеров произвольных предметов и объектов в качестве отправной точки используются соответствующие дефиниции.

Пример 1. Фация. За сто с лишним лет, прошедших с тех пор как А. Грессли ввел в геологию понятие «фация», оно стало одним из наиболее широко используемых в геологической науке и практике. В геологии нефти и газа этот предмет исследований стал особенно популярным в последние годы по мере того, как нефтегазопроисводческий процесс все более ориентируется на неантиклинальные ловушки, различные типы которых ассоциируются с определенными палеогеографическими (палеогеоморфологическими) условиями. И вместе с тем почти с момента появления данного понятия вокруг него ведется интенсивная дискуссия. К настоящему времени опубликовано более ста различных дефиниций фации. Представляется возможным подразделить их по основным исходным посылкам на пять групп. Эти посылки (I—V) и наиболее типичные определения (1—8) приведены ниже [37]:

I. Порода (осадок, слой), выделяемая на основании присущих ей вещественных характеристик.

1. Д. В. Наливкин. Ископаемая фация — осадок (горная порода), на всем своем протяжении обладающий одинаковым литологическим составом и заключающий в себе одинаковую фауну и флору.

2. В. П. Маркевич. Некоторый объем осадка или породы, характеризующейся сходным комплексом палеонтологических, петрографических или физико-химических признаков, обусловленных тектоническими, физико-химическими, биотическими и географическими условиями образования осадка.

3. Т. П. Давыдова, Ц. Л. Гольдштейн. Порода (и отложения из которых они произошли), отличающиеся от пород (и отложений) того же возраста.

II. Порода (осадок, слой), выделяемая на основании сочетания вещественных характеристик и базирующегося на них умозаключения об условиях формирования породы.

4. В. Е. Хаин. Тип осадочной породы, возникающий в определенных физико-географических условиях.

5. Г. Ф. Крашенинников. Комплекс отложений, отличающихся составом и физико-географическими условиями образования от соседних отложений того же стратиграфического отрезка.

III. Совокупность признаков породы (осадка), характеризующих среду осадконакопления.

6. Н. С. Шатский. Совокупность литологических, палеонтологических, а также тектонических особенностей осадков, характеризующих палеогеографические, физико-географические условия осадков.

7. К. Данбар, Дж. Роджерс. Общий облик пород, литологический, биологический (а в более широком понимании также структурный, тектонический, метаморфический), отражающий условия окружающей среды, при которых образовались породы.

IV. Физико-географические условия.

8. Н. М. Страхов. Среда отложения пород со всеми ее особенностями (рельефом, химическим режимом, органическим миром).

9. Н. В. Логвиненко. Обстановка осадконакопления (современная или древняя), овеществленная в осадке или породе.

V. Изменение свойств отложений.

10. А. Грессли (по Н. Б. Вассовичу). Совокупность видоизменения отложений, выражающаяся в том или ином петрографическом, геогностическом или собственно палеонтологическом их отличии.

Рассмотрение существующих дефиниций фации со всей очевидностью показывает, что, во-первых, под один термин «фация» различные группы исследователей подводят принципиально различные понятия, отражающие несопоставимые предметы и объекты (порода, обстановка, комплекс признаков). Во-вторых, на основе любого из существующих определений фации объективно выделить последнюю не представляется возможным.

Возьмем для примера наиболее конкретное из всех (материальный наблюдаемый объект с минимальным числом конкретно указанных признаков) определение 1, данное Д. В. Наливкиным. Оно связывает строгой зависимостью литологию и фауну. В природе оба этих признака, конечно, взаимосвязаны, но не функциональной зависимостью, когда какое-либо изменение одного влечет строго определенное изменение другого. Взаимосвязь литологического состава и фауны опосредована через реакцию (причем различную) на изменение тех или иных условий. Поэтому быстрая смена литологического состава может происходить при незначительном изменении состава фауны, и наоборот. Кроме

того, каждый из указанных признаков является комплексным. Фауна в данном осадке (породе) может состоять из нескольких или даже многих родов и видов (различных фораминифер, остракод, пеллеципод и т. п.), и изменение роли каждого из родов и видов в пределах осадка может происходить неодинаково ввиду различной реакции их на изменение условий.

Литологический состав, в свою очередь, включает зернистость, минеральный состав и другие признаки, характеризующиеся своими закономерностями изменения. Для каждого в отдельности частного признака пределы для суждения об «одинаковости» могут быть, конечно, определены (заданы), но решение такой задачи для суммы признаков, каждый из которых изменяется по своим законам, невозможно. Поэтому на практике вопрос о выделении фации, проведении ее границ может решаться не на основе объективных критериев, а лишь путем субъективного мнения о том, в достаточной ли степени изменилась порода, чтобы отнести ее к другой фации.

Вышесказанное в еще большей степени относится к таким трактовкам, в которых фация рассматривается не как материальный объект, а как совокупность признаков или свойств отложения (отчленение) имеет смысл, лишь когда формально зафиксированы признаки, на основе которых проводится это сравнение, разделение. В отношении фации это условие не соблюдено. Так, в определении 2, наиболее близком к рассмотренному выше, речь идет о комплексе петрографических, палеонтологических и физико-химических признаков, и каждый может понимать сколь угодно широко этот комплекс в целом и каждую группу признаков в частности. А в определении 3 набор признаков вообще не оговорен, и потому любое минимальное отличие может рассматриваться как основание для выделения независимых фаций.

Не лучше положение, если в качестве одного из признаков или единственного признака принимается физико-географическая обстановка. В последнем случае на понятие «фация» в полной мере переносится дополнительная неопределенность, связанная с неоднозначностью понимания элементов физико-географической обстановки, в частности масштабы. В результате в одной и той же геологической ситуации разными исследователями выделяются перекрывающиеся разномасштабные фации, например дельтовая, субаэральной равнины, дельтовой протоки, прируслового вала дельтовой протоки и др. Кроме того, физико-географическая палеообстановка является не реальным, т. е. фиксируемым признаком, а умозаключением, основанным на комплексе признаков, надежная и однозначная палеогеографическая интерпретация которых не всегда возможна.

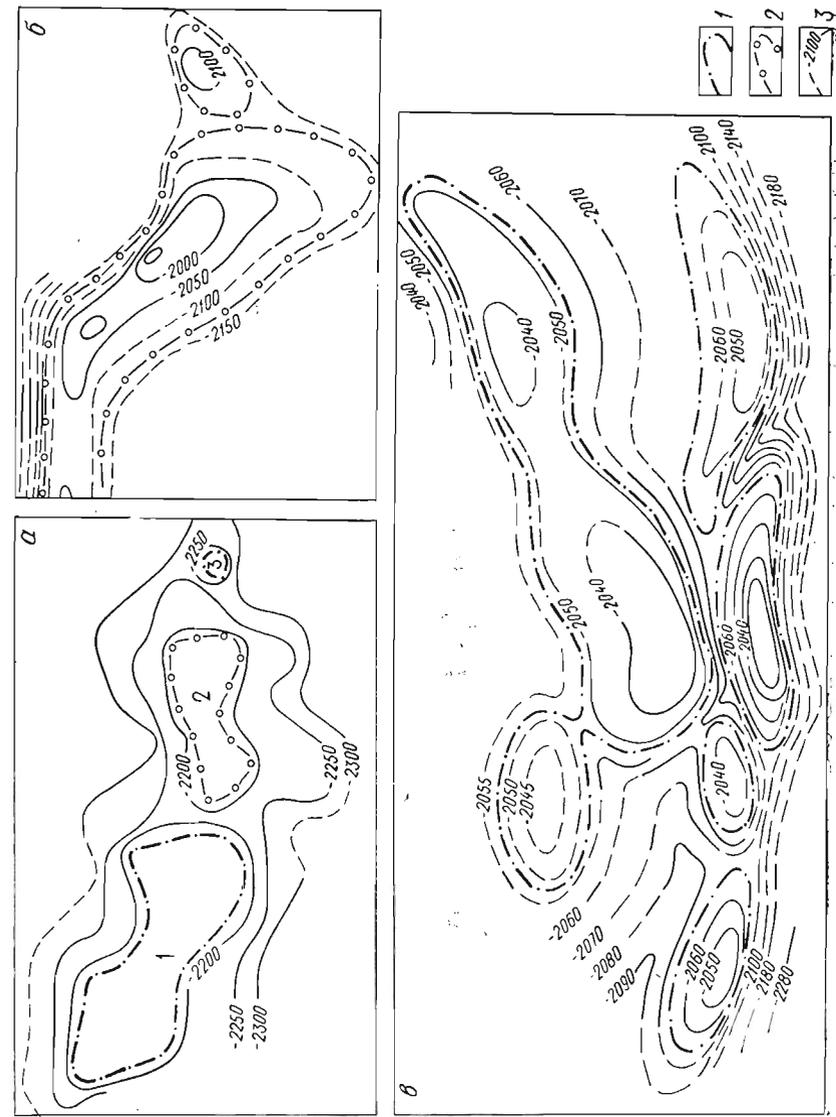
Таким образом, пока не сформулировано определение фации, из которого вытекало бы правило (алгоритм) однозначного выделения соответствующих объектов на практике. Поэтому под термином «фация» выделяются типично произвольные предмет и объ-

екты. И именно в ощущаемой неудовлетворенности таким положением кроется одна из причин многолетней дискуссии о том, что же такое «фация».

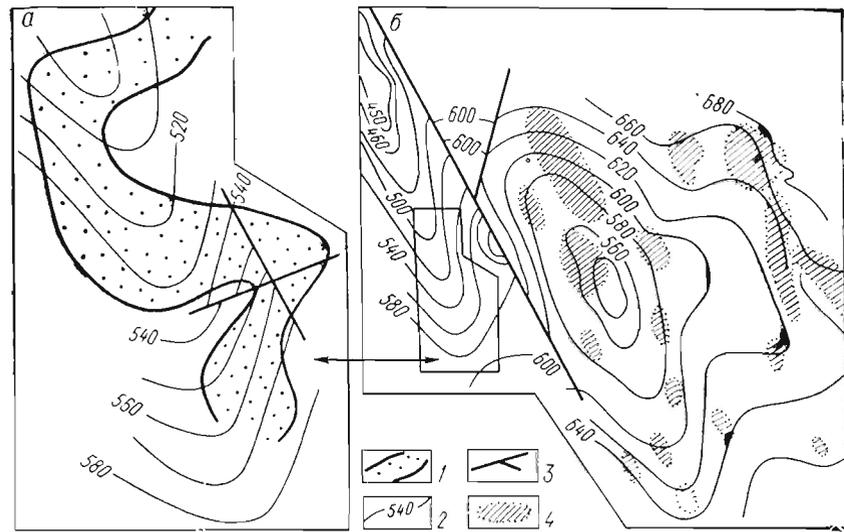
Пример 2. Месторождение нефти и газа. В настоящее время преобладающая часть существующей информации по геологии минерального сырья так или иначе увязывается с таким узловым предметом исследования, как месторождение. Хранение многочисленных разнообразных геологических параметров организовано таким образом, что они выступают в качестве характеристик месторождения. Вместе с тем в вопросе выделения месторождений наблюдается положение, весьма сходное с рассмотренным в отношении фации.

Ниже мы рассмотрим существующее положение на примере месторождения нефти и газа, но оно в равной мере свойственно и геологии других видов минерального сырья. Подробный логический анализ дефиниций понятия «месторождение нефти и (или) газа» приведен в работе [43]. Этот анализ со всей очевидностью показывает, что, во-первых, существуют принципиальные расхождения в подходах к ограничению месторождения нефти и газа, понимаемого то как структурный элемент, то как участок земной коры, то как совокупность залежей. Во-вторых, из фации существующей дефиниции не вытекает правило связи (закон композиции) залежей, объединяемых в месторождение. Именно в силу этих причин процедура выделения месторождения ничем не регламентирована, неизбежным следствием чего является выделение различных по числу, масштабу и композиции месторождений нефти и газа в сходных геологических условиях. Несопоставимость выделяемых в геологической практике месторождений столь разительна, что нет необходимости в приведении здесь анализа понятийной основы и вполне достаточно ограничиться иллюстрацией конкретных примеров. Последние взяты из справочников по месторождениям нефти и газа СССР [26] и зарубежных стран [25].

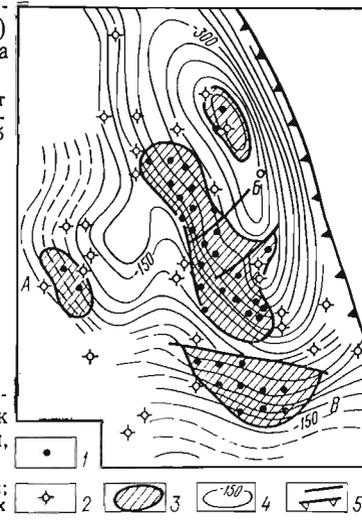
Наиболее распространенная причина выделения несопоставимых месторождений нефти и газа связана со стремлением приурочивать месторождение к локальному структурному элементу. Однако поскольку дефиниции понятия «месторождение» не накладывают формальных ограничений на масштаб локального структурного элемента, в этом вопросе неизбежен субъективизм. Поэтому при выделении месторождений в одних случаях под локальным элементом понимают какое-либо одно осложнение более крупной складки и выделяют столько месторождений, сколько имеется осложнений и приуроченных к ним залежей (рис. 8, а), в других случаях — всю эту складку, охватывающую ряд более мелких осложнений (залежей), и выделяют одно месторождение (рис. 8, б). Не менее часты случаи, когда единое месторождение охватывает произвольную (не связанную соподчиненностью между собой или по отношению к более крупной структуре) группу складок. Примером может служить Пономаревское месторождение (рис. 8, в).



**Рис. 8.** Примеры соотношения локальных поднятий и выделяемых месторождений нефти и газа [43]:  
 а — группа месторождений: 1 — Восточно-Метовское, 2 — Восточно-Веселовское, 3 — Восточно-Веселовское (Западная Сибирь); б — Ленинградское месторождение (Красноярский край); в — Пономаревское месторождение (Оренбургская область).  
 1 — контур нефтеносности; 2 — контур газоносности; 3 — изогипсы по кровле продуктивных горизонтов



**Рис. 9.** Структурный план и положение залежей на Войвожском (а) и Нибельском (б) месторождениях Тимано-Печерского бассейна [26]:  
 1 — нефтеносные отложения палеорусл (пласт 1г); 2 — изогипсы по кровле пласта 1а; 3 — нарушения; 4 — нефтеносные песчаные линзы пласта 1б



**Рис. 10.** Схема расположения залежей месторождения Вест-Сейдж-Крик (А), Сейдж-Крик (Б) и Норт-Дивер (В) бассейна Биг-Хорн, США [31]:  
 1 — продуктивные скважины; 2 — сухие скважины; 3 — залежи; 4 — изогипсы кровли пенсильванских отложений (в м); 5 — разрывы

Другая распространенная причина, обуславливающая несопоставимость выделяемых месторождений, — стремление провести структурный подход к ограничению (выделению) месторождений в условиях, когда формирование и размещение объединяемых залежей никак не контролировалось структурным элементом, в пределах которого эти залежи размещаются ныне. Например, залежи в линзовидных коллекторах, в частности, приуроченных к погребенным руслам, рифам и барам, участкам трещиноватости и т. п.,

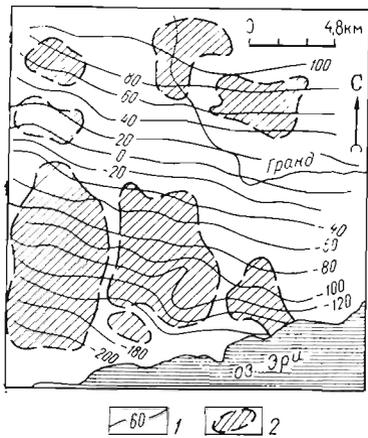


Рис. 11. Схема расположения залежей нефти на месторождении Халдиманд (Канада) [31]:  
1 — изогипсы поверхности размыва; 2 — залежи нефти

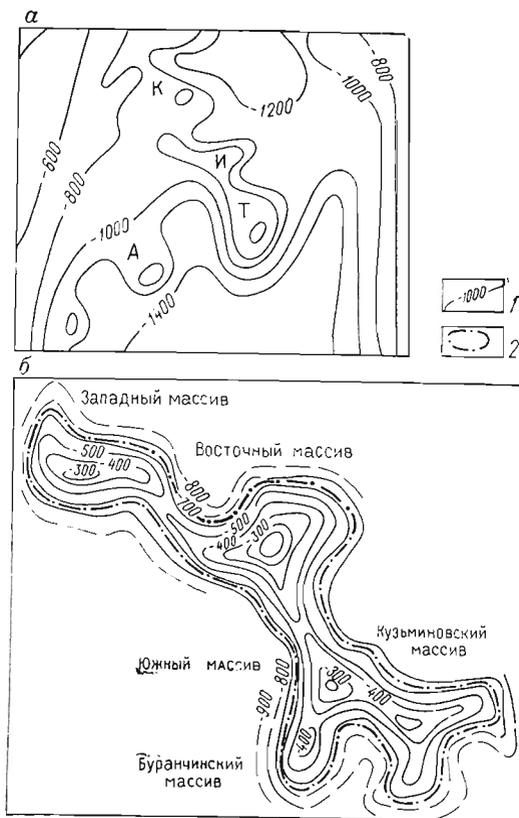


Рис. 12. Схема Ишимбаевского месторождения нефти (Башкирская АССР):  
а — положение в региональном плане [26]. Месторождения: К — Кусялиуловское, И — Ишимбаевское, Т — Термень-Елгинское, А — Аллакаевское; б — соотношение структурного плана и контура нефтеносности [31]. 1 — изогипсы по поверхности сакмарско-артинских отложений; 2 — контур нефтеносности

часто возникают задолго до того, как в данном месте сформировался наблюдаемый ныне единый структурный элемент (единая тектоническая структура). Примеры таких залежей приведены на рис. 9. В таких случаях объединение залежей в месторождение проводится на совершенно искусственной основе (зачастую случайного сонахождения, а не происхождения) и потому лишено того геологического (генетического) содержания, на основе которого только и можно выявить природные закономерности. Многими геологами эта искусственность отчетливо осознается, в связи с чем не менее часто они в подобных ситуациях не обращают внимания на современный структурный план, и тогда или откровенно произвольно (по принципу близости) объединяются залежи в место-



Рис. 13. Пример соотношения структурного плана продуктивного горизонта, контура нефтегазоносности и выделяемых месторождений нефти [26].

Месторождения: I — Туймазинское, II — Александровское, III — Бавлинское, IV — Северо-Бавлинское (Башкирская ССР).  
1 — изогипсы по кровле пашинского горизонта девона; 2 — контур нефтеносности; 3 — нарушения

рождение, или каждая такая залежь выступает как самостоятельное месторождение (рис. 10).

Совершенно аналогичное положение с выделением месторождений наблюдается в тех случаях (например, в условиях моноклиналей), когда выделить локальный структурный элемент вообще затруднительно.

Иллюстрацией может служить месторождение Халдиманд (Канада, Онтарио), связанное с участками повышенной трещиноватости и кавернозности в интервале, залегающем непосредственно под поверхностью несогласия (рис. 11).

Принципиально различные подходы к выделению месторождений часто проявляются и в случае, если одна залежь распространяется на несколько поднятий. Ишимбаевское месторождение (рис. 12) отличается от соседних месторождений той же зоны (например, Кусялиуловского, Аллакаевского) тем, что охватывает не один, а несколько самостоятельных рифовых куполов. Это вызвано тем, что на эти купола распространяется единая залежь нефти, т. е. приоритет отдан принципу «одна залежь не должна

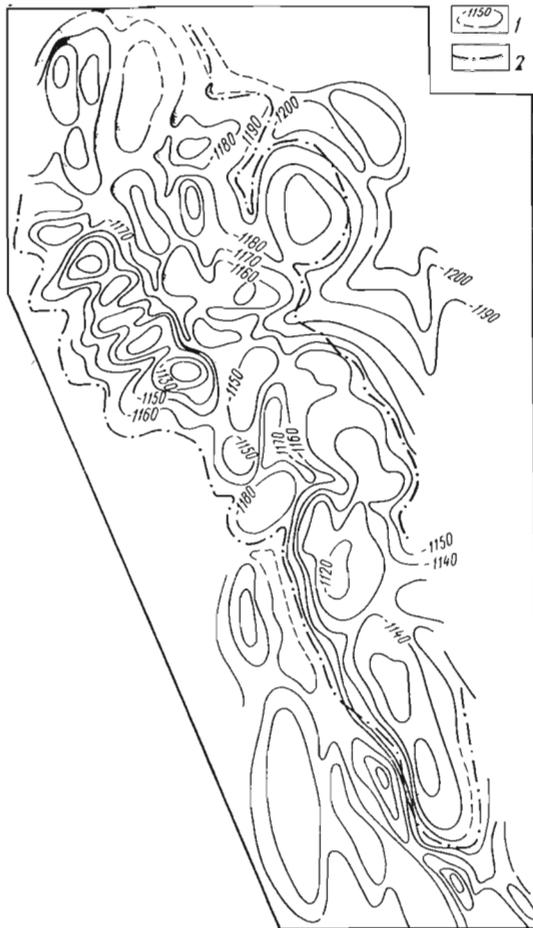


Рис. 14. Схема положения залежей Арлано-Дюртюлинской зоны нефтегазоаккумуляции (Баш-АССР) [26]:

1 — изогипсы по терригенной толще нижнего карбона; 2 — контур нефтегазоносности

быть в пределах разных месторождений». В других случаях приоритет отдается структурному фактору, даже при наличии единой залежи. Примерами могут служить Туймазинское, Бавлинское, Северо-Бавлинское и Александровское месторождения, на площади которых развита единая залежь нефти в пласте  $D_1$  пашийского горизонта (рис. 13), а также Арлано-Дюртюлинская зона (рис. 14).

Приведенные примеры разнообразия принципов (вернее, беспринципности) выделения месторождений нефти и газа касались случаев относительного соответствия структурных планов в разрезе осадочного чехла. Если же наблюдается несоответствие планов, то возникают дополнительные расхождения. Наиболее наглядно это проявляется в случаях, когда в осадочном чехле выражено несколько структурных этажей. Такая картина, напри-

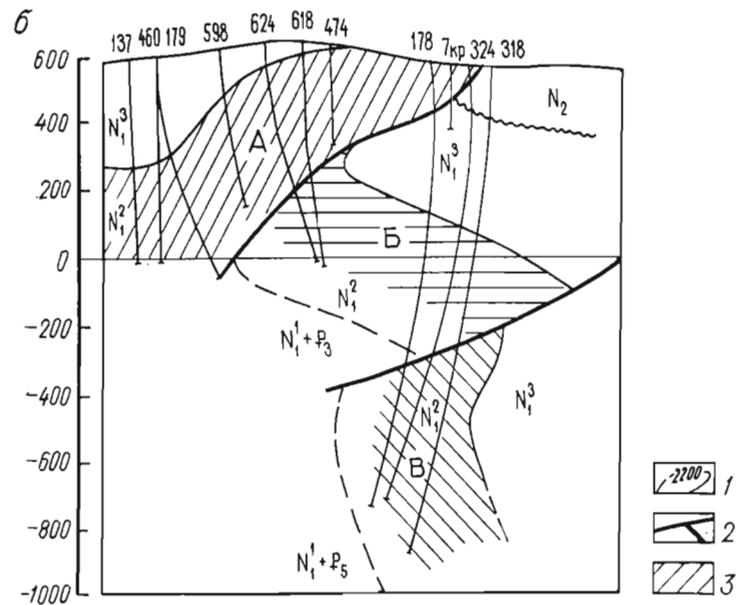
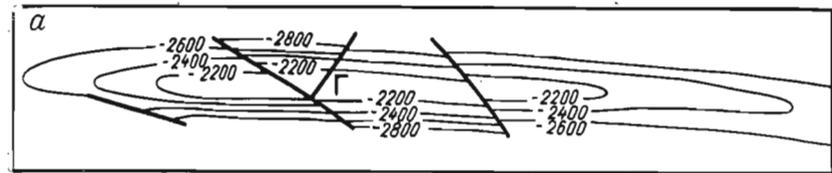


Рис. 15. Структурный план кровли продуктивных отложений верхнего мела (а) и схематический разрез (б) зоны Малгобек-Алиюрт (Грозненская область) [26]:

А — месторождение Борис; Б и В — соответственно месторождения первого, второго поднадвигов Северо-Вознесенской антиклинали; 1 — изогипсы по кровле продуктивных отложений верхнего мела, м; 2 — нарушения; 3 — продуктивная зона неогеновых отложений (пласты I—XVI)

мер, характерна для Терского антиклинория (Чечено-Ингушская АССР). Здесь выделяются, с одной стороны, самостоятельные месторождения Малгобек-Вознесенское, Алхазовское и Алиюртовское, которым по нижнему структурному этажу (мел) соответствует единая антиклиналь и единая залежь (рис. 15, а), а верхний — неогеновый структурный этаж разбит на многочисленные блоки, к большинству которых приурочены группы разновозрастных продуктивных горизонтов. Многие такие блоки в свою очередь выделяются как самостоятельные месторождения (рис. 15, б). В тех случаях, когда наблюдается соответствие структурных планов, месторождение на данной площади выделяется в объеме всего продуктивного интервала разреза, при несоответствии планов приходится на одном и том же участке выделять месторождения одно над другим.

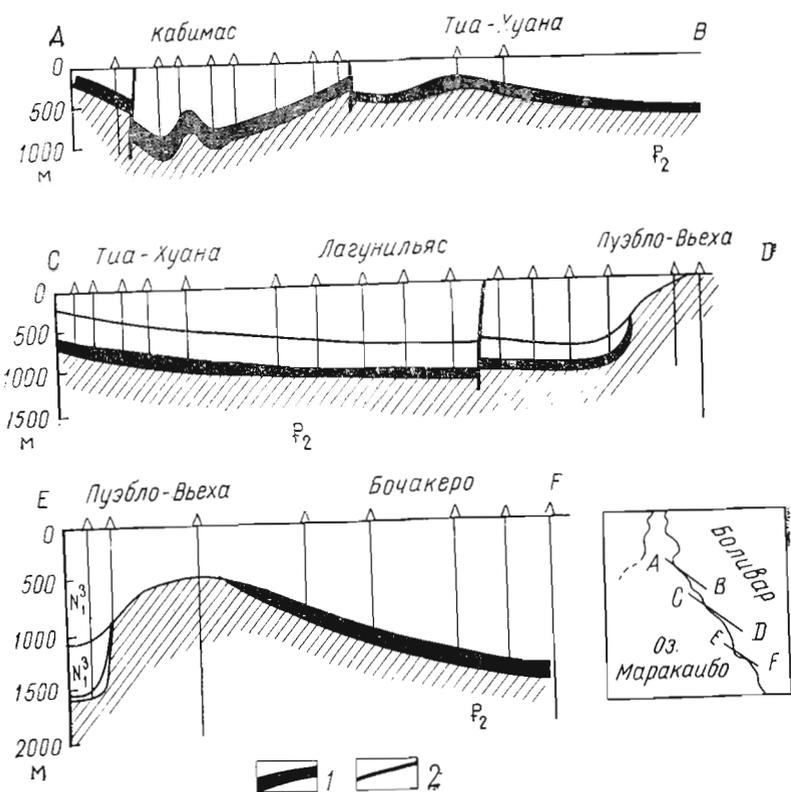


Рис. 16. Геологический разрез зоны месторождений Боливар — Прибрежный (Маракаибский бассейн, Венесуэла) [25]:

1 — залежи нефти; 2 — разрывные нарушения

Важно также отметить, что обособить по геологическим параметрам отдельное месторождение в пределах зоны или области нефтегазонакопления в ряде случаев крайне затруднительно.

Характерный пример — одна из богатейших в мире зон нефтегазонакопления — Боливар-Коастал (Маракаибский бассейн, Венесуэла). Здесь выделяется ряд месторождений, из которых Лагунильяс, Тиа-Хуана и Бокачиро относятся к категории гигантских (запасы свыше 500 млн. т). Эти месторождения были выявлены вначале как отдельные, но в процессе доразведки выяснилось, что между ними нет участков, лишенных залежей (рис. 16). Они практически сливаются в одно сверхгигантское месторождение, которое из-за его масштабов сочли более правильным называть зоной нефтегазонакопления. Выделяемые в этой зоне месторождения представляют собой не обособленные по геологическим параметрам совокупности залежей или участки с обособленными совокупностями залежей, а части единого поля нефтеносности, условно выделенные в соответствии с тем, как складывалась исто-

рия нефтепоисковых работ в этом районе. В значительной мере сходная картина наблюдается в пределах Арлано-Дюртиюлинской зоны нефтегазонакопления (см. рис. 14). Не случайно в справочниках по месторождениям Советского Союза [26] на профилях и картах показаны не отдельные месторождения, а вся эта зона в целом.

Приведенные примеры достаточно наглядно показывают, что вследствие неудовлетворительного логического оформления (на уровне определения понятия) такого предмета исследования, как месторождение нефти и (или) газа, соответствующие объекты выделяются на практике неоднозначно (произвольно) и потому в большинстве случаев несопоставимы.

Число произвольных предметов исследования в геологии не ограничивается приведенными выше двумя примерами. Такие предметы широко используются практически во всех отраслях геологии. И это очень важный, но не единственный фактор, существенно сказывающийся на корректности обобщения геологических данных, которое всегда осуществляется на статистической основе. Несопоставимость объектов обуславливает несопоставимость характеризующей их объектно-привязанной информации. Несопоставимая же информация по несопоставимым объектам не может служить надежной базой для выявления закономерных (природных) связей, лежащих в основе любой геологической задачи.

## 2.2.2. Методологические основы построения геологических классификаций

Любой фрагмент рассматриваемой (объектно-привязанной) геологической информации, как правило, не автономен, не независим, а выступает в качестве характеристики того или иного геологического объекта. При решении большинства геологических задач возникает необходимость комплексной обработки геологической информации не по одному, а по нескольким типам геологических объектов. Возможность корректного соотнесения информации, характеризующей разнотипные геологические объекты, зависит не только от того, насколько соблюдено рассмотренное выше правило однозначности выделения каждого конкретного геологического объекта, но и в не меньшей степени от того, насколько правильно соотнесены между собой типы этих объектов в рамках соответствующей геологической классификации (типологии). Формально-логическое несовершенство последней — вторая по значению причина, часто обуславливающая малоудовлетворительный результат обработки данных, в частности, с применением ММ и ЭВМ.

Классификация служит одним из основных средств упорядочения, систематизации и обобщения фактического материала, характеризующего классифицируемые предметы исследования. Поэтому в геологии, как и в любой другой сфере научно-производст-

венной деятельности, практически каждый фрагмент информации прямо или косвенно связан с той или иной классификацией. Так, представление о типах пород, залежей, ловушек, месторождений, тектонических элементов, складчатости, несогласий непосредственно базируется на соответствующих классификациях.

Под классификацией понимается результат логической операции (классифицирования), при которой все предметы, однородные в каком-либо отношении и образующие вследствие этого целостную совокупность, распределены по классам, отражающим видовые отличия. Поскольку классификация является результатом логической операции, естественно, что ее построение регулируется логическими правилами и требованиями. Если последние не соблюдены, то классификация построена на некорректной логической основе и не в состоянии выполнять свою систематизирующую функцию. Основным регулятив построения классификации — разработанные в формальной логике *правила деления объема понятия*.

Рассмотрение этих правил, представляющих собой методологическую основу классифицирования, необходимо предварить некоторыми элементарными сведениями из теории классификации. В каждой классификации различаются делимое, основание деления и члены деления. *Делимое* — совокупность однородных в определенном отношении предметов (понятий), т. е. объем некоторого общего (родового) понятия. *Основание деления* — это признаки, по которым предметы распределяются по классам. Классы (а также подклассы и их дальнейшие подразделения — виды, подвиды и т. п.), рассматриваемые по отношению к той совокупности, из которой они непосредственно получены при операции деления, называются *членами деления*. Понятия классов, подклассов и т. д. — это таксономические категории или *таксоны*, т. е. общие понятия различных ступеней деления исходной совокупности. Каждой такой ступени деления, т. е. каждому таксону, отвечает свое (отличное от других таксонов) основание деления. Этих кратких пояснений вполне достаточно для понимания правил деления объема понятия (их всего 4), которые заключаются в следующем [19].

1. *На одном и том же таксономическом уровне* (т. е. на одной и той же ступени деления) *должно использоваться одно и то же основание*. Нарушение этого правила свойственно почти всем геологическим классификациям. Например, в большинстве классификаций зон нефтегазоаккумуляции выделяются, наравне со структурными, рифовые зоны нефтегазоаккумуляции. В первом случае основанием служит наличие структурной дислокации слоев, а во втором — генетический тип коллектора. Подобно этому в классификациях осадочных пород, как правило, на одном уровне выделяются терригенные (основание — источник поступления материала) и карбонатные (основание — минеральный состав) породы или песчаники (основание — размерность зерен) и известняки (основание — минеральный состав).

2. *Деление должно быть соразмерным*. Это означает, что общий объем всех членов деления должен равняться объему делимого понятия, т. е. в перечень по какому-нибудь признаку видовых понятий данного родового понятия непременно должны быть включены все виды, ни на один меньше, ни на один больше. В качестве примера несоблюдения этого правила можно указать на отсутствие «гидродинамических» ловушек в большинстве классификаций ловушек нефти и газа. Такое нарушение правила соразмерности в логике известно под термином «неполное деление». Противоположное нарушение правила соразмерности, когда в число видов включаются и не относящиеся к данному роду, называется «слишком широким делением». Пример такой логической ошибки — весьма распространенное включение в род «терригенные породы» всех глин, хотя известно, что некоторые глины являются хемогенными, и они не могут рассматриваться в составе терригенных пород.

3. *Члены деления должны взаимно исключать друг друга*. Согласно этому правилу каждый отдельный предмет должен находиться в объеме какого-либо подразделения классификации и ни в коем случае не должен входить частично в объем другого подразделения. Несоблюдение этого приводит к логической ошибке — «пересечение классов». Например, в классификации залежей нефти и газа М. Ф. Мирчинка наряду с залежами, ограниченными дизъюнктивными нарушениями, выделяются нарушенные залежи, приуроченные к моноклиналиям. При таком разделении залежь, ограниченная нарушением, может быть с равным основанием включена в два из указанных подразделений классификации. В качестве еще одного примера можно указать на выделение такой породы, как кальцирудит, представленной карбонатными обломками песчаной и более грубой размерности, сцементированными кальцитом. В Геологическом словаре [6, т. I] указывается, что кальцирудит рассматривается или как грубообломочный известняк, или как известняковый грубообломочный песчаник. Таким образом, кальцирудит может с равным основанием относиться к разным подразделениям классификации осадочных пород. Эти примеры показывают, что нарушение рассматриваемого правила может быть (хотя и не всегда) следствием несоблюдения правила единства основания на одном таксономическом уровне деления.

4. *Деление должно быть непрерывным*. Члены деления должны быть понятиями не только соподчиненными, но и обязательно непосредственно низшими по отношению к понятию предшествующего таксона. Нарушение этого правила называется «скачком в делении». Так, понятие «осадочные породы» делится на понятие «обломочные породы», «органогенные породы», «хемогенные породы», а эти понятия — на более частные понятия, например, хемогенные породы — на «кремнистые», «железистые», «карбонатные» и др., а они, в свою очередь, — на конкретные породы — доломит, известняк и т. д. Если в классификации следующим уровнем деления понятия «осадочные породы» будет, например, «доломит», ми-

ная «хемогенные породы» и «карбонатные породы», то это и есть типичный «скачок в делении».

Кроме правил деления объема понятия к построению классификации часто предъявляется ряд дополнительных требований, из которых наиболее важными представляются следующие.

а. *Необходимость строгого определения понятий, на которые опирается классификация.* Это требование имеет в виду приведенные формулировки используемых понятий в соответствии с указанными выше требованиями к выделению предметов исследования.

б. *Необходимость выбора в качестве оснований деления таких характеристик, которые являются технически или, в крайнем случае, принципиально поддающимися однозначной констатации и, кроме того, характерны не для всех предметов классифицируемой совокупности.* Если признак нельзя констатировать или если он неделимый, т. е. такой, какой в равной мере есть у всех предметов исходной совокупности, то классифицирование невозможно.

Каждый предмет обладает бесконечным числом свойств и поэтому одни и те же предметы в принципе можно классифицировать неограниченным числом способов, в зависимости от избранных оснований деления. Поскольку «у природы, без человека нет критериев предпочтительности одного свойства перед другим» [10, с. 66], то обычно выбирают те, которые позволяют наилучшим образом решать поставленные перед классификацией задачи. Поэтому нет никаких ограничений на выбор оснований деления. Единственное, но принципиально важное и безусловное ограничение, исключающее произвол при построении любой классификации, — необходимость соблюдения указанных выше формально-логических требований. Однако, несмотря на простоту и очевидную правомерность этих требований, они корректно соблюдены лишь в небольшом числе геологических классификаций.

Во-первых, это связано с тем, что в соответствии с предложенными С. С. Розовой типами классификаций значительное число последних относится к категории описательных, т. е. к самой зачаточной, по существу, преднаучной форме классификации. Примером может служить любая из существующих классификаций элементов нефтегеологического районирования. В них начальным классом является ловушка, залежь, месторождение или зона нефтегазонакопления, а конечным — нефтегазоносный бассейн или провинция, группа бассейнов или провинций или пояс нефтегазонакопления [31]. Здесь лишь констатируется факт существования данных классов исследуемого объекта и никак не отражен принцип подразделения на классы в зависимости от обуславливающих факторов, т. е. основание деления просто не сформулировано.

Во-вторых, в большинстве тех классификаций, где сформулированы факторы, обуславливающие специфику классов и более частных таксонов, не соблюдены правила деления объема понятия. Ниже мы проиллюстрируем это положение на нескольких примерах.

Пример 3. Литологические классификации.

Практически в любой из используемых в литологии классификаций на одном и том же таксономическом уровне использованы различные основания деления. При этом сам набор признаков, используемых для разделения, в принципе сохраняется, но значение их (а соответственно и вынесение на тот или иной таксономический уровень) для разных групп оценивается различно. Для примера можно рассмотреть всем известную классификацию осадочных пород Л. Б. Рухина. В ней на наиболее общем таксономическом уровне выделено пять групп основных типов осадочных пород: пирокластические, обломочные, глинистые, химические и биохимические, каустобиолиты. Уже на этом уровне для пирокластических пород основанием выделения является вулканогенный генезис исходного материала, для обломочных и глинистых — размерность частиц, для химических и биохимических пород — способ разрушения материнских пород, а для каустобиолитов — преимущественно углеводородный состав. На следующем таксономическом уровне пирокластические породы дифференцируются по соотношению осадочного и вулканогенного материала (вулканические туфы и туфобрекчии, туффиты, туфогены), обломочные породы — по градациям размерности частиц (грубообломочные, песчаные, алевритовые), химические и биохимические породы — по минеральному составу (алюминистые, железистые, марганцевистые, кремнистые и т. д.). Что же касается глинистых пород, то ближайшие члены их деления отвечают по уровню общности не второму, а более частным уровням деления других групп пород.

Такое положение неизбежно приводит к «пересечению классов». Так, согласно Л. Б. Рухину, к глинистым относятся породы, состоящие более чем на 50 % из частиц размером  $<0,01$  мм и содержащие не менее 30 % частиц размером  $<0,001$  мм. Породы, содержащие более 40 % глинистых частиц, следует называть, в зависимости от размера частиц, глинистыми алевролитами, глинистыми песчаниками и т. п. Из этого становится очевидным отсутствие основания для разделения обломочных и глинистых пород на первом таксономическом уровне, поскольку если исходить из размерности частиц, то глины должны занять место одного из членов деления обломочных пород на втором таксономическом уровне. По существу, все геологи так и поступают на практике, рассматривая «на равных» песчаник и глину, а не обломочные породы и глину.

Невыдержанность основания деления на первом таксономическом уровне приводит к возможности отнесения одной и той же породы к различным подразделениям классификации. Так, в геологическом словаре [6] песчаник известковистый грубообломочный рассматривается как синоним грубообломочного известняка — кальцирудита.

Невыдержанность основания деления часто наблюдается и в простейших одноуровневых литологических классификациях. Например, Н. В. Логвиненко [23] среди первичных структур хемогенных пород выделяет в качестве равнозначных: микрозернистые,

пелитоморфные (основание — размерность частиц), оолитовые и пизолитовые или бобовые (сразу несколько оснований — форма, внутренняя структура, размер), псевдооолитовые (основание — степень перекристаллизации). Подобных примеров можно привести множество. Такое положение приводит к искажению результатов совместной обработки данных, характеризующих различные типы пород, обуславливает замену общих методов анализа специализированными по типам пород.

Пример 4. Классификация ловушек нефти и газа. Поскольку и в настоящее время технология прямых поисков скоплений нефти и газа еще несовершенна, методика нефтегазописковых работ ориентирована на предварительное выявление вероятной ловушки нефти и газа, т. е. геологических условий, в которых возможно удержание подвижных углеводородных флюидов. И уже затем бурением устанавливается наличие или отсутствие в ловушке скопления нефти и газа. Поэтому осуществление нефтегазописковых работ немыслимо без представления об известных и вероятных типах ловушек, и это представление базируется на классификации последних. Вполне естественно, что этому вопросу уделяется особое внимание, выразившееся, в частности, в том, что начиная с 30-х годов советскими и зарубежными исследователями предложено несколько десятков общих<sup>1</sup> классификаций ловушек. В дополнение к общим предложено не меньшее число частных классификаций отдельных групп ловушек по различным признакам.

Несмотря на обилие общих классификаций ловушек, принцип подразделения на классы практически во всех из них остается неизменным. Этот принцип заключается в разделении всей совокупности ловушек на классы структурных, неструктурных и комбинированных. Наблюдающиеся отклонения имеют принципиальный характер и заключаются в следующем; 1) неструктурные ловушки разделяются на два самостоятельных класса — литологические и стратиграфические; 2) комбинированные — на ряд классов: структурно-литологические, структурно-стратиграфические, литолого-стратиграфические; 3) структурные — на два класса: складчатых дислокаций и разрывных дислокаций; 4) для неструктурных ловушек употребляются иные термины, например, «неантиклинальные» или «стратиграфические» (последний распространен среди американских геологов).

Рассмотрение существующих общих классификаций ловушек с изложенных позиций показывает, что подавляющее большинство из них не согласуется с перечисленными правилами деления объема понятия. Прежде всего не соблюдено первое из этих правил,

<sup>1</sup> Под общей классификацией понимается такая классификация, которая охватывает всю совокупность классифицируемых объектов, в данном случае ловушек. В отличие от нее частные классификации охватывают лишь определенные группы классифицируемых объектов (например, литологические, стратиграфические, гидродинамические, палеогеоморфологические ловушки).

касающееся единства основания деления. Сформулировать критерий, положенный в основу выделения классов (основание деления) в таких классификациях, не представляется возможным, поскольку он не является общим, выдержанным, а свой для каждого класса. В связи с этим не случайно, что у различных авторов содержание одних и тех же понятий (стратиграфические, литологические и структурные ловушки) не совпадает, несколько классов объединяются в один или, наоборот, один класс разделен на несколько самостоятельных классов и т. д. Подобные замечания могут быть сделаны и по отношению к подразделению классов на подклассы, группы.

Перечень всех ловушек в указанных классификациях не охватывает полностью объем понятия «ловушка» (второе правило). В частности, в них нет места так называемым «гидродинамическим ловушкам». Механическое подключение их в виде самостоятельного класса к выделенным классам структурных, стратиграфических, литологических и комбинированных ловушек было бы очередным нарушением первого правила деления объема понятия (требование единства основания деления). Весьма показательны, что при уже весьма значительном числе статей, посвященных «гидродинамическим» ловушкам, круг рассматриваемых в них вопросов ограничивается теоретическим обоснованием возможности их образования, демонстрацией конкретных примеров, но ни в одной из работ не показано место «гидродинамических» ловушек в общей классификации. То же самое можно сказать и в отношении значительного числа «палеогеоморфологических» ловушек.

Несоблюдение первого правила часто влечет за собой и нарушение третьего. Так, во многих классификациях, наряду с «рифовыми», выделяются линзовидные ловушки. Но куполовидный риф, заключенный в толще глин и мергелей, представляет собой линзовидную ловушку, и в этом случае налицо явное «перекрывание классов».

Отмеченное формально-логическое несовершенство приводит к тому, что не все ловушки однозначно укладываются в рамки классификации, и побуждает геологов к созданию новых классификаций, в большей мере учитывающих особенности известных им (но различных) наборов ловушек. В этом заключается основная причина того, что при столь большом числе классификаций ловушек пока нет общепринятой. Такое положение еще более ограничивает возможность получения корректного результата обработки данных по ловушкам, поскольку искажение результата происходит не только из-за формально-логического несовершенства отдельной классификации, но и потому, что при широком обобщении приходится учитывать ловушки, зафиксированные другими геологами, исходившими из других классификаций. В частности, по этой причине, несмотря на обширный накопленный фактический материал, пока не удается обосновать статистически оптимальный комплекс поисковых признаков (и соответственно комплекс фикси-

рующих эти признаки методов поиска) для большинства так называемых «неструктурных» ловушек нефти и газа.

Пример 5. Типизация органогенных построек. Примеры 3 и 4 касались логических нарушений в рамках систематизированной совокупности объектов, объединяемых общностью того или иного (родового) признака. Но во многих случаях при решении геологических задач приходится иметь дело с информацией по объектам, хотя и обладающим общностью в каком-либо отношении, но тем не менее представляющим хаотическую, логически неорганизованную совокупность. Эта хаотичность обусловлена тем, что отдельные типовые разновидности объектов изначально выделялись автономно, вне логической увязки (по основаниям деления) с другими разновидностями родственных объектов. Здесь организация объектов и характеризующей их информации находится еще на доклассификационном уровне, которому свойственны предельные нарушения принципа однозначности выделения объектов и всех правил деления объема понятия. С таким положением неизбежно сталкиваешься, например, при широком (глобальном) обобщении данных по многим палеогеоморфологическим объектам, к которым приурочены неантиклинальные ловушки нефти и газа, в частности, по органогенным постройкам [37].

Сказанное не означает, что пока еще не выработано классификации органогенных построек. Такие классификации есть, и их несколько. Но проблема заключается в том, что характеризующие информацией объекты (отдельные их типы) зафиксированы в терминах не какой-либо одной классификации (типологии), а независимо и несогласованно.

Термин «органогенная постройка» объединяет широкий набор морфологически и генетически разнородных и неоднозначно понимаемых (выделенных) предметов исследования, таких как рифы, биогермы, биостромы, карбонатные банки, каждый из которых, в свою очередь, подразделяется на разновидности. Для иллюстрации существующего положения вполне достаточно ограничиться таким предметом исследования, как риф.

Во-первых, под этот термин разные исследователи подводят принципиально разные предметы и объекты. Четко обозначается по крайней мере пять различных подходов к пониманию рифа [37]: а) выраженные в рельефе и вышедшие на уровень волнолома карбонатные тела, созданные колониальными или нарастающими организмами с жестким скелетом, т. е. каркасными организмами; б) выраженные в рельефе волноустойчивые и потенциально волноустойчивые тела, созданные любыми организмами на месте их обитания; в) мощные, ограниченные по латерали массивы чистых или преимущественно чистых пород, сложенные карбонатным материалом любой природы — органической или неорганической; г) комплекс генетически взаимосвязанных разнородных пород, включающий собственно органогенную постройку и контактирующие с ней разновозрастные эквиваленты; д) выражен-

ные в рельефе геологические тела, образованные из обломков организмов, не за счет их жизнедеятельности, а путем чисто механической аккумуляции (подобно тому, как формируются бары). Примерами могут служить органогенные постройки в юрских отложениях Западной Сибири, так называемые «устричные рифы» залива Белокси (США), онколито-оолитовые рифы в Главном доломите Цехштейна (ПНР) и др.

Во-вторых, типы «рифов» выделяются разными исследователями по совершенно разным основаниям — комбинациям признаков, число которых достигает 20 [37]. Так, выделяются рифы береговые, отмелевые, краевые (т. е. по отношению к палеогеографической зональности), барьерные (по наличию обратного склона, обособляющего лагуну), атолловые (по форме в плане близкой к кольцевой), патч-риффы (по форме и размерам), а также морфологические, биогенные, экологические, стратиграфические и др.

В таких условиях не удивительно, что, например, такой хорошо изученный нефтегазоносный карбонатный массив, как Капитен (шт. Техас, США), описан в литературе то как барьерный риф (на том основании, что он образует протяженное линейное тело с крутым склоном в сторону бассейна), то как линейная органогенная банка (по причине отсутствия в его теле каркасных организмов), то как стратиграфический риф.

Таково положение и с многими другими палеогеоморфологическими объектами, представляющими интерес в нефтегазопроисхождении. Исследованиями многочисленных коллективов геологов у нас в стране и за рубежом пока не установлены обоснованные комплексы поисковых признаков и методов поиска неантиклинальных ловушек, приуроченных к различным палеогеоморфологическим объектам. Причина заключается в том, что проблема эта не только содержательная (геологическая), но и информационная. Корректно обобщить многоаспектную информацию по значительному числу логически не увязанных типов объектов просто нельзя. Эта проблема может быть решена только на основе предварительного логического упорядочения совокупности типов объектов, информация по которым подлежит обобщению, и привлечения современных средств обработки информации.

### 2.2.3. Согласованность объектов и признаков

Многие выделяемые, в том числе и наиболее распространенные, предметы и объекты геологии связаны генетической преемственностью. В качестве примера из геологии нефти и газа можно указать систему «коллектор — природный резервуар — ловушка — залежь». Каждое последующее звено в этой генетической цепочке создается на базе предыдущего за счет присовокупления принципиально важных новых свойств. Так, если коллектор дополняется таким свойством, как экранирование слабопроницаемыми породами (хотя бы сверху), то он превращается в качественно новое с точки зрения геологии нефти и газа состояние — природ-

ный резервуар. Последний еще не определяет наличие ловушек. Для того чтобы образовалась ловушка, на природный резервуар должны наложиться важнейшие дополнительные условия, например, локальный антиклинальный изгиб кровли, седиментационная (риф, бар) или эрозионная выпуклость кровли, изгиб латерального экрана и др. Чтобы образовалась залежь, в ловушку должны поступить углеводороды (что происходит далеко не с каждой ловушкой).

Конечно, генетическая связь между звеньями указанной генетической цепочки самая тесная. Но каждому выделяемому предмету исследования свойственны только ему присущие функции и соответственно признаки, оправдывающие его выделение. Так, определяющим для породы-коллектора являются фильтрационные свойства, тесно связанные с вещественным составом и генезисом породы. Именно поэтому многими исследователями выделяются фациально-генетические типы коллекторов. Определяющими для природного резервуара являются особенности гидравлического режима (гидравлическая независимость, направление потока флюидов, перепад напоров), обусловленные качеством покрышки и латеральных экранов, числом последних. Если бы это было не так, то понятие о природном резервуаре было бы вообще излишним и можно было ограничиться понятием о породе-коллекторе. Подобно этому определяющая характеристика для ловушки — механизм улавливания УВ, обусловленный морфологической выраженностью ловушки, а для залежи — состав углеводородных флюидов, запасы УВ, характер контакта с пластовой водой и др.

Уже тот факт, что одинаковые (по составу флюидов, объему, запасам) залежи могут быть в различных типах ловушек и разные залежи — в однотипных ловушках, что одни и те же типы ловушек и коллекторов приурочены к разнородным природным резервуарам, показывает неправомерность переноса характеристик с одного предмета исследований на другой, даже если между ними тесная генетическая связь. Тем не менее на практике такое положение встречается нередко. Так, говорят о литологических, стратиграфических, антиклинальных залежах (а не ловушках), о рифовых, дельтовых, баровых ловушках (а не породах-коллекторах) и т. п. На последнем из этих примеров мы и покажем ниже, какие издержки могут происходить из такого подхода при обработке данных.

Пример 6. «Рифовые» и «дельтовые» ловушки. Геологоразведочной практикой доказано большое значение древних рифогенных и дельтовых образований как вместилищ скопленений нефти и газа. Это сделало их целенаправленным объектом нефтегазопроисковых работ во многих регионах нашей страны и за рубежом. Оправданный практический интерес к этим объектам породил постановку задачи выявления рифовых и дельтовых ловушек, которые в некоторых классификациях возведены в ранг специфических типов ловушек.

Рифовый и дельтовый генезис соответствующих образований

обуславливает такие специфические (с точки зрения нефтегазонасности) их особенности, как относительно высокие коллекторские свойства, относительно благоприятные условия аккумуляции УВ (вследствие частого непосредственного контакта материнских отложений и коллекторов) и характер распространения (геометрию в плане). По этим параметрам правомерно говорить о рифогенных и дельтовых коллекторах, а с учетом генетической преемственности и о природных резервуарах. Термины же «рифовые» и «дельтовые» ловушки являются дезориентирующими, поскольку никакой специфики ловушек, обусловленной рифовым или дельтовым генезисом коллектора, не наблюдается.

При любом генезисе коллектора тип ловушки определяется сочетанием двух параметров природного резервуара: морфологией кровли (локальной морфологической аномалией на общем фоне) и формой распространения в плане. Факторы, обуславливающие морфологическую аномалию кровли, весьма разнообразны. Например, положительная выгнутость (выпуклость) кровли природного резервуара может быть связана с антиклинальным изгибом слоев (тектонический генезис), иметь седиментационную (куполовидные рифы, бары) или эрозионную (эрозионные выступы, останцы) природу, вызываться эпигенетическими изменениями пород (ловушки уплотнения, облекания) или сочетанием нескольких из указанных факторов. В свою очередь, форма площадного распространения природного резервуара может быть обусловлена как первичной (седиментационной) геометрией коллектора определенного генезиса (в том числе рифогенного, дельтового), так и сочетанием этого коллектора с проницаемыми образованиями в контактирующих с ним (латерально и по разрезу) образованиях. В зависимости от конкретных сочетаний всех указанных факторов, с коллекторами однотипного генезиса могут быть связаны разнотипные ловушки. Проиллюстрируем это на примере сначала рифовых, а затем дельтовых коллекторов.

По форме распространения в плане погребенные рифы могут быть подразделены на две большие группы: зональные (например, барьерные и краевые рифы, в том числе кольцеобразные — атоллы) и локально-изометричные (рифы мелководного шельфа и шельфового склона). Если в контакте с погребенным зональным рифом находятся исключительно слабопроницаемые породы, то параметры такого рифа и параметры природного резервуара практически совпадают. В этом случае ловушки нефти и газа могут создаваться прежде всего в участках относительного локального поднятия кровли, образующегося в результате действия тектонического фактора — наложения на зональный риф антиклинали, периклинали или структурного носа (а, б и в на рис. 17). Среди выявленных ловушек, приуроченных к зональным рифам, таких подавляющее большинство. Это связано, по-видимому, не столько с объективным их преобладанием, сколько с тем, что современные методы поиска приспособлены к выявлению именно таких ловушек.

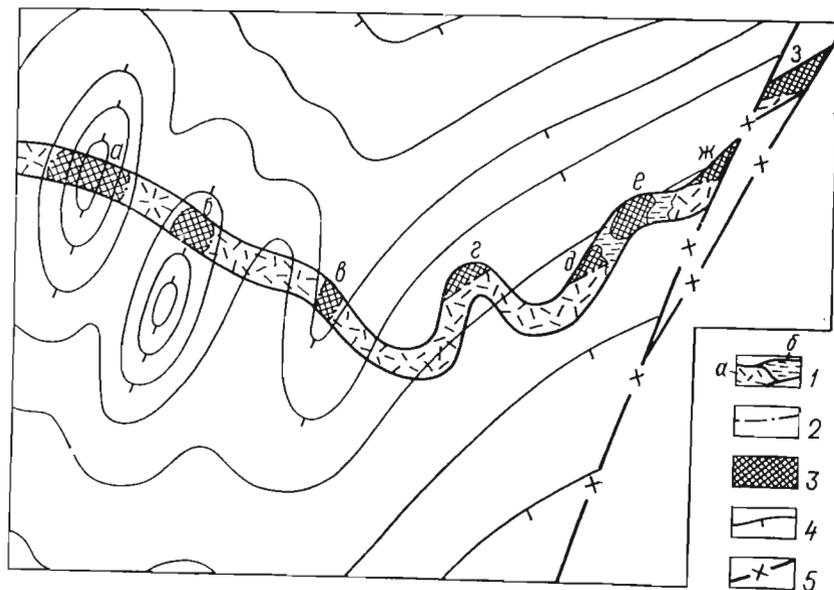


Рис. 17. Морфологические типы ловушек нефти и газа, связанных с зональным природным резервуаром [37]:

1 — зональный природный резервуар с высокими (а) и низкими (б) коллекторскими свойствами; 2 — положение контакта УВ — вода в плане; 3 — залежь УВ; 4 — структурные изолинии; 5 — разломы

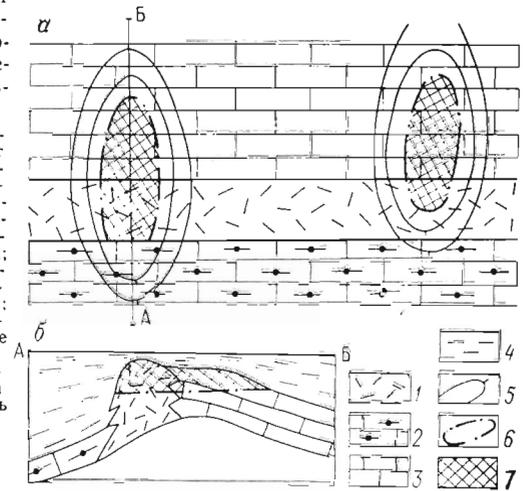
В условиях моноклинали (местный структурный план слабо дифференцирован) ловушка во всесторонне экранированном зональном рифе может образоваться только в результате его изгиба выпуклостью по восстанию слоев (тип г на рис. 17) или выклинивания зонального природного резервуара (в данном случае рифа) по восстанию слоев, в том числе в результате локального ухудшения коллекторских свойств (тип д на рис. 17) или тектонического экранирования (тип ж на рис. 17). В результате локального ухудшения коллекторских свойств возможно образование и обычных линзовидных ловушек (тип е на рис. 17).

Если примыкающие к зональному рифу и одновозрастные ему мелководно-шельфовые известняки также представляют собой коллектор, то и риф, и эти известняки составляют единый природный резервуар. В таких случаях положение ловушки определяется не зональным рифом, а местным структурным планом. В зависимости от последнего ловушка и залежь могут лишь частично и в разной степени быть связанными с рифом (рис. 18), а могут вообще оказаться за его пределами.

В случае локально-изометричных рифов, хотя они уже сами по себе обеспечивают локальное поднятие кровли, необходимое для образования ловушки, возможности аккумуляции УВ в них также существенно зависят от соотношения с контактирующими

Рис. 18. Возможные смещения ловушки (залежи) нефти и газа в случае контакта органогенной постройки с проницаемыми «зарифовыми» одновозрастными известняками:

а — положение ловушек в плане; б — профильный разрез вкострости органической постройки. 1 — зональная органогенная постройка; 2 — «предрифовые» относительно глубоководные известняки, одновозрастные с органогенной постройкой; 3 — зарифовые мелководно-шельфовые известняки, одновозрастные с органогенной постройкой; 4 — толщина компенсации рифового рельефа (в данном примере сложена слабопроницаемыми породами); 5 — структурные изолинии; 6 — положение контакта «УВ — вода» в плане; 7 — залежь УВ



образованиями. Если покрывающие отложения слабопроницаемы, а латеральные эквиваленты (одновозрастные с рифом) представлены породами-коллекторами, то риф вместе с контактирующими латеральными эквивалентами входит в состав единого природного резервуара и объем возможной ловушки определяется относительно превышением рифа (рис. 19, а). Если такое превышение невелико, то при высоких перепадах напоров пластовых вод условия для аккумуляции УВ могут оказаться неблагоприятными, и несмотря на наличие погребенного рифа, ловушка может отсутствовать. К тому же нередки случаи, когда в состав природного резервуара входят не только риф и его латеральные эквиваленты, но и покрывающие слои. Тогда ловушка может частично выходить за пределы рифа в вышележащие слои (рис. 20), а иногда и полностью оказаться за его пределами. Существенно иное положение создается, когда риф заключен в слабопроницаемые породы (рис. 19, б), представляя собой обычную линзовидную ловушку.

Подобное разнообразие типов ловушек наблюдается и в палеодельтовых природных резервуарах. В условиях субаэральной дельтовой равнины породы-коллекторы обычно связаны с дельтовыми протоками и потому часто имеют зональную форму распространения в плане. Если отложения такой протоки заключены со всех сторон в слабопроницаемых, преимущественно глинистых образованиях поймы, то образуется зональный природный резервуар, которому могут быть свойственны все те типы ловушек, которые указаны на рис. 17. Но не менее часто отложения отдельных разновозрастных дельтовых протоков налагаются друг на друга, образуя гидродинамически единый природный резервуар относительно широкого площадного развития. В таких условиях обра-

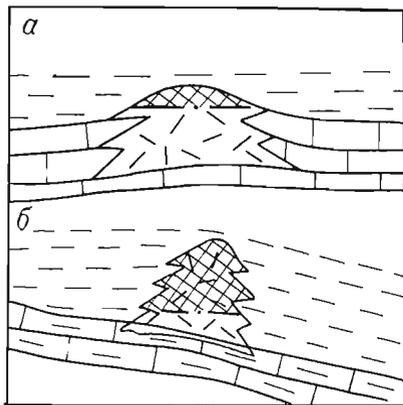


Рис. 19. Типовые схемы локально-изометричных органогенных построек шельфовой равнины (а) и шельфового склона (б):

1 — органогенная постройка; 2 — провинцаемые; мелководные известняки; 3 — слабопронцаемые глинистые известняки и мергели; 4 — глины; 5 — контакт «УВ—вода»; 6 — залежь

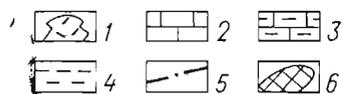
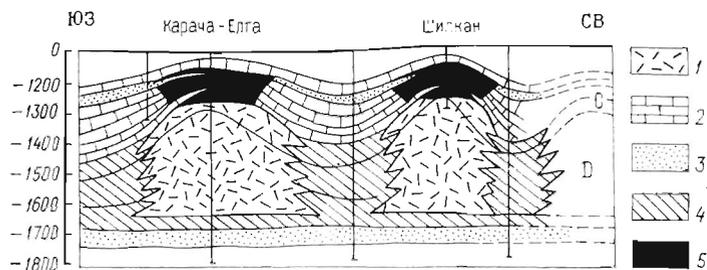


Рис. 20. Геологический разрез месторождений Шилкан и Карача-Елга Камско-Кинельской впадины. По В. Г. Кузнецову, 1975.

1 — органогенная постройка; 2 — мелководно-шельфовые слонстые известняки; 3 — терригенные отложения; 4 — кремнисто-глинисто-карбонатные породы; 5 — залежи нефти



зуются обычные антиклинальные ловушки, а у границ природного резервуара — ловушки того же морфологического типа, что показаны на рис. 18, а. Для авандельтовых частей палеодельты характерно широкое площадное развитие коллектора. Это обусловлено тем, что песчаный материал, принесенный дельтовыми протоками, плащеобразно развевается волнами и вдольбереговыми течениями. Ловушки образуются преимущественно в результате антиклинального изгиба коллектора.

Таким образом, с одной стороны, к погребенным рифам и палеодельтам приурочены ловушки, весьма разнообразные по морфологии, генезису, возможностям и методам поиска, а с другой стороны, с генетически различными коллекторами (рифовыми, дельтовыми, баровыми и др.) могут быть связаны принципиально однотипные ловушки. В таких условиях любое обобщение, базирующееся на понимании рифовых и дельтовых ловушек как самостоятельных типов, неизбежно приведет к некорректности результата обработки данных.

## 2.3. СОДЕРЖАТЕЛЬНАЯ ПОСТАНОВКА ГЕОЛОГИЧЕСКИХ ЗАДАЧ

Разработка геологических задач начинается с их содержательной постановки. Последняя заключается прежде всего в указании набора исходных данных (показателей определенных объектов, значения которых будут рассматриваться), задании цели, которая должна быть достигнута, и формулировании на теоретическом или эмпирическом уровне представлений о функциональных и логических связях между заданными и искомыми свойствами или объектами.

Логические связи, подразумеваемые между используемыми в задаче предметами исследования и их признаками, часто в явном виде не выражены и достоверность их далеко не всегда очевидна. Но поскольку именно эти связи служат основой для последующей более глубокой формализации модели решения (т. е. для математического описания задачи), они в преобладающей степени определяют достоверность получаемого при решении результата. Поэтому геологу — постановщику задачи очень важно иметь представление о том, с позиций каких логических требований следует подходить к содержательной постановке. Эти позиции, основанные на опыте методологического анализа моделей решения многих задач геологии, отражены на рис. 7 и проиллюстрированы ниже.

### 2.3.1. Соответствие предметов исследования решаемой задаче

Поставить задачу — это прежде всего задать триаду: цель — предметы (объекты) — необходимая информация. Даже если поставленная задача базируется на необъективированной (координатно-привязанной) информации, например проведение картирования (геохимического, геологического, геофизического), то, как было показано в разд. 1.2, в этой задаче задействован в неявном виде тот или иной предмет исследования, с учетом которого была выбрана система наблюдения — масштаб съемки и набор регистрируемых параметров. В разд. 2.1.1 отмечалось, что предметы исследования геологии не были преподнесены самой природой, а конструировались геологами исходя из потребностей практики, и что многие из них исторически обусловлены. Следствием, вытекающим из рассмотренной и разделяемой авторами модельно-целевой концепции, является право исследователя выбирать из имеющихся и даже конструировать новые предметы исследования, наиболее полно соответствующие той или иной задаче. Уже сама возможность выбора, во-первых, свидетельствует об отсутствии жесткой предопределенности предмета исследования той или иной задаче и, во-вторых, ставит вопрос об оптимальности выбора предмета.

Важность выбора, наиболее соответствующего задаче предмета исследования, хорошо иллюстрируется многочисленными примерами

ми из истории науки и практики. Один из таких ярких примеров — становление эволюционной теории Ч. Дарвина [13]. И до Дарвина в биологии были высказаны эволюционные представления, однако они формулировались относительно таких предметов, которые не позволили придать этим представлениям всеобъемлющий характер. Успех дарвинизма объясняется не в последнюю очередь удачным выбором предмета исследования, относительно которого формулировались законы теории. Понятие «вид», избранное Дарвиным в качестве основного, было известно и раньше. Но именно к этому времени оно приобрело устойчивость и определенность, что дало возможность четкого разграничения видов, возможность увидеть их сходство и различия.

Пример этот относится к области науки, но поскольку научные теории не самоцель, а средство совершенствования практической деятельности, выбор предмета исследования непосредственно сказывается и на реализации практических задач. Применительно к геологии минерального сырья, последнее наиболее остро проявляется в кардинальной научной и практической проблеме познания (и использования) закономерностей размещения скоплений полезных ископаемых. В качестве предмета, относительно которого ищутся и описываются закономерности размещения минерального сырья в недрах, используется, как правило, месторождение полезного ископаемого. Но, как было показано в разд. 2.2.1 (пример 2), геологические объекты более низкого уровня — залежи нефти и газа, рудные тела, и др., объединяются в месторождения часто на основе сонахождения (что весьма удобно для организации совместной их разработки), а не происхождения. Это делает выделяемые месторождения несопоставимыми, скорее технологическими, нежели природными, объединениями. А природные закономерности, знание которых лежит в основе прогнозирования и поиска полезных ископаемых, базируются только на знании происхождения (парагенетических связей). Поэтому месторождение (при современной практике их выделения), будучи оптимальным предметом в задаче организации разработки скоплений УВ, не представляет собой оптимального предмета для решения задачи выявления закономерностей размещения скоплений полезных ископаемых.

Выявляемые «закономерности» размещения плохо сопоставимых месторождений неизбежно обладают невысокими прогностическими возможностями, что не может не сказаться на геологоразведочной практике. Для решения этой задачи требуются иные предметы исследования, для которых обеспечиваются однозначность выделения и соответственно сопоставимость (залежь нефти, рудное тело и др., а также их совокупности, объединяемые на основе парагенетических связей) [28]. Таким образом, степень соответствия предмета исследования поставленной задаче в значительной степени определяет корректность содержательной постановки задачи и, как следствие, достоверность результата ее решения.

В ряде случаев истоки использования малоудачного для той или иной задачи предмета исследования диктуются не какими-либо природными факторами, а коренятся в истории становления соответствующей прикладной науки. Например, геология нефти и газа выростала из недр геотектоники, и потому она до сих пор для решения своих задач пытается приспособить, использовать предметы исследования тектоники. Кульминацией такого подхода явилась в свое время антиклинальная теория поисков. Не случайно определения таких вроде бы чисто нефтегеологических предметов исследования, как «месторождение нефти и газа», «зона нефтегазоаккумуляции», «нефтегазоносный бассейн» [28, 43], прямо или косвенно базируются на понятии «тектоническая структура». То, что такой подход нерационален, что он заводит в тупик, уже было проиллюстрировано на примерах Арлано-Дюртиюлинской зоны (см. рис. 14) и района Боливар-Коастал (см. рис. 16), в пределах которых оказалось невозможным выделить месторождения на структурно-тектонической основе. Именно потому, что каждой задаче должны соответствовать свои предметы исследования, использование тектонических предметов для целей геологии нефти и газа не всегда возможно. С методологической точки зрения гораздо более корректно не подменять нефтегазопоскоковые предметы исследования общегеологическими (в частности, общетектоническими), а иметь собственно нефтегазопоскоковые предметы, которые можно было бы сопоставлять с общегеологическими, и на этой основе выявлять закономерности размещения соответствующих нефтегазопоскоковых объектов в зависимости от различных геологических условий.

### 2.3.2. Связность параметров

Обязательный элемент содержательной постановки задачи — указание связей между заданными и искомыми свойствами или объектами. Эти связи, независимо от того, являются ли они жесткодетерминированными или имеют вероятностную природу, в принципе должны быть закономерными, инвариантными, т. е. устойчивыми, повторяющимися при принципиальном сохранении внешней среды. В противном случае постановка и модель решения задачи оказываются жестко связанными с определенными фиксированными условиями, и любое их изменение потребовало бы адаптации постановки и модели решения задачи. В то же время вопрос о выборе связей — один из самых проблематичных, поскольку представление о связях отражает имеющийся опыт, который, как правило, несет в себе значительный элемент субъективизма. Это особенно касается геологии, где опыт трудновоспроизводим экспериментально и соответственно зачастую труднопроверяем. Поэтому в основе содержательной постановки геологических задач часто лежат сомнительные, так называемые «внешние» связи, т. е. такие, которые не опираются на внутренние, *сущест-*

венные характеристики объектов, не обусловлены необходимостью этих объектов.

Показательным примером внешней связи, часто используемой в геологии нефти и газа, является распределение ловушек (залей, месторождений, запасов) по стратиграфическим комплексам, структурным элементам, интервалам глубин, фациям и т. п., описываемое как геологические закономерности. Внешний (сильно опосредованный и не обусловленный необходимо) характер таких связей проистекает из того, что в определении ловушки (залей, месторождения, запасов) стратиграфическая, структурная, глубинная, фациальная приуроченности в числе существенных признаков не значатся. Ловушки (залей, месторождения, запасы) могут быть приурочены к разнообразным стратиграфическим и глубинным интервалам, структурным элементам и фациям. И если их число более тяготеет к тому или иному интервалу, элементу, фации, то это вызвано иными, более глубокими факторами, и именно относительно них и следует формулировать связи, которые будут уже связаны внутренними, т. е. действительно закономерными. Внешние же связи не обладают важнейшим свойством, оправдывающим само их выделение, — инвариантностью. И действительно, можно привести много примеров того, что описанные для какого-либо района «закономерности» распределения ловушек по стратиграфическим интервалам, структурным элементам и другим геологическим условиям противоречат соответствующим фактическим данным по другим районам, даже с весьма сходными геологическими условиями.

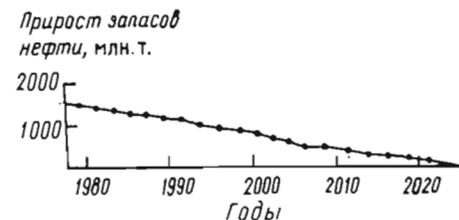
Другой пример широкого использования чисто внешних связей — использование фактора времени в задачах долгосрочного прогнозирования показателей поисково-разведочного процесса. Однако, прежде чем непосредственно перейти к описанию этого примера, следует привести некоторые предваряющие пояснения.

А. Э. Конторовичем и другими сибирскими геологами с успехом внедрена в практику планирования геологоразведочных работ на нефть и газ задача многовариантного расчета на ЭВМ объемов поисково-разведочного бурения, необходимого для обеспечения планируемого прироста запасов нефти и газа. При этом учитываются также степень начальной перспективности и степень разведанности конкретного участка недр, выраженные количественно величиной начальных суммарных извлекаемых ресурсов углеводородов  $Q$  и коэффициентом разведанности ресурсов —  $K_p$ :  $Q = D + Z + P$ ;  $K_p = (D + Z) / Q$ , где  $D$  — накопленная добыча;  $Z$  — разведанные запасы ( $A + B + C_1$ );  $P$  — прогнозные ресурсы.

Учитывается и еще ряд показателей, но их влияние на результат расчета значительно меньше, и в данном контексте (чтобы не усложнять основную идею) их можно опустить.

Если правомерность описанного выше подхода не подвергается сомнению, то столь же правомерна и постановка обратной задачи — прогнозирование прироста разведанных запасов в зависимости от затраченного объема поисково-разведочного бурения

Рис. 21. Кривая освоения ресурсов нефти и газа (прирост запасов) региона Ближний и Средний Восток [17]



при известных степени начальной перспективности  $Q$  и разведанности  $K_p$ .

Решение обратной задачи базируется на учете той же связи между теми же параметрами. Именно знание этой связи обеспечивает приемлемую степень надежности прогноза. Вместе с тем прогноз прироста разведанных запасов и добычи принято, и это совершенно оправданно, увязывать с определенными временными рубежами (датами). Это, с одной стороны, обеспечивает весьма наглядное обозрение результативности геологоразведочного процесса, а с другой — совершенно необходимо для межотраслевого планирования, в котором «состыковка» производимой и потребляемой различными отраслями продукции увязывается во временной динамике. Таким путем формируется достаточно представительная статистика динамики запасов и добычи по годам, и возникает соблазн воспринимать эту динамику как отражение определенной закономерной связи.

Соответственно строятся графики зависимости прироста запасов (или изменения  $K_p$ ) от времени (рис. 21). На этих графиках форма кривой на ретроспективу соответствует фактической динамике, а на перспективу — экстраполируется исходя из тех или иных общих соображений (с учетом того, что на начальной стадии освоения ресурсов прирост запасов должен расти, а по мере исчерпания их — уменьшаться и т. п.).

Нетрудно показать, что связь прироста запасов от времени чисто внешняя, крайне опосредованная. Очевидно, что если взять экстремальную ситуацию, например, в бассейне, являющемся объектом прогнозирования, вообще какой-то период не бурить скважины, то время будет идти, а прироста запасов не произойдет. Равномерно текущее время здесь по существу опосредует объемы поисково-разведочного бурения, изменяющиеся хотя и во времени, но в зависимости не от него, а от суммы разнообразных факторов (например, целесообразность переброски части выделенных финансовых, материально-технических ресурсов в ином бассейне, где получены более важные результаты геологоразведочных работ). В таких условиях надежность полученного результата не может быть обеспечена.

Среди весьма многочисленных опубликованных примеров геолого-экономических задач, базирующихся на прогнозировании прироста запасов (и добычи) нефти и газа во времени, пример, приведенный на рис. 21, избран потому, что он наглядно иллю-

стрирует влияние совершенно случайных, плохо предсказуемых и независимых от времени факторов. Так, Ирано-Иракский военный конфликт, начавшийся в 1982 г., привел к практическому прекращению поисково-разведочных работ (в первую очередь, бурения) в этих двух крупнейших нефтедобывающих странах Ближне-Средневосточного региона. Кроме того, начавшееся сокращение потребления нефти в мире и обусловленное этим снижение мировых цен на нефть также повлияли на сокращение объемов поисково-разведочных работ. Все это привело к тому, что фактический прирост запасов нефти в этом регионе в середине 80-х годов оказался существенно меньшим, чем это вытекает из приведенной кривой. Если бы этот прогноз базировался на реальной связи с тем или иным показателем (например, объем поисково-разведочного бурения), то надежность прогноза была бы несравненно выше. Поэтому при постановке содержательных задач надо стремиться ориентироваться не на внешние, а на существенные и потому закономерные связи.

Выше был приведен пример слабой взаимосвязи учитываемых параметров. Но при содержательной постановке геологических задач часто приходится учитывать и противоположную возможность — влияние на результат решения взаимозависимости анализируемых параметров. Последнее хорошо видно на примере широко используемого в геологии аппарата оценки информативности признаков. Он применяется в разнообразных геологических задачах. Одной из таких задач является обоснование комплекса поисковых (диагностических) признаков геологического объекта. Многие из объектов геологии (в частности, так называемые «неантиклинальные» ловушки нефти и газа) описываются очень большим числом параметров, которые анализируются при их выделении. Вместе с тем для регистрации значений каждого признака требуется применение определенного метода наблюдения. В итоге это обуславливает громоздкий комплекс поисковых методов, что удорожает и затрудняет поиски геологического объекта и потому малоприемлемо для промышленной технологии поиска. Встает вопрос о минимизации поисковых признаков (и соответственно оптимизации комплекса методов поиска) на основе оценки информативности признаков и отбора наиболее информативных.

И. Д. Савинским в результате анализа геологических задач, использующих аппарат оценки информативности признаков, было показано, что степень информативности любого отдельного признака сильно зависит от выбранной совокупности признаков. Один и тот же признак может оказаться высокоинформативным в одном наборе признаков и малоинформативным в другом. Поэтому, учитывая взаимозависимость признаков, принципиально сказывающуюся в широком диапазоне геологических задач, следует особенно строго подходить к выбору параметров задачи.

### 2.3.3. Согласованность данных по условиям измерения и масштабу

Обязательность согласованности однотипных и одновременно рассматриваемых данных по условиям измерения очевидна и вряд ли нуждается в дополнительном обосновании. В геологии признание этого проявляется во введении категорий запасов и ресурсов полезных ископаемых, в требовании кондиционности геологических карт<sup>1</sup>. При этом, наряду с решением прямой задачи (обеспечение идентичных условий измерения), в практике геологического анализа часто приходится решать и обратную задачу — обеспечивать сопоставимость данных, полученных изначально при различных условиях измерения. В качестве примеров можно указать на приведенные<sup>2</sup> пластовые давления, приведение силы тяжести к уровню моря, пересчеты дебитов скважин, замеренных при разных штуцерах, и др.

В большинстве случаев требование согласованности данных по условиям измерения стараются соблюдать, особенно применительно к непосредственно измеряемым или изначально вычисляемым параметрам. Но нередко встречаются ситуации, когда отсутствие такой согласованности не столь очевидно, завуалировано. В основном это касается характеристик, являющихся производными от непосредственно замеряемых или изначально вычисляемых (см. пример 7). Если такая ситуация не распознана и соответственно на обрабатываемые (в процессе решения геологической задачи) данные не введена поправка на различия в условиях измерения, то это не может не сказаться негативно на достоверности результата решения задачи.

Пример 7. Плотность ресурсов нефти и газа<sup>3</sup>. При определении стратегии геологоразведочных работ на нефть и газ в качестве исходного и фундаментального используется понятие «суммарные ресурсы нефти и газа». Большинство иных параметров, учитываемых при планировании и оценке таких результатов геологоразведочных работ, как коэффициент разведанности ресурсов, плотность ресурсов, геологическая эффективность, находятся в прямой или косвенной связи с понятием «суммарные ресурсы нефти и газа». Согласно последним методическим указаниям по количественной оценке прогнозных ресурсов нефти, газа и конденсата, утвержденным Мингео СССР, Миннефтепромом СССР и Мингазпромом СССР, «суммарные ресурсы» объединяют две резко различающиеся между собой группы. В первую из них

<sup>1</sup> Если точность (полнота) наблюдений в пределах хотя бы небольшой части картируемой площади не выдерживается, то карта для всей площади составляется в масштабе, соответствующем обоснованности по наименее изученному участку.

<sup>2</sup> К определенной горизонтальной плоскости (уровень моря, водонефтяной контакт и т. п.).

<sup>3</sup> В равной мере все рассмотренное в данном примере относится (приложимо) и к другим видам полезных ископаемых.

Таблица 1

Структура НСИР условных объектов оценки, млн. т

Показатели	Объект оценки I	Объект оценки II	Показатели	Объект оценки I	Объект оценки II
$Q$	150	150	$Z$	80	10
$D$	30	0	$P$	40	140

входят выявленные запасы уже обнаруженных, разведанных и вовлеченных в разработку залежей (соответствует категориям А—С<sub>1</sub>), а во вторую — невыявленные, еще не обнаруженные ресурсы, лишь предполагаемые, прогнозируемые на основании геолого-геофизических данных и сложившихся представлений о геологическом строении и нефтегазоносности оцениваемых регионов (ориентировочно категории С<sub>2</sub>—Д). Если учесть накопленную добычу, то вместе с указанными двумя группами они образуют начальные суммарные извлекаемые ресурсы (НСИР) нефти и газа.

Таким образом, в общем случае величина НСИР, которую ниже будем обозначать через  $Q$ , структурируется, т. е. состоит из совокупности накопленной добычи  $D$ , текущих разведанных запасов  $Z$ , прогнозных ресурсов  $P$ , т. е.  $Q = D + Z + P$ .

Составляющие НСИР элементы существенно не равнозначны по условиям измерения (вычисления), а следовательно, и по достоверности. Наименее достоверна величина прогнозных ресурсов. По оценкам, выполненным другими исследователями, относительная погрешность оценки  $P$  в сторону завышения достигает 0,5, а в сторону занижения — 0,1. Несколько выше достоверность оценки  $Z$ : относительная погрешность в сторону завышения составляет 0,3, а в сторону занижения — 0,2. Величина  $D$  определяется без погрешностей. С учетом этого вполне естественно, что для того или иного геологического тела достоверность оценки НСИР существенно зависит от их структуры, т. е. соотношения равнодостоверных составляющих. Иными словами, вероятность подтверждения на практике сделанной оценки НСИР тем выше, чем меньшую относительную роль в структуре  $Q$  играет  $P$ . Так, если взять два объекта (литолого-стратиграфические комплексы или бассейны в целом) с равными величинами оценки НСИР, но с разной структурой последних (табл. 1), то, основываясь на различиях в достоверности оценки разведанных запасов и прогнозных ресурсов, легко провести вычислительный эксперимент и показать, что вероятность подтверждения оценки НСИР для первого объекта много выше, чем для второго.

Таким образом, с одной стороны, всеми признается, что отдельные составляющие НСИР несопоставимы вследствие принципиальных различий в условиях измерения (вычисления), а следовательно и достоверности. Понимание этого проявляется, в частности, в том, что величина  $P$  практически не подвергается динамической

коррекции в зависимости от динамики  $D$  и  $Z$ , а пересматривается время от времени независимо на основе качественного учета результатов геологоразведочных работ. С другой же стороны, совершенно не учитывается, что несопоставимость  $Z$  и  $P$  по условиям измерения делает несопоставимыми  $Q$  с различающимися структурами. Во многих типовых задачах и геологических построениях оценки НСИР с резко различными структурами последних выступают «на равных». Эта несопоставимость НСИР в полной мере переносится и на производные параметры. Одним из последних является широко используемая плотность НСИР, обозначаемая через  $\rho$ :  $\rho = Q/V$ , где  $V$  — объем объекта оценки НСИР.

Плотность НСИР выступает в качестве меры изначальной перспективности различных геологических тел и является основным показателем, отражаемым на картах ресурсов нефти и газа. При этом районы и литолого-стратиграфические оценочные комплексы с близкими плотностями НСИР изображаются как равнозначные, независимо от того, какова структура НСИР. В таких случаях равнодостоверно оцененными оказываются районы, где преобладающая часть НСИР разведана, и районы, где разведанность НСИР ничтожна (значения  $Q$  и  $P$  близки). Фактически это означает нарушение непреложного для геологического картирования принципа кондиционности со всеми вытекающими последствиями для достоверности результата обработки соответствующих данных.

Вместе с тем имеется реальная основа для выполнения применительно к НСИР и производным параметрам требования согласованности данных по условиям измерения. Эту основу создает переход от традиционной оценки НСИР к вычислению такой вероятностной характеристики, как доверительное значение НСИР —  $\hat{Q}$ . Последнее означает, что в данном объекте оценки НСИР, последние с вероятностью  $p$  будут иметь значение не ниже  $\hat{Q}$ . Обычно принимается  $p = 0,9$ . При переходе от  $Q$  к  $\hat{Q}$  вводится коррекция на различия в структуре НСИР. Вследствие этого значения  $\hat{Q}$  для различных объектов оценки НСИР сопоставимы (так как приведены к одному и тому же значению  $p$ ), что обеспечивает корректность сравнительной оценки группы объектов оценки НСИР по нефтегазоносному потенциалу. Технология получения доверительных значений НСИР приведена в работе [30].

Столь же неочевидной, завуалированной, часто является и несогласованность масштабов объектов, процессов, параметров и временных интервалов, характеризующих совместно обрабатываемыми данными. Примеры неучета этого из практики содержательной постановки геологических и геолого-экономических задач весьма многочисленны. Так, весьма часто содержательная постановка задачи опирается на те или иные зависимости не с конкретными величинами разведанных извлекаемых запасов, а с условно принятыми градациями крупности месторождения (залежи, рудного тела). В частности, согласно действующей инструкции Государственной комиссии по запасам полезных ископаемых при Совете

Таблица 2

## Категории крупности месторождений (залежей) нефти и газа

Месторождение	Нефть, млн. т	Газ, млрд. м <sup>3</sup>	Месторождение	Нефть, млн. т	Газ, млрд. м <sup>3</sup>
Мелкое	10	10	Крупное Уникальное	30—300	30—500
Среднее	10—30	10—30		300	500

Министров СССР приняты всего четыре категории крупности месторождений нефти и газа (табл. 2).

Соответствующие им интервалы значений, обусловленные не только и не столько геологическими, сколько технико-экономическими факторами, очень широки и неравномерны (например, 0—10 и 30—500). Поэтому учесть влияние тех или иных геологических факторов на величину месторождения или залежи (например, мощности покрывки, амплитуды или площади ловушки, значения пористости и др.) часто невозможно, поскольку степень их влияния по масштабу меньше, чем это может быть зафиксировано при принятых градациях. Ситуация в этом случае подобна той, какая создается при слишком большом выбранном сечении структурной карты: многие структуры, амплитуда которых меньше величины этого сечения, не проявляются. Если же постановщики соответствующих задач, не получив требуемого результата, переходили к понижению масштаба градаций крупности (например, вводили равномерное шкалирование с интервалом 5—10 млн. т), то во многих случаях те или иные зависимости проявлялись достаточно четко.

Наиболее типичный случай несогласованности по масштабу времени — неучет временного лага. Для иллюстрации этого воспользуемся частично примером, приведенным в разд. 2.3.2. Там отмечалось, что при относительной однородности того или иного участка недр по степени перспективности в нефтегазоносном отношении (плотности начальных суммарных извлекаемых ресурсов) существует зависимость между накопленным объемом поисково-разведочного бурения и величиной начальных запасов нефти и газа, а с учетом коэффициента разведанности ресурсов  $K_p$  — также между объемом бурения за определенное время и приростом запасов. Если зависимость между указанными величинами в конкретных районах определялась по итогам каждого года (годовой объем поисково-разведочного бурения и годовой прирост запасов), то разброс данных оказывался столь большим, что установить наличие зависимости не удавалось. Когда же стали использовать значения, суммированные по пятилетиям, то зависимость становилась явно видимой. Это связано с тем, что средняя продолжительность опосредования и разведки месторождения — около 5 лет. Именно в течение соизмеримого промежутка времени сказываются закономерные связи. При годичном же масштабе вре-

мени превалировало влияние факторов, имеющих меньшую периодичность проявления.

#### 2.4. СОГЛАСОВАНИЕ ИСХОДНЫХ ДАННЫХ С ТРЕБОВАНИЯМИ ФОРМАЛЬНОГО (МАТЕМАТИЧЕСКОГО) АППАРАТА ОБРАБОТКИ

Содержательная постановка геологической задачи, как правило, характеризуется недостаточной степенью формализации для выбора и применения формального математического аппарата. Для этого постановку задачи необходимо перевести на более строгий уровень формализации, называемый математической моделью задачи. Суть этого заключается в переводе зависимостей, выраженных в содержательной постановке на языке геологии, на язык абстрактных математических соотношений, применительно к которым и выбирается затем оптимальный вычислительный метод. Последний предъявляет к исходным данным свои дополнительные требования, несоблюдение которых может оказать существенное негативное влияние на качество содержательного результата обработки даже при полной корректности содержательной постановки и математической модели задачи. В настоящем разделе мы ограничимся характеристикой трех наиболее универсальных требований к исходным данным (см. рис. 7), накладываемым формальным аппаратом обработки. Обеспечение соблюдения этих требований — одна из функций геоинформатики.

##### 2.4.1. Обеспечение статистической однородности

Геологические объекты часто характеризуются значительной неоднородностью строения и (или) неравномерной изученностью. Неучет этого обстоятельства, проявляющийся обычно в том, что по данному объекту берется усредненное значение той или иной характеристики, может привести к существенному отклонению от реальности геологического результата, получаемого при решении задачи, даже если содержательная постановка и математическая модель ее решения принципиально верны. Проиллюстрируем это положение на следующем примере.

Пример 8. Использование усредненных показателей по нефтегазоносным бассейнам при выявлении эмпирических закономерностей. Планирование геологоразведочных работ на нефть и газ в значительной мере базируется на прогнозировании величины ожидаемого прироста запасов нефти и газа. Последнее, в свою очередь, опирается на зависимость, существующую между плотностью начальных суммарных ресурсов нефти и газа, суммарным объемом проведенного поисково-разведочного бурения и достигнутыми при этом результатами, выраженными в виде величины начальных запасов нефти и газа<sup>1</sup> или числа открытых залежей различных категорий

<sup>1</sup> Начальные запасы — это результат суммирования значений накопленной добычи и текущих запасов.

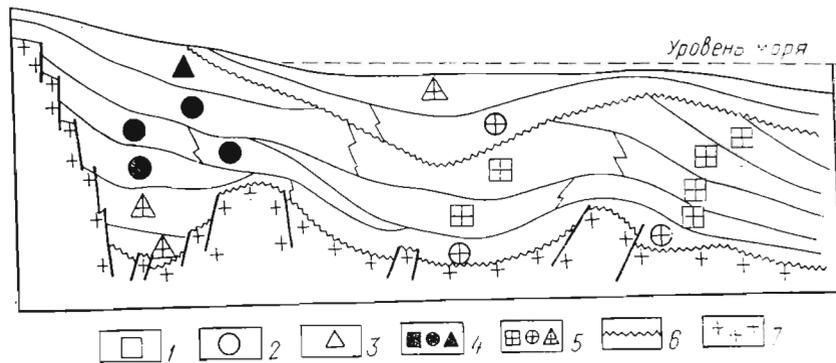


Рис. 22. Геологический разрез нефтегазоносного бассейна, характеризующегося неравномерностью распределения ресурсов и изученности поисково-разведочным бурением.

Литолого-стратиграфические оценочные комплексы: 1, 2, 3 — соответственно с высокой, средней и низкой плотностью ресурсов; 4 — охваченные поисково-разведочным бурением; 5 — не изученные поисково-разведочным бурением; 6 — поверхность несогласия; 7 — кристаллический фундамент

крупности. Поэтому многие исследователи составляют и уточняют на региональном и глобальном уровнях эмпирическую зависимость между указанными величинами. Эта зависимость представляется, как правило, в форме графически выраженных кривых.

В качестве объектов, относительно которых формулируются указанные зависимости, выступают чаще всего нефтегазоносные бассейны или области. Объекты, по которым берутся соответствующие эмпирические данные, весьма многочисленны и сильно различны по степени неоднородности строения и неравномерности изученности поисково-разведочным бурением. На рис. 22 приведена принципиальная схема одного реального, существенно неоднородного по перспективам нефтегазоносности бассейна. Бурением в нем охвачены лишь несколько литолого-стратиграфических комплексов, каждый из которых может рассматриваться как относительно однородный. Оптимальным решением в такой ситуации явилось бы получение искомой зависимости для каждого изученного бурением относительно однородного комплекса. Но если для последних плотность начальных суммарных ресурсов и результаты геологоразведочных работ обычно легко установить, то покомплексное распределение метража бурения для многих районов<sup>1</sup> восстановить крайне трудно.

Чтобы преодолеть эту трудность, используют обычно более доступные данные по объемам бурения в бассейне в целом. Соответственно приходится пользоваться и усредненным по бассейну в целом значением плотности начальных суммарных ресурсов УВ и относить к бассейну в целом результаты геологоразведочных ра-

<sup>1</sup> Статистические эмпирические зависимости необходимо обосновать на весьма представительном материале.

бот, характеризующие фактически лишь несколько литолого-стратиграфических комплексов. Поскольку значение плотности суммарных ресурсов УВ по отдельному комплексу может сильно (в схематизированном на рис. 22 конкретном случае — почти на порядок) отличаться от усредненного значения соответствующего показателя по бассейну в целом, то очевидно, что полученная таким путем элементарная (в графическом виде представленная точкой) зависимость далека от реальности. Подобные данные составляют значительную долю той исходной информационной основы, на которой базируются эмпирические зависимости, формулируемые относительно сложных геологических объектов. В этом одна из основных причин того, что прогностические возможности многих таких зависимостей невысоки.

Рассмотренный пример отражает тот нередкий случай, когда увеличение количества исходных данных, вследствие их некачественности, вызывает прямо противоположный эффект — снижение качества получаемого геологического результата. Поэтому целесообразнее опираться на данные, хотя и меньшие по количеству, но отвечающие требованию статистической однородности.

#### 2.4.2. Представительность и независимость исходных данных

В разд. 2.3.3. затрагивался вопрос о закономерном характере связей, которые должны лежать в основе содержательной постановки задачи. Если такая связь надежно установлена и является жесткодетерминированной, то на ее основе корректный результат может быть получен и при весьма скромной представительности данных. Но геологических задач, базирующихся на жесткодетерминированных связях, мало. Гораздо чаще суть геологического исследования состоит в установлении связей между объектами, процессами и (или) их характеристиками. При этом используемые в геологии связи по самой сути геологоразведочного процесса имеют в основном статистическую вероятностную природу. Это связано с тем, что из-за краткости жизни человека по сравнению с длительностью большинства геологических процессов (например, формирования скоплений полезных ископаемых), непосредственно наблюдать эти процессы нельзя и приходится фиксировать лишь их следы на многих геологических объектах, а затем проводить обобщение на статистической основе. В свою очередь, поскольку статистические связи и закономерности действуют в области массовых явлений, их выявление требует статистического анализа массового материала в отношении как числа, так и разнообразия случаев.

В идеале желательно было бы проанализировать всю совокупность объектов, процессов, в отношении которых формулируются искомые закономерные связи. На языке математической статистики такая полная совокупность называется генеральной. Получение генеральной совокупности применительно к объектам геологии, как правило, малореально, и потому почти всегда прихо-

дится довольствоваться данными по доступной (опробуемой) совокупности. И качество получаемого результата обработки данных в значительной мере определяется тем, как получена эта выборка, в какой мере она отражает генеральную совокупность. В подходе к получению выборки и заключается суть требований представительности и независимости данных. Эти два требования в значительной мере взаимосвязаны, но для удобства изложения рассматриваются последовательно. Эти требования — один из частных случаев более общего системного требования, известного как «метод полной группы» [27].

В практике геологических исследований требование, чтобы выборка адекватно отражала генеральную совокупность, часто не соблюдается. Нередки случаи, когда исследователи пытаются установить закономерные связи на ограниченном числе избранных, хорошо известных им объектов, а затем распространяют их (хотя и с оговоркой «на примере...») на все объекты того же типа. «Считается почему-то само собой разумеющимся, что солидное, детальное и многоаспектное изучение одного-двух месторождений или регионов во всех случаях предпочтительнее, чем знание всех однотипных объектов в некоем заданном отношении. Стоит ли, однако, доказывать, что при игнорировании хотя бы эмпирической полной группы результат научного обобщения несколько ценнее, например, представленный дикаря о мировой архитектуре, сформированных на основе глубокого проникновения в устройство хижин его племени» [27, с. 50]. Недостаточная представительность (по количеству и разнообразию объектов) исходных данных наряду с отмечавшимися в разд. 2.2. несопоставимостью однотипных объектов и логической некорректностью типизаций — одна из главных причин невысоких предсказательных возможностей, полученных на этой основе связей, которые неизбежно оказываются незакономерными.

Одно из следствий недостаточной представительности (избирательности) исходных данных — обилие устанавливаемых связей. Например, закономерности размещения скоплений нефти и газа устанавливаются различными исследователями (и входят затем в содержательные постановки геологических задач) в связи с той или иной системой разломов или флексур, зонами высоких градиентов мощностей, амплитудой ловушек, временем формирования последних, теми или иными фациями, степенью унаследованности структурных планов, качеством и (или) мощностью покрывки, направлением миграции флюидов, распространением материнских отложений, геометрической зональностью, степенью эпигенетического изменения пород-коллекторов, приуроченностью к той или иной фазе цикла седиментации и многими другими факторами. Интересно, что почти все указанные связи (закономерности), независимо друг от друга, описаны почти для каждого более или менее крупного нефтегазоносного региона. Но, поскольку они установлены, как правило, на недостаточно представительном исходном материале, эти нередко противоречащие друг другу

«закономерности» часто оказываются несостоятельными в большинстве районов того же региона. Эта несостоятельность неизбежно переносится затем и на результаты решения геологических задач, в содержательной постановке которых задействованы также статистически непредставительные связи. Поэтому обеспечение достаточной представительности исходных данных — необходимая предпосылка корректности любого геологического обобщения, осуществляемого на статистической основе.

Требование, чтобы выборка включала сравнительно большое число случаев, характеризующихся тем или иным разнообразием, — необходимое, но еще недостаточное условие корректного отражения ею генеральной совокупности. Необходимо также, чтобы выборка была случайной; все элементы генеральной совокупности должны иметь по возможности одинаковую вероятность попадания в выборку, т. е. быть независимыми. Назначение этого требования состоит в устранении влияния исследователя, который намеренно или ненамеренно может отдавать предпочтение одним членам генеральной совокупности перед другими, в результате чего распределение признака в выборке обязательно будет отличаться от распределения в генеральной совокупности. Характеристики, определяемые по результатам анализа таких выборочных совокупностей, всегда будут содержать систематическую погрешность, т. е. тенденциозное отклонение от оцениваемой величины в одну сторону [11].

На основе учета необходимых предпосылок применения математической статистики разработан ряд способов формирования случайной выборки. Описание и примеры этих способов приведены во многих работах по применению математических методов и ЭВМ в геологии. Важно отметить, что при соблюдении требования независимости исходных данных методами математической статистики может быть определен необходимый объем исходных данных, позволяющий оценить ту или иную характеристику генеральной совокупности с требуемой точностью и надежностью.

## 2.5. НЕКОТОРЫЕ МЕТОДИЧЕСКИЕ ПРИЕМЫ ОРГАНИЗАЦИИ И ФОРМАЛИЗАЦИИ ИСХОДНЫХ ГЕОЛОГИЧЕСКИХ ДАННЫХ

Охарактеризованные выше логические требования в совокупности и взаимосвязи образуют основу того методологического фундамента, на котором базируются (должны базироваться) формализация и организация исходных геологических данных<sup>1</sup>. В настоящем разделе мы специально выделим два методических приема организации и формализации данных, обеспечивающие как корректность результата обработки последних, так и возможность их анализа исходя из различных теоретических предпосылок. Это приемы: а) информационного приоритета элементарных предме-

<sup>1</sup> Некоторые дополнительные требования к организации и формализации данных, предъявляемые ЭВМ, рассмотрены в гл. 3.

тов (объектов) и б) независимого поэлементного описания объектов. Несмотря на принципиальную простоту этих приемов, значимость их с точки зрения геоинформатики трудно переоценить.

### 2.5.1. Информационный приоритет элементарных предметов геологии

Каждая наука и базирующаяся на ней практика ограничивают соответствующую им относительно самостоятельную и замкнутую область явлений, могущую быть представленной в виде композиции эмпирически элементарных предметов исследования (например, атомов, клеток и т. п.). В науке термин «элементарный предмет» не означает, что последний неделим ни по какому признаку, что он наиболее простой (это положение хорошо иллюстрируется примером весьма сложных элементарных частиц в физике). Понятие об элементарных предметах указывает лишь на последний предел, до которого доходит анализ той или иной конкретной области исследований, т. е. в пределах данной области эти предметы рассматриваются как исходные и неделимые. В частности, в рамках проблемы закономерностей пространственного размещения скоплений минерального сырья, такими элементарными предметами исследования и объектами являются ловушка и залежь в геологии нефти и газа, рудное тело в рудной геологии и т. п.

Сущность рассматриваемого приема заключается в том, что при подборе и любом способе хранения исходных геологических данных (от личных картотек до компьютерных баз данных), а также при содержательной постановке геологических задач должна быть обеспечена привязка максимально возможной части данных к уровню элементарных предметов. Практически значения тех или иных характеристик сложных геологических объектов, состоящих из совокупности элементарных, должны, если это принципиально возможно, получаться из значений соответствующих характеристик элементарных объектов или, по крайней мере, сопровождаться такими значениями. Например, путем операций над значениями характеристик элементарных объектов, в частности залежей или рудных тел, всегда может быть получено большинство интегральных характеристик по месторождению в целом (разведанные запасы, глубина продуктивного интервала, состав сырья и др.). Это значит, что потеря информации при максимальной ее привязке к элементарным объектам не происходит, но при этом открываются определенные возможности повышения надежности получаемого геологического результата.

Принципиальная важность такого подхода обусловлена, во-первых, отмечавшейся в разд. 1.2 жесткой зависимостью значений характеристик объекта от способа его выделения. А поскольку для многих объектов геологии способ их выделения пока не формализован, это приводит к несопоставимости выделяемых объектов и соответственно к существенному искажению результатов, полученных при аналитической обработке данных. Такое положение свойственно как сложным (пример 2), так и элементарным

(пример 1) объектам, но для последних в целом острота данной проблемы несравненно меньше. Так, при наличии проиллюстрированной разительной несопоставимости выделяемых месторождений (нефти, газа, руд и др.) границы составляющих их элементарных объектов (залежей, рудных тел и др.) проводятся различными геологами почти однозначно. Поэтому вовлечение в аналитическую обработку данных, характеризующих элементарные объекты, обеспечивает в большинстве случаев более корректный результат. Во-вторых, при отмечавшейся неформализованности (индивидуальности) выделения многих, особенно сложных объектов геологии, принцип, положенный в основу выделения, часто трудно поддается расшифровке даже при знании структуры данного сложного объекта (см. рис. 8, в, 9, 11 и др.).

В большинстве случаев геологу, стремящемуся использовать информацию по тем или иным (не им выделенным) объектам, ни принцип их выделения, ни тем более их структура не известны. В таких случаях использование информации, характеризующей такой объект в целом, с точки зрения иной теоретической концепции крайне затруднено и в любом случае чревато серьезными ошибками. Наличие информации по более или менее однозначно выделенным элементарным объектам позволяет различным исследователям независимо анализировать геологическую ситуацию с различных позиций, проводя ту или иную комбинацию элементарных объектов при образовании сложных. Это существенно расширяет возможность обработки (операциональность) геологической информации, намного продлевает срок ее актуальности в условиях закономерно изменяющихся со временем теоретических (в частности, поисковых) концепций.

### 2.5.2. Независимое поэлементное описание геологических объектов

Хотя данный методический прием, в отличие от рассмотренного выше, и не связан с иерархическими соотношениями предметов (элементарные ↔ сложные), но по самой своей сути имеет сходную с ним методологическую основу. Он также заключается в разложении описания неоднозначно понимаемого предмета (объекта) на однозначно понимаемые элементы. Важность этого обусловлена тем, что в геологии термин часто обозначает не какую-то элементарную характеристику, а то или иное сочетание таких характеристик или их значений. При этом создается ситуация, когда сами элементарные (исходные) характеристики понимаются всеми более или менее однозначно, а сочетания их, обозначаемые различными исследователями одним и тем же термином часто не совпадают. И наоборот, при одинаковых сочетаниях элементарных параметров характеризуемые ими предметы обозначаются различными терминами. Поэтому при независимом поэлементном описании, т. е. разложении обозначаемых термином сочетаний элементарных характеристик на сами элементарные характери-

ки и независимой фиксации значений последних, достигает ряд принципиальных преимуществ, поскольку обеспечивается:

1) однозначность понимания и единая схема (алгоритм) описания для всей совокупности классифицируемых объектов, а тем самым сопоставимость данных, исключающая искажение результатов их обработки;

2) компактность описания и хранения, а соответственно простота и простота обработки данных, так как число фиксируемых независимо элементов, могущих образовывать сочетания, неизмеримо меньше числа возможных их сочетаний;

3) открытость структуры информации для любого пользователя, поскольку такая информация не связана жестко с теми или иными теоретическими представлениями и в то же время доступна для анализа с позиций любого пользователя;

4) возможность составления логически корректного фиксированного набора возможных значений каждого отдельного показателя. Такой набор называется *классификатором*. Естественно, что при составлении классификатора по единичному элементарному признаку гораздо легче соблюсти логические правила деления объема понятия (см. разд. 2.1.3), чем в случае традиционных для геологии характеристик объекта, отражающих часто те или иные сочетания элементарных признаков.

Ниже мы поясним вышесказанное на нескольких примерах. При этом для большей иллюстративности мы специально будем отталкиваться от рассмотренных выше примеров 3 и 4 (разд. 2.2.2), отражающих негативный опыт, с тем чтобы показать, как его можно избежать на основе независимого поэлементного описания.

Пример 9. Формализованное поэлементное описание осадочных пород. В рассмотренном выше примере 3 было показано, что при выделении типов (а соответственно и при описании) осадочных пород на одном таксономическом уровне используются разные признаки (основания деления) и их сочетания, хотя общий набор признаков, характеризующих разные типы пород, в целом сохраняется. Практическое следствие этого — применение различных схем описания для разных групп пород, что затрудняет совместную аналитическую обработку характеризующих их данных. При этом те или иные литологические характеристики представлены в описаниях пород опосредованно, т. е. отражаются терминами, обозначающими не сами эти характеристики, а предметы, которым они присущи. Это хорошо видно на примере соотношения характеристик «размерность зерен» и «минеральный состав» и включающего их понятия «тип породы». Так, нефтегазонасыщенные горизонты в подавляющем большинстве случаев представлены двумя типами пород — песчаниками и известняками. Практически во всех моделях аналитической обработки данных по нефтегазонасыщенным горизонтам эти типы пород выступают на одном таксономическом уровне. В то же время песчаники выделяются по размерности частиц, а известняки — по минеральному составу

(кальций), и потому, например, породы кальцитового состава и песчаной размерности относятся то к песчаникам (известковистым), то к известнякам (песчанистым). Это — нарушение первого правила деления объема понятия (см. разд. 2.2.2), что не может не сказаться на корректности получаемого при обработке результата.

Таких и гораздо более сложных подобных ситуаций много, и это негативно сказывается на качестве результата обобщения (обработки) данных. Снять эту проблему вообще или существенно ослабить негативный эффект можно, если, отказавшись от традиционных терминов, обозначающих типы пород, или наряду с ними, перейти непосредственно к стоящим за ними литологическим параметрам, в первую очередь — минеральному составу и размерности зерен. Независимая фиксация значений этих параметров позволяет точно, полно и компактно отразить представленную традиционными терминами, информацию. Так, если о породе известно только, что это песчаник, то в графу «размерность частиц» записывается соответствующая градация размерности (или конкретное значение преобладающей размерности), а графа «минеральный состав» остается незаполненной. Если известно также, что песчаник кварцевый, то в графу «минеральный состав» записывается «кварц». Порода, обозначаемую только термином «известняк», можно отобразить записью «кальцит» в графе «минеральный состав», а графа «размерность частиц» останется незаполненной. Если же нужно отметить, также, что известняк, например, пелитоморфный, то в последнюю графу записывается «пелитовая».

В качестве набора возможных значений, записываемых в графу «минеральный состав», может быть использована, например, совокупность стандартных индексов или названий минералов, а в качестве классификатора для показателя «размерность частиц» также можно воспользоваться принятыми в литологии традиционными терминами, обозначающими конкретные градации численных величин: «псаммитовая», «пелитовая», «мелкая», «средняя», «крупная», «грубая» и др. Варьируя зафиксированными в классификаторах значениями размерности и минерального состава, можно отобразить практически все разнообразие осадочных пород, в том числе соли, доломиты, глины, алевролиты, ангидриты и т. п.

Проинтерпретировать такую разложенную на элементы информацию на языке исходных терминов всегда можно, и соответственно при этом не происходит существенных потерь информации (большой детальности описания должен соответствовать больший набор независимо фиксируемых параметров или большая детальность классификаторов). При таком подходе исключается также возможность описанной выше ситуации, когда одна и та же порода может быть охарактеризована неодинаково, например, то как известняк грубообломочный (кальцирудит), то как песчаник грубообломочный известковый. Фиксация значений «Са» и «грубая»

соответственно в графах «минеральный состав» и «размерность зерен» определяет ситуацию однозначно.

Для простоты восприятия сущности поэлементного формализованного описания осадочных пород мы оперировали предельно ограниченным набором (всего два) независимо фиксируемых параметров. При реальном формализованном описании пород их, естественно, больше. Ниже в несколько сокращенном виде отражен набор параметров, на основе которых специалистами ряда геологических организаций нашей страны и некоторых стран — членов СЭВ осуществлено формализованное описание осадочных пород в разнообразных неантиклинальных ловушках нефти и газа.

При формализованном описании осадочных пород использовались следующие классификаторы<sup>1</sup> значений параметров:

#### Классификатор параметра «минеральный состав»

Кварц	Серпичит
Халцедон	Хлориты
Опал	Кальцит
Оксиды и гидроксиды железа	Доломит
Оксиды и гидроксиды марганца	Арагонит
Плагиоклазы	Сидерит
Ортоклаз	Целестин
Микроклин	Ангидрит
Мусковит	Гипс
Биотит	Галит
Монтмориллонит	Калийно-магнезиальная соль
Каолинит	Фосфаты
Гидрослюда	Цеоциты
Глауконит	Пирит
Иллит	Марказит

#### Классификатор параметра «внутренняя структура компонентов»

Монокристаллическая	Сферолитовая
Поликристаллическая	Концентрическая
Аморфная	Сгустковая
Радиальная	Органическая
Лучисто-агрегатная	Волокнистая
Радиолитовая	Сетчатая
Акснолитовая	Инкрустационная

#### Классификатор параметра «размерность компонентов»

Грубая (>1,0 мм)	Мелкая (0,25—0,05 мм)
Крупная (1,0—0,5 мм)	Тонкая (0,5—0,005 мм)
Средняя (0,5—0,25 мм)	Криптозернистая (<0,005 мм)

#### Классификатор параметра «степень механической обработки»

Угловатая окатанность	Отсутствие следов механической обработки
Полуугловатая окатанность	
Округлая окатанность	

<sup>1</sup> Ряд классификаторов приведен с выборочным набором значений — в объеме, необходимом для понимания сущности подхода. Состав классификаторов может изменяться в зависимости от требуемой детальности описания.

Классификатор параметра «тип упаковки (контактов)»  
 Бесконтактный  
 Соприкосновения (точечного и линейного)  
 Приспособления (по извилистым, изогнутым и микростилолитовым границам)

Классификатор параметра «способ размещения компонентов в породе»

Равномерное упорядоченное (ориентированное)  
 Равномерное неупорядоченное (беспорядочное)  
 Неравномерное упорядоченное (сегрегационное)  
 Неравномерное неупорядоченное (пятнистое)  
 Пятнистое каемчатое  
 Пятнистое крустификационное  
 Пятнистое пленочное

#### Классификатор параметра «макроструктура»

Массивная  
 Брекчиевидная  
 Линзовидная  
 Горизонтально-слоистая  
 Градационная  
 Косослоистая однонаправленная  
 Косослоистая разнонаправленная  
 Косослоистая изогнутая

#### Классификатор параметра «текстура поверхностей наслоения»

Знаки ряби асимметричные  
 Знаки ряби симметричные  
 Отпечатки (кристаллов, соли, льда, ходов и следов организмов)  
 Трещины усыхания  
 Следы растворения и перекристаллизации (стилолитовые швы, сутуры, фунтиковы структуры)  
 Слепки (гieroглифы)  
 Отсутствие следов, знаков

#### Классификатор параметра «окраска»

Светлая  
 Темная  
 Пестроцветная сплошная  
 Пестроцветная пятнами  
 Красноцветная сплошная  
 Красноцветная пятнами

#### Классификатор параметра «включения»

Конкреции фосфатные  
 Конкреции кремнистые  
 Конкреции железистые  
 Конкреции железо-марганцевые  
 Конкреции сульфатные  
 Каустобиолиты сапропелевые  
 Каустобиолиты гумусовые

Поскольку выделение неантиклинальных ловушек основывается в преобладающей степени именно на литологических параметрах (они лежат в основе палеогеографических реконструкций), то требуется обеспечить возможность детального описания. На большом числе эталонных неантиклинальных ловушек установлено, что эта достаточно сложная задача может быть решена на основе весьма ограниченного набора элементарных параметров. Ниже приведен пример поэлементного формализованного описания породы, слагающей погребенную органогенную постройку (и приуроченную к ней ловушку) на Южно-Сосновском месторождении Припятского прогиба.

Каркасные (форменные) компоненты	
Минеральный состав, %	Доломит, 62 / кальцит, 26 / ангидрит, 8 //
Структура компонентов	Сгустковая / сетчатая //
Преобладающая размерность, %	Средняя, 58 / мелкая, 31 / тонкая, 8 //
Степень механической обработки	Округлая //
Тип упаковки	Бесконтактный //
Способ размещения	Беспорядочный / пятнистый //
Некаркасные компоненты	
Минеральный состав	Кальцит / доломит / ангидрит / пирит //
Структура компонентов	Поликристаллическая //
Преобладающая размерность	Тонкая / криптозернистая //
Степень механической обработки	—
Способ размещения	Беспорядочный //
Содержание, %	—
Эпигенетический цемент	
Минеральный состав	Доломит //
Способ размещения	Беспорядочный //
Содержание, %	—
Макроструктура	Массивная / брекчиевидная //
Текстура поверхностей наслоения	Растворение //
Окраска	Светлая //
Включения (название и минеральный состав)	Конкреции сульфатные //

Это описание представляет собой свертку всей информации, заключенной первоначально в весьма пространном традиционном литологическом описании, но, в отличие от последнего, в подготовленной к обработке и сопоставимой с другими подобными описаниями форме. В то же время эта форма без каких-либо затруднений может быть снова трансформирована в традиционное описание, причем под тем или иным углом зрения.

Следует подчеркнуть, что выбор параметров для формализованного поэлементного описания часто представляет собой результат решения нетривиальной творческой задачи, поскольку обеспечить действительную формализацию (логическую непротиворечивость) с помощью традиционно используемых в геологии характеристик не всегда возможно. Поэтому приходится вычленять из традиционных и вводить некоторые более элементарные нетрадиционные параметры. Но поскольку последние, как правило, весьма просты, они обычно используются геологами. В качестве иллюстрации сказанного хотелось бы обратить внимание на отсутствие в структуре приведенного формализованного описания таких традиционных характеристик, как «форма зерен» и «вторичные изменения».

Отсутствие первого из них связано с тем, что форма зерен не является универсальной характеристикой. Она широко используется для обломочных пород и обычно не рассматривается применительно к породам хемогенного и биогенного генезиса. Расширять толкование этой характеристики и понимать ее как форму частиц (в качестве общего признака всех типов пород) нецелесообразно, поскольку форма частиц — сложный признак, определяемый как внутренней структурой компонентов, так и условиями их

накопления (степенью механической обработки). Например, компонентам концентрической и радиальной структур (оолиты, сферолиты и т. д.) соответствует первично округлая или овальная форма, частицам монокристаллической структуры, таким как зерна кварца, полевого шпата, — идиоморфная форма, компонентам аморфной структуры — неправильная или округлая форма и т. п. На это накладываются условия накопления компонентов, фиксируемые посредством указания степени механической обработки, отражающей последующие изменения формы компонентов. Поэтому сочетанием значений таких параметров, как структура компонентов и степень механической обработки, удается полностью описать и то, что традиционно понимается как форма зерен.

Вторая из рассматриваемых традиционных характеристик — «вторичные изменения», базируется сразу на нескольких показателях — минеральном составе, внутренней структуре компонентов и характере размещения их в породе, и может быть описана сочетанием значений этих параметров. Например, регенерация песчаных зерен отображается фиксацией кварца в цементе (параметр «минеральный состав»), а степень регенерации — его количеством в цементе (параметр «содержание»). Отмечаемое каемчатое размещение регенерационного кварца (параметр «способ размещения») позволяет отличить его от первичноседиментационного цемента того же состава, но распределенного в породе беспорядочно или пятнами. Подобным образом и многие другие часто используемые традиционные комплексные характеристики не исключаются из рассмотрения, а фиксируются в иной форме — с разложением на элементы, сочетания которых эти характеристики отражают.

Пример 10. Формализованное поэлементное описание типа ловушки нефти и газа. Тип ловушки принято жестко связывать с той или иной классификацией последних. В то же время реализация такого подхода в геологической практике сопряжена со значительными трудностями, обусловленными, как было показано при рассмотрении примера 4, во-первых, обилием существующих классификаций (при отсутствии общепринятой) и, во-вторых, логическими издержками при их построении. Это предопределяет отсутствие общей основы для сравнения типов ловушек, выделяемых различными исследователями, и отрицательно сказывается на качестве результатов обработки соответствующих данных. Рассмотрение совокупности используемых классификаций показало, что при всем их разнообразии в основе типизации ловушек лежат следующие основные геологические факторы: 1) роль тектонических дислокаций (выделяются структурные или антиклинальные и неструктурные или неантиклинальные, а также комбинированные ловушки); 2) генезис коллектора (выделяются, например, рифовые, дельтовые и т. п. ловушки); 3) геометрия коллектора (при сплошном развитии коллектора выделяются антиклинальные ловушки, при локализованном — зональные линзовидные); 4) генезис латерального экрана

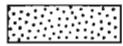
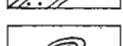
Параметр	Классификатор	Графическое изображение
Степень латерального экранирования (геометрия) природного резервуара	Экранирование отсутствует	
	Одностороннее (фронтальное) экранирование	
	Двустороннее (зональное) экранирование	
	Трехстороннее (торцевое) экранирование	
	Всестороннее (линзовидное) экранирование	
Морфология латерального экрана (зоны)	Субпрямолинейная	
	Изгиб в направлении подъема слоев	
	Изгиб в направлении падения слоев	
Генезис латерального экрана	Собственно выклинивание	
	Фациальное замещение	
	Примыкание к преседиментационному эрозионному экрану (в том числе накопление в эрозионном врезе)	
	Стратиграфическое несогласие (постседиментационное)	
	Нарушение	
Морфология кровли природного резервуара	Эпигенетическое изменение	
	Окисление нефти	
	Моноклираль ненарушенная	
	Моноклираль нарушенная	
	Изгиб вверх ненарушенный	
	Изгиб вверх нарушенный	
	Изгиб вниз ненарушенный	
Изгиб вниз нарушенный		
Генезис морфологической аномалии кровли природного резервуара	Тектонический	
	Первично-седиментационный	
	Эрозионный	
	Облекание, уплотнение	

Рис. 23. Параметры поэлементного описания ловушки нефти и газа

(стратиграфические ловушки, когда экраном является поверхность несогласия; литологические ловушки, если экранирование обусловлено фациальным замещением, эпигенетическими изменениями или собственно выклиниванием коллектора; структурные ловушки в случае экранирования нарушением и т. п.).

В каждой отдельной классификации ловушек эти факторы учитываются с разной степенью детальности и в разных сочетаниях, но, как правило, с несоблюдением первого правила деления объема понятия (с невыдержанностью основания на одном таксономическом уровне). Последнее обусловлено главным образом комплексностью большинства этих факторов: традиционно принятые для них градации значений отражают разнообразные сочетания нескольких элементарных параметров. Поэтому оказалось принципиально возможным вычленить последние и описать морфологический тип ловушки<sup>1</sup> не терминами, отражающими те или иные сочетания элементарных параметров, а независимой фиксацией значения каждого элементарного параметра (рис. 23, 24). Такой подход был использован при формировании банка данных по неантиклинальным ловушкам нефти и газа, создаваемого на основе международного научно-технического сотрудничества стран — членов СЭВ в области научно-технической информации в геологии и нефтегазовой геологии и геофизики.

Таким путем, без привязки к какой-либо конкретной классификации обеспечивается сопоставимость данных по типам ловушки. В то же время любой геолог по формализованному описанию может четко проинтерпретировать тип ловушки с точки зрения той или иной разделяемой им классификации ловушек.

### 2.5.3. Типизация геологических задач

В процессе геологоразведочной деятельности приходится решать весьма широкий круг задач, подавляющая часть которых хорошо группируется в два класса: собственно геологические, т. е. такие, решение которых приводит к накоплению знания о геологическом строении недр (включая размещение полезных ископаемых) и оптимизации геологоразведочного процесса. Интерес к типизации задач указанных классов стимулируется внедрением ММ и ЭВМ и обусловлен весьма важными практическими соображениями. Сведение великого множества рассматриваемых автономно и решаемых по-разному задач к ограниченному числу типов позволило бы объединить разрозненные и потому недостаточно эффективные усилия постановщиков и разработчиков технологий решения задач, стандартизировать методы их решения, обосновать

<sup>1</sup> Такие параметры, как литология и генезис коллектора (см. разд. 2.1.3), характеризуют породу-коллектор, а не ловушки, но при желании они могут быть включены как дополнительные элементы описания.

№ п/п	Примеры	Степень латерального экранирования коллектора	Морфология латерального экрана	Генезис латерального экрана	Морфология кровли коллектора	Генезис морфологической аномалии кровли коллектора	Графическая модель	
							в плане	в разрезе
1	Любая антиклинальная складка	Отсутствует	Нет	Нет	Выпуклость	Тектонический		
2	Любая структура уплотнения (над погребенным рифом, баром, эрозийным останцом)	Отсутствует	Нет	Нет	Выпуклость	Облекание		
3	Месторождение Джаяя (Индонезия)	Отсутствует	Нет	Нет	Выпуклость	Седиментационный		
4	Месторождение Касим (Индонезия)	Фронтальное	Прямолинейная	Нарушение	Выпуклость	Седиментационный		
5	Месторождение Эдсон (Западная Канада)	Фронтальное	Изгиб по востанию	Несогласие	Моноклираль	Нет		
6	Залежи в Хадженском районе Краснодарского края (СССР)	Фронтальное	Изгиб по востанию	Примыкание	Моноклираль	Нет		
7	Месторождение Ист-Тексас (США)	Фронтальное	Изгиб по падению	Несогласие	Выпуклость	Тектонический		
8	Месторождение Алида (США)	Фронтальное	Прямолинейная	Несогласие	Выпуклость	Тектонический		
9	Месторождение Реймерс (США)	Зональное	Прямолинейная	Примыкание	Выпуклость	Тектонический		
10	Месторождение Тилбари (Канада)	Линзовидное		Замещения	Выпуклость	Седиментационный		

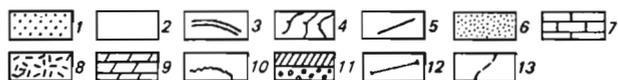


Рис. 24. Примеры поэлементного описания ловушки нефти и газа:

1 — порода-коллектор; 2 — неколлектор; 3 — латеральный экран; 4 — изогипсы; 5 — нарушение; 6 — песчаник; 7 — известняк; 8 — органогенный известняк; 9 — доломит; 10 — перерыв в осадконакоплении или размыв; 11 — залежи нефти (а) и газа (б); 12 — направление разреза; 13 — водонефтяной или газовой контакту

оптимальные программные средства, повысив таким образом уровень достоверности результатов решения и технологии геологоразведочных работ в целом.

Типология геологоразведочных задач позволила бы также оптимальным образом организовать хранение исходной для этих задач ГИ в базах данных.

Типизация геологоразведочных задач существенно облегчила бы геологам, использующим в своей работе традиционные неформализованные подходы, освоение логического и математического аппарата, применяемого в геологических задачах. Геолог в состоянии охватить и воспринять арсенал математических и логических средств не сам по себе, а через типовые задачи своей предметной области.

Естественно, что проблеме типизации геологических задач было уделено значительное внимание. Однако действительно формализованная типизация геологоразведочных задач на языке, доступном геологам, опирающимся на традиционные средства, выработанная не была. Движение в указанном направлении ограничилось в лучшем случае констатацией и формальной фиксацией эмпирически установленного набора междисциплинарных (в основном математических) методов и процедур, лежащих в основе логики геологических задач. Для обозначения и разграничения наиболее широко используемых в геологии междисциплинарных процедур были использованы в основном такие общенаучные термины, как «прогноз», «диагноз», «конструирование», «классификация», «установление мер сходства и различия», «оценка связей» и т. п. Добавление к этому слов «геологических объектов» принципиально ничего при этом не меняет. Это далеко не равносильно типизации геологических задач, поскольку не были формально зафиксированы связи междисциплинарных процедур с геологической предметной областью (геологической информацией).

Наряду с попытками общей типизации задач геологии происходил процесс специализации на отдельных обособленных классах геологических задач, для обозначения которых автономно вводились различные термины. В совокупности это привело к парадоксальной ситуации: в математической геологии, занимающейся внедрением формальных подходов, семантическая расплывчатость терминов оказалась не меньшей, чем в традиционном языке геологии. В качестве примера можно указать на термин «задачи геологического прогнозирования». Благодаря прекрасным математическим разработкам А. Н. Бугайца в области рудной геологии этот термин получил очень широкое распространение. Им обычно обозначаются задачи, конечная цель которых — прогноз пространственного положения (локализация) геологического объекта. Но по определению, прогнозирование — это «предсказание, вероятность по своей природе» [19]. Совершенно очевидно, что подавляющее большинство геологических заключений вероятно по своей природе, и потому под термин «геологическое прогнозирование» подходят в равной мере многие принципиально различные геоло-

Таблица 3

## Технологические типы геологических задач, базирующихся на предметном знании

Тип ГИ на входе	Процедура	Тип ГИ на выходе	Примеры содержательной постановки геологических задач
НИ <sub>1</sub>	Интерполяция и экстраполяция	НИ <sub>2</sub>	Построение карт геологических параметров Изучение изменчивости геологических характеристик по площади
НИ <sub>2</sub>	Выделение аномалий	ОИ <sub>1</sub>	Выделение участков локального уменьшения мощности горизонта (для целей реконструкции положения палеоподнятий)
ОИ <sub>1</sub>	Измерение параметров объекта	ОИ <sub>2</sub>	Оптимизация разведочной сети
ОИ <sub>2</sub>	Интерполяция и экстраполяция	ОИ <sub>3</sub>	Изучение пространственных изменений характеристик объекта
ОИ <sub>3</sub>	Моделирование объекта	АОИ	Определение количественных характеристик (интервалов значений и их частоты) параметров объектов определенного типа
АОИ	Эмпирическое обобщение	ИИ <sub>1</sub>	Оптимизация описания объектов данного типа (для кадастров объектов, сравнительного анализа, установления комплекса поисковых признаков и т. п.)
НИ <sub>2</sub> +ИИ	Геологическая интерпретация	ОИ <sub>1</sub>	Прогнозирование пространственного положения (локализация) геологических объектов
НИ <sub>2</sub> +ИИ	Классификация (районирование по комплексу признаков)	ОИ <sub>1</sub>	Районирование территории по степени перспективности (построение карт перспективности)

гические задачи (например, предсказание не только положения геологического объекта, но и его типа, свойств, состояния в настоящем и будущем и др.).

Принципиальную основу для формализованной типизации геологических задач может представлять информационная модель геологоразведочной деятельности, отраженная на рис. 1, 2, 4, 5. В разд. 1.2 отмечалась возможность построения многочисленных классификаций (типизаций) одних и тех же предметов в зависимости от того, какие выбираются основные деления. Выбор же этот определяется той практической целью, для реализации которой создается классификация. В контексте данной работы естественно ориентироваться на те цели и задачи, которые должны решаться средствами геоинформатики. Поскольку последняя определяется как машинизированная технология сбора, хранения, обработки и использования геологической информации, вполне правомерно в качестве основания для типизации геологических задач использовать различия в технологии их решения. Эти различия определяются типом ГИ на входе задачи (подзадачи),

## Технологические типы геологических задач по выявлению связей между геологическими объектами и (или) характеристиками

Тип входной ГИ	Процедура	Тип. выходной ГИ	Примеры содержательной постановки геологических задач
НИ <sub>1</sub>	Оценка связей	ФИ <sub>1</sub>	По кернам скважин района X оценить влияние глубины залегания глинистой толщи Y на удельную плотность глин  На основе кернового материала установить характер зависимости коллекторских свойств известняков горизонта K в районе Z от вторичных изменений (доломитизации, вторичной трещиноватости)
НИ <sub>2</sub>	То же	ФИ <sub>1</sub>	Установить степень соответствия структурных планов двух разновозрастных горизонтов  Выяснить, как отражается изменение мощности соленосной толщи района B в гравиметрическом поле
ОИ <sub>1</sub>	»	ФИ	Оценить приуроченность объектов одного типа к объектам другого типа, например, погребенных барьерных рифов к региональным палеофлексурам (ФИ <sub>2</sub> ) или залежей нефти района N к структурам древнего заложения (ФИ <sub>1</sub> )
ОИ <sub>2</sub>	»	ФИ <sub>1</sub>	Для залежи нефти Y выяснить (по кернам и каротажным кривым) наличие зависимости между изменчивостью литологии коллектора и коэффициентом его нефтенасыщенности  По пробам из шурфов и обнажений установить зависимость концентрации элемента X на месторождении Y от типа пород
ОИ <sub>3</sub>	»	ФИ <sub>1</sub>	Выяснить тенденцию смещения свода антиклинального поднятия с глубиной (на основе сопоставления серии структурных карт поднятия по различным горизонтам)

Тип входной ГИ	Процедура	Тип выходной ГИ	Примеры содержательной постановки геологических задач
АОИ	Оценка связей	ФИ <sub>2</sub>	Установить, какие типы ловушек нефти и газа характерны для природных резервуаров определенного (например, руслового) генезиса  Выделить наиболее информативные (диагностические, поисковые) признаки объектов данного типа
НИ <sub>2</sub> +АОИ <sub>1</sub>	То же	ФИ <sub>1</sub>	Выяснить, проявляются ли в гравиметрическом поле данного района погребенные органогенные постройки
ОИ <sub>3</sub> +АОИ	»	ФИ	Установить, проявляются ли фациальные различия песчаных горизонтов в форме кривых каротажа (в частности, СП)

процедурой обработки и типом ГИ на выходе. Следует отметить, что в этой триаде истоки первого и последнего элементов (информация) лежат в геологической предметной области, а средний элемент (процедура) является междисциплинарным, общим для многих предметных областей.

Если число возможных традиционных содержательных постановок геологических задач практически неограниченно (из-за обилия видов полезных ископаемых и свойственной им специфики типов выделяемых предметов, понятийно-терминологической базы, закономерностей размещения объектов и методик их поиска), то число типов ГИ и типовых процедур обработки информации относительно невелико. Поскольку к тому же к каждому типу ГИ приложимо лишь ограниченное число процедур, то число возможных их сочетаний (т. е. технологических типов геологических задач) вполне обозримо и конечно. В свою очередь, каждой процедуре обработки информации соответствует определенный ограниченный набор математических методов. Поэтому связь геологической задачи и соответствующего ей математического аппарата может быть определена через указание процедуры (процедур) обработки информации.

Переход от любой из разнообразных традиционных содержательных постановок геологических задач к указанию технологического типа задачи достаточно формализован и прост для геолога, так как для этого последнему достаточно иметь представление о типах ГИ и процедурах ее преобразования (см. рис. 1, 2, 4, 5).

С указанных позиций совокупность геологических задач может

Технологические типы геологических задач по прогнозированию на основе связей между геологическими объектами и (или) характеристиками

Тип входной ГИ	Процедура	Тип выходной ГИ	Примеры содержательной постановки геологической задачи
НИ <sub>1</sub> +ФИ <sub>1</sub>	Прогнозирование по связям	НИ <sub>1</sub>	Оценить удельную плотность глин в конкретных точках исходя из глубин залегания глинистой толщи У  Осуществить прогноз коллекторских свойств известняков горизонта К в скважинах по данным о доломитизации и вторичной трещиноватости
НИ <sub>2</sub> +ФИ <sub>1</sub>	То же	НИ <sub>2</sub>	Построить структурную карту не вскрытого скважинами участка горизонта исходя из изученного структурного плана вышележащего горизонта  По гравиметрическим данным составить прогнозную карту изопакит соленосной толщи в той части района, где она не пройдена скважинами
ОИ <sub>1</sub> +ФИ	»	ОИ <sub>1</sub>	Наметить участки поиска барьерных рифов исходя из вероятной приуроченности их к региональным палеофлексурам  Наметить первоочередные объекты для поисково-разведочного бурения на нефть и газ исходя из приуроченности залежей преимущественно к структурам древнего заложения
ОИ <sub>2</sub> +ФИ <sub>1</sub>	»	ОИ <sub>2</sub>	С целью подсчета запасов нефти залежи У осуществить прогноз коэффициента нефтенасыщенности на основе данных о литологии коллектора в тех точках (скважинах), где соответствующие прямые определения не проводились
ОИ <sub>3</sub> +ФИ <sub>1</sub>	»	ОИ <sub>3</sub>	Обосновать положение поисковой скважины в сводовой части поднятия по не вскрытому скважинами перспективному горизонту с учетом тенденции смещения свода антиклинального поднятия с глубиной

Продолжение табл. 5

Тип входной ГИ	Процедура	Тип выходной ГИ	Примеры содержательной постановки геологической задачи
АОИ <sub>1</sub> +ФИ	Прогнозирование по связям	АОИ	<p>Спроектировать систему расстановки скважин для поисков скоплений нефти и газа, приуроченных к погребенному палеоруслу, с учетом специфики предполагаемых типов ловушек</p> <p>Обосновать систему поисков объектов определенного типа на основе регистрации поисковых признаков объектов данного типа</p>
НИ <sub>2</sub> +ФИ <sub>1</sub>	То же	ОИ <sub>1</sub>	Наметить по гравиметрическим данным участки возможного развития погребенных органогенных построек
ОИ <sub>3</sub> +ФИ	»	АОИ	Провести фаціальную интерпретацию (осуществить прогнозирование генезиса) песчаных горизонтов разреза по кривым каротажа (в частности, ПС)

быть четко подразделена на две большие группы, базирующиеся прямо или косвенно на предметном и на функциональном знании.

В табл. 3 сведены типы задач, реализуемых посредством какой-либо одной процедуры преобразования информации (монопроцедурные задачи). Но многие традиционные геологические задачи являются полипроцедурными, т. е. их реализация требует последовательного решения нескольких монопроцедурных задач. В качестве примера можно указать на такие содержательные постановки геологических задач, как установление типа геологического объекта или подсчет запасов полезных ископаемых.

В табл. 4 и 5 приведены технологические типы геологических задач, основанные на функциональном знании. Реализация задач этой группы базируется в основном на двух процедурах — оценке связей и прогнозировании на основе связей. Соответственно этому в рамках указанной группы задач выделяются две подгруппы. Между задачами этих подгрупп есть известный элемент преемственности. Чтобы отразить это обстоятельство, в табл. 4 и 5 в качестве примера использованы родственные содержательные постановки задач. Следует отметить, что в настоящее время в подавляющем большинстве типов геологических задач этой группы задействована первичная (эмпирическая) функциональная информация (ФИ<sub>1</sub>), что отражает недостаточный уровень развития теоретической геологии.

Подобным образом могут быть зафиксированы технологические типы задач оптимизации геологоразведочного процесса. В послед-

нем случае появляется еще один вид информации — априорная (АИ). В качестве примера последней можно назвать задание ограничений по метражу бурения и затратам, ориентацию на достижение определенной величины прироста запасов и т. п. В задачах оптимизации геологоразведочного процесса широко используется также деятельностная информация (ДИ).

Рассмотренный выше технологический подход к формализованной типизации задач геологической отрасли позволяет как геологу, заинтересованному в результатах решения задачи, так и разработчику математической модели задачи и специалисту по автоматизированным системам обработки ГИ легко ориентироваться при формировании оптимальной технологии решения задачи.

## ОРГАНИЗАЦИЯ ГЕОЛОГИЧЕСКИХ ДАННЫХ В АВТОМАТИЗИРОВАННЫХ СИСТЕМАХ

Все более широкое применение электронно-вычислительной техники в практике обработки геологической информации требует уточнения представлений о функциях и возможностях ЭВМ. Известно, что ЭВМ является универсальным инструментом, позволяющим получать качественно новые результаты в основном за счет огромной скорости обработки больших объемов информации и пригодным к использованию в любой сфере научной и производственной деятельности. Однако ЭВМ сама по себе «не знает» проблем, задач и специфики этих сфер деятельности. Для того чтобы в любой сфере, в частности геологической, использовать возможности ЭВМ, необходимо заложить в нее не только информацию о предметной области — сведения о геологических объектах и связях между ними, геологических параметрах и их возможных значениях, геологических задачах, но и математические формулировки задач, описание способов получения результата (алгоритмы) и описание алгоритмов на специальном языке (программы), без которых ЭВМ не решит поставленных задач. Поэтому на базе ЭВМ создаются автоматизированные системы, в рамках которых осуществляется решение геологических задач.

Автоматизированная система (АС) представляет собой сложное взаимодействие связанных между собой компонентов: собственно ЭВМ; комплекса программ, организующих работу ЭВМ; специальным образом организованной исходной информации; комплекса программ, обеспечивающих различные действия над информацией (управление информацией); комплекса программ, реализующих вычислительные процессы в соответствии с избранным математическим методом и др.

От того, насколько соответствуют друг другу компоненты АС, зависит эффективность применения ЭВМ. Если, например, неудачно организована исходная информация (т. е. не решены корректно рассмотренные в предыдущих главах проблемы информатики и неверно подобран комплекс программ управления), то система, использующая самую совершенную ЭВМ, будет обладать меньшими возможностями, чем система, базирующаяся на менее мощной ЭВМ, но с лучшей организацией информации и более адекватным управляющим программным комплексом.

Таким образом, применение ЭВМ подразумевает достаточно сложную многоступенчатую технологию, базирующуюся на знаниях как вычислительной техники и кибернетики, так и предметных (в данном случае геологических). Эта технология требует в первую очередь корректной постановки геологических задач, формализации процедур их решения (автоматизировать можно

лишь формализуемые, сводимые к элементарным действиям) и соответствующей логической организации исходной геологической информации с учетом требований ЭВМ к форме ее представления.

Именно уровнем выполнения этих технологических операций определяются корректность результатов автоматизированного решения геологической задачи и возможности, открывающиеся перед геологом-пользователем. Последнему принципиально важно иметь представление об организации (формализации) информации в АС вообще и в конкретной АС в частности, поскольку лишь на этой основе геолог может оценить, что ему следует ожидать от применения данной системы. В том, чтобы вооружить геолога — потенциального пользователя АС — основой для формирования такого представления, и состоит назначение данной главы.

### 3.1. ОБЩИЕ ПРИНЦИПЫ ОРГАНИЗАЦИИ ОБРАБОТКИ ИНФОРМАЦИИ В СИСТЕМАХ

#### 3.1.1. Типы систем по их назначению

В основе любой деятельности, содержащей работу с информацией, лежат процедуры ее накопления, хранения, коррекции и выборки, обеспечивающие дальнейшее использование информации. Все эти процедуры по своей сути и в принципе формальны, поэтому, естественно, в первую очередь автоматизировались именно они. Системы, выполняющие такие операции, называются автоматизированными информационными (АИС) или автоматизированными информационно-поисковыми (АИПС) системами.

В зарубежной литературе термин «информационная система» иногда употребляется в более узком смысле, для обозначения АС, ориентированных на удовлетворение произвольных (заранее не определенных) запросов, и потому допускающих длительный поиск информации. По времени, в течение которого реализуется запрос на поиск данных, антиподом информационных АС выступают АС оперативной обработки, предназначенные для удовлетворения заранее определенных (стандартных) типов запросов с жесткими временными ограничениями на поиск (работают в реальном режиме времени). Однако вопросы оперативности обработки информации здесь специально не рассматриваются, поскольку характер основной массы задач геологической предметной области часто не требует работы АС в реальном режиме времени (например, оперативной корректировки быстро меняющихся данных или практически моментальных ответов на запросы), а также потому, что обсуждается организация исходной информации в АС и возможности ее использования, и именно этот аспект положен в основу типизации автоматизированных систем.

В зависимости от того, какая единица информации обрабатывается системой, автоматизированные информационные или информационно-поисковые системы подразделяются на документо- и фактографические [41].

В автоматизированных документографических информационно-поисковых системах (АДИПС) в качестве единицы ввода, обмена, выдачи используется порция информации, называемая документом и содержащая рассматриваемую как единое целое совокупность фактов. Каждый отдельный факт, имеющийся в документе, не выделяется и не обрабатывается системой. Документами являются статьи, рефераты, таблицы, карты и т. д.

Основная функция АДИПС — поиск хранящихся в ней необходимых пользователям документов. Обычно употребляются упрощенные способы поиска с использованием ключевых слоев (дескрипторов), приближенно отражающих содержание документа. В качестве дескрипторов выступают слова и специальные термины, которые наиболее характерны для данной предметной области и соответствующего документального фонда (для выбора дескрипторов используются тезаурусы<sup>1</sup>). Например, для предметной области геологии ключевыми словами могут служить названия элементов политико-административного, географического и геологического районирования, полезных ископаемых, разделов геологии и смежных наук, методов геологических исследований, типов документов и т. д.

Ключевые слова документа, совокупность которых образует его поисковый образ, выделяются специалистами службы научно-технической информации при помещении документов в информационные массивы АДИПС или самой ЭВМ в процессе ввода. Нахождение требуемых документов осуществляется сопоставлением поискового образа каждого документа с ключевыми словами запроса.

Из вышеприведенной характеристики следует, что АДИПС может выдать по запросу лишь документ в целом, но не конкретные фактические сведения, в нем содержащиеся, даже если эти сведения в определенной степени формализованы. Например, если документ представляет собой таблицу, содержащую сведения о ежегодной динамике добычи полезных ископаемых (нефть, газ, уголь и др.), то запросить и получить можно лишь всю таблицу, а не отдельное число, характеризующее величину добычи того или иного полезного ископаемого в определенный год.

Фактографические системы по сравнению с документографическими требуют большей степени структуризованности и формализованности исходных сведений: вся информация должна быть разложена на элементы, каждый из которых имеет наименование. Представленная в таком виде информация называется данными, а каждый ее элемент — элементом данных.

Приведенное определение понятия «данные» (свойства данных рассматриваются в следующих разделах главы) часто не соответствует, а иногда прямо противоположно тому, что принято назы-

<sup>1</sup> Тезаурус — словарь, содержащий перечень слов, отобранных в результате анализа текстов по какой-либо предметной области и систематизированных в соответствии с принятой системой (например, по частоте встречаемости).

вать геологическими данными. Например, к геологическим данным относят описание керна скважины:

1210—1212 м — песчаник серый и темно-серый, кварцевый, от мелко- до крупнозернистого, косослоистый, слабоцементированный, пиритизированный.

Однако описание керна в традиционной форме не является данными для фактографической системы, а представляет собой документ и может быть введено и хранимо лишь в АДИПС. Для того чтобы эта информация стала данными для фактографической системы, ее необходимо разложить на элементы, поименовав каждый из них:

Интервал бурения: от 1210  
до 1212 м  
Тип породы: песчаник  
Цвет: серый, темно-серый

Минеральный состав: кварц, пирит  
Размерность зерен: мелкая, средняя, крупная  
Слоистость: косая  
Степень цементации: слабая

Кроме того, для каждого элемента данных необходимо задать совокупность допустимых значений (формально определить область значений), т. е. в приведенном примере перечислить допустимые названия типов пород, цвета, слоистости и др., или для числовых данных определить граничные значения (здать возможные минимальные и максимальные значения глубин интервалов отбора керна).

Таким образом, в автоматизированных фактографических информационных системах (АФИС) единицей ввода, обмена и выдачи является отдельный поименованный факт — элемент данных, принимающий значения из заранее формально определенной области. Примеры элементов данных: площадь, запасы, глубина, тип породы, минеральный состав и их значения.

По сравнению с АДИПС фактографические системы не только позволяют более адекватно описать предметную область и соответственно более эффективно обеспечить информационные потребности пользователей, но и дают возможность автоматизированного решения аналитических (вычислительных, логических) задач, связанных с различными преобразованиями информации. Именно это, а также широкое распространение в настоящее время позволило в данной работе акцентировать внимание на АФИС.

По назначению АФИС можно подразделить на справочные, обеспечивающие выдачу данных по запросам пользователей (обработка выданной информации выполняется самими пользователем), и справочно-аналитические, выполняющие как справочные, так и аналитические функции. В последнем случае в роли пользователей данных выступают прикладные программы, реализующие решение аналитических задач.

### 3.1.2. Технологическая структура систем

Функционирование АС предусматривает наличие разнообразного обеспечения — обеспечивающих подсистем (рис. 25).

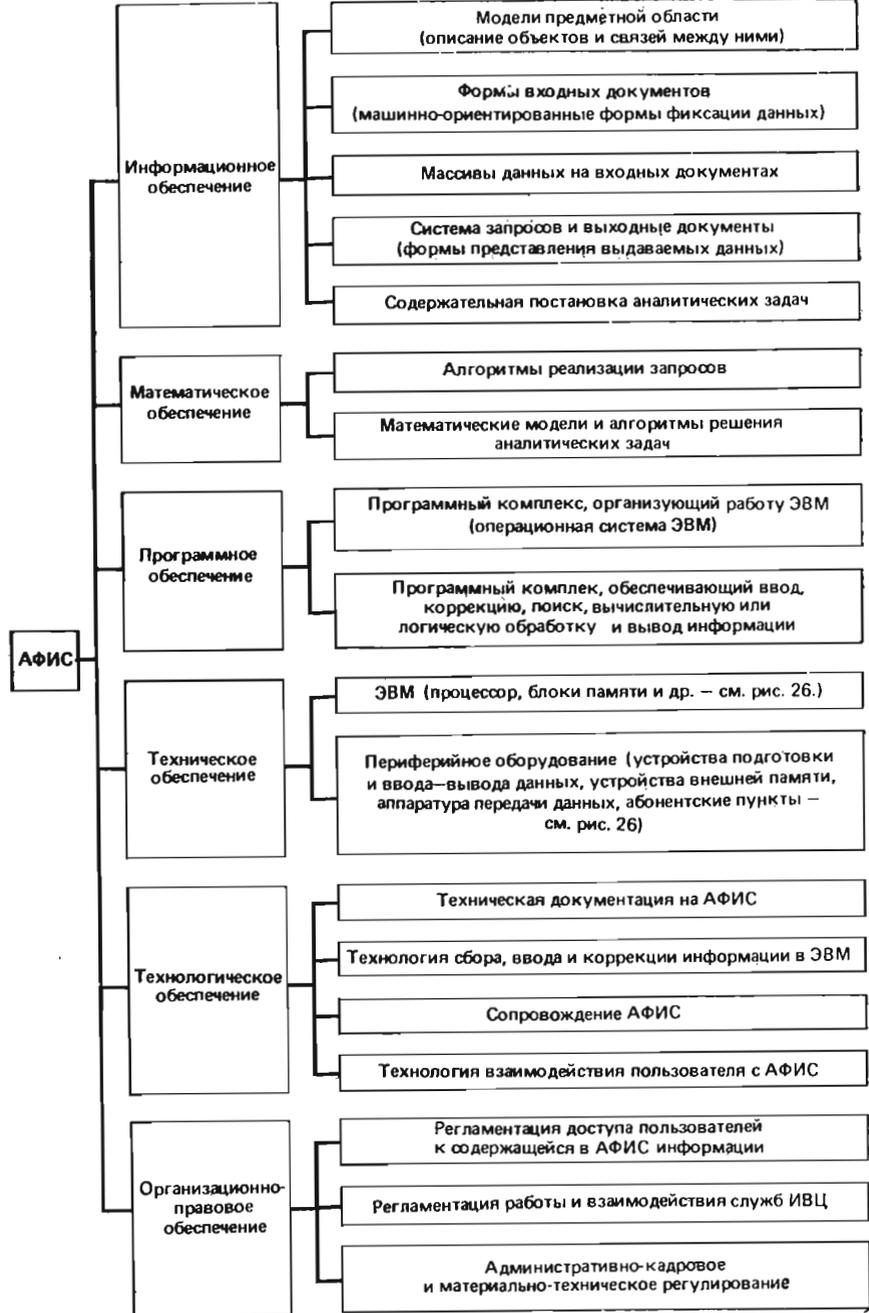


Рис. 25. Обеспечивающие подсистемы (виды обеспечения) АФИС

Информационное обеспечение состоит в описании предметной области, при котором задаются типы объектов (административная единица, нефтегазоносный бассейн, металлогеническая провинция, месторождение, залежь, скважина и т. д.), характеристики этих объектов (площадь, климатические условия, мощности и литологические параметры пород, типы полезных ископаемых, глубины залегания залежей, методы поисков и разведки и др.) и взаимоотношения между объектами (залежь входит в состав месторождения, месторождение — в область, республику или страну и в то же время в бассейн или металлогеническую провинцию и т. д.).

Информационное обеспечение включает также: набор системно увязанных, специально разработанных форм (таблиц, бланков) для записи информации, подлежащей вводу в АФИС, — так называемые машинно-ориентированные входные документы; массивы данных на входных документах, составляющие информационный фонд АФИС; набор типовых информационно-справочных запросов на поиск данных, сформулированных на естественном или близком к нему языке, но с соблюдением определенных правил (сюда включаются, в частности, запросы, по которым формируются в требуемом порядке исходные данные для решения аналитических задач); формы представления результатов обработки информации — так называемые выходные документы; содержательные постановки аналитических задач (см. разд. 2.3).

Кроме того, в информационное обеспечение входят средства представления информации — системы знаков, терминов, индексов, имен собственных, которыми можно пользоваться при фиксации значений показателей. Средства представления информации должны удовлетворять общим для различных АС требованиям: быть унифицированными (обеспечивать однозначность представления одних и тех же сведений), понятными для пользователей; обеспечивать полноту и детальность представления информации, допускать возможность дополнения и изменения без перестройки их общей структуры; применяемые знаки (буквы, цифры, знаки препинания, служебные символы) должны соответствовать возможностям используемых технических средств.

Подробнее структура информационного обеспечения по всем аспектам и требования, предъявляемым к нему, рассмотрены в разд. 3.2—3.6.

Математическое обеспечение предназначено для описания языком математики процедур реализации информационно-справочных запросов и решения аналитических задач. Математическое обеспечение включает математические модели аналитических задач и алгоритмы решения аналитических и информационно-справочных (выдача данных по запросам) задач. При этом модели и алгоритмы могут быть как весьма универсальными — пригодными для использования в разнообразных содержательных задачах (например, вычисление средних и дисперсии распределения, коэффициентов корреляции, которые часто включаются в ма-

тематическое обеспечение самой ЭВМ), так и узко ориентированными на решение какой-то определенной аналитической или информационно-справочной задачи.

Программное обеспечение представляет собой комплекс машинных программ, написанных на различных алгоритмических языках, а также описаний и инструкций по их применению, обеспечивающих работу АС. Различают общее и специальное программное обеспечение. Первое прилагается к любой современной ЭВМ. Основная часть общего программного обеспечения — операционная система, организующая и обеспечивающая выполнение вычислительных обрабатываемых процессов. Главные функции операционной системы: включение в работу других программ и участие во вводе данных в ЭВМ; координация использования ресурсов (оборудования) ЭВМ; планирование и управление вычислительным процессом. Кроме операционной системы, общее программное обеспечение включает также средства поддержания программ в рабочем состоянии, средства трансляции программ (перевода программ решения задач, записанных на разных машинно-независимых языках программирования, на язык, понятный ЭВМ); средства контроля, диагностики и коррекции работы ЭВМ.

Специальное программное обеспечение характеризует скорее возможности АС, а не ЭВМ. Оно включает комплекс программ управления информацией, осуществляющих ввод, коррекцию, пополнение, поиск и выдачу данных; программы управления теледоступом (обеспечение связи с удаленными дисплеями, абонентскими пунктами); программы доступа к экрану дисплея; программы, реализующие решение информационно-справочных и аналитических задач; вспомогательные средства, ускоряющие, стабилизирующие и контролирующие работу указанных выше программ.

Техническое обеспечение — это технические устройства вычислительного центра, с помощью которых функционирует АС. Основным элементом технического обеспечения — современная ЭВМ (или многомашинный комплекс ЭВМ), составные части которой выполняют следующие функции: процессор (или несколько процессоров) осуществляет все вычислительные (арифметические, логические) операции; оперативное запоминающее устройство служит для записи, хранения и быстрой выборки данных и программ при расчетах, проводимых в процессоре; постоянное запоминающее устройство предназначено для постоянного хранения и выборки ряда внутримашинных программ, в том числе программ операционной системы; каналы (интерфейсы) контролируют связь между ЭВМ и периферийным оборудованием.

Основными частями периферийного оборудования являются (рис. 26): устройства внешней памяти (или долговременные запоминающие устройства для записи, хранения и считывания данных) на магнитных носителях информации, емкость которых значительно больше, чем емкость оперативного запоминающего

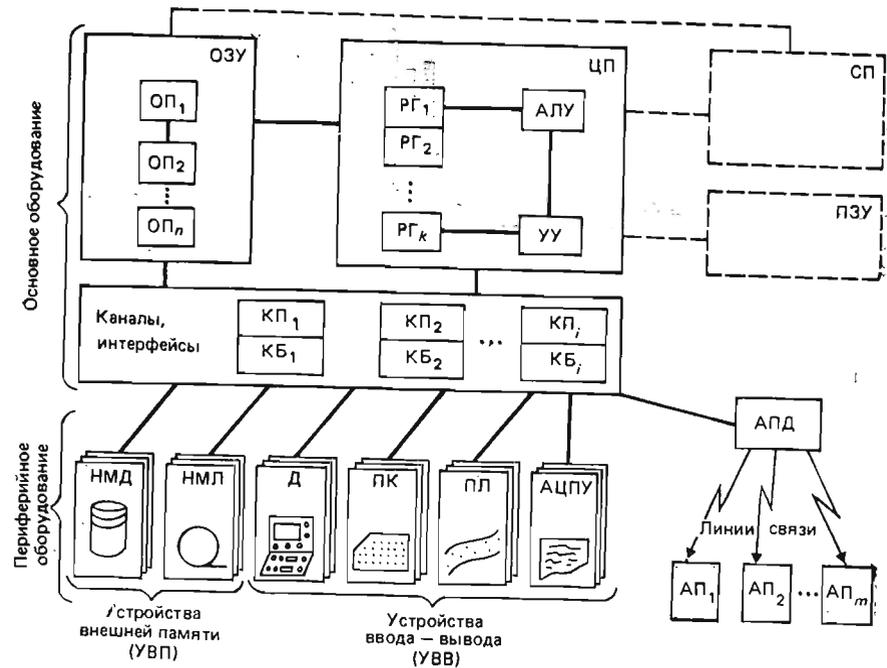


Рис. 26. Типовая структура современного вычислительного комплекса (ЭВМ): ЦП — центральный процессор; АЛУ — арифметико-логическое устройство; УУ — устройство управления ЦП и всеми операциями, производимыми ЭВМ; РГ — регистры-ячейки сверхбыстрой памяти (для хранения данных, над которыми выполняется текущая операция, и результата ее выполнения); ОЗУ — оперативное запоминающее устройство; ОП — блок оперативной памяти (блок ОЗУ); ПЗУ — постоянное запоминающее устройство; СП — специальный процессор (реализующий нестандартные вычисления); КП — каналный (интерфейсный) процессор; КБ — каналный буфер (для промежуточного хранения вводимых-выводимых данных); НМД — накопители на магнитных дисках; НМЛ — накопители на магнитных лентах; Д — дисплей — устройство ввода-вывода данных с отображением на экране; ПК — устройство ввода-вывода данных на перфокартах; ПЛ — устройство ввода-вывода данных на перфолентах; АЦПУ — алфавитно-цифровое печатающее устройство для вывода данных на бумагу; АПД — аппаратура передачи данных; АП — абонентский пункт

устройства, но ввод из них данных в ЭВМ осуществляется медленнее; устройства ввода-вывода данных<sup>1</sup>; абонентские пункты (видеотерминалы) — комплексы технических средств, включающие приемно-передающие устройства и устройства ввода-вывода и обеспечивающие непосредственную работу с ЭВМ одновременно многих пользователей, удаленных от ЭВМ.

Кроме того, в состав технического обеспечения входят устройства подготовки данных: перфораторы (для занесения данных на перфокарты и перфоленты) и устройства подготовки данных на магнитных носителях, которые могут представлять целые комплексы с собственной небольшой ЭВМ, выполняющей предварительную обработку и контроль за вводимыми данными.

<sup>1</sup> Все носители информации, с которых она вводится в ЭВМ (перфокарты, перфоленты, магнитные диски, магнитные ленты), называются машинными носителями.

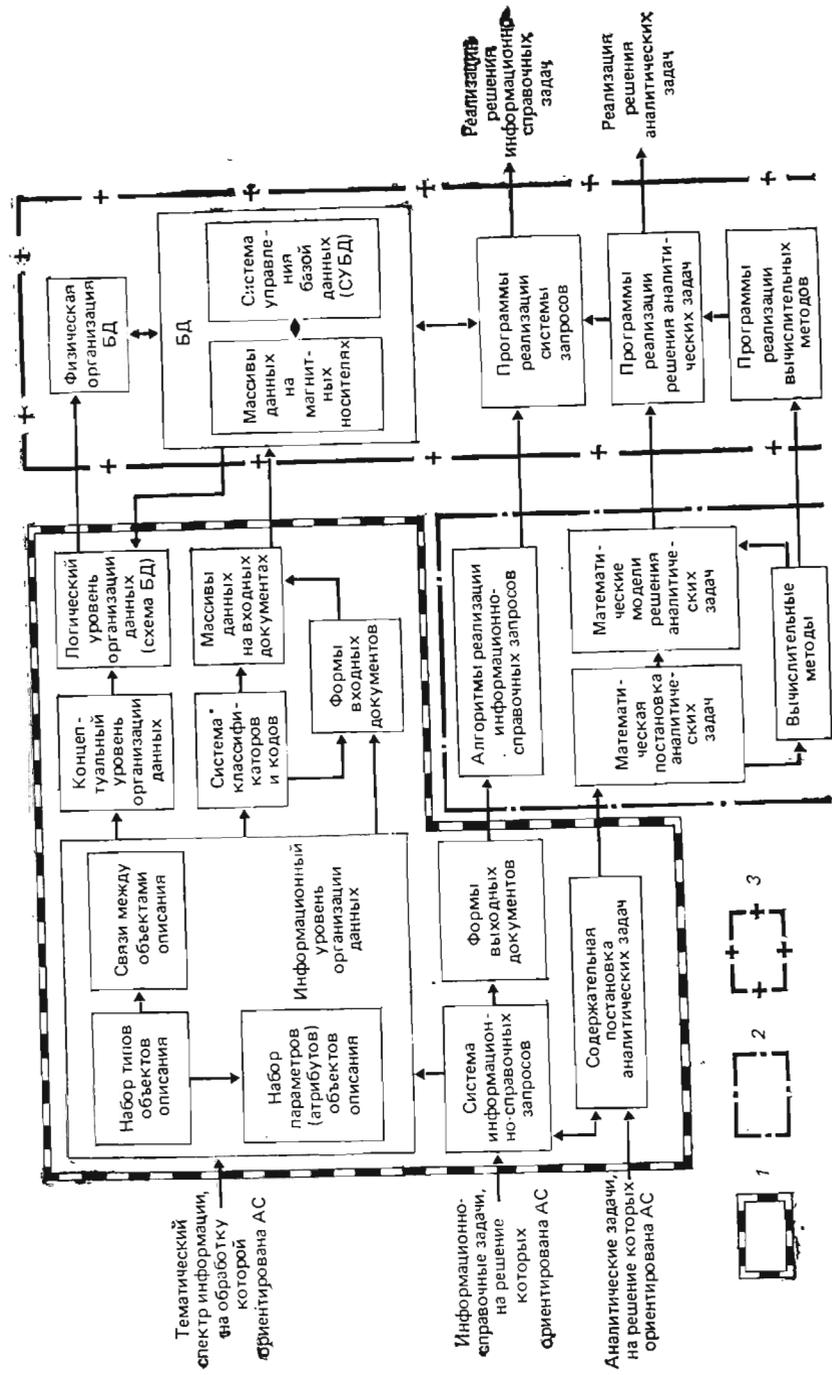


Рис. 27. Структура информационного обеспечения АС и его окружения:  
1 — информационное обеспечение; 2 — математическое обеспечение; 3 — программное обеспечение

Технологическое обеспечение включает набор документов, регламентирующих все элементы процесса функционирования АС и связи этих элементов — инструкции по заполнению входных документов, подготовке данных на машинных носителях, комплектованию АС информацией (периодическая поступление входных документов, порядок их обработки, порядок и способ переноса информации на машинные носители, а также ввода и коррекции ее в ЭВМ), по контролю и коррекции работы всех звеньев АС, способам расширения информационного фонда и программного обеспечения, порядку и способам общения пользователей с АС и выдачи результатов обработки данных.

Организационно-правовое обеспечение состоит в юридическом закреплении норм, которыми руководствуются все специалисты, связанные в своей деятельности с АФИС. Оно включает, в частности, регламентацию избирательного доступа различных пользователей к информации, должностные инструкции и другие материалы, определяющие порядок взаимодействия различных служб информационно-вычислительного центра, нормы и способы материально-технического снабжения и др.

Из материала, рассмотренного в предшествующих разделах данной главы, очевидно, что ЭВМ и автоматизированная система (АС) — не синонимы (первая представляет собой очень важный, но не единственный компонент второй) и что геологическая информация с момента получения ее геологом в результате собственных исследований или изъятия из каких-либо документов до момента, когда геолог нажатием кнопки получит ее на экране дисплея или в виде распечатки АЦПУ, проходит сложный путь. Значительная и очень важная часть этого пути, без которой невозможно применение ЭВМ для обработки геологической информации, приходится на этап, который назван информационным обеспечением АФИС (рис. 27). На этом этапе проводится последовательная формализация геологической информации — ее логическая организация, без которой невозможно ее физическая организация в среде ЭВМ. В разработке информационного обеспечения необходимо участие геолога — потенциального пользователя АФИС.

### 3.1.3. Информационно-технологические аспекты общения пользователя с АС

Принято различать 2 принципиальных режима общения с автоматизированной системой (АС), реализующей решение информационно-справочных или информационно-аналитических задач — *пакетный* и *диалоговый*, или интерактивный. Различие между ними заключается в том, что сама идеология и соответственно технология пакетного режима не предусматривает какого-либо непосредственного взаимодействия между человеком и ЭВМ в процессе решения задачи. Пользователь обычно сдает на информационно-вычислительный центр (ИВЦ) задание и получает результат решения в виде распечатки на АЦПУ. Диалог же человека с АС

подразумевает интерактивный, непосредственный «обмен посланиями между пользователем и АС в соответствии с условными языком диалога и формой диалога для достижения определенной задачи» [12, с. 42]. Язык и форма диалога могут быть весьма разнообразными: от ответа на вопросы в терминах «да — нет» до взаимодействия на ограниченном естественном языке (в том числе с речевым вводом-выводом) и обмена графическими образами (с помощью графического дисплея со световым пером). Но при любом языке и форме диалога решение задачи организовано таким образом, что оно не может быть реализовано без непосредственного включения в процесс решения задачи информации, исходящей от пользователя.

Важно отметить, что диалоговый режим часто отождествляют с человеко-машинной технологией решения задач. Человеко-машинная технология чаще всего базируется на диалоговых средствах, но отождествление их неправомерно, поскольку при этом происходит смешение двух существенно различающихся показателей — «режим (технология) общения с АС (ЭВМ)» и «режим (технология) решения задачи».

Появление человеко-машинной технологии решения задач обусловлено тем обстоятельством, что процедура получения вывода (заключения) на основе обработки соответствующей информации далеко не всегда поддается формализации. Поэтому реализация указанной процедуры не может быть возложена только на ЭВМ. «Понимающую» лишь формализованный язык, а требует органического включения в процесс решения задачи плохоформализуемых опыта и интуиции человека (исследователя). Проиллюстрированные в гл. 2 характерные для геологии неоднозначность понимания предметов, выделения и описания объектов, интерпретации связей между различными предметами и объектами геологии, неполнота описания информационных ситуаций, заставляющая часто принимать решения на субъективной основе личного опыта (например, широко используемый метод геологических аналогий), и другие подобные факторы не позволяют пока формализовать процедуру решения многих как информационно-справочных, так и аналитических геологических задач. Это предопределяет тот большой интерес, который проявляется в геологии к человеко-машинной технологии. Особенно эффективна эта технология в задачах, связанных с получением нового знания, образованием информации, до того отсутствовавшей в системе, что характерно для сложных справочных запросов (см. разд. 3.4.3) и для аналитических задач.

Суть человеко-машинной технологии заключается в том, что происходит как бы разделение труда между человеком (геологом) и ЭВМ. При этом пользователь получает принципиальную возможность воздействовать на процесс удовлетворения своих плохоформализуемых информационных потребностей, т. е. возможность варьировать условиями поиска данных, осуществлять свободный выбор направлений поиска, используя полученную на пре-

дыдущих этапах информацию. Для аналитических задач это позволяет воздействовать на процесс решения, задавая (или изменяя) параметры, выбирая один из альтернативных вариантов решения и (или) вычислительной процедуры (математический метод), принимая решение о завершении задачи (в случае удовлетворяющего результата) или о необходимости дополнительной итерации (повторного решения) с изменением параметров и др. Таким образом, «все наиболее трудоемкие операции формализованного характера возлагаются на ЭВМ, а задачи творческого характера (постановка задачи, разработка принципов выбора способов действия по достижению целей, принятие решений на всех этапах исследований) — на исследователя» [45, с. 24].

Такая технология решения информационно-справочной и аналитической задачи может быть реализована не только в диалоговом, но и в пакетном режиме общения с АС. Ведь процесс решения любой геологической задачи можно разбить на звенья, поддающиеся и не поддающиеся формализации, и построить решение таким образом, что формализуемые звенья реализуются АС в пакетном режиме. Пользователь же получает распечатку, отражающую результат реализации такого звена, обдумывает его и, исходя из этого, выбирает последующий ход решения (задает или изменяет параметры, модифицирует модель решения, выбирает вычислительный метод и др.), поручая АС решение в пакетном режиме следующего звена в соответствии со своим выбором (установкой). Таким образом, многократно общаясь с АС в пакетном режиме, пользователь может реализовать человеко-машинную технологию решения конечных задач в полном функциональном объеме. Все предусмотреть заранее нельзя. Пакетный режим вполне оправдан, если пользователю требуется значительное время на оценку промежуточного результата и принятие решения. В таких случаях для пользователя гораздо предпочтительнее поработать с распечатанными на АЦПУ промежуточными результатами работы АС без спешки, в удобное для него время и в удобном месте, вместо того чтобы спешить и нервничать перед «ждущим» экраном дисплея. С другой стороны, в ряде случаев время, требуемое на осуществление ЭВМ вычислительной процедуры, велико, и пользователь вынужден бездействовать у экрана дисплея, ожидая окончания вычислений, что существенно снижает производительность его труда. С позиции человеко-машинной технологии противопоставление диалогового режима пакетному по функциональным возможностям вряд ли правомерно, а эффективность того и другого зависит в первую очередь от класса автоматизируемых задач.

Основное и принципиальное различие между пакетным и диалоговым режимом, из которого проистекают все другие особенности, определяется интенсивностью обратной связи. Подобно тому, как появление телефона принципиально изменило возможности общения (и соответственно деятельности) людей по сравнению с перепиской, развитие диалоговых средств сыграло поистине

революционизирующую роль в возможностях общения человека с ЭВМ и с АС. Интенсивность обратной связи в режиме диалога по сравнению с пакетным для задачи средней трудоемкости выше примерно на два порядка. Этот количественный скачок вызвал коренные качественные изменения — обеспечил возможность решения в оптимальные (с точки зрения пользователя) сроки множества трудноформализуемых задач, в которых реальная или моделируемая ситуация предполагает осуществляемое пользователем широкое варьирование целевых установок, понятийной основы (предметы и связи), разнообразных параметров задачи. Затраты времени и труда при решении таких задач в пакетном режиме (а соответственно и ощущение ответственности за правдивость принятого решения) столь велики, что не стимулируют обращение пользователя к АС.

В диалоговом режиме этот психологический барьер легко преодолевается, поскольку быстрота взаимодействия позволяет даже путем простого итерационного перебора параметров провести сравнительную оценку получаемых результатов и сделать достаточно обоснованный выбор. В свою очередь, преимущество в интенсивной обратной связи предъявляет повышенные требования к средствам обращения за информацией — языку запросов, который должен обладать такими качествами, как простота, доступность пользователю-специалисту предметной области, а программы, реализующие запросы, должны обеспечивать быстрый ответ. Для аналитических задач это означает гораздо более жесткие требования к модели задачи, к выбору вычислительного метода, которые должны быть такими, чтобы свести к минимуму простоты (ожидания) в диалоге, сводящие на нет преимущества диалогового режима.

С внедрением человеко-машинной технологии, использующей диалоговый режим общения, всю большую остроту приобретают вопросы адаптации АС к требованиям массового пользователя. В последние годы в развитии АС, расширении их функциональных и технологических возможностей достигнуты впечатляющие успехи, которые в принципе позволяют адаптировать АС применительно к требованиям любого пользователя. Тем не менее многие человеко-машинные АС создавались и продолжают создаваться с тем расчетом, что непосредственный доступ к АС в состоянии получить лишь пользователь, имеющий значительные познания в области машинной технологии. Однако, как показывает опыт, на встречном пути адаптации широкого потенциального пользователя к возможностям, предоставляемым такими АС, результаты оказались более чем скромными. Пользователь АС не универсален и не однороден. Специальные исследования этого вопроса, проведенные рядом ведущих зарубежных специалистов по разработке АС [12], привели к появлению нескольких классификаций пользователей. Последние различаются прежде всего по уровню владения машинной технологией — от неподготовленного пользователя до прикладного и системного программиста.

Действительно, если предлагаемая система для реализации информационно-справочной функции требует формирования запроса на данные на одном из универсальных языков программирования (типа PL-1, ФОРТРАН), то естественно, что непосредственный диалог с АС доступен лишь человеку, владеющему знаниями прикладного программиста. То же можно сказать и об аналитической АС, предназначенной для интерпретации геолого-геофизических данных и ориентируемой на возможность, например, «...конструирования пользователем дополнительных вычислительных схем по сравнению с указанными в системе в явном виде...» [33, с. 9]. Для таких АС технология решения задач должна предусматривать наличие посредника между ЭВМ и геологом, что отрицательно сказывается на эффективности как решения, так и внедрения АС.

Большинство действующих и разрабатываемых АС в геологии предъявляют непомерно высокие требования и к уровню подготовки пользователя в его профессиональной (геологической) области. Так, в одной из основополагающих работ по методологии создания АС в геологии ставится вопрос о том, что АС должна предоставлять пользователю возможность «уточнить или сформировать заново модель, включая определение понятий, в том числе и тех, которые не имеют явного выражения в системе, но являются результатом композиции уже заданных элементов модели» [33, с. 11]. Очевидно, что основная масса геологов, обращающихся к АС главным образом с целью использовать предлагаемые системой уже готовые модели решения уже конкретных задач, не в состоянии воспользоваться такой возможностью.

Влияние профессионального геологического опыта на эффективность использования человеко-машинных АС хорошо видна на примере АС, ориентированных на обработку картографической (координатно-привязанной) геологической информации с целью прогноза положения поисковых объектов, в частности рудных и нефтяных месторождений [38, 45]. Типовая ситуация здесь заключается в том, что геолог в режиме диалога с АС указывает поисковые признаки прогнозируемого объекта, в соответствии с которыми АС выдает карту с очерченными участками предполагаемого положения объектов данного типа. Методологическая основа задания комплекса поисковых признаков (особенно в слабоизученных районах) — метод геологических аналогий. Такой подход требует от геолога детального знания большого числа возможных конкретных аналогов, из которых должен быть выбран наиболее адекватный данной геологической ситуации. Этот выбор, предопределяющий геологическую эффективность (достоверность) прогноза, целиком зависит от эрудиции и опыта геолога.

Естественно, что возможности общения различных групп пользователей с АС существенно различны, и реальное человеко-машинное взаимодействие возможно только, если средства и форма общения, предоставляемые АС, соответствуют уровню того или иного пользователя. Быстро расширяющееся в последние годы внедрение математических методов и ЭВМ в научную и практи-

ческую геологическую деятельность, разработка средств общения с АС на хорошо понятном для геолога профессиональном языке посредством ставших привычными экрана и клавиатуры пишущей машинки с отображением на телеэкране способствуют преодолению психологического барьера, мешавшего геологу воспринимать работу с АС как элемент технологии собственно геологической деятельности. Сейчас мы находимся на начальной стадии процесса приобщения массового геолога к использованию в своей работе проблемно-ориентированных АС. От того, насколько такие АС будут отвечать требованиям геологов, не имеющих специальных знаний в области компьютерной технологии (их принято называть конечными пользователями), зависит интенсивность и эффективность внедрения ММ и ЭВМ в геологии. Типовые характеристики конечных пользователей, проанализированные в работе [12], сводятся в основном к следующему:

а) АС нужна им для выполнения специфических задач, и они будут оценивать ее в соответствии с ее пригодностью (простотой) для обслуживания этих задач;

б) у них нет знаний и навыков по работе с ЭВМ вообще и с данной АС в частности, и они не стремятся к получению таких знаний, пытаются минимизировать необходимое обучение;

в) они предпочитают не проявлять собственную инициативу, а пассивно подчиняться тому, чтобы их вели через систему.

С учетом вышеизложенного представляется принципиально важным, чтобы на выходе человеко-машинную АС отличали следующие свойства:

четкая проблемная ориентированность (дифференциация конечных пользователей по специализации);

системно-управляемый диалог, в котором предложения по стратегии достижения цели исходят от АС, а по тактике — от конечного пользователя<sup>1</sup> (многоходовой пакетный режим решения также рассматривается как диалог, но крайне замедленный);

минимизация обучения за счет использования профессионального языка соответствующей (геологической) предметной области;

простота структуры решаемых задач, снижающая требования к квалификации конечного пользователя и «простота» в диалоге. Простота возникает из-за сложности быстрой оценки промежуточных ответов ЭВМ или длительности вычислительных процедур.

Такая ориентация АС на потребности конечного пользователя вовсе не означает каких-либо ограничений для более подготовленных пользователей. Для этого человеко-машинная АС должна иметь несколько уровней и соответственно несколько точек входа в АС. Представляется оптимальной трехуровневая структура АС, где кроме уровня конечного пользователя были бы структурно обособлены уровни разработчика задачи и разработчика собст-

<sup>1</sup> Взгляд на жестко фиксированный диалог как на нечто тривиальное при современных возможностях АС, отражает недостаточный учет специфики массового конечного пользователя и потому представляется неправомерным.

венно АС (программно-математического обеспечения). При таком подходе каждому такому уровню отвечали бы определенные группы пользователей. На уровне разработчика задачи (комплекс задач) пользователем является обычно геолог-специалист по методологии и методике конкретной области геологии, например, оценка ресурсов, методика поисков и разведка месторождений того или иного вида минерального сырья, экономика геологоразведочных работ и др. Он, как правило, в кооперации с системным аналитиком (для информационно-справочных задач) или прикладным математиком формирует стратегию решения геологических задач, в частности осуществляет содержательную постановку системы запросов или аналитических задач, разработку для последних математических моделей и выбор вычислительного метода, определение формы таблиц исходной информации и представление результатов решения. На этом уровне должны быть предоставлены широкие возможности создавать и многократно модифицировать стратегию решения задачи применительно к различным информационным потребностям и ситуациям. Единственное накладываемое ограничение заключается в том, что все разработки должны быть встроены в общую структуру задач таким образом, чтобы были соблюдены сформулированные выше требования к АС применительно к возможностям, предоставляемым конечному пользователю.

Интересы и возможности прикладного (проблемного) и системного программистов должны реализовываться на уровне разработчика собственно АС. Предоставляемые им на этом уровне средства должны позволять оперативно оптимизировать вычислительную процедуру и физическую организацию данных в соответствии с изменениями в содержательном и математическом описании задач, легко адаптировать систему применительно к постоянному расширению круга встраиваемых и решаемых задач, проводить диагностику АС для оценки эффективности организации данных в АС.

### 3.2. ОСНОВЫ ОРГАНИЗАЦИИ БАЗ ГЕОЛОГИЧЕСКИХ ДАННЫХ

На первых этапах применения ЭВМ в организации данных для автоматизированной обработки преобладал подход, который условно можно назвать позадачным. Он состоит в том, что для каждого приложения<sup>1</sup> создаются свои массивы информации. Из всего множества сведений о предметной области выделяется необходимая конкретному приложению совокупность, т. е. взгляд на информационное обеспечение определяется и ограничивается конкретной задачей (рис. 28). Такой подход характерен тем, что организацией данных занимаются прикладные программисты, не привлекая специалистов в данной предметной области, например геоло-

<sup>1</sup> Под приложением массива информации (массива данных) понимается решение на его основе отдельной справочной или аналитической задачи.

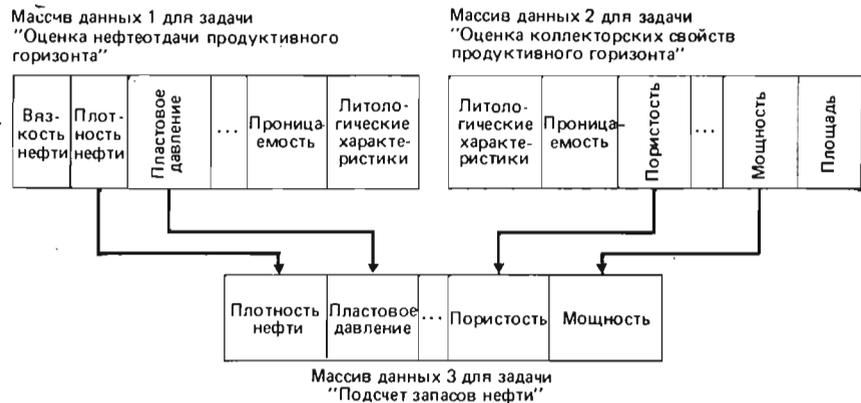


Рис. 28. Организация исходных данных при позадачном подходе

гов. Программисты сами выбирают вид и способ представления данных, причем описание массивов информации входит в тексты самих программ обработки (решения задач), т. е. программы связаны с данными. Это значит, что если организация данных каким-либо образом изменится, необходимо соответствующим образом изменить программы. И наоборот, если изменяются программы, то часто нужно по-иному организовывать данные, т. е. уже введенные данные либо заново подготавливать и вводить, создавая модифицированные массивы, либо разрабатывать специальные программы перекачки (переорганизации) данных из старых массивов в новые. Как ни старались разработчики учесть возможные изменения требований к обработке, им это не удавалось. По мере накопления опыта в создании и эксплуатации соответствующих систем было установлено, что требования к обработке изменяются случайным, непредсказуемым заранее образом [25].

Более того, при наличии нескольких приложений используемые ими массивы содержали, наряду со специализированными, общие для них данные. Хранение одних и тех же элементов данных в различных массивах приводит, во-первых, к неэффективному использованию ресурсов ЭВМ и, во-вторых, к дублированию процесса ввода и к необходимости параллельной корректировки в разных массивах. Это создает предпосылки к образованию противоречивой информации о предметной области, в связи с чем результаты, полученные приложениями по одним и тем же данным, но из скорректированных и нескорректированных массивов, будут не соответствовать друг другу.

Проиллюстрировать последнее можно на массивах 1 и 2 (см. рис. 28), сформированных соответственно для задач «Оценка нефтеотдачи продуктивного горизонта» и «Оценка коллекторских свойств продуктивного горизонта». Они содержат общие элементы данных: проницаемость и литологические характеристики, которые в процессе доразведки могут изменять свои значения. Если

не обеспечена параллельная корректировка данных в этих массивах (т. е. значения литологических характеристик и проницаемости в одном массиве будут изменены, а в другом останутся без изменений или будут изменены ошибочно), то это приведет к тому, что использование результатов решения задач 1 и 2 для задачи 3 исказит результат решения последней, так как при решении задач 1 и 2 в расчет принимались разные значения одних и тех же данных.

В рамках позадачного подхода затруднено решение задач, для которых необходима совместная обработка сведений разного характера, содержащихся в изолированных, неодинаково организованных массивах. Например, практически неосуществимо автоматизированное решение задачи «Подсчет запасов нефти» (см. рис. 28), так как для этого требуется образование массива данных 3 (необходимые данные отмечены стрелками) из данных массивов 1 и 2, по-разному сформированных прикладными программистами для каждого из частных приложений.

Отмеченные недостатки позадачного подхода и, кроме того, общность основных процедур информационного обеспечения АС привели к централизованному подходу в организации данных, оформившемуся для фактографических систем в концепцию баз данных (БД). Она характеризуется тенденцией рассмотрения данных отдельно от приложений на уровне упоминавшихся информационных процедур: накопления, хранения, коррекции и выборки информации.

### 3.2.1. Концепция базы данных

Если попытаться обобщить все определения, то под базой данных (БД) понимают, во-первых, специальным образом логически связанные между собой данные о предметной области, записанные на машинные носители информации (магнитные ленты, магнитные диски и др.), и, во-вторых, пакет прикладных программ, называемый системой управления базой данных (СУБД) и обеспечивающий накопление, хранение, коррекцию, поиск и выдачу информации. Требования к организации БД следующие:

БД должна быть интегрированной, т. е. содержать данные для многих приложений и пользователей;

данные не должны зависеть от их использования конкретными приложениями, т. е. должны допускать модифицирование старых и построение новых приложений;

БД должна легко модифицироваться при реорганизации и расширении хранимых сведений о предметной области;

прикладные программы (программы решения задач) не должны изменяться при модификации БД;

доступ к данным должны иметь только лица с соответствующими полномочиями.

Концепция предполагает централизованное управление всеми обрабатываемыми данными предметной области, а именно: все

данные концентрируются в одном универсальном хранилище с общими универсальными обслуживающими механизмами. При этом устраняется избыточность<sup>1</sup> данных (дублирование), характерная для позадачного подхода.

При централизованном управлении эффективней используются машинные носители информации; устраняется противоречивость (несовместимость) хранимых данных за счет невозможности асинхронных ввода и коррекции данных; осуществляется однородность ввода общих для разных задач данных, т. е. оптимизируется один из самых трудоемких процессов автоматизированной технологии; данные могут быть совместно используемы, т. е. допускается рассмотрение любых из общностей; создаются предпосылки для стандартизации, необходимой при обмене данными между разными АС; централизованно контролируется доступ к данным.

Одним из главных достоинств концепции БД является то, что этот подход сделал возможным разработку пакетов прикладных программ, составляющих системы управления базами данных (СУБД), которые могут использоваться для создания АФИС в самых различных предметных областях.

При реализации концепции БД решаются две основные системные задачи. Первая заключается в обеспечении возможности запрашивать и отыскивать информацию без трудоемкого написания программ на обычном (универсальном) языке программирования. Решение этой задачи основано на том, что система управления базой данных оперирует специальным языком более высокого уровня, чем обычный универсальный язык программирования. В современных СУБД он приближается к естественному. Тенденция создания языков общения с ЭВМ, близких к естественному и ориентированных на пользователя — не программиста, дает возможность конечному пользователю — специалисту в той или иной предметной области непосредственно, без участия программиста общаться с автоматизированной системой — самому формулировать запросы и получать ответы.

Вторая системная задача состоит в обеспечении независимости данных от использующих их программ, с тем чтобы можно было, с одной стороны, добавлять и перестраивать данные, не переделывая прикладных программ, и, с другой — модифицировать и разрабатывать новые прикладные программы без реорганизации всей БД, а только с добавлением данных.

Независимость данных и программ их обработки является принципиальным моментом в концепции БД, так как одно из наиболее важных свойств БД состоит в возможности ее постоянного изменения (дополнения, коррекции, удаления данных разного типа). Для обеспечения независимости необходим аппарат, позволяющий обращаться к данным без указания в программах характеристик

<sup>1</sup> Точнее, излишняя избыточность, так как для улучшения некоторых функциональных характеристик БД «минимальная или управляемая» избыточность может присутствовать.

их представления в ЭВМ (машинных характеристик). Этим аппаратом и является логический уровень представления (логическая модель) данных (см. рис. 27). При таком взаимоотношении данных и программ изменение БД не повлияет на обрабатываемые программы. Изменится лишь логическая модель данных. И наоборот, модификация старых и включение новых приложений (задач и прикладных программ) не приведет к реорганизации всей БД, поскольку формирование совокупности данных для этих приложений осуществляется СУБД на основе хранимого определения данных.

Обеспечение независимости данных и прикладных программ резко повышает возможности программиста в удовлетворении требований пользователя-геолога к накоплению, хранению, коррекции данных, расширению БД, а также к решению информационно-справочных и аналитических задач.

### 3.2.2. Средства моделирования системы объектов предметной области в базе данных

В основе логической организации данных лежит модельный принцип: данные организуются в соответствии не с потребностями отдельных приложений, а с общими представлениями о предметной области как о ее модели. Именно это обеспечивает принципиальную возможность использования одних и тех же данных в разных задачах и реализацию задач комплексного анализа данных. В общем случае модель предметной, в частности геологической, области прямо или опосредованно базируется на понятиях объекта описания, его характеристик (атрибутов) и связей между объектами. Ниже рассмотрены принципы, лежащие в основе моделирования предметной области в базе данных. При этом используются различные подходы к логической организации данных с последовательным возрастанием степени формализации последних, что обуславливает определенную типизацию моделей предметной области (моделей данных).

По типу структуры данных модели подразделяются на иерархические, или древовидные (теоретико-графовые), сетевые (теоретико-графовые), реляционные (теоретико-множественные); по уровню организации данных — на информационные, концептуальные, логические (схемы БД).

Далее в соответствующих разделах приведены более подробные описания типов моделей данных в контексте интеграции геологической информации в АФИС.

В рамках концепции БД сосуществуют два уровня видения объектов. Во-первых, это объекты предметной области (ОПО) — геологические объекты, причем их выделение во многом определяется локальными задачами, решаемыми тем или иным специалистом-геологом. Так, петрограф в качестве объекта описания может выделить «шлиф», литолог — «породу», геолог-нефтяник оперирует такими объектами, как «ловушка», «коллектор», «по-

крышка», «залежь», «месторождение нефти и газа», «нефтегазоносный бассейн» и др., а геолог-рудник — объектами «рудное тело», «магматическое тело», «рудная зона», «металлогеническая провинция» и т. п. Это значит, что в геологической предметной области может быть выделено большое число весьма разнообразных и разномасштабных объектов (см. гл. 1). Каждый из них характеризуется своим набором характеристик (атрибутов), которые могут быть как уникальными, так и повторяющимися для разных типов ОПО. Например, геохимическими показателями могут характеризоваться объекты «порода», «коллектор», «покрышка»; показатели запасов полезного ископаемого могут быть характеристиками объектов «залежь», «месторождение», «страна» и т. д.

Во-вторых, это объекты базы данных (ОБД), под которыми понимаются рассматриваемые как единое целое совокупности атрибутов<sup>1</sup>, являющихся характеристиками некоторых аспектов объектов предметной области. ОБД не адекватны объектам реального мира (ОПО), поскольку отражают только те аспекты последних, которые находятся в компетенции данной АС. Например, такой объект, как «месторождение», в АС по ресурсам полезных ископаемых будет представлен характеристиками динамики запасов и добычи полезного ископаемого, а в АС по геологии месторождений — характеристиками геотектонического положения, разреза, соотношения залежей и вмещающих пород и т. д. Но в том и другом случае совокупность атрибутов (ОБД) «месторождение» не будет полностью адекватна набору характеристик месторождения как объекта предметной области (геологического объекта).

Более того, ОБД не обязательно должен однозначно соответствовать традиционному ОПО. Например, тот же геологический объект — «месторождение» может быть представлен в БД совокупностью таких ОБД, как «ловушка», «залежь», «разрез», «природный резервуар» и др., каждый из которых отражает свою тематически связанную совокупность атрибутов.

С формальной точки зрения совокупность может состоять из одного атрибута, т. е. каждый элемент данных<sup>2</sup> можно рассматривать как отображение некоторого абстрактного объекта. Однако на практике выделяют ОБД, представляющие набор характеристик, которые принято рассматривать во взаимосвязи, как совокупно характеризующие тот или иной традиционно анализируемый аспект объекта предметной области.

Отдельный элемент данных или даже отдельный ОБД малоинформативен и часто вообще не имеет смысла, если не соотнесен с другим ОБД. Например, описание породы-покрышки ничего не дает, если его нельзя соотнести с описанием ловушки, а последнюю с месторождением. Поэтому данные, зафиксированные в БД,

требуют совместного рассмотрения. Для этого в БД должны быть представлены связи, являющиеся отражением принятых соответствий объектов реального мира. Механизм представления связей в БД базируется на выделении следующих типов соответствий:

«один к одному», когда одному объекту соответствует только один объект (одна ловушка — одна залежь углеводородов);

«один ко многим», когда одному объекту соответствует несколько объектов (одна металлогеническая провинция — много месторождений, один природный резервуар — много ловушек углеводородов);

«многие к одному», когда нескольким объектам соответствует один (является обратным к соответствию «один ко многим»);

«многие ко многим», когда одному объекту ставится в соответствие несколько объектов и, кроме того, он сам может соответствовать нескольким объектам одновременно. Это наиболее общий случай соответствий в геологии (один химический элемент в составе многих минералов и, наоборот, в составе одного минерала многие химические элементы; один нефтегазоносный бассейн может распространяться на многие страны, и, наоборот, в пределах одной страны могут располагаться многие нефтегазоносные бассейны).

В литературе по базам данных для обозначения типов соответствий приняты следующие удобные обозначения: 1 : 1 (один к одному), 1 : М (один ко многим), М : 1 (многие к одному) и М : М (многие ко многим). Типы 1 : М и М : 1 отличаются направлением связей, т. е. направлением перехода от объекта к объекту при проходе по их структуре в БД. Необходимость выделения обоих этих типов обусловлена тем, что в большинстве СУБД поддерживается переход лишь в одном направлении. В СУБД, обеспечивающих переход в обоих направлениях, характеристика направленности связей не используется и для них М : 1 и 1 : М равнозначны.

В рамках каждого из этих типов между одними и теми же объектами могут существовать различные зависимости. Например, два пласта могут соотноситься между собой как элементы ловушки углеводородов (коллектор-покрышка), но между теми же пластами имеется и временная зависимость (подстилающий и покрывающий). Использование различных зависимостей при обработке данных значительно расширяет сферу приложений, однако для этого необходимо, чтобы СУБД имела аппарат именованной разновидности однотипных связей.

Связи могут устанавливаться как между объектами разных типов, так и между однотипными объектами, но не все СУБД обеспечивают эту возможность. Это следует учитывать и геологу при оценке возможностей конкретной АС (использующей конкретную СУБД), и разработчику АС при выборе СУБД в соответствии с требованиями геологов — постановщиков задач и потенциальных пользователей.

<sup>1</sup> В литературе по БД они называются записями об объекте.

<sup>2</sup> Элемент данных, рассматриваемый в качестве характеристики какого-либо объекта, называется атрибутом этого объекта.

### 3.2.3. Типы структур данных

Организация данных при моделировании предметной области базируется на одном из трех подходов к их структуризации (см. разд. 3.2.2). Поскольку каждая БД и СУБД ориентирована на определенный подход, то типы структур данных БД и СУБД совпадают и могут быть выделены иерархические (древовидные), сетевые и реляционные структуры данных.

Первые два подхода сравнимы по методу представления структур, так как имеют общий формализм представления — графы, в третьем подходе используется теоретико-множественный аппарат математических отношений. В соответствии с этим выделяются два класса моделей данных: теоретико-графовые и теоретико-множественные [7].

В теоретико-графовых моделях вершинами графа представляются либо объекты, либо элементы данных, ребрами графа — их взаимосоответствия (связи). В зависимости от ограничений на вид графа, выделяются иерархические и сетевые структуры.

Иерархические (древовидные) структуры представляются древовидным графом (рис. 29, а), отражающим иерархию (соподчиненность) элементов, соответствующих узлам графа. На самом верхнем уровне иерархии имеется только один узел — корень («нефтегазоносный бассейн»). Каждый другой узел (кроме корня) связан только с одним узлом на более высоком уровне, называемым исходным для данного узла, который сам называется порожденным. Например, узел «месторождение» связан на более высоком уровне только с узлом «нефтегазоносный бассейн», который служит для него исходным. В то же время узел «месторождение»

является порожденным по отношению к узлу «нефтегазоносный бассейн».

Данные исходного узла — общие для всех порожденных узлов, что соответствует свойственной иерархической структуре связи типа «один ко многим». Например (см. рис. 29, б), данные, характеризующие то или иное месторождение в целом (геотектоническое положение, метод подготовки площади под бурение и т. д.), являются общими для всех залежей этого месторождения, что отражает иерархические отношения между объектами (на одном месторождении — три, две, одна залежи, но невозможно, чтобы одна залежь распространялась более чем на одно месторождение).

Принципиально для иерархического подхода то, что каждая совокупность данных, отображаемых порожденным узлом, не может существовать без совокупности, предшествующей ей в иерархии, ибо изолированные узлы не допускаются.

Этот подход к организации данных наиболее перспективен для предметных областей, в которых существуют общепринятые, устоявшиеся иерархические классификации объектов. В приведенном на рис. 29 примере иерархия объектов предметной области (НГБ → месторождение → залежь) соблюдена. Однако любое нарушение иерархии ОПО, естественно, затрудняет использование иерархических моделей, так как они накладывают достаточно жесткие ограничения на типы связей между объектами, не разрешая связи типа М : М.

Если же в указанную выше группу объектов требуется ввести объект «природный резервуар» (рис. 30), то иерархический подход становится неэффективным. На одном месторождении или разведочной площади может быть несколько природных резервуаров (ПР), но в то же время один ПР может распространяться на несколько месторождений и разведочных площадей, т. е. объекты «месторождение — ПР», и «разведочная площадь — ПР» находятся в отношении М : М, а такие связи невозможны в рамках иерархической структуры. В этом случае неадекватное применение иерархической модели ведет либо к избыточности (дублированию) самих данных, либо к избыточности описания данных. Первое происходит, когда объект (природный резервуар, см. рис. 30, а) выступает как порожденный для разных исходных объектов (месторождений, разведочных площадей), что требует дублирования атрибутов этого порожденного объекта столько раз, сколько исходных он имеет. Например, если использовать иерархическую модель (рис. 31, а) для организации данных по объектам «месторождение — ПР», «разведочная площадь — ПР», то элементы данных, характеризующие ПР в целом (имя ПР, стратиграфическое положение, фациальная приуроченность, основные типы слагающих пород и др.), приходится дублировать столько раз, на скольких месторождениях выделен этот природный резервуар. Так, в примере на рис. 31, б общие данные дублируются по ярактинскому ПР 3 раза (на Ярактинском и Марковском месторождениях и Погаповской площади), по парфеновскому ПР и осинскому ПР —

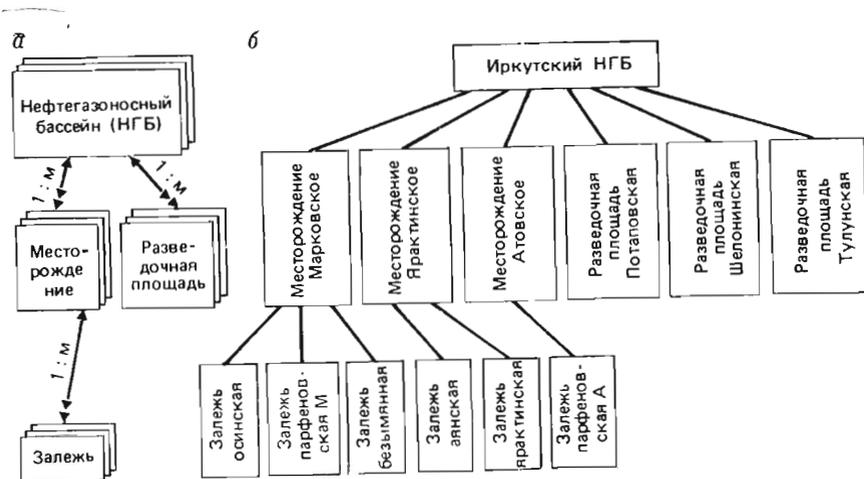


Рис. 29. Иерархическая (древовидная) структура нефтегеологических объектов: а — схема; б — экземпляр схемы

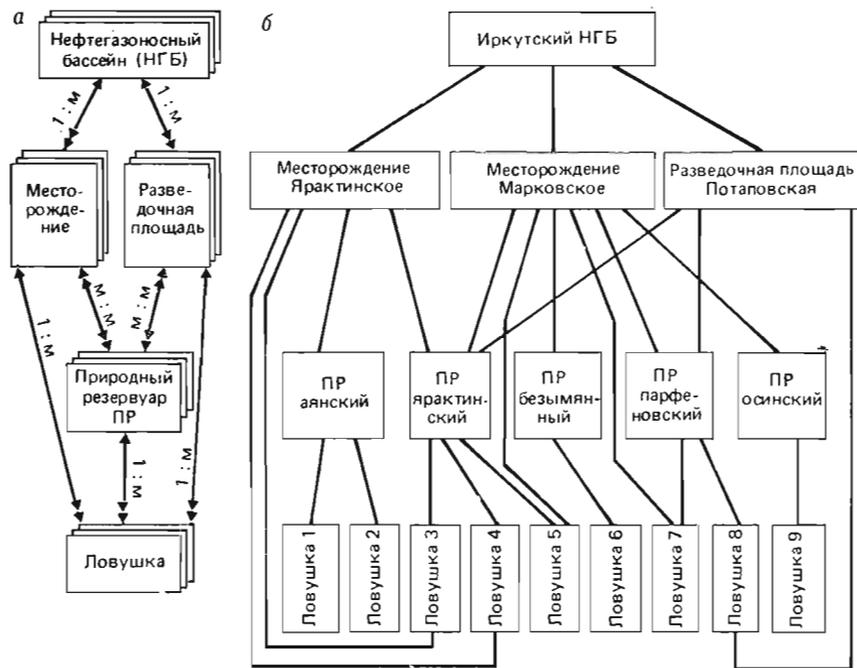


Рис. 30. Сетевая структура нефтегеологических объектов:  
а — схема; б — экземпляр схемы

по 2 раза (на Марковском месторождении и Потаповской площади).

Избыточность описания данных возникает тогда, когда однотипные (однородные по совокупности атрибутов) объекты имеют связь с разнотипными объектами более высокого уровня. В таких случаях, чтобы соблюсти жесткие требования иерархии (для каждого объекта не может быть нескольких исходных одновременно), однотипные по сути объекты приходится искусственно относить к разным типам и представлять разными узлами, а соответственно и дублировать описание (названия атрибутов и ряд других характеристик). В качестве примера можно привести такой тип объекта, как скважина, которая может быть связана с месторождением и через него с нефтегазоносным бассейном, но может быть связана и непосредственно с бассейном (если она пробурена вне пределов месторождения). Соблюсти требования (ограничения) иерархической структуры данных можно, если тип объекта «скважина» заменить двумя разными типами объектов, например «скважина на месторождении» и «скважина вне месторождения», продублировав для их описания одну и ту же совокупность атрибутов, что и вызывает избыточность описания.

Иерархическое представление, как правило, предопределяет и жесткую иерархическую структуру доступа, т. е. возможность

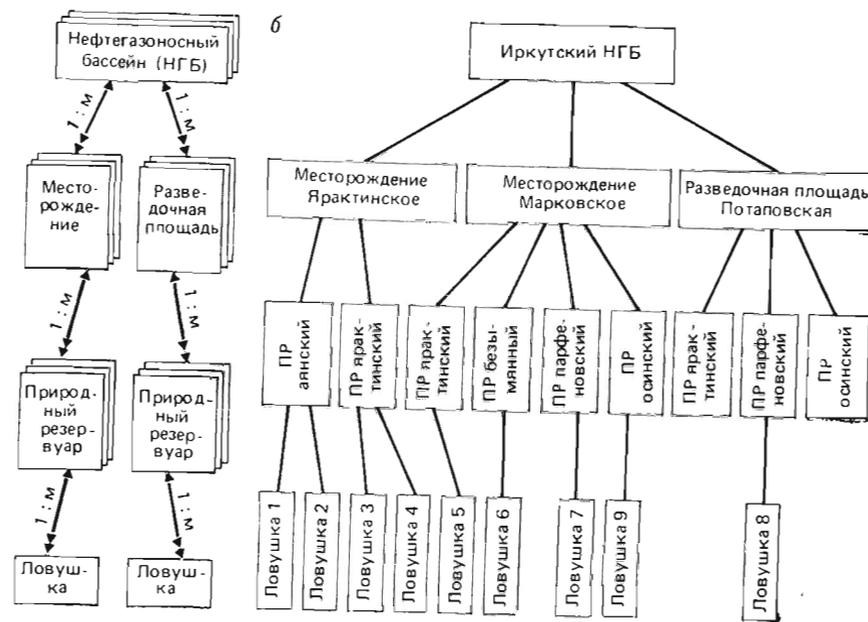


Рис. 31. Пример представления сетевой структуры объектов, приведенной на рис. 30, в виде иерархической с избыточностью:

а — схема; б — экземпляр схемы

перехода по структуре объектов в БД только в рамках одной ветви. Переход от одной ветви иерархии к другой достаточно не эффективен, поскольку требует возвращения на общий исходный узел. Например, если в древовидной структуре объектов с двумя ветвями «бассейн→месторождение нефти→залежь» и «бассейн→природный резервуар→коллектор→порода» необходимо получить для некоторой залежи характеристики коллектора, то доступ к этим данным в базе данных возможен лишь при проходе по первой ветви до общего исходного узла («бассейн»), а затем по второй ветви к объекту «коллектор». Поэтому, чтобы как-то снять отмеченную проблему неэффективности поиска при сохранении связи типа «один ко многим», приходится в описании объектов, выстраиваемых в одну простую ветвь иерархии, включать несвойственные им характеристики, которые логичнее включать в другие ветви (например, пользуясь вышеприведенным примером, в описание залежи — характеристики коллектора).

Проблемы, обусловленные ограничениями иерархического подхода, частично удастся разрешить применением сетевых моделей.

Сетевые структуры (см. рис. 30) по сравнению с иерархическими более общие и допускают более свободную структуру связей: отдельный узел может иметь произвольное число не только порожденных узлов, но и непосредственно исходных (например,

на рис. 30, а узел «природный резервуар» имеет два исходных узла: «месторождение» и «разведочная площадь»). Это дает возможность прямого представления соответствия типа «многие ко многим», часто встречающегося в геологической предметной области. Моделируется это соответствие связями «многие ко многим», либо связями «один ко многим» в обоих направлениях (например, на рис. 29, а связи М : М могут быть заменены связями 1 : М в двух направлениях), что невозможно в рамках иерархической структуры, запрещающей узлы, являющиеся одновременно исходными и порожденными относительно друг друга.

Сетевые структуры представляют значительно большие возможности моделирования геологической предметной области. Во-первых, их применением частично снимается проблема дублирования данных (на рис. 30, б видно, что общие характеристики каждого природного резервуара описываются один раз, а связями указывается, к каким месторождениям они относятся). Следует отметить, что сетевую модель можно представить в иерархическом виде, допустив избыточность данных за счет дублирования порожденных узлов, имеющих несколько исходных. Это наглядно видно при сравнении рис. 30 и 31.

Во-вторых, с помощью сетевых структур легко представим объекты пересечения нескольких иерархических классификаций. Например, пересечение иерархий геологических объектов (бассейн — месторождение — ...) и политико-административных объектов (республика — область — ...) представляется, скажем, объектом «часть бассейна в пределах республики» и т. д.

В-третьих, повышается эффективность доступа к данным, поскольку возможности задания связей между объектами и их атрибутами (см. рис. 30, б) не ограничены.

Однако отсутствие ограничений на представление связей, открывая широкие изобразительные возможности, часто приводит к усложнению структуры данных, что делает ее трудновоспринимаемой [1]. На рис. 32 приведен пример усложнения структуры, обусловленного увеличением числа объектов, по сравнению со структурой, отраженной на рис. 30. Если к этим объектам добавить объекты политико-административного и макроландшафтного (территория, акватория) районирования, то возникают структуры, в которых очень трудно разобраться. А поиск информации по запросам геологов — пользователей АС требует детального знания логической структуры базы данных. Сложность этой структуры, естественно, вызывает трудности в реализации запросов. В настоящее время еще нет основанных на сетевой модели АС, эффективно реализующих непредусмотренные заранее запросы.

Следует отметить также, что как сетевые, так и иерархические модели рассчитаны на представление бинарных отношений узлов, т. е. отображают соответствия только между двумя объектами. На практике же может возникнуть необходимость описания отношений большей степени. Например, при описании условий локализации оруденения для указания соответствия между рудным

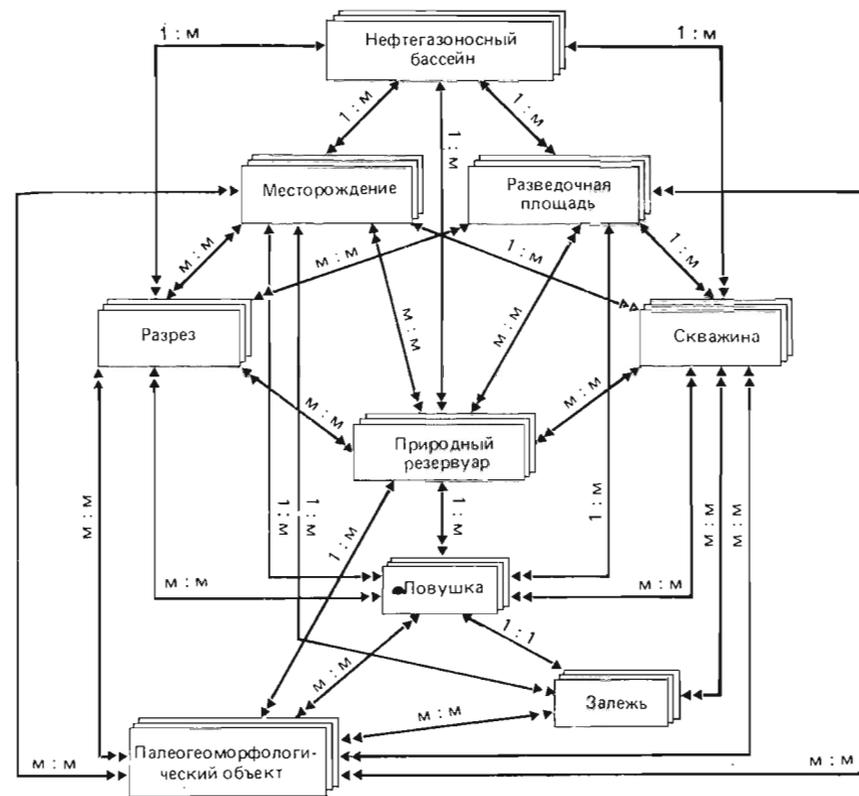


Рис. 32. Пример усложнения структуры нефтегеологических объектов при увеличении их числа (по сравнению с рис. 30).

телом и вмещающими породами необходимо использовать отношение вида «находится в окружении», степень которого может быть больше двух (т. е. во взаимосвязи может находиться больше двух объектов). Разложение этого отношения на совокупность бинарных (для применения иерархической или сетевой модели) не всегда возможно, если имеющаяся информация недостаточна. Кроме того, для сетевых и иерархических моделей не характерна типичная для геологии табличная форма представления данных, в связи с чем обработка таких данных значительно усложняется (требуется введение специальных системных механизмов и программ обработки). Все это обуславливает возрастающий интерес к реляционному подходу при структуризации геологических данных [36].

Реляционные структуры. Реляционный подход основан на математической теории отношений. Главное его отличие от сетевого и иерархического подходов состоит в том, что первичен не объект предметной области, отображаемый совокупностью раз-

Отношение 1: совокупности элементов данных «Характеристика месторождения»

Месторожде- ние (разве- дочная пло- щадь)	Геотектоническое положение	Категория крупности	Имя элемента данных							Классы ловушек нефти и газа		
			Число зале- жей общее	Число зале- жей нефтяных	Число зале- жей газонф- тяных	Число зале- жей газокон- денсатных	...	Число про- дуктивных горизонтов	Возраст про- дуктивных горизонтов		Верхняя гра- ница интерва- ла залегания продуктивных горизонтов, м	Нижняя гра- ница интерва- ла залегания продуктивных горизонтов, м
Значение элемента данных												
Ярактинское	Антеклиза докемб- рийской платформы	Средняя	4	4	—	—	...	2	V, C	3100	3150	Литологические
Марковское	Антеклиза докемб- рийской платформы	Мелкая	3	1	1	1	...	3	V, C	3000	3350	Литологические
Потаповская	Антеклиза докемб- рийской платформы	—	—	—	—	—	...	1	V	3250	3280	Литологические

Таблица 7

Отношение 2: совокупности элементов данных «Характеристика природного резервуара»

Месторожде- ние (разве- дочная пло- щадь)	Природный резервуар	Типы пород коллектора	Имя элемента данных							Типы пород покрывающих отложений	Типы пород латерального экрана
			Мощность коллектора, макс., м	Мощность коллектора мин., м	Пористость коллектора открытая макс., %	Пористость коллектора открытая мин., %	...	Генезис кол- лектора по обстояткам осаждения	Генезис кол- лектора по постседимен- тационным факторам		
Значение элемента данных											
Ярактинское	Аянский	Песчаник	5	0	10	1	...	Мелкий шельф	—	Аргиллит, доломит	Аргиллит, доломит
Ярактинское	Ярактинский	Песчаник	40	0	30	2	...	Дельтовая	—	Доломит, аргиллит	Аргиллит, алевролит, песчаник
Марковское	Ярактинский	Песчаник	20	3	20	1	...	Глубокий шельф	Трещиноватость	Аргиллит, алевролит	Аргиллит, алевролит
Марковское	Парфеновский	Песчаник	25	0	30	1	...	Мелкий шельф	—	Доломит	Аргиллит, алевролит, песчаник
Марковское	Осинский	Известняк	30	10	20	2	...	Мелкий шельф	Трещиноватость	Соль	Известняк, доломит
Потаповская	Ярактинский	Песчаник	15	2	15	1	...	Глубокий шельф	—	Аргиллит, алевролит	Аргиллит, алевролит
Потаповская	Парфеновский	Песчаник	20	0	18	1	...	Мелкий шельф	Трещиноватость	Аргиллит, алевролит, доломит	Аргиллит, алевролит, песчаник

Таблица 8

Отношение 3: совокупности элементов данных «Характеристика ловушки»

Месторождение (разведочная площадь)	Природный резервуар	Ло- вушка	Степень латерального экранирования	Имя элемента данных			Морфология кровли коллектора	Происхождение морфологической аномалии кровли коллектора	Типы палеогеомор- фологических объектов
				...	Происхождение латерального экрана	...			
Значение элемента данных									
Ярактинское	Аянский	Л-1	Всестороннее	...	Выклинивание	Моноклиаль	Тектоническое	—	
Ярактинское	Аянский	Л-2	Всестороннее	...	Выклинивание	Моноклиаль	Тектоническое	—	
Ярактинское	Ярактинский	Л-3	Всестороннее	...	Фаціальное замещение, постседиментационное из- менение	Моноклиаль	Тектоническое	Дельтовая протока	
Ярактинское	Ярактинский	Л-4	Всестороннее	...	Фаціальное замещение, постседиментационное из- менение	Выпуклость	Седиментацион- ное	Латеральный и усть- евой бары аван- дельты	
Марковское	Ярактинский	Л-5	Всестороннее	...	Фаціальное замещение	Моноклиаль	Тектоническое	—	
Марковское	Парфенов- ский	Л-7	Всестороннее	...	Фаціальное замещение, постседиментационное из- менение	Выпуклость	Седиментацион- ное	Бар внедельтовой акватории	
Марковское	Осинский	Л-9	Всестороннее	...	Фаціальное замещение	Выпуклость	Седиментацион- ное	Органогенная по- стройка	
Потаповская	Парфенов- ский	Л-8	Всестороннее	...	Фаціальное замещение, постседиментационное из- менение	Выпуклость	Седиментацион- ное	Бар внедельтовой акватории	

нообразных элементов данных, а набор однородных совокупностей элементов данных с одними и теми же именами, называемый отношением (табл. 6, 7, 8). Структура данных представляется системой отношений. Отношение можно изображать как двухмерную таблицу с одинаковой структурой строк, что удобно для таблично представляемой информации. Каждая строка таблицы содержит значения совокупности объединенных по смыслу элементов данных, определяющей отношение. В рамках реляционных моделей не только представление данных имеет табличную форму, но и вся обработка данных, получение промежуточных и конечных результатов сводится к манипулированию элементами (строками и столбцами) таблиц. Поисквые процедуры основываются на реляционной алгебре или реляционном исчислении [32], что позволяет реализовать в АФИС заранее непредусмотренные запросы.

При реляционном подходе не используется понятие объекта, что очень важно для предметных областей, в которых существует несколько точек зрения или концепций на определение системы объектов (например, для геологии). Это позволяет частично ослабить негативное влияние неоднозначности понимания и соотношения типов объектов (см. разд. 2.2). Категория связи здесь также не используется, а установление соответствия между данными происходит по совпадению значений элементов данных с одинаковыми именами, принадлежащих различным отношениям. Так, установление соответствия данных табл. 6 данным табл. 7 проводится по совпадению имен собственных (значение элемента данных) месторождений (имя элемента данных), а данных табл. 7 данным табл. 8 — по названиям месторождения и природного резервуара.

Правило установления соответствия между данными различных отношений может быть использовано для составления новых отношений из исходных. Они состоят из комбинацией значений элементов данных тех строк исходных отношений, между которыми установлено соответствие. Так, из табл. 6, 7, 8, можно составить новое отношение (таблицу), которое будет представлять собой совокупность всех или выборочных элементов данных из этих таблиц для каждого из месторождений, имена которых совпадают в исходных таблицах. Допустим, можно составить отношение, в котором Марковское и Ярактинское месторождения будут описаны одинаково всеми имеющимися в табл. 6, 7, 8 характеристиками, либо для Марковского месторождения составить отношение из элементов данных: геотектоническое положение (см. табл. 6), типы пород и генезис коллектора (см. табл. 7), типы палеогеоморфологических объектов (см. табл. 8), а для Ярактинского — из элементов данных: категория крупности (см. табл. 6), мощность и пористость коллектора (см. табл. 7), степень латерального экранирования и морфология кровли коллектора (см. табл. 8).

Поскольку степень исходных и порожденных отношений на логическом уровне никакими условиями не ограничивается, то при реляционном подходе не возникает проблем представления отно-

шений степени больше двух, что трудно осуществить в рамках иерархической и сетевой моделей данных.

Реляционный подход является наиболее естественным, простым, гибким, а главное, формальным, но и он имеет ряд ограничений [32]. Наиболее существенное из них обусловлено тем, что ключи (имена, определяющие различные отношения) при прохождении по связанной цепочке отношений в значительной степени повторяются. Так, если обратиться к табл. 6, 7, 8, ключами для данных являются: в табл. 6 — имя месторождения, в табл. 7 — имена месторождения и природного резервуара, а в табл. 8 — имена месторождения, природного резервуара и ловушки. Это несет в себе определенный элемент избыточности и может привести к образованию громоздких ключей. Кроме того, теория отношений, лежащая в основе реляционного подхода, не предполагает сложных конструкций значений элементов данных (например, множественных значений — перечислений), в связи с чем в реляционных СУБД отсутствуют специальные системные (программные) средства представления таких конструкций. Представление этих конструкций в терминах отношений менее эффективно по сравнению с системными средствами структурирования данных, заложенными в СУБД, ориентированных на иные типы структур (иерархические, сетевые). Тем не менее преимущества, предоставляемые реляционным подходом, с лихвой окупают его ограничения.

Следует отметить, что использование в реляционном подходе отношений в качестве исходной основы вовсе не означает полной несопоставимости с видением предметной области через понятия «объект» и «связи». В терминах теоретико-графового представления (на котором базируются иерархическая и сетевая модели данных) каждую строку таблицы можно представить как отображение некоторого объекта, а правила сопоставления строк различных таблиц — как связи между этими объектами.

Так, каждая строка табл. 6 может отображать объект «месторождение», табл. 7 — объект «природный резервуар», табл. 8 — объект «ловушка углеводородов». По совпадающим значениям соответствующих элементов данных (именам ловушки, природного резервуара, месторождения) можно установить, что объект «ловушка» связан с объектом «природный резервуар» связью типа 1 : M (в одном природном резервуаре может быть несколько ловушек: в аянском — 2, в ярактинском — 3 и т. д.), а между объектами «природный резервуар» и «месторождение» существует связь типа M : M (например, на Марковском месторождении — 3 природных резервуара: ярактинский, парфеновский, осинский, но в то же время ярактинский природный резервуар развит на двух месторождениях: Марковском и Ярактинском, а парфеновский — на Марковском месторождении и Потаповской разведочной площади).

### 3.2.4. Уровни организации данных в АФИС

ЭВМ не способна «вникать» в содержательный смысл информации, а может лишь осуществлять те или иные операции со зна-

ками. Именно это делает возможным применение ЭВМ в самых разнообразных сферах человеческой деятельности. Но это означает и необходимость представить подлежащую обработке информацию на уровне знакового формализма. Чтобы выйти на этот уровень от представлений об информации специалиста соответствующей предметной области (в данном случае, геолога), требуется реализовать сложную многоступенчатую технологию преобразования информации (представления данных). Выделяется три последовательных уровня организации данных: информационный, концептуальный и логический. Каждый последующий из них характеризуется большей степенью формализации по сравнению с предыдущим. На первых двух уровнях происходит первичная организация информации, составляющей так называемое внешнее представление данных. Термин «внешнее» здесь означает, что это представление формируется вне связи с СУБД.

Информационный уровень. На этом уровне проводится первичное обобщение представлений о предметной области отдельных групп специалистов этой области, для которых создается данная АС. Каждую такую группу может интересовать своя система объектов. Различия в системах используемых предметов и объектов характерны прежде всего для специалистов разных отраслей геологии — нефтегазовой, рудной, угольной и т. п. Но и в рамках одной отрасли геологии разные группы специалистов работают с предметами и объектами различного уровня. Например, специалисты по общей оценке перспектив нефтегазоносности оперируют в основном объектами, представляющими элементы нефтегазогеологического районирования (нефтегазоносный бассейн или провинция, зона нефтегазонакопления, литолого-стратиграфический комплекс и т. д.), а геологов, специализирующихся на выявлении неантиклинальных ловушек нефти и газа, интересуют в первую очередь такие объекты, как ловушки, палеогеоморфологические элементы и т. п. Используются и общие для всех или нескольких групп объекты (тектонические элементы, стратиграфические единицы, породы и др.). Кроме того, информация может собираться по объектам политико-административного (регион, республика, область и т. д.) и физико-географического (территория, акватория) делений, а также по их пересечениям (например, акватория части бассейна, расположенной в пределах данной республики).

У разных групп геологов — потенциальных пользователей АС, как правило, не совпадают требования к полноте описания одних и тех же объектов (их интересуют разные наборы характеристик объектов). Нередки случаи расхождения во взглядах на взаимное соответствие (соподчиненность) предметов разного уровня (например, рудного месторождения и рудной зоны).

Уяснение информационных потребностей каждой группы потенциальных пользователей АС, выявление степени согласованности (и рассогласованности) их взглядов на информационное отражение предметной области составляют сущность информационного

уровня организации данных. Итогом его являются зафиксированные совокупности рассматриваемых типов объектов, представления о взаимосответствии между ними, наборы показателей, характеризующих каждый тип объекта.

Основа (инструмент) проектирования АС на информационном уровне — обследование информационных потоков в рамках автоматизируемой сферы деятельности потенциальных пользователей. При обследовании, помимо выявления систем объектов, их взаимосвязей и показателей, определяются и такие важные для проектирования АС характеристики информационных потоков, как интенсивность, однородность, периодичность поступления и возможный предельный объем данных и др. Очень важно на информационном уровне представления данных учесть возможное расширение информационных потребностей в обозримой перспективе. Например, если предполагается решение в дальнейшем тех или иных новых задач и имеется представление о задействованной в моделях решения этих задач информации (параметрах модели), то целесообразно включить ее в состав данных уже на информационном уровне их организации. Чем глубже будут проработаны перспективы расширения информационной базы АС, тем проще реализация этих расширений впоследствии.

Концептуальный уровень. На этом уровне организации данных осуществляется интеграция всех пользовательских представлений о предметной области на основе единого подхода к структуризации данных, в результате чего создается концептуальная модель предметной области.

Возможны два пути построения концептуальной модели: от отдельных характеристик к их совокупностям, которые могут отображать объекты, — восходящее проектирование, и от объектов к их атрибутам (характеристикам) — нисходящее проектирование.

При восходящем проектировании на основе представлений специалистов данной предметной области выделяются характеристики, которые предлагается интегрировать в БД в качестве элементов данных, и анализируются зависимости между ними, в результате чего они объединяются в совокупности, которые могут соотноситься с типами объектов предметной области. Как правило, этим занимается системный аналитик — специалист по созданию БД, формирующий структуру базы данных при консультации со специалистом данной предметной области.

При нисходящем проектировании сначала выделяются типы объектов предметной области, а затем определяются те их характеристики (атрибуты), которые предлагается ввести в БД как совокупности элементов данных, отображающие ОПО, т. е. как объекты БД. При этом может оказаться, что различные типы ОПО имеют общие группы атрибутов (например, бассейн, месторождение, залежь характеризуются запасами и добычей полезного ископаемого), которые разумно выделить в отдельные совокупности. В результате этого описание таких типов ОПО будет

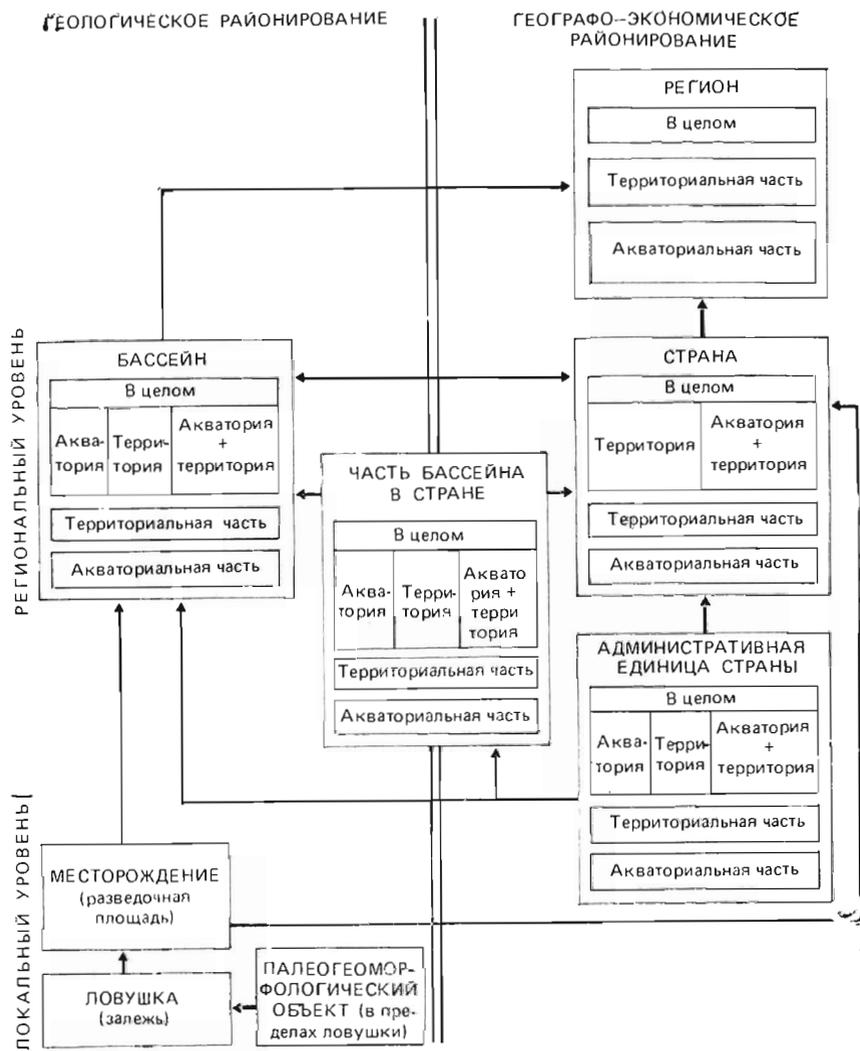


Рис. 33. Пример представления структуры объектов предметной области (ОПО) на концептуальном уровне

в БД разложено на общие и уникальные совокупности атрибутов, являющиеся объектами ОБД. Это значит, что отдельный ОПО будет представлен в базе данных несколькими ОБД. Обычно такое построение концептуальной модели делает специалист предметной области с помощью системного аналитика.

Оба пути построения концептуальной модели могут использоваться при организации информации по любой предметной области. Выбор одного из них зависит от того, насколько общепринято

(согласовано) объектное членение и состав характеристик объектов различных типов в данной предметной области.

В геологии, где информация привязана, собирается и обрабатывается относительно определенных систем объектов, пусть даже не удовлетворяющих формально-логическим требованиям, более предпочтителен второй путь построения концептуальной модели — от объектов к их атрибутам. Однако расхождения во взглядах на соподчиненность геологических объектов и состав описывающих их атрибутов (см. разд. 3.2.2) могут вызвать необходимость использования первого пути, по крайней мере в некоторых областях геологии.

Но в том и другом случае на концептуальном уровне элементами модели предметной области будут либо объекты базы данных (совокупности атрибутов), отображающие необходимые пользователю типы ОПО, либо совокупности элементов данных, соотносимые с ОПО.

Для отражения взаимоотношений между элементами модели могут использоваться как теоретико-графовые, так и теоретико-множественные представления. При объектном взгляде на предметную область формирование концептуальной модели, как правило, начинается с представления структуры ОПО (рис. 33). На рис. 33 четко видно, данные какого раздела геологической предметной области будут обрабатываться АФИС, какие объекты и по каким основаниям выделены и как они соотносятся между собой, а также дано обособление двух уровней (регионального и локального) объектов, различающихся по характеру описания и взаимосвязей. На этой схеме четко зафиксированы устанавливаемые связи между объектами, что предопределяет возможные направления поиска данных в АФИС. Так, если не задана связь между палеогеоморфологическим объектом и нефтегазоносным бассейном, то получить информацию о принадлежности первых ко вторым непосредственно будет нельзя. Реализация такого запроса потребует перехода по всей цепочке промежуточных связей с перебором данных по ловушкам, месторождениям и бассейнам.

Концептуальная схема данных дает возможность сделать заключение о типе структуры данных, наиболее адекватно описывающей предметную область. Использование иерархической структуры здесь невозможно (или возможно лишь с большим дублированием данных), так как некоторые порожденные узлы (объекты «часть бассейна в стране», «месторождение») имеют по два исходных узла (страна и «бассейн»). По-видимому, в данном случае более подходит сетевая или реляционная структура. Окончательное решение этого вопроса требует анализа совокупности атрибутов объектов концептуального уровня представления данных.

Если при таком анализе выявляется, что многие объекты имеют общие группы характеристик, то для уменьшения избыточности описания их можно выделить в отдельные совокупности элементов данных (ОБД) и использовать для описания различных типов ОПО. Возникает структура данных (рис. 34), где типу ОПО

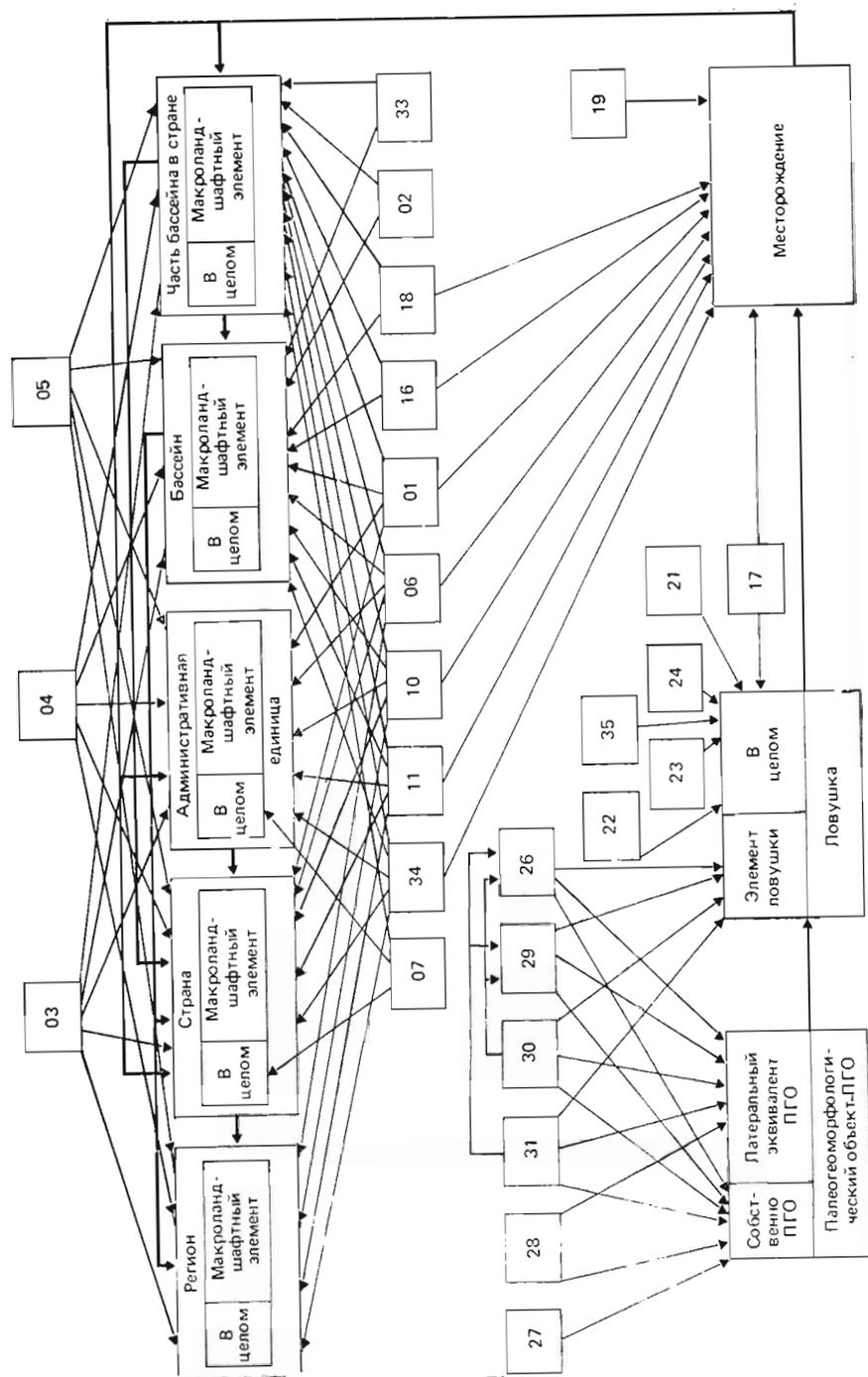


Рис. 34. Пример модели данных на концептуальном уровне (подсистема «Ресурсы нефти и газа» АФИС-Геоинформ).

01—35 — тематические группы параметров (объекты базы данных — ОБД, входные документы — экраны маски): 01 — условия проведения работ (площадь, ..., природная зона, ..., среднегодовые температуры, ..., количество осадков, ..., инфраструктура, ...); 02 — общая геологическая характеристика (тектоническое положение, ..., инфраструктура, ..., объем осадочного выполнения, ..., возраст фундамента, ..., типы пород, ...); 03 — ресурсы, добыча нефти и газа (ресурсы начальные суммарные геологические, ..., то же извлекаемые, ..., добыча накопленная общая, ...); 04 — распределение ресурсов, запасов, добычи (по интервалам глубин, ..., по стратиграфическим интервалам, ...); 05 — распределение месторождений нефти и газа (по категориям крупности, ..., по состоянию освоения, ...); 06 — динамика запасов и добычи нефти и газа (ежегодные запасы текущие извлекаемые, ..., добыча общая), 07 — динамика использования ресурсов нефти и газа (ежегодные потребление, ..., экспорт, ..., импорт, ...); 10 — динамика фонда действующих эксплуатационных скважин (ежегодное число скважин фонтанирующих, ..., насосных, ...); 11 — объем бурения (ежегодное число и метраж скважин лонковых, ..., разведочных, ..., всего, ..., продуктивных, ...); 16 — характеристика разреза (ч. 1) (по интервалам; глубина кровли, ..., мощность, ...); 17 — потенциал месторождения, залежи нефти и газа (тип по фазовому состоянию флюидов, ..., запасы начальные нефти+канденсата, ..., газа свободного, ...); 18 — характеристика разреза (ч. 2) (по интервалам: типы пород, ..., фауна, ...); 19 — общая характеристика месторождения (дата открытия, ..., методы подготовки под ПРБ, ..., продуктивные горизонты, ..., классы ловушек, ...); 21 — характеристика залежи (глубина, ..., пластовое давление, ..., режим, ..., плотность нефти, ..., газа, ...); 22 — состав нефти (фракции по температуре кипения, ..., содержание серы, ..., кокса, ...); 23 — состав флюидов (компоненты и их содержание в воде, газе свободном, ..., растворенном в нефти, ...); 24 — характеристика ловушки (стратиграфическое положение, ..., степень латерального экранирования, ..., генезис латерального экрана, ..., морфология кровли, ...); 26 — литолого-физические характеристики породы (ч. 1) (тип породы, ..., минеральный состав, ..., размерность, ..., каркасных и некаркасных компонентов); 27 — палеогеоморфологический объект (ч. 1) (тип палеогеоморфологического объекта, латеральные эквиваленты, ..., характер контактов, ...); 28 — палеогеоморфологический объект (ч. 2) (типы пород, ..., мощность, ..., ширина, ..., тип поперечного профиля, ...); 29 — литолого-физические характеристики породы (ч. 2) (макроструктура, ..., окраска, ..., пористость, ..., эпигенетический цемент, ...); 30 — геохимическая характеристика (соединения, ..., микроэлементы, ..., их содержания, ...); 31 — палеонтологическая характеристика (фауна, ..., флора, ..., сохранность, ...); 33 — параметры и результаты расчета ресурсов нефти и газа ( $N_b$  участка и толщи, ..., давление пластовое, ..., давление насыщения, ..., коэффициенты концентрации свободного газа, ...); 34 — объем геофизических работ (ежегодные объемы сейсморазведки, ..., гравиразведки, ..., в отрядо-месяцах, ..., тыс. км ...); 35 — характеристика природного резервуара (типы пород и мощность коллектора, ..., покрышки, ..., генезис коллектора, ..., проницаемость, ...)

не всегда прямо соответствует тип ОБД, а первый конструируется из вторых (например, ОПО «месторождение» составляется из ОБД «общая характеристика месторождения», «характеристика разреза», «динамика запасов и добычи», «динамика объемов бурения и геофизических работ», «условия проведения работ»).

При таком представлении предметной области усложняются связи, так как необходимо представлять не только связи между ОПО, но и связи, по которым из нескольких ОБД конструируется их описание. Но поскольку число ОБД, из которых составляется описание ОПО, может быть и больше двух (в приведенном выше примере число ОБД, из которых собирается ОПО «месторождение», равно 5), а в графах одной связью соотносятся только два объекта, то может возникнуть необходимость использовать для моделирования предметной области реляционную структуру, которая снимает проблему представления отношений степени больше двух (см. разд. 3.2.3), хотя с некоторыми трудностями эта проблема может быть решена и в рамках сетевой структуры.

При реляционном подходе предметная область описывается не пообъектными совокупностями элементов данных, а наборами однородных совокупностей элементов данных (отношениями — см. разд. 3.2.3). В эти совокупности в качестве элементов данных могут включаться типы ОПО и их имена, и тогда структура дан-

Таблица 9

## Совокупности элементов данных «Запасы и добыча»

Запасы нефти	Запасы газа	Добыча нефти	Добыча газа	Год	Тип ОПО	Имя ОПО
					Регион Бассейн Страна Месторождение	

ных, представленная на рис. 34, может быть изображена в виде таблиц. Реляционное представление фрагмента этой структуры показано в табл. 9—11.

Таким образом, на концептуальном уровне создается модель данных, адекватно описывающая предметную область и служащая исходной основой для выбора типа СУБД и формирования модели данных на логическом уровне.

**Логический уровень.** На этом уровне модель данных переводится с языка геологических представлений, которыми оперирует геолог, на язык, воспринимаемый СУБД. Эта работа выполняется системным программистом, и он тем быстрее, проще и корректнее делает ее, чем детальнее и правильнее разработана концептуальная модель предметной области.

Каждая отдельная СУБД настроена, как правило, на какой либо один тип структуры данных. Поэтому выбор СУБД определяется в первую очередь структурой данных, которой моделируется предметная область на концептуальном уровне. Поскольку существуют разные СУБД, поддерживающие одну и ту же структуру данных, то выбор конкретной СУБД зависит и от принятого в данной СУБД формализованного логического языка описания данных (ЯОД), представляющего собой совокупность инструкций, по которым соответствующие программы СУБД формируют в ЭВМ массивы информации. В результате описания структуры данных концептуального уровня на ЯОД конкретной СУБД на выходе логического уровня формируется так называемая схема базы данных, которая вводится в ЭВМ. Эта схема обеспечивает настройку ЭВМ на «понимание» конкретной предметной области.

Таблица 10

## Совокупности элементов данных «Использование ресурсов»

Потребление нефти	Потребление газа	Экспорт нефти	Экспорт газа	Импорт нефти	Импорт газа	Год	Тип ОПО	Имя ОПО
							Страна	

Таблица 11

## Совокупности элементов данных «Характеристика бассейна»

Геотектоническое положение	Объем осадочного выполнения	Возраст фундамента	Число структурных этажей	Тип ОПО	Имя ОПО
				Бассейн	

## 3.3. ТРЕБОВАНИЯ, ПРЕДЪЯВЛЯЕМЫЕ АС К ФОРМЕ ПРЕДСТАВЛЕНИЯ ИСХОДНЫХ ДАННЫХ

Соблюдение формально-логических требований к представлению геологических данных, кроме расширения возможностей по анализу, обобщению и выявлению новых закономерностей, создает предпосылки для автоматизации обработки информации. ЭВМ не способна соотносить знаки с отображаемой ими действительностью, а оперирует лишь формой знакового выражения, т. е. правила оперирования основаны исключительно на процедурах манипуляции формой представления знаков в ЭВМ. Соответственно любая содержательная обработка информации должна описываться в терминах оперирования формой знаков. Поэтому применение ЭВМ выдвигает дополнительные требования к представлению (формализации) исходного материала, из которых наибольшее значение для геолога-пользователя АС имеют представление информации в виде данных и идентифицируемость объектов базы данных.

Понимание этих требований и накладываемых ими ограничений на возможности, предоставляемые АС геологу-пользователю, необходимо последнему для эффективного общения с АС.

## 3.3.1. Представление информации в виде данных

Выше было дано определение данных как информации, представленной в виде отдельных поименованных фактов — элементов данных, области значений которых формально определены. Наиболее важные характеристики данных — имя элемента данных, значение элемента данных и тип элемента данных. Ниже мы рассмотрим сначала сущность этих характеристик, а затем требования, предъявляемые к ним АС.

Смысл первых двух характеристик очевиден: использование имен позволяет описывать состав и структуру массивов данных и осуществлять поиск значений элементов данных.

Несколько подробнее рассмотрим содержание характеристики «тип элемента данных». В большинстве АС предусматривается использование нескольких типов данных, которые различаются по форме знакового выражения значений (символьные и числовые) и по структуре значений. Каждый из типов имеет свою форму ма-

шинного представления, свои возможности по обработке, требует различных ресурсов памяти ЭВМ и программного обеспечения.

Символьными<sup>1</sup> называют такие данные, в которых каждое значение представляется строкой символов, воспринимаемых ЭВМ как единый знак. Состав или алфавит допустимых символов определяется конкретным программным и техническим обеспечением. Как правило, он включает буквы латинского алфавита, символы «кириллицы», арабские цифры, знаки пунктуации и ряд специальных символов, которые используются в различных комбинациях, например, «PZ 2» (для обозначения верхнепалеозойского возраста) или «БУ-75» (тип бурового станка) или «поисково-разведочная» (тип скважины) и т. п. Хотя в число символов входят цифры, но они в составе символьных данных имеют тот же статус, что и буквы, знаки препинания, т. е. никакие арифметические и логические действия над ними не возможны. Содержательная обработка символьных данных (в том числе и цифр, рассматриваемых как символы) базируется на единственной операции, состоящей в сравнении знаков на совпадение.

В числовых данных значением может быть только число (целое и вещественное), т. е. цепочка цифр и один разделитель целой и дробной части для вещественного (дробного) числа. Над числами выполняется более широкий спектр действий содержательной обработки, включающий арифметические и логические действия.

Иногда, применительно к некоторым СУБД со специальными механизмами содержательной обработки, выделяются такие типы данных, как булевы (логические) и текстовые. Булевы данные могут принимать только два значения — «да» и «нет» (или «истина» и «ложь»), которые кодируются соответственно единицей и нулем. Над ними выполняются операции булевой алгебры. Область применения булевых данных в геологии относительно ограничена, известные приложения касаются в основном задач распознавания и классификации.

Текстовые данные являются разновидностью символьных, отличаясь неформализованностью области значений. Соответственно возможности обработки таких данных крайне ограничены. К текстовым данным вполне применимы процедуры, обеспечивающие хранение и выдачу, однако никакой содержательной их обработки система не проводит, и это должен делать сам пользователь (т. е. пользователь сам анализирует выданные текстовые данные, выделяет значения, сравнивает их на совпадение и т. п.).

По структуре значений различают данные с простыми (элементарными) и множественными (составными) значениями, образованными из перечисления нескольких элементарных. В качестве примеров элементарных значений можно назвать «28» (величина пористости, %), K2 (возраст пород), «песчаник» (тип породы) и т. п. Но в геологии при указании значения того или иного пока-

зателя часто невозможно (с точки зрения геолога) ограничиться элементарным значением. Так, значением показателя «минеральный состав» (породы), как правило, является перечисление (кварц, полевой шпат, ...). Подобно этому показатель «возраст продуктивных горизонтов» (на месторождении, в бассейне) часто требует перечисления или указания интервала (например, P—T). Это и есть примеры весьма распространенных в геологической практике множественных значений.

Если для геолога проблемы различения элементарных и составных значений просто не существует, то АС для этого требуется иметь в составе СУБД специальные достаточно сложные программные средства. Если их нет, то АС будет воспринимать как элементарные (т. е. как логически неделимые) и составные значения. А это сильно сказывается на возможности получения пользователем информации по запросам. Для примера обратимся к такой типичной характеристике месторождений нефти и газа, как тип пород, слагающих продуктивный горизонт. Если геологу потребуется получить справку — на каких месторождениях продуктивный горизонт сложен известняками, то в числе выданных названий будут отсутствовать те месторождения, на которых в строении продуктивного горизонта наряду с известняками участвуют и другие типы пород (например, известняк, доломит). Поскольку вторичная доломитизация характерна для многих слоев известняка, геолог получит существенно неполную информацию, даже если в базе данных содержатся все необходимые сведения для исчерпывающего ответа на запрос. Связано это с тем, что для АС, не имеющих специальных средств доступа к элементам сложных значений, запись «известняк, доломит», воспринимается не как сочетание элементарных значений «известняк» и «доломит», а как независимое от них самостоятельное, логически неделимое значение. Указывать же в запросе возможные сочетания типов пород (достаточно многочисленные) крайне неэффективно, так как это требует детального знания соответствующего классификатора, затрудняет формализацию запросов.

Вышеприведенные примеры составных значений касались перечисления элементов одного уровня. Но весьма часто встречаются данные, составные значения которых имеют разноуровневую структуру (когда элементарные значения нижнего уровня выступают как уточняющие характеристики элементарных значений верхнего уровня). Например, минеральный состав породы может описываться не только названиями минералов, но и указанием их процентного содержания (кварц — 40, полевые шпаты — 50, слюда — 10), или значение показателя «геотектоническое положение» может иметь трехуровневую структуру (платформа, древняя, аптеклиз). В таком случае элементарные части этого составного значения не могут рассматриваться как равноправные, и нужен специальный программный механизм не только их различения, но и взаимосопоставления.

Поэтому специфика геологической предметной области, в кото-

<sup>1</sup> Иногда в литературе используется термин «строковые данные», или «строковый тип».

рой широко используются составные значения, имеющие часто сложную структуру, требуют использования СУБД, обладающих механизмами обработки составных значений.

Требования к именам элементов данных. Поскольку поиск элемента данных в памяти ЭВМ осуществляется по его имени, очевидно, что это имя должно быть уникальным, т. е. различные элементы данных не должны иметь одинаковые имена. Это главное требование к именам выглядит очевидным и на первый взгляд акцентирование на нем внимания представляется излишним. Действительно, если области значений показателей различны, например, область значений одного показателя — 0—50 %, а другого — белый, черный, серый, красный и т. п., то здесь имена показателей (в данном случае «пористость коллектора» и «цвет пород») неизбежно будут различными.

Однако в геологической предметной области необходимость соблюдения этого требования часто вызывает значительные сложности, обусловленные тем, что нередки случаи, когда для элементов данных, различающихся по смыслу, области значений оказываются одинаковыми. Так, во всех отраслях геологии широко используется понятие геологического возраста с областью значений, отвечающей стандартной геохронологической шкале, в рамках которой выделяются и обозначаются общепринятыми индексами эры, периоды, эпохи, века и др. В практической геологии значение возраста имеет смысл не само по себе, а применительно к определенному объекту — интервалу разреза, продуктивному горизонту, структурному этажу, складчатому сооружению, фундаменту платформы и т. п. ЭВМ не способна различать такие характеристики по смыслу или контексту, поэтому при формализации геологической информации соответствующим элементам данных необходимо присвоить различные имена: возраст геотектонического элемента, возраст продуктивного горизонта, возраст структурного этажа, т. е. сделать имена уникальными.

Второй пример касается такой характеристики, как «тип пород». При описании разреза геологу бывает важно зафиксировать не только типы пород, но и их соотношения. Так, с точки зрения геолога-нефтяника разрезы, сложенные преимущественно песчаниками с прослоями глин и преимущественно глинами с прослоями песчаников, принципиально различны, хотя в обоих случаях участвуют одни и те же типы пород. Различие между такими разрезами при невозможности указать мощности различных пород или их процентное содержание в разрезе обычно отражается указанием основных пород и прослоев. Поэтому при преобразовании этой информации в данные необходимо выделить два элемента данных с именами: типы пород основные и типы пород прослоев (но с одинаковыми областями значений, заданными классификатором типов пород). Если же использовать один элемент данных с именем просто «типы пород», то при соответствующем запросе дифференцировать указанные выше различные типы разреза не удастся. Такой способ образования уникального имени соответ-

ствует рассмотрению имени как единой, логически неделимой величины. Обеспечивая точность ответа на запрос, он тем не менее не всегда удобен, так как несколько усложняет формулировку запросов. Так, если требуется найти разрезы, в которых есть известняки (неважно, слагают ли они разрез целиком или только образуют прослой), то запрос (при указанном выше способе образования имен) придется сформулировать так: выдать разрезы, где типы пород основные — известняки и типы пород прослоев — известняки. Для того чтобы упростить язык запросов, в некоторых СУБД предусматриваются механизмы составных имен: у нескольких элементов данных есть общее имя и уточняющие его имена. В нашем примере это будет выглядеть так: типы пород

основные, т. е. образуются три элемента данных с именами: типы прослоев пород, типы пород основные, типы пород прослоев. Соответственно возможны три простых запроса на выдачу разрезов: 1) в строении разреза участвуют известняки (типы пород — известняки); 2) разрез целиком или в основном сложен известняками (типы пород основные — известняки); 3) разрез включает прослой известняков (типы пород прослоев — известняки).

Некоторые СУБД допускают синонимию, т. е. использование нескольких имен для обозначения одного элемента данных. Например, можно использовать одновременно названия «типы пород продуктивного горизонта» и «литологический состав продуктивного горизонта», «возраст свиты» и «стратиграфическое положение свиты», или названия одного и того же элемента данных на разных языках (при формировании многоязычной АФИС).

Важно также отметить часто встречающиеся ограничения: на длину имени, определяющие максимально возможное число символов в имени (обычно до 8 или 16). Это очень жесткое ограничение в условиях геологии, которой свойствен описательный язык. Многие имена данных просто не укладываются в такие ограничения, а искусственно образованные сокращения имен или непонятны геологу, или допускают неоднозначность толкования; запрещение использования в качестве первого символа имени цифры и ряда определенных специальных литер, например знаков препинания. Эти ограничения вытекают из требований используемых языков программирования.

Требования к значениям и типам элементов данных

В наиболее общем виде требование к значениям формулируется как требование формализованности определения области значений. Соблюдение этого требования обуславливает общность признаков, определяющих механизмы и возможности обработки, качество получаемого при этом результата. Если конкретизировать это общее требование, то (для отдельного элемента данных) оно состоит в логически корректном задании области значений; в обеспечении единства системы измерения, определяющей область значений; в соблюдении соответствия значений типам, обработка

которых обеспечивается СУБД; в соблюдении требования типовой однородности области значений.

Задание области значений элемента данных заключается в определении минимальной и максимальной возможной величины значения для числовых данных и перечня допустимых значений для символьных данных.

При этом для числовых данных необходимо определить также характер используемых чисел. Например, для элемента данных «открытая пористость» область возможных значений определяется вещественными числами от 0 до 50, а для элемента данных «количество залежей» — только целыми числами от 0 до 40.

Для символьных данных область значений определяется, как правило, заданием классификатора, в котором перечисляются допустимые значения элемента данных. Например, для элемента данных «макротекстура породы» область значений описывается классификатором типов слоистости, для элемента данных «литологический состав горизонта» — классификатором типов пород и т. д. В качестве значений классификатора могут выступать имена собственные (например, классификаторы стран, нефтегазоносных бассейнов, металлических провинций и др.), термины, химические символы (классификатор элементов), символы геологического возраста и другие элементы естественного и профессионального языка. По своей логической структуре классификаторы подразделяются на две группы — базирующиеся на перечислении (упоминавшиеся классификаторы стран, бассейнов, а также некоторые из названных ниже общепромышленных классификаторов, отражающие геологическую деятельность) и базирующиеся на научной классификации. К последним относятся (должны относиться) классификаторы параметров, характеризующих геологическое строение недр. Естественно, что таких классификаторов большинство. Рассмотренные в гл. 2 методологические основы построения научных классификаций, известные как правила деления объема понятия, в полной мере относятся и к классификаторам второй группы. Естественно, что все отрицательные моменты, связанные с логической некорректностью геологических классификаций (см. разд. 2.2.2), неизбежно переносятся на классификаторы АС. Ниже это будет проиллюстрировано на примере классификатора «Тектонические структуры».

Классификаторы призваны обеспечить не только сопоставимость данных в рамках определенной АС, но и возможность обмена данными между различными АС без перекодировки данных. В идеале целесообразно иметь единый набор взаимосогласованных классификаторов, из которого для каждой АС (соответственно предметной области и назначению) формировалась бы та или иная совокупность классификаторов и полнота (детальность) каждого из них. Но пока это нереально, для используемых классификаторов установлены три уровня общности: общесоюзный, отраслевой и локальный [42]. Общесоюзные классификаторы (например, «Страны мира и территории», «Полезные ископаемые

и подземные воды», «Химические элементы») обязаны использовать все АС страны, работающие с соответствующими данными. Отраслевые классификаторы обязательны для всех АС соответствующей отрасли. Для АС геологической направленности разработано более 40 отраслевых классификаторов, в том числе:

Этапы и стадии геологоразведочных работ;  
Работы съемочные и поисковые;  
Работы гидрогеологические и инженерно-геологические;  
Работы геофизические;  
Геофизические методы;  
Работы горнопроходческие;  
Работы буровые;  
Горные выработки и буровые скважины;  
Системы разработки месторождений полезных ископаемых;

Тектонические структуры;  
Геохронологические подразделения;  
Породы магматические;  
Породы метаморфические и метасоматические;  
Породы осадочные;  
Структуры горных пород;  
Текстуры горных пород;  
Формы тел горных пород;  
Цвет горных пород.

Локальные классификаторы разрабатываются применительно к задачам конкретной АС. В качестве примеров локальных классификаторов АФИС — «Неантиклинальные ловушки нефти и газа» можно указать такие, как «Палеогеоморфологические объекты», «Степень латерального экранирования природного резервуара», «Категории крупности залежей» и др. Среди классификаторов практически любой АС локальные составляют подавляющее большинство. Это сильно сужает возможность непосредственного обмена данными между АС. Такое положение отражает как состояние языка геологии (в частности, частое отсутствие общепринятых классификаторов, проиллюстрированное в гл. 2 на примере классификаций ловушек), так и отсутствие единой информационной политики в отрасли.

Следует отметить, что качество классификаторов в огромной степени влияет на эффективность подготовки данных на машинно-ориентированных входных документах, на возможность обработки данных и качество результатов их обработки. Это накладывает исключительно высокую ответственность на разработчиков классификаторов, которые обычно создаются проектировщиками информационного обеспечения АС в тесном сотрудничестве с геологами — владельцами исходных данных и потенциальными или действительными пользователями данной АС. Данное обстоятельство еще раз подчеркивает важность знания геологами изложенных в гл. 2 требований к построению классификаций (часто геологи по инерции настаивают на использовании привычных, но заведомо логически некорректно построенных классификаторов).

К сожалению, приходится констатировать, что этим требованиям часто не отвечают классификаторы не только локального, но в ряде случаев и отраслевого уровня, что наглядно видно на примере такого наиболее употребительного классификатора, как «Тектонические структуры». Во-первых, он включает перечень типов структур, состоящий из 316 наименований. Соответственно

геологу, заполняющему входной документ, надо просмотреть весь этот перечень (и не раз, так как в конце уже не помнишь, что было в начале), чтобы выбрать из него название типа структуры, наиболее отвечающего наблюдаемой геологической ситуации. Запомнить содержание такого классификатора и точные формулировки названий просто невозможно, особенно учитывая, что геологу приходится работать с десятками классификаторов. Это резко повышает затраты труда и времени геолога при заполнении входных документов, увеличивает долю нетворческой деятельности.

Во-вторых, составляющие данный классификатор названия типов тектонических структур часто громоздки, например, «структура тектоническая унаследованная», «расположение складок кулисообразное», «останец тектонического покрова» и т. п. Записывать на входных документах такие названия сложно, тем более что их может быть несколько в одном документе (например, если описываемый нефтегазоносный бассейн расположен в пределах нескольких тектонических структур, в частности древней платформы и передового прогиба кайнозой). Кроме того, представление громоздких значений требует больших объемов памяти на машинных носителях, что снижает эффективность системы. Использование же вместо текстовых значений соответствующих им числовых кодов, приведенных в классификаторе, влечет за собой дополнительные сложности:

коды требуют расшифровки на выходе, что увеличивает время обработки и усложняет программы;

ошибки в числовых кодах очень трудно диагностировать и поэтому современные АС, как правило, коды не используют; при текстовых записях ошибки, которые возможны и даже неизбежны, достаточно легко выявляются как визуально, так и автоматически с помощью специальных программ, своряющих побуквенно вводимые записи с текстами классификаторов;

запомнить порядковые числовые коды еще сложнее, чем значения из слов естественного языка.

В-третьих, и это самое главное, данный классификатор не согласуется с методологическими основами построения научных классификаций. В нем нарушены все четыре правила деления объема понятия (см. разд. 2.2.2). Прежде всего, не выдержаны не только основания деления, но и не ясны сами таксономические уровни деления. Так, выделяются на равных складка коробчатая (по форме) и складка скальвания (по генезису), движения колебательные (по ритмичности) и движения посторогенные (по фазе складчатости), складчатость унаследованная (по тенденции) и складчатость платформенная (по контрастности) и т. п. Отношения соподчиненности (масштабность) не зафиксированы: в числе структур, наряду с такими объектами, как кора земная, мантия верхняя, астеносфера, присутствуют седло складки, киль синклинали, центриклиналь и др. Одна и та же часть недр по субъективному мнению заполняющего входной документ может быть соотнесена сразу с несколькими типами выделяемых структур. Например, крупное поднятие в осадочном чехле склона древней плат-

формы в равной мере относится к таким выделяемым структурам, как слой осадочный, платформа, платформа древняя, платформа материковая, чехол платформ, плита, антеклиза, складчатость платформенная.

К тектоническим структурам, судя по упоминанию в классификаторе, отнесены самые неожиданные здесь объекты, такие как залежь нефтяная, залежь рудная, ловушка нефтяная, ловушка рудная, движения эвстатические, инверсия и др. Этот отраслевой классификатор не имеет отношения к классифицированию и типологии, а представляет собой в некоторой мере систематизированный перечень терминов, имеющих то или иное (часто весьма слабое) отношение к тектонике (а не тектоническим структурам) и почерпнутых из многочисленных, часто противоречащих друг другу тектонических классификаций. При таком построении классификатора неизбежны сильные искажения геологических выводов, получаемых на основе статистической обработки (обобщения) соответствующих данных. Вполне естественно, что в такой ситуации разработчики информационного обеспечения АС часто оказываются вынужденными пренебречь прогрессивной идеей отраслевого классификатора и опираться на локальный классификатор, лишенный столь очевидных недостатков.

Еще одно требование к значениям заключается в том, что форма значения должна быть абсолютной (например, глубина залегания 1525 м; размерность зерен — средняя), а не относительной (например, глубина залегания — ниже горизонта X, размерность зерен — крупнее, чем в пачке У), так как при этом неопределима область значений.

Важным является требование единства системы измерения значений каждого отдельного элемента данных, причем система измерений понимается здесь широко. Для числовых данных — это фактически система фиксированных единиц измерения. Так, для элементов данных «пластовая температура» и «глубина залегания» могут быть установлены системы измерения в градусах Цельсия и в метрах соответственно. Если же требуется фиксировать эти данные и в других единицах, скажем, градусах Фаренгейта и футах, то необходимо ввести и соответствующие элементы данных, которые в этом случае будут такими: «пластовая температура в °С» и «пластовая температура в °F», «глубина залегания в м» и «глубина залегания в ф.».

Для символьных данных роль системы измерения по существу играет принцип классифицирования, и корректность процедуры классифицирования означает и единство системы измерения. Для предметов и объектов, выделение которых базируется на понятиях, характеризующихся неопределенностью, например фация (см. пример, разд. 2.2.1), формально задать область значений, а соответственно и систему измерения просто нельзя, что не может не сказаться негативно на качестве результатов аналитической обработки данных.

В геологической предметной области используются весьма раз-

нообразные по значениям типы данных: символьные, числовые, простые, составные и др. Между тем конкретная СУБД содержит программы, обеспечивающие обработку определенного, фиксированного набора типов данных. Этот набор определяет и ограничивает допустимые типы значений, поэтому все предназначенные для обработки данные должны соответствовать тому или иному «разрешенному» типу, что должно учитываться при формировании системы показателей. Возможности приспособления к требованиям СУБД рассмотрим на уже упоминавшемся примере элемента данных «минеральный состав», имеющего, как правило, множественное значение (например, кварц, полевые шпаты, слюды). Если СУБД не имеет средств анализа элементов составного значения, то при организации базы данных возможны следующие альтернативы: отказаться от поиска по каждому отдельному значению (так как система оперирует с составным значением как с единым неделимым знаком) или разрешить запись лишь одного, преобладающего значения, что позволит производить содержательную обработку, но ведет к снижению количества и качества информации. Но можно преобразовать эти данные в вид: минеральный состав 1 — кварц, минеральный состав 2 — полевые шпаты, минеральный состав 3 — слюды, т. е. в три элемента данных, но каждый с простым значением, допускающим обработку. При этом не происходит потери информации, но имена элементов данных приобретают нетрадиционный, формальный вид, к которому нетрудно привыкнуть.

Можно расширить список типов данных, разрешенных СУБД, но для этого надо создать алгоритмы обработки и соответствующие прикладные программы. При этом используемая система должна обеспечивать эффективное присоединение этих прикладных программ, что также необходимо учитывать при выборе СУБД.

При формировании информационного обеспечения необходимо учитывать, что система рассматривает область значений каждого элемента данных как однородную по типу значений, т. е. все значения одного элемента данных должны относиться к одному типу, который и присваивается всей области значений. Тип области значений определяет возможности их обработки. Например, если все значения элемента данных «глубина залегания» числовые (1000, 2000, 3000 и т. д.), то вся область значений является числовой, и над значениями могут выполняться такие операции содержательной обработки, как арифметические и логические действия. Однако если хотя бы одно значение имеет в знаковом выражении лишь один нечисловой символ (<1000, больше 6000, 1000—1500), то и вся область значений не может рассматриваться как числовая (нарушена ее типовая однородность) и должна быть отнесена к символьной (поскольку числовые данные — частный случай символьных), в связи с чем становятся невозможными арифметические и логические действия над значениями.

Большинство современных СУБД обеспечивают обработку данных с фиксированной длиной (количество символов) значений, т. е. для области значений каждого показателя требуется уста-

новить максимально необходимое и допустимое количество символов в значениях. При задании максимальной длины значений приходится учитывать два условия. Первое — необходимость обеспечить возможность представления самого длинного числа или значения классификатора. Второе условие связано с тем, что резервирование памяти ЭВМ под область значений показателя проводится по самому длинному значению. А поскольку нередки случаи, когда конкретные значения имеют длину меньше заданной максимальной, происходит нерациональное расходование памяти. Для того чтобы уменьшить последнее, нужен компромисс при задании максимальной длины значений (например, введение в классификаторы сокращенной формы записи длинных значений). Более предпочтительна обработка данных с переменной длиной значений (установление максимальной длины не требуется), однако существующие СУБД редко обеспечивают такую возможность.

### 3.3.2. Идентифицируемость объектов базы данных

Для обработки данных АФИС должна иметь возможность идентификации<sup>1</sup> каждого конкретного объекта. Без этого поиск информации по конкретным объектам и выдача ее по запросам невозможны. Совокупность атрибутов каждого типа объектов должна включать один или несколько атрибутов, по которым проводится идентификация объекта. Совокупность таких атрибутов (в частном случае — один атрибут) называется идентификатором. Например, идентификатором для объекта «залежь нефти» могут служить имя (название) включающего ее месторождения и стратиграфическое положение залежи в разрезе; для объекта «месторождение» — имена этого месторождения, включающих его элементов административного (республика, область) и (или) геологического районирования (нефтегазоносный бассейн, металлогеническая провинция и т. п.), а также тип минерального сырья (углеводороды, руды и т. д.). Идентификатор образует ключ<sup>2</sup>, на основе которого формируется адрес конкретного объекта и осуществляется поиск характеризующих его данных в памяти ЭВМ.

Из определения типа объекта вытекает требование однородности описания, заключающееся в том, что все объекты этого типа должны описываться одной и той же совокупностью атрибутов (иначе, по определению это будут объекты разных типов). Частное следствие требования однородности — неизменность правила формирования идентификатора (простого или составного, состоящего из нескольких атрибутов) для всех объектов данного типа. Это значит, что для каждого типа объекта должно быть

<sup>1</sup> Под идентификацией объекта понимается его «различаемость» среди других объектов, «узнаваемость» ЭВМ по ограниченному набору атрибутов-идентификаторов.

<sup>2</sup> Ключами называются соответствующие идентификаторам элементы данных при отображении объекта в ЭВМ.

определено, из скольких и каких атрибутов состоит его идентификатор, так как в СУБД предусмотрен только один алгоритм доступа к данным по каждому типу объекта базы данных. Кроме того, для каждой конкретной СУБД существует ограничение на максимальное количество идентификаторов.

К значению идентификатора предъявляется жесткое требование уникальности, т. е. синонимия имен недопустима. При достаточно большом количестве объектов описания в БД вполне могут встретиться синонимичные (совпадающие по форме знакового выражения) имена. Для составных идентификаторов имя каждого атрибута может не быть уникальным, однако их комбинация этим свойством должна обладать. Например, в разных областях страны или в разных странах могут встретиться месторождения с одинаковым названием. Но если идентификатор типа объекта «месторождение» состоит из атрибутов «имя месторождения» — «имя страны», т. е. является составным, то комбинация имен значений этих атрибутов будет удовлетворять требованию уникальности.

Значения идентификаторов, с одной стороны, должны быть уникальными для АС, с другой — общепринятыми для геологов, так как именно по ним формируется запрос и осуществляется доступ к информации об объекте. Соответственно они должны быть известны широкому кругу пользователей. Однако ряд объектов геологической предметной области не имеет собственных общепринятых имен, которые можно было бы использовать в качестве значений идентификатора. Например, объект «часть бассейна в пределах республики» не имеет собственного имени, но его можно составить из имен республики и бассейна, тогда оно будет и уникальным, и общепринятым, и может служить значением составного идентификатора этого объекта.

Сложнее, если для идентификации некоторого объекта нельзя подобрать общепринятые имена других объектов иного уровня. Например, на многозалежном месторождении нефти (газа) каждая залежь часто не имеет собственного имени, а обозначается обычно названием (возрастом) того продуктивного горизонта, к которому она приурочена. Если на месторождении в одном и том же продуктивном горизонте заключено несколько залежей, то с формальной точки зрения они будут иметь одинаковые имена, что противоречит требованию уникальности идентификатора. Реакция АС на это будет заключаться в том, что при попытке ввести данные по каждой из залежей возникает одна из двух следующих ситуаций: а) данные по первой из залежей вводятся нормально, а по последующим будут не приняты системой, так как в ней уже имеются данные по объекту данного типа с тем же именем, б) система будет принимать данные по всем залежам, но при вводе данных по каждой последующей залежи автоматически будут изыматься данные по предыдущей (АС будет воспринимать данные по последующей залежи как корректировку данных по предыдущей). Поэтому, чтобы ввести в БД данные по всем залежам, необходимо придать им уникальные имена, обеспечив тем самым

уникальность их идентификаторов. Это может быть достигнуто приданием залежам чисто искусственных имен, например «1», «2» и т. д., или «западная», «восточная», «центральная» и т. п. Таким образом, составной идентификатор отдельной залежи будет в рассмотренном случае состоять из имени месторождения, имени (возраста) продуктивного горизонта и искусственно введенного имени залежи.

При таком подходе решается проблема уникальности идентификатора объекта, но возникает другая проблема, заключающаяся в том, что такие уникальные имена известны лишь тому, кто их придумал. Многие пользователи не могут назвать это имя в запросе на информацию. Современные системы способны оказать пользователю помощь в этом вопросе, предоставляя ему посредством режима справок по структуре и составу базы данных (так называемый режим «меню») перечни имен объектов, описанных в базе данных (см. разд. 3.4.3).

#### 3.4. ОРГАНИЗАЦИЯ ИНФОРМАЦИОННОГО ОБЕСПЕЧЕНИЯ АФИС

Организовать информационное обеспечение АФИС — значит реализовать все его элементы, показанные на рис. 27: построить модель данных (пройти путь все возрастающей степени структуризации и формализации информации от информационного до логического уровня организации данных); выбрать СУБД адекватную модели данных; разработать единую систему входных документов (форм фиксации исходной информации) и систему классификаторов и кодов; сформировать массивы данных на входных документах в соответствии с устанавливаемыми правилами (инструкцией) их заполнения, сформировать систему типовых информационно-справочных запросов (и соответствующие типы выходных документов, в которых фиксируются результаты реализации запросов), осуществить содержательную постановку аналитических задач и сценария их решения.

Общаясь с АС, конечный пользователь — геолог самое непосредственное отношение имеет именно к информационному обеспечению. Он в той или иной мере, прямо или косвенно участвует (должен участвовать) в реализации всех элементов информационного обеспечения, но особенно велика его роль в формировании системы входных документов, классификаторов и массива данных. Здесь возникают проблемы, связанные, во-первых, с тем, что исходная геологическая информация часто не удовлетворяет требованиям к представлению данных и их организации (см. разд. 3.2, 3.3) и, во-вторых, с необходимостью выработки компромисса между особенностями геологической информации, с одной стороны, и правилами ее представления и ограничениями, накладываемыми ЭВМ и СУБД — с другой.

### 3.4.1. Особенности геологической информации

На современном этапе эффективность геологических исследований и успешность геологоразведочных работ в большей степени определяются тем, насколько полно используется накопленная в результате геологической деятельности информация, как отечественная, так и зарубежная. Все возрастающий интерес к последней обусловлен тем, что в широко применяемом в геологии методе аналогий в качестве аналогов используются геологические объекты различных регионов мира, а также тем, что геологические структуры «не признают» государственных границ. В связи с этим, наряду с общей спецификой геологической информации [42], при формировании базы данных приходится учитывать подчас существенные различия в принятых в разных странах системах описания геологических объектов и измерения (фиксации) значений их характеристик.

Специфика геологической информации обусловлена низкой степенью формализованности геологического языка, что отчетливо обнажается на концептуальном уровне организации данных и предопределяет важность формализации исходной информации.

Это конкретно проявляется в следующем:

отсутствие логически корректно оформленных понятий, обеспечивающих однозначность понимания и выделения однотипных геологических объектов (см. пример 2, гл. 2);

подведение под один термин существенно различных предметов исследования геологии (см. примеры 1, 5, гл. 2);

обозначение однотипных предметов различными терминами, что детально рассмотрено на примерах песчаных аккумулятивных тел в дельтовых акваториях в работе [37];

описание однотипных объектов разными (частично совпадающими) наборами показателей, что весьма характерно, например, для так называемых неантиклинальных ловушек нефти и газа [37];

описание объектов разных типов частично перекрывающимися наборами показателей (например, запасами и добычей полезных ископаемых характеризуются залежь, месторождение, страна и др.; разрез может быть характеристикой разведочной площади, месторождения, нефтегазоносного бассейна и его частей и т. д.). Это ведет к избыточности (дублированию) описаний, нерациональному использованию памяти ЭВМ;

использование признаков, которые понимаются неоднозначно субъективно, так как они не сопровождаются заданием меры. Так, например, характер контакта слоев часто описывается такими значениями, как постепенный, резкий, размеры геологического тела — крупное, мелкое, среднее и т. п., а перспективность объекта — как высокая, низкая, средняя. Естественно, что одному исследователю представляется резким, мелким, низким, другим может определяться как постепенный, средний, высокий;

различиями методов определения значений показателя (например, пористости и проницаемости по керну и промыслово-геофизи-

ческим данным абсолютного возраста пород — калий-аргоновым и рубидий-стронциевым методами и др.), что ведет к необходимости фиксировать, хранить и обрабатывать несколько значений одного показателя объекта, которые не образуют составного значения — перечисления (...и...и, см. разд. 3.3), а являются альтернативными (...или...или), в связи с чем не могут использоваться программные механизмы обработки составных значений;

использование в качестве фактографической информации, которая таковой по-существу не является, а представляет собой умозаключение на основе действительно фактографической информации. Умозаключение может быть и неоднозначным. Например, часто указываются условия формирования различных геологических объектов, которые фактически являются выводом, базирующимся на анализе значений того или иного набора регистрируемых вещественных характеристик.

Привлечение зарубежной геологической информации выдвигает дополнительные проблемы. Это связано, во-первых, с тем, что в разных источниках информации часто содержатся несовпадающие значения какого-либо показателя по одному и тому же объекту (например, величина запасов и добычи по месторождению, стране). Возникает необходимость фиксировать несколько значений, указывать приоритетность одного из них (поскольку в модели аналитической обработки может участвовать лишь одно значение), фиксировать несколько источников информации. Во-вторых, различия в национальных системах измерения и классификаций (например, в единицах измерения длины, объемов геофизических работ, классификациях категорий ресурсов и запасов, скважин, марок углей и др.) нарушают требование единства системы измерения значений одного элемента данных. Поэтому необходимы решения по способам фиксации, хранения и обработки такой информации. Из-за того, что такая информация трудносопоставима без определенных преобразований, возникает необходимость осуществления операций по приведению данных к единой шкале, без чего невозможны их суммирование и совместная обработка.

В-третьих, обычная неполнота исходных данных по объектам низшего уровня вынуждает искусственно вводить обобщающие данные как характеристики объектов более высокого уровня. Эти обобщающие данные, в принципе, избыточны, так как их можно вычислять по исходной совокупности при условии, что она полная. Например, из-за обычного отсутствия описания всех месторождений какого-либо бассейна и соответственно тех или иных данных по ним, в частности запасов и добычи, приходится эти параметры независимо вводить и при характеристике бассейна в целом, вместо того чтобы получать их суммированием по отдельным месторождениям.

### 3.4.2. Формирование массива данных

Технология формирования массива данных в ЭВМ состоит в получении информации непосредственным измерением и описанием

параметров геологических объектов в полевых или лабораторных условиях, а также извлечением ее из различных опубликованных и неопубликованных источников, представлении информации в виде данных, занесении этих данных в макеты входных документов и переносе с них на машинные носители информации (перфокарты, магнитную ленту) и, наконец, в непосредственном вводе данных в ЭВМ. Последние звенья этой технологической цепи несколько различны для пакетного и диалогового (интерактивного) режимов общения с АС. При диалоговом режиме в принципе можно обойтись без заполнения бумажных бланков входных документов, используя, например, так называемую масочную технологию. Суть ее заключается в том, что макет (маска) входного документа вызывается на экран дисплея, и заполнение макета данными осуществляется также непосредственно на экране.

Реализация технологии формирования массива данных обеспечивается разработкой системы входных документов, классификаторов, кодов для унификации фиксации символьных данных, инструкции по заполнению входных документов, а также инструкцией по подготовке данных на машинных носителях. Геологи, непосредственно участвующие в формировании массива данных, должны строго руководствоваться указанными нормативными материалами. Последний из них предназначен не столько для геологов, сколько для работников информационно-вычислительного центра, где АС эксплуатируется в пакетном режиме.

Разработка состава и формулировка названий показателей (имен элементов данных) в соответствии с принципами и приемами формализации исходной геологической информации (см. гл. 2) и требованиями к форме ее представления (см. разд. 3.3) позволяет решить большую часть проблем, связанных с неформализованностью геологического языка (см. разд. 3.4.1). Состав показателей во входных документах должен достаточно полно описывать предметную область в соответствии с разработанной моделью данных, что обеспечивает возможность широкого комплекса приложений (задач), и обязательно включать атрибуты-идентификаторы объектов предметной области.

В целях экономии ресурсов памяти ЭВМ желательно фиксировать во входных документах только собственно фактографическую исходную информацию, а для получения обобщающих характеристик (или интерпретаций) включать в систему алгоритмы их генерации (например, алгоритм интерпретации условий осадконакопления по литолого-физическим, геохимическим, палеонтологическим параметрам породы). Однако в связи с отмечавшейся возможной неполнотой исходных данных (особенно по зарубежным объектам) приходится включать во входные документы необходимое количество обобщающих показателей. По этой же причине, а также потому, что не всегда удается приведение различных единиц измерения одного и того же показателя к единой системе (это требование к представлению данных), необходимо предусмотреть возможность фиксации его значений в различных единицах изме-

рения. Чтобы не нарушать указанное требование, приходится искусственно вводить вместо одного показателя два (см. 3.3.1). Чтобы зафиксировать объемы гравirazведки в отрядо-месяцах и в физических точках, нужно образовать два показателя с разными наименованиями: «объемы гравirazведки в отрядо-месяцах» и «объемы гравirazведки в физических точках».

Включение в состав показателей идентификаторов объектов обеспечивает привязку данных, содержащихся во входном документе, не только к типу ОПО, отображенному в модели данных, но и к конкретному геологическому объекту.

При разработке набора, формы и структуры входных документов необходимо учитывать ряд требований, выполнение которых снимает некоторые негативные моменты, связанные с особенностями геологической (в том числе зарубежной) информации и способствует упрощению заполнения входных документов, а также более эффективному вводу данных и использованию ресурсов ЭВМ. Эти требования состоят в обеспечении минимального дублирования заносимых во входные документы показателей, максимальной заполненности входных документов, минимизации числа входных документов, соответствия формы и структуры входных документов удобству заполнения и возможности прямого ввода данных (машинно-ориентированные формы входных документов).

Существуют различные подходы к организации системы входных документов. В ряде АС каждый показатель (атрибут какого-либо объекта) фиксируется в отдельном входном документе. Этим достигается максимальная заполненность последнего при наличии значения показателя (ибо при его отсутствии документ вообще не заполняется). Но при этом идентификаторы, привязывающие данный показатель к объекту, дублируются столько раз, сколько показателей имеет объект, и столько же раз вписывается наименование показателя. Если же наименование показателя внесено в документ заранее, то количество входных документов резко возрастает, так как оно соответствует числу показателей, описывающих предметную область.

Возможна система объектно-специализированных входных документов, когда каждому типу ОПО соответствует отдельный документ, содержащий максимально полный набор атрибутов, описывающих объект. Таким образом минимизируется число входных документов (оно равно числу типов объектов) и дублирование идентификаторов объектов, но зато увеличивается дублирование показателей, так как многие объекты имеют общие показатели. Кроме того, уменьшается заполненность документов, так как далеко не для всех показателей имеются значения.

По-видимому, оптимальной является система входных документов, сочетающая как объектно-специализированные, так и унифицированные, т. е. содержащие наборы показателей, описывающих объекты нескольких типов, документы. Именно такая система документов сформирована при разработке по ВНИИзарубежгеологии информационного обеспечения АФИС-Геоинформ (рис. 35).





что обеспечивает упрощение заполнения входных документов. Такой формой является табличное представление показателей, т. е. сведение близких по смыслу показателей в таблицы. Они весьма удобны и привычны для геолога, так как и при традиционной обработке геологической информации, т. е. без применения автоматизированных систем, преобладающая часть ее сводится им в таблицы. Однако непосредственный ввод данных с традиционных для геологии таблиц в ЭВМ, как правило, невозможен. Поэтому при пакетном режиме общения с АС данные из таблиц приходится переписывать в специальные бланки, с которых уже осуществляется перенос данных на машинные носители. В связи с этим во многих современных АФИС стремятся разработать такие входные документы, которые, с одной стороны, близки к традиционным формам сбора и фиксации данных и потому удобны для геологов, а с другой стороны, структурированы таким образом, что позволяют осуществлять непосредственный ввод с них данных в ЭВМ. При этом, как правило, требуется написание специальных прикладных программ ввода, но такой подход эффективней, чем постоянная служба по переносу данных, так как требует не систематических, а одноразовых затрат. Такими компромиссными могут быть документы со специальной структурой.

Структура входных документов может быть различной (см. рис. 35, 36), но, как правило, в них постоянно есть две основные части: в одной содержатся данные, описывающие конкретный ОПО (наименования атрибутов и их значения), а в другой — идентификаторы объекта, к которому эти данные относятся, и типа документа (который определяет тематический аспект набора показателей). В документе обычно содержатся специальные указания, помогающие фиксировать разнообразную информацию и облегчающие ее ввод и последующий поиск в базе данных.

Рассмотрим основные элементы структуры входных документов на примере АФИС-Геонформ (см. рис. 35). Указанная АФИС ориентирована на доступную любому геологу технологию ввода данных через экран дисплея. В связи с этим и форма входных документов идентична экранному изображению (маске), предназначенному для ввода данных. На форме входных документов сказываются соответствующие размеру экрана дисплея ограничения по компоновке и числу полей, в которых указаны названия показателей и полей для записи значений этих показателей.

Все атрибуты-идентификаторы объединены в заголовок документа, предшествующий таблице данных. В нем фиксируется тип входного документа (M01, M11...), тип объекта описания (см. рис. 33), а также «адрес» объекта описания. Адрес объекта не только привязывает данные к конкретному геологическому объекту указанием имени (названия) последнего в соответствующем поле. Поскольку элементы адреса отображают состав и структуру типов ОПО, они устанавливают и взаимоотношения между объектами. Так, указание в адресе документа, описывающего характеристики месторождения, не только имени месторождения, но и

имен бассейна и страны, отображает связи «месторождение — бассейн», «месторождение — страна», «бассейн — страна».

Таблица данных представляет собой одно- или двумерную таблицу, в которой указаны наименования атрибутов (показателей), системы измерения и записи их значений (размерность показателя) и выделены поля для записи значений. В заголовок документа введено несколько полей для специальных указаний.

Поле «использование данных» предназначено для обеспечения возможности фиксации, ввода и хранения нескольких значений одного и того же показателя, полученных из разных источников или разными методами измерения. В этом случае на объект заполняется несколько экземпляров одного и того же документа. В том экземпляре, где зафиксировано приоритетное значение, в данном поле делается соответствующее указание, например, записывается «обработка» (ОБ), а в остальных экземплярах с другими значениями — иные указания, например «архив-1», «архив-2», (A1, A2) и т. д., если таких значений несколько.

Для того чтобы иметь возможность описывать геологические объекты, не вошедшие в структуру объектов, указанную на рис. 31 (например, отдельные тектонические зоны бассейна, части месторождения и т. д.), в заголовок введено поле «№ листа...». При одинаковых идентификаторах объекта порядковый номер, записанный в это поле, будет отличать экземпляры одного и того же документа, описывающие разные части одного объекта. Поле «дата» предназначено для указания года, которым датируются данные, зафиксированные в документе.

Для некоторых специальных указаний, фиксация значений которых неформализована, введено специальное поле «библиография и примечания» (см. рис. 35, а), где могут быть записаны источники информации, любые замечания заполняющего (например, оценка достоверности данных и др.).

Во всех полях для записи значений идентификаторов, данных, специальных указаний изображением линейки для записи введены ограничения на длину значения (максимальное количество знаков), указано положение запятой, разделяющей целую и дробную части вещественных чисел. Введенные ограничения на длину записи согласованы с длинами записываемых значений из классификаторов и служат одним из средств контроля правильности заполнения. Подобная структура входных документов приближает их к машинно-ориентированному виду, что позволяет проводить ввод данных в ЭВМ непосредственно с этих документов.

При заполнении входных документов соблюдаются единые правила, которые оформляются в виде «Инструкции по заполнению входных документов». Эти правила включают описание структуры входных документов конкретной АФИС; системы знаков, которые можно использовать при записи значений; системы классификаторов и кодов для записи нечисловых значений; способы записи значений различных показателей.

Вопросы разработки классификаторов для задания областей

значений символьных данных рассматривались в разд. 3.3. Здесь отметим лишь, что они должны быть составлены для всех атрибутов, значения которых не являются числовыми, в том числе и для атрибутов-идентификаторов (классификаторы стран, бассейнов, административных единиц, месторождений и т. д.). Для тех идентификаторов, которые не имеют общепринятых имен (например, названия ловушки, залежи нефти и газа), необходимо задать правила их образования. В качестве классификаторов могут использоваться и уже существующие общепризнанные классификации (периодическая система элементов Менделеева, частично геохронологическая таблица и др.).

Способы записи значений показателей могут быть различными в зависимости от описываемой предметной области и от используемого программного и технического обеспечения. Но во всех случаях указывается: максимальное число знаков для записи значений каждого показателя; способ записи значения, когда число знаков в нем меньше максимально установленного; место записи знака «—» в отрицательных числах; максимальное число перечисляемых элементарных значений в составных значениях и (в необходимых случаях) знак-разделитель, которым одно элементарное значение отделяется от другого (запятая, косая черта либо иной знак) и др. Необходимо отметить также, что при записи значений показателей во входные документы могут использоваться лишь те символы, которые имеются в алфавите ЭВМ. Это необходимо иметь в виду и при разработке или подборе классификаторов. Так, для использования в качестве классификатора геохронологической таблицы принятые в ней индексы палеогена (P) и кембрия (C), обозначенные знаками, не содержащимися ни в русском, ни в латинском алфавите и не имеющими стандартных кодов в ЭВМ, необходимо заменить на знаки, обеспечиваемые ЭВМ, например, палеоген — PG, кембрий — CM.

Только строгое буквальное соблюдение единых правил при заполнении входных документов позволит сформировать массив данных таким образом, что будут обеспечены как правильность и быстрота ввода, так и корректность обработки данных при удовлетворении информационно-справочных запросов и автоматизированном решении аналитических задач.

### 3.4.3. Формирование информационно-справочных запросов к АС

Обработка запросов на поиск и выдачу информации — основной режим работы АФИС. Конкретные запросы, которые могут быть сформулированы на содержательном (геологическом) уровне, столь же разнообразны, как и профессиональные интересы различных геологов-пользователей АФИС, и поэтому просто не поддаются перечислению и систематизации. Тем не менее анализ опыта применения АФИС в геологической предметной области позволяет выделить наиболее типовые запросы, систематизировать их и представить в формализованном виде (не на содержатель-

ном, а на технологическом уровне). Прежде всего типовые запросы могут быть условно подразделены на три группы: I — запросы по составу и структуре базы данных; II — запросы на данные, характеризующие объекты предметной области; III — запросы на числовую обработку данных. Ниже рассматриваются содержание и специфика запросов каждой группы.

I. Запросы по составу и структуре базы данных. Возможности различных пользователей корректно сформулировать запрос (а это обязательная предпосылка получения запрашиваемой информации) зависят от степени знакомства и опыта работы пользователя с данной конкретной АФИС. Пользователя, впервые обращающегося к АФИС, интересует прежде всего, какие типы геологических объектов (ОПО) охарактеризованы в базе данных, каков набор показателей, описывающих каждый тип ОПО, какими классификаторами (названия и состав каждого из них) фиксируются значения показателей. Это все информация, отражающая структуру базы данных. Только зная структуру БД, геолог-пользователь может составить представление о том, какой фрагмент, срез предметной области отражает (информационно моделирует) данная АФИС. Это предопределяет круг принципиально возможных запросов. Для того чтобы эти запросы были привязаны к конкретным геологическим объектам, желательно знать имена (идентификаторы) объектов определенных типов, например, нефтегазоносных, угленосных бассейнов, металлогенических провинций, месторождений нефти, угля, руд и т. п., информация по которым имеется в базе данных. Ведь прежде чем затрачивать время на формулировку запроса на поиск конкретных (предусмотренных системой) данных по месторождению X или бассейну Y или области Z, лучше убедиться, что информация по этим ОПО введена в АФИС. Это уже запросы по составу базы данных.

Обращение к АС с запросами по составу и структуре базы данных важно не только для того, чтобы представить себе, какую информацию можно получить от данной АС, но и для того, чтобы общаться с последней на понятном ей языке. Любая АФИС может воспринять запрос, только если он сформулирован на том же языке (в тех же именах и терминах), каким описана в ней предметная область. Это совершенно естественно, поскольку в основе поиска данных лежит операция сличения на похожесть текста запроса с текстом описания предметной области на логическом уровне в базе данных (имена элементов данных — показателей, значений их в классификаторах). Если пользователь запрашивает те или иные данные, например, по Волго-Уральской нефтегазоносной области, а она описана в БД частями — Прикаспийская впадина, Тимано-Печорский бассейн и др., то он получит ответ, что таких данных не имеется, хотя фактически они есть и могут быть получены при иной формулировке запроса. Такая же ситуация возникает, если запросить данные, например, по возрасту продуктивного горизонта на нефтяном месторождении X, если этот показатель описан как «стратиграфическая приуроченность».

Для того чтобы удовлетворить потребность в информации по составу и структуре БД, СУБД должна обладать специальными программными сервисными средствами. Если их нет, то прямое общение геолога-пользователя с АФИС невозможно или сильно затруднено. В таких случаях администратор базы данных, хорошо знающий ее состав и структуру, принимает от геолога запрос на информацию (не формализованный) и сам переводит его на язык, доступный восприятию данной АФИС. Но большинство современных СУБД оснащено специальными программными средствами, обеспечивающими режим «меню». Так называется форма общения пользователя с АФИС, когда последняя представляет пользователю полные списки своих возможностей.

Существуют различные варианты этого режима, но для восприятия структуры и наполнения баз геологических данных представляется оптимальным «меню» древовидной структуры, применяемое в случаях, когда списки имен объектов и наименований показателей очень большие. В этом случае списки упорядочивают в виде древовидного (иерархического) разбиения на классы, подклассы, группы и т. п. Например, список месторождений, охарактеризованных в АФИС, может превышать многие сотни (а залежей, рудных тел — тысячи). Просмотр их трудоемок и утомителен. В этом случае они могут быть сгруппированы по бассейнам (провинциям), областям, республикам и т. п. Геологу, желающему быстро уточнить наличие и точное название в БД интересующего его месторождения, целесообразно просмотреть более краткие списки сначала бассейнов, а затем месторождений, относящихся к выбранному бассейну, сузив таким образом область поиска. Принципиально также решается вопрос и с названиями показателей, которые предварительно были сгруппированы в тематически однородные и потому удобные для просмотра группы.

Таким образом, запросы по составу и структуре базы данных сводятся к следующим основным четырем типам:

выдать перечень типов объектов предметной области (ОПО), описанных в БД;

выдать перечень имен ОПО конкретного типа (например, залежь или месторождение или бассейн и др.), по которым имеются данные в БД;

выдать перечень показателей, характеризующих ОПО конкретного типа;

выдать перечень имен ОПО конкретного типа, по которым в БД имеется информация по указываемым показателям (или названию группы показателей)<sup>1</sup>;

<sup>1</sup> Входные документы составляются часто таким образом, что каждый из них объединяет набор показателей, характеризующих какой-либо один тематический аспект и потому рассматриваемых совместно (в общем контексте). Соответственно выделяемые в древовидном меню группы показателей часто отражают содержание того или иного входного документа. Поэтому частным случаем запроса данного типа является запрос на выдачу перечня имен ОПО конкретного типа, по которым в БД вводился входной документ № X.

выдать классификатор значений показателя.

II. Запросы на данные, характеризующие объекты предметной области. Независимо от того, какого типа структуру данных использует АФИС (в том числе и реляционную), геолог-пользователь все равно мыслит категориями предметно-объектной организации мира и в соответствии с этим формирует свои запросы. Поэтому формирование запросов этой группы заключается в указании комбинации реквизитов ОПО и значений этих реквизитов. Если посмотреть на это на уровне не конкретных реквизитов, а типов последних, то возможный набор этих типов, участвующих в формировании запросов, относительно невелик. Он включает:

*OB* — тип объекта предметной области. Набор типов ОПО, по которым способна выдавать информацию данная АФИС, устанавливается на концептуальном уровне организации данных.

*A* — адрес (идентификатор) объекта предметной области, на основании которого осуществляется различие конкретного ОПО среди других ОПО. Адрес обязательно включает, наряду с другими возможными параметрами ОПО, уникальное имя (имена) собственное.

*P* — совокупность поименованных характеристик ОПО (в частности, тип входного документа), по которым осуществляется поиск (отбор) последних.

*Q* — ограничения, накладываемые на величину значений характеристик ОПО. Эти ограничения сводятся к трем ситуациям: =, <, >. Например, интересуют такие ОПО, как месторождения (*OB* — месторождение) с добычей нефти (*P* — добыча нефти) более 30 млн. т/год (*Q* > 30 млн. т/год).

*D* — ограничения, накладываемые на значения реквизитов по дате, к которой это значение относится (для ряда показателей, значения имеют смысл лишь применительно к определенной дате: например, нет добычи, запасов минерального сырья вообще, а есть лишь на конкретную дату — год).

*Z* — совокупность поименованных характеристик ОПО, значения которых должны быть выданы по запросу.

Указанными типами реквизитов исчерпываются в основном наиболее простые типовые запросы рассматриваемой группы. Но на практике весьма велика доля более сложных запросов, в формулировке которых участвует несколько типов ОПО, находящихся в той или иной взаимосвязи. В качестве примера можно привести следующий запрос: выдать имена металлогенических провинций (нефтегазоносных или угленосных бассейнов), в пределах которых расположены месторождения с добычей руд (нефти, угля) более 10 млн. т/год. При таком запросе, чтобы получить данные (имена) по тому типу ОПО, что нас интересует (провинция, бассейн), необходимо прежде обратиться к ОПО иного типа — месторождению, проанализировать его характеристики, значения последних и уже через них выйти на непосредственно интересующие нас ОПО. Условием назвать ОПО, по которому (которым) затребованы итоговые данные, *основным*, а ОПО, через который (которые) определяются параметры основного ОПО, — *транзитивным*. С учетом транзитивных ОПО общее число типов реквизитов, могущих участвовать в запросе, увеличивается, а типизация запросов базируется на отображении той или иной комбинации ти-

Номер типа запроса	Основной объект				Транзитивный объект				В результате реализации запроса выдаются		
	Тип <i>OB</i>	Адрес <i>A</i>	Наименования показателей <i>P</i>	Ограничения		Тип <i>ob</i>	Адрес <i>a</i>	Наименования показателей <i>p</i>		Ограничения	
				на значения показателей <i>Q</i>	по дате <i>D</i>					на значения показателей <i>q</i>	по дате <i>d</i>
1	+	+								Значения <i>Z</i>	
2	+	+								»	
3	+	+		+						»	
4	+		+							»	
5	+		+	+						»	
6	+		+							<i>A</i> (имена)	
7	+		+							»	
8	+				+					Значения <i>Z</i>	
9	+				+					»	
10	+							+		»	
11	+							+		»	
12	+							+		<i>A</i> (имена)	
13	+							+		»	

## Примеры запросов на поиск данных

№ типа запроса	Формулировка типа запроса	Пример запроса на содержательном уровне
1	Выдать всю имеющуюся информацию по идентификатору объекта ( <i>OB</i> , <i>A</i> )	Выдать все, что известно ( <i>Z</i> ) по месторождению ( <i>OB</i> ) Лансоник ( <i>A</i> )
2	Выдать совокупность данных или отдельное данное по нескольким (в частном случае — одному) объектам по: идентификаторам объектов ( <i>OB</i> , <i>A</i> ) наименованиям показателей ( <i>Z</i> )	Выдать по бассейнам ( <i>OB</i> ) Северо-Каспийскому и Тимано-Печерскому ( <i>A</i> ) данные по максимальной глубине фундамента, типам рельефа и инфраструктуре ( <i>Z</i> )
3	То же, но с ограничениями по дате, накладываемыми на значения показателей	Выдать по месторождениям ( <i>OB</i> ) Саликово и Михнево ( <i>A</i> ) данные по добыче нефти и количеству эксплуатационных скважин ( <i>Z</i> ) в 1985 г. ( <i>D</i> )
4	Выдать совокупность данных или отдельное данное ( <i>Z</i> ) по: идентификатору типа объектов ( <i>OB</i> ) наименованиям показателей объектов того же типа ( <i>P</i> ) ограничениям, накладываемым на значения этих показателей ( <i>Q</i> )	Выдать тип залежей (по фазовому состоянию углеводородов ( <i>Z</i> ) по месторождениям ( <i>OB</i> ), в которых глубина продуктивных горизонтов ( <i>P</i> ) превышает 5 км ( <i>Q</i> )
5	То же, но дополнительно с ограничениями по дате	Выдать максимальную глубину продуктивных горизонтов ( <i>Z</i> ) по месторождениям ( <i>OB</i> ), на которых добыча нефти ( <i>P</i> ) превышала 10 млн. т ( <i>Q</i> ) в 1985 г. ( <i>D</i> )
6	Выдать перечень идентификаторов объектов ( <i>A</i> ) по: идентификаторам типа этих объектов ( <i>OB</i> ) наименованиям показателей, характеризующих эти объекты ( <i>P</i> ) условиям, накладываемым на значения показателей ( <i>Q</i> )	Выдать имена ( <i>A</i> ) бассейнов ( <i>OB</i> ), в которых максимальная глубина фундамента < 4 км и имеются крупнейшие месторождения нефти ( <i>P</i> , <i>Q</i> )
7	То же, но с дополнительным ограничением по дате	Выдать имена ( <i>A</i> ) бассейнов ( <i>OB</i> ), в которых добыча свободного газа ( <i>P</i> ) составляла > 50 млрд. м <sup>3</sup> ( <i>Q</i> ) в 1985 г. ( <i>D</i> )

№ типа запроса	Формулировка типа запроса	Пример запроса на содержательном уровне
1	2	3
8	Выдать совокупность данных или отдельное данное по нескольким (в частности — одному) объектам по: идентификаторам типов объектов (OB) идентификаторам конкретных транзитивных объектов (ob, a)	Выдать данные о возрасте продуктивных горизонтов (Z) на месторождениях (OB) Северо-Каспийского и Тимано-Печорского (a) бассейнов (ob)
9	То же, но с дополнительными ограничениями по дате	Выдать данные о добыче нефти (Z) в 1985 г. (D) на месторождениях (OB) Северо-Каспийского (a) бассейна (ob)
10	Выдать совокупность данных или отдельное данное по нескольким (в частности — одному) объектам по: идентификаторам типа этих объектов (OB) идентификаторам типа связанных с ними транзитивных объектов иного уровня (ob) наименованиям показателей транзитивных объектов (p) условиям, накладываемым на значения показателей транзитивных объектов (q)	Выдать данные по мощностям осадочного выполнения (Z) бассейнов (OB), в которых расположены месторождения нефти (ob) с начальными извлекаемыми запасами (p) > 100 млн. т (q)
11	То же, но с дополнительными ограничениями по дате (d)	Выдать данные по суммарным начальным ресурсам нефти (Z) областей, в которых расположены месторождения нефти (ob) с добычей нефти (p) более 5 млн. т (q) в 1985 г. (d)
12	Выдать перечень идентификаторов объектов (A) по: идентификаторам типа этих объектов (OB) идентификаторам типа связанных с ними транзитивных объектов иного уровня (ob) наименованиям показателей транзитивных объектов (p) условиям, накладываемым на значения показателей транзитивных объектов (q)	Выдать имена (A) бассейнов (OB), в которых расположены месторождения (ob) с глубинами продуктивных горизонтов (p) более 5 км (q)

№ типа запроса	Формулировка типа запроса	Пример запроса на содержательном уровне
1	2	3
13	То же, но с дополнительными ограничениями по дате (d)	Выдать имена (A) бассейнов (OB), в которых имеются месторождения (ob) с добычей свободного газа (p) в 1985 г. (d) более 1 млрд. м <sup>3</sup> (q)

пов реквизитов. Параметры типовых запросов приведены в табл. 12 и 13.

Следует отметить, что для простоты восприятия принципа формализации типов запросов мы рассматривали выделенные типы реквизитов как единые и неделимые. На практике иногда приходится некоторые из них делать комплексными. В частности, как видно из входных документов, приведенных на рис. 33—36, тип ОПО определяется там по сочетанию значений не только OB (в указанном выше содержании), но и дополнительного параметра FU — физико-географические условия (территория, акватория). Это позволяет обособить такой тип ОПО, как отдельно территория и акватория бассейна (области республики, страны). Последнее важно в связи с бурным развитием морских геологоразведочных работ на нефть и газ.

III. Запросы на числовую обработку данных. В рамках информационно-справочной функции АФИС возможно весьма ограниченное число типов запросов этой группы. Основные из них два:

подсчет числа объектов предметной области, отвечающих определенным условиям (например, выдать количество нефтегазоносных бассейнов, в которых имеются месторождения нефти с разведанными запасами > 30 млн. т и со стратиграфической приуроченностью продуктивных горизонтов к кайнозою);

суммирование значений характеристик объектов, отвечающих определенным условиям (например, выдать общие разведанные запасы газа всех бассейнов, расположенных в пределах древних платформ).

Среди запросов к АФИС на числовую обработку, кроме вышеуказанных, наиболее часты такие, как подсчет среднего, подсчет процентов, расчет на одну единицу и др.

Из приведенных примеров видно, что, прежде чем проводится собственно счет, осуществляется поиск ОПО на той же основе (по соответствующей запросу комбинации типов реквизитов), что при реализации запросов группы II.

Отнесение запросов на числовую обработку к информационно-справочным, а не аналитическим, несколько условно. Это связано с тем, что в результате реализации этих запросов получается новая информация, что характерно для аналитической функ-

ции. Но поскольку такие запросы теснейшим образом связаны с подлинно информационно-справочными, элементарны с вычислительной точки зрения и возможности их реализации заложены в большинстве современных СУБД, то их принято рассматривать в рамках информационно-справочной функции АС.

Форма представления результатов реализации информационно-справочного запроса называется выходным документом. Многие из них представляют собой стандартные элементы периодически подготавливаемых справок, сводок, отчетов (например, о состоянии разведанных запасов и добычи полезных ископаемых, о количестве, категориях крупности, состоянии освоения месторождений минерального сырья). Такие выходные документы заранее макетируются (их вид и содержание согласовывается с пользователем), а запрос на них оформляется как стандартный, т. е. программа его реализации включается в число стандартных программ АС. При каждом повторении запроса этого типа уже не требуется составлять программу его реализации, а достаточно лишь вызвать соответствующую стандартную программу и сообщить ей изменяющиеся каждый раз переменные параметры запроса.

Выше уже указывалось, что в качестве пользователя АФИС может выступать прикладная программа, реализующая решение аналитической задачи. Эта программа уже предопределяет, какие данные и в каком порядке должны быть представлены на ее входе. Запросы на данные, исходные для той или иной аналитической задачи, принципиально ничем не отличаются от рассмотренных выше типов запросов. Как правило, формирование массива данных для аналитической задачи требует последовательной реализации запросов нескольких типов.

#### 3.4.4. Информационные аспекты автоматизированного решения аналитических задач

Если при реализации информационно-справочной функции АС формируются выходные структуры элементов данных без преобразования значений последних, то аналитическая функция АС направлена на образование новых значений элементов данных. Эти новые значения получаются на основе преобразования исходных данных с учетом установленных или предполагаемых зависимостей между исходными и искомыми данными, заключенных в самих процедурах обработки. Для этого АС должна иметь средства программного и математического обеспечения, выходящие за рамки универсальной СУБД, реализующей в основном информационно-справочную функцию, и специальным образом организованное информационное обеспечение, рассмотрению структуры которого посвящен настоящий раздел.

При формировании информационного обеспечения автоматизированного решения аналитических задач выполняются разработки, которые по последовательности условно можно подразделить на

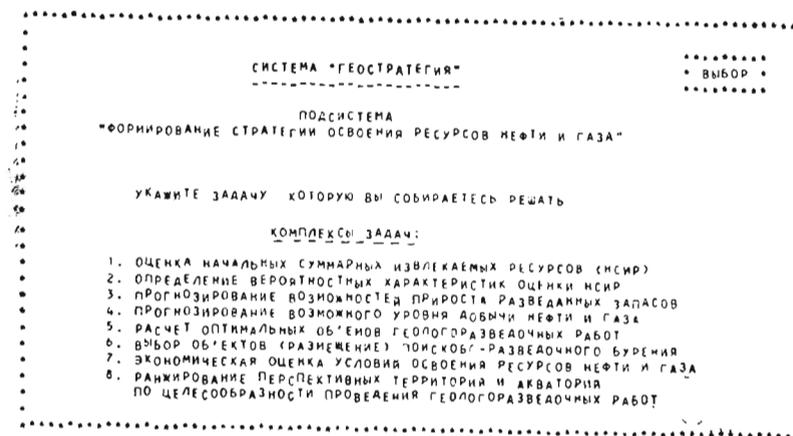


Рис. 37. Сценарная маска автоматизированной системы

две стадии: формирование состава и моделей решения задач и разработка сценариев решения задач.

Первая стадия включает прежде всего обоснование комплекса содержательно взаимосвязанных геологических (геолого-экономических задач), на решение которых ориентирована АС аналитической обработки данных. Слишком широкий тематический охват решаемых задач делает АС менее эффективной по многим позициям. Поэтому обычно предпочтение отдается четко проблемно-ориентированным АС, имеющим вполне определенных постоянных<sup>1</sup> пользователей, которые нацелены на достижение конкретного конечного результата.

В качестве примера такой проблемно-ориентированной АС можно назвать разработанную во ВНИИзарубежгеологии диалоговую человеко-машинную систему «Геостратегия», ориентированную на решение комплекса аналитических задач по формированию оптимальной стратегии освоения ресурсов полезных ископаемых. Полное содержательное обеспечение подсистемы «формирование стратегии освоения ресурсов нефти и газа» включает восемь тесно взаимосвязанных комплексов задач (рис. 37). На примере АС мы хотели бы проиллюстрировать реализацию стремления к взаимосогласованности отдельных комплексов задач, в совокупности обеспечивающих достижение цели.

Хотя в данной АС безусловно обеспечивается возможность независимого входа в любой из комплексов задач, общая структура содержательного обеспечения такова, что между отдельными комплексами задач сохраняется значительный элемент преемственности. Так, входными данными для комплекса задач № 2 — «вероят-

<sup>1</sup> АС обычно включают задачи, подлежащие многократному, а не разовому решению.

ностная оценка ресурсов нефти и газа» являются результаты реализации комплекса задач № 1 «оценка начальных суммарных ресурсов нефти и газа». В свою очередь, результат решения комплекса задач № 2 используется при решении практически всех других комплексов задач. Подобно этому, при прогнозировании возможного уровня добычи нефти и газа (комплекс задач № 4) в качестве исходной основы используется результат прогнозирования прироста запасов нефти и газа (комплекс задач № 3), а этот прогноз, в свою очередь, определяется применительно к тому выбору объектов поисково-разведочного бурения, который сделан при решении комплекса задач № 6. Этот же выбор в значительной степени базируется на результатах решения комплексов задач № 1 и 2. При реализации комплекса задач № 8 — «ранжирование по целесообразности проведения геологоразведочных работ на нефть и газ» используются результаты, полученные в процессе решения всех других комплексов задач.

Учитывая, что содержательная ориентация АС аналитической обработки геологических данных (комплексов решаемых задач) весьма разнообразна, ниже мы ограничимся рассмотрением общей для каждой задачи технологии информационного обеспечения автоматизированного решения.

Информационное обеспечение отдельной задачи начинается с ее содержательной постановки, сущность которой кратко охарактеризована в разд. 2.3. Она является основой для разработки математической модели задачи, определяющей в свою очередь выбор вычислительного метода решения. Следует отметить, что содержательная постановка и математическое описание задачи, как и само решение, для многих задач могут быть осуществлены и вне связи с АС вообще и с какой-либо конкретной АС в частности. Модели всех или большинства аналитических геологических задач, на решение которых ориентирована конкретная АС, в принципе могут быть заимствованы извне. Поэтому, если говорить о технологии информационного обеспечения узко применительно к АС аналитической обработки, то она начинается со второй стадии — с разработки сценария решения задачи. Состав осуществляемых на этой стадии разработок зависит от режима общения пользователя с АС и режима решения задачи (см. разд. 3.1.3). Наиболее ограничен он при простом (без человеко-машинной технологии) пакетном режиме: вопрос сводится в основном к выбору формы представления исходных данных и формы представления результатов решения на содержательном уровне, т. е. проинтерпретированных в терминах геологического языка. В этом случае переход от первого ко второму полностью скрыт от конечного пользователя и организуется в рамках математического и программного обеспечения. В наиболее полном виде информационное обеспечение организуется при диалоговом режиме общения, реализующем человеко-машинную технологию решения задачи. При этом возникает необходимость в разработке сценария диалога.

Выбор формы представления исходных данных заключается в разработке макета таблицы (для многих задач — это несколько типов таблиц), куда заносятся исходные для задачи данные в заранее predetermined составе и порядке расположения. Если комплекс аналитической обработки связан специальной программой-интерфейсом с управляемой СУБД базой геологических данных, то макет таблицы каждого типа фактически представляет собой макет выходного документа АФИС, пользователем которого выступает прикладная программа, реализующая решение аналитической задачи. Если такой связи нет, то макет таблицы заполняется пользователем перед сдачей заказа на решение (при пакетном режиме) или непосредственно в процессе решения (при диалоговом режиме). В последнем случае заполнение проводится на экране, куда выводится макет таблицы. Но в любом случае входы алгоритмов и программ решения задачи жестко привязаны к форме представления исходных данных.

При диалоговом режиме общения с АС выбор формы представления исходных данных есть частный случай более общей задачи разработки сценария диалога. В соответствии с обоснованным в разд. 3.1.3. принципом системно-управляемого диалога, ориентированного на слабоподготовленного конечного пользователя, интерактивное взаимодействие последнего с АС должно быть заранее спланировано, а результат этого планирования оформляется в виде сценария решения, который представляется в форме взаимосвязанных набора сценарных масок и графа диалога.

Решение задачи прежде всего должно быть расчленено на части, реализуемые АС и пользователем. При этом на АС возлагаются не только те или иные фрагменты собственно решения, но и определенные (и не менее важные) функции, связанные с тем, что в системно-управляемом диалоге АС принадлежит более активная роль по сравнению с конечным пользователем. Это значит, что АС должна направлять конечного пользователя в рамках заранее намеченной канвы решения, своевременно сообщая ему информацию, необходимую для понимания сущности задачи, ожидаемых от него действий (информации), характера и способа принятия требуемых решений. Соответственно пользователь взаимодействует с АС, не только воздействуя непосредственно на процесс решения (путем выбора из альтернатив хода решения, сообщения системе какой-либо необходимой для решения информации), но и в сопутствующем решению режиме обучения.

Диалоговый режим накладывает на это одно чисто техническое ограничение. Поскольку диалог АС и пользователя осуществляется через экран дисплея, то информация, как сообщаемая АС пользователю, так и исходящая от последнего, должна быть разбита на смысловые фрагменты, по возможности уместимые на экране. Такие фрагменты называются *сценарными масками*. Оговорка «по возможности» сделана в связи с тем, что иногда приходится сталкиваться с единой в смысловом отношении информацией, не уместяющейся на экране. Таковы чаще всего крупные

таблицы, имеющие большое число строк и столбцов. Разбиение такой информации (таблицы) на самостоятельные фрагменты в соответствии с изобразительными возможностями (площадью) экрана<sup>1</sup> приводит к отрыву значений показателей от наименований последних и (или) имен объектов, характеризующих значениями показателей, что в конечном итоге делает информацию трудновоспринимаемой или вообще не воспринимаемой пользователем. Для таких случаев разработаны специальные программные средства, позволяющие «перелистывать» неумещающуюся информацию, т. е. смещать ее вправо-влево и вверх-вниз, выводя на экран ту или иную ее часть по желанию пользователя.

Сценарные маски в зависимости от смыслового содержания предопределяют три формы диалогового общения пользователя с АС.

**Восприятие информации.** Каждая соответствующая маска сообщает пользователю ту или иную информацию, не предполагая какой-либо ответной реакции (односторонне направленная информация). Таковы, например, сценарные маски, знакомящие пользователя с содержательной постановкой и математической моделью задачи, объясняющие, какая информация и в каком виде от него потребуется, каковы стадии решения задачи (рис. 38)<sup>2</sup>.

**Выбор альтернатив.** Значительное число сценарных масок предлагает пользователю осуществить тот или иной выбор из перечисленного возможного набора действий. Уже в самом начале работы с АС пользователю приходится выбирать ту или иную задачу (см. рис. 37), затем осуществлять выбор из того или иного набора условий или действий (рис. 39). Этот выбор практически осуществляется постановкой принятого в данной АС знака (например, звездочки) в специально выделенном поле экрана. На рис. 37, 39 такие поля обозначены точками слева от наименования возможных альтернатив.

**Запись.** Соответствующие маски предполагают запись в специально отведенные в них поля исходящей от пользователя и предусмотренной сценарием информации. Это может быть заполнение таблиц исходных данных, ввод значений по каким-либо дополнительным параметрам (ограничениям) задачи и другие случаи включения данных непосредственно в процессе решения. На рис. 40 приведен пример сценарной маски из задачи «Выбор объектов поисково-разведочного бурения на нефть и газ», где пользователю предлагается сделать запись — указать основные ограничения задачи — по метражу и затратам.

Полный набор сценарных масок одной задачи охватывает все допустимые варианты решения, число которых может быть значительным. Каждый конкретный вариант, определяемый (составляе-

<sup>1</sup> 20—24 строки, по 80 знаков в строке (считая пробелы между словами или цифровыми значениями).

<sup>2</sup> Здесь и далее примеры сценарных матриц взяты из сценариев решения задач, включенных в упоминавшуюся выше человеко-машинную систему «Геостратегия».

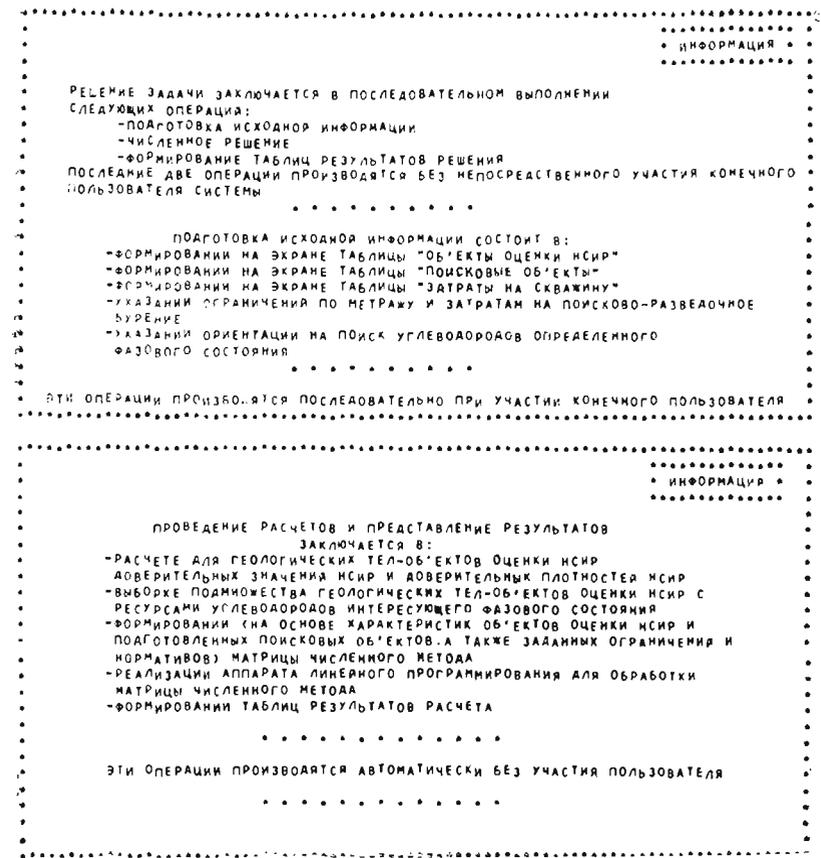


Рис. 38. Примеры сценарных масок, предназначенных для информирования (обучения) пользователя

мый) пользователем в процессе решения, путем осуществляемого многократно по ходу решения выбора из возможных альтернатив, предопределяет те или иные набор и последовательность сценарных масок, т. е. при реализации единичного варианта решения в диалоге участвуют далеко не все сценарные маски. Соответствие между каждым возможным вариантом решения и отвечающими ему набором и последовательностью сценарных масок отражается в виде *графа диалога*. Последний представляет собой обычный древовидный граф, подобный тому, что строится при отображении иерархической структуры данных (см. рис. 29, 31). Вершинам его соответствуют конкретные сценарные маски. Реализация различных вариантов решения задачи соответствует при этом переходу по различным ветвям древовидного графа.

Наряду с формированием системы сценарных масок и графа диалога, важным элементом сценарной стадии информационного

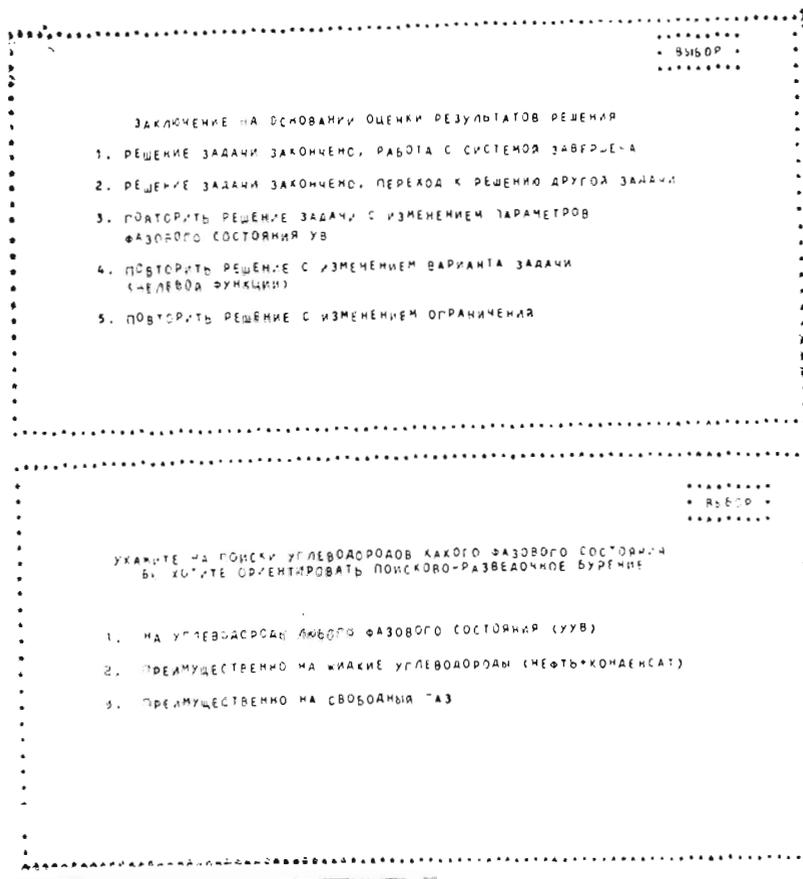


Рис. 39. Примеры сценарных масок, предлагающих выбор альтернативы

обеспечения автоматизированного решения аналитических задач является разработка формы представления результатов решения задачи на языке геолога-пользователя. Результат реализации вычислительной процедуры представлен обычно большим набором чисел, содержательно истолковать которые конечный пользователь, как правило, не в состоянии. Это доступно лишь прикладному математику, знакомому с моделью решения задачи. Возникает необходимость представить результат решения в форме, понятной конечному пользователю и в виде, пригодном для осмысления им результата в чисто содержательном (геологическом) аспекте. Для этого необходимо, во-первых, из всей совокупности цифровых данных (отчета АС о реализации вычислительной процедуры) выделить те, что могут быть содержательно интерпретированы, т. е. отделить те, что характеризуют технологичность, оптимальность реализации вычислительной процедуры, и, во-вторых, поставить им в соответствие названия показателей, значениями которых они

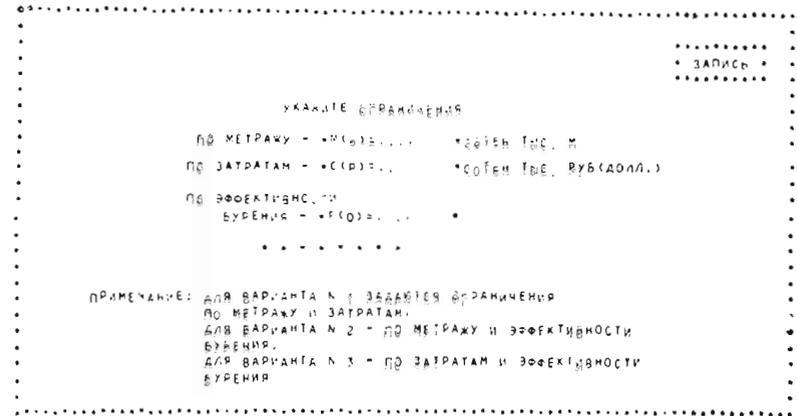


Рис. 40. Пример сценарной маски, предназначенной для включения пользователем данных в процессе решения задачи

являются. Итогом этой работы является макет таблицы (или нескольких таблиц) результатов решения задачи. Заполнение их результирующими данными производится автоматически.

Все итоги разработки сценария в той или иной форме трансформируются затем в рамках проблемно-ориентированных средств комплекса программно-математического обеспечения АС. В частности, АС должна хранить знание о содержании и форме каждой сценарной маски, включая макеты таблиц исходных данных, реализовывать логику графа диалога (иметь программы, управляющие строго определенной последовательностью смены масок при различных вариантах задачи) и др. Сценарий решения задачи служит исходной основой для формирования соответствующих программно-математических средств на уровне разработчика АС.

Таково принципиальное содержание информационного обеспечения автоматизированного решения аналитических задач, ориентированного на конечного пользователя — геолога. Функции последнего в общении (через дисплей) с диалоговой АС аналитической обработки данных заключается в следующем:

- выбор интересующей задачи из комплекса задач, решаемых АС;
- ознакомление с содержательной постановкой и математическим описанием задачи и организацией ее решения (см. рис. 37). Пользователь, знакомый с задачей, уже решавший ее на данной АС, может эту стадию миновать;
- формирование массива исходных данных (заполнение таблиц на экране или обращение за данными к сопряженной АФИС);
- ввод по ходу решения указываемых системой дополнительных параметров (ограничений) решения и выбор альтернатив частных вариантов решения;
- анализ и оценка результата решения задачи<sup>1</sup> и принятие решения о завершении или повторении решения с изменением тех или иных параметров (см. рис. 39).

<sup>1</sup> Организация самого вычислительного процесса не является функцией конечного пользователя и потому должна быть скрыта от него. После указания исходящей от пользователя управляющей информации ему представляется для осмысления содержательно интерпретированный результат реализации вычислительной процедуры.

Специфика рассмотренной проблемы формализации геологических данных, а в более общем плане и геоинформатики в целом — многоаспектность, тесная связь с широким комплексом методологических, организационных и психологических проблем, решение которых часто требует выхода за пределы геологической предметной области. В этих условиях практическая отдача от приложения средств и методов геоинформатики зависит не только от многочисленных частных (хотя и исключительно важных) методических подходов, но и в не меньшей степени от общей концепции развития геоинформатики. Важно акцентировать внимание на двух наиболее принципиальных положениях, которые по убеждению авторов должны лежать в основе такой концепции и которыми руководствовались авторы при работе над этой книгой.

1. Прогресс на магистральном пути интенсификации и повышения эффективности геологоразведочной деятельности в конечном итоге определяется возможностями повышения надежности получаемых геологических выводов (прогнозов), оптимальностью принимаемых на их основе решений. Судя по тематике научно-исследовательских и опытно-конструкторских разработок в области геологии, эта кардинальная цель может быть достигнута посредством, во-первых, совершенствования технических средств и методов изучения (регистрации параметров) недр и, во-вторых, широкой математизацией и компьютеризацией геологоразведочной деятельности. Развитие средств и методов изучения недр обусловлено современным состоянием информационной базы геологии, в частности предшествующим предметным знанием (выше уже отмечалось, что система понятий определяет систему наблюдений). Все это справедливо по отношению к математическим методам и ЭВМ в геологии: система понятий — предметов исследования и заданных связей между ними определяет эффективность системы обработки данных. Без правильной логической организации (формализации) последних самое широкое внедрение математических методов и ЭВМ выльется в подмену подлинного научно-технического прогресса его чисто внешними атрибутами. Поэтому альтернативы организации геологических знаний на более строгой логической основе, обеспечивающей, в частности, возможность корректной формализации данных, не существует. И первое принципиальное положение, которое авторы стремились раскрыть в данной работе, заключается в том, что одна из важнейших предпосылок научно-технического прогресса — это развитие (организационное, тематическое и функциональное обособление) геоинформатики. До настоящего времени эта важнейшая сфера геологоразведочной деятельности недооценивалась.

2. Одна из важнейших функций геоинформатики, которой непо-

средственно посвящена данная работа, — разработка требований к логической организации геологической информации и соответствующих методических рекомендаций. Учитывая объем и традиционный способ хранения геологических данных, реализация этих рекомендаций возможна лишь силами геологов, непосредственно создающих и обрабатывающих геологические данные традиционными для геологии средствами, т. е. геологов самых разных специальностей, как правило, не имеющих необходимых знаний по методологии науки, информатике, прикладной математике и компьютерной технологии. Восприятие сущности и значимости геоинформатики таким геологом не может быть привнесено извне — с позиции других (точных) наук, а должно быть сформировано в недрах самой геологии, исходить из сознания насущных потребностей ее развития. Иных путей сделать основы геоинформатики достоянием массового геолога нет. Поэтому второе принципиальное положение заключается в необходимости ориентации основ геоинформатики на возможности восприятия именно массовым геологом.

Стремление следовать этому положению наложило определенные ограничения как на форму изложения, так и на содержание настоящей работы. Прежде всего авторы сознательно стремились к изложению на описательном уровне, соответствующем профессиональному языку геологии, даже тех материалов, которые традиционно представляются на более строгом языке формальной логики и математики. Во-вторых, авторы не стремились к исчерпывающей полноте охвата проблемы и предельной глубине проработки поднятых вопросов, чтобы множество частных не заглохло цельное видение проблемы и основных средств и методов ее разрешения. Тем не менее, изложенные материалы вполне достаточны не только для понимания проблемы на концептуальном уровне, но и для оказания геологам практической помощи в совершенствовании организации исходных геологических данных, что является необходимой предпосылкой повышения надежности результатов их обработки (в том числе с использованием математических методов и автоматизированных систем).

## СПИСОК ЛИТЕРАТУРЫ

1. *Атре Ш.* Структурный подход к организации баз данных (пер. с англ.). Финансы и статистика, 1983.
2. *Бергер М. Г., Вассоевич Н. Б.* Геологическая терминология. М. Изд-во МГУ, 1974.
3. *Бойко В. В., Савинков В. М.* Проектирование информационной базы автоматизированной системы на основе СУБД. М., Финансы и статистика, 1982.
4. *Воронин Ю. А., Еганов Э. А.* Методологические вопросы построения автоматизированных систем управления в геологии. — В кн.: Методологические проблемы геологии. Киев, Наукова думка, 1972, с. 91—99.
5. *Воронин Ю. А., Еганов Э. А.* Методологические вопросы применения математических методов в геологии. Новосибирск, Наука, 1974.
6. *Геологический словарь.* т. 1, II. М., Недра, 1973.
7. *Горбатов В. А., Павлов П. Г., Четвериков В. Н.* Логическое управление информационными процессами. М., Энергоатомиздат, 1984.
8. *Горский Д. П.* Определение. М., Мысль, 1974.
9. *Готт В. С., Семенюк Э. П., Урсул А. Д.* Категории современной науки. М., Мысль, 1984.
10. *Груза В. В.* Методологические проблемы геологии. Л., Недра, 1977.
11. *Дементьев Л. Ф.* Математические методы и ЭВМ в нефтегазовой геологии. М., Недра, 1983.
12. *Денинг В., Эссиг Г., Маас С.* Диалоговые системы «Человек — ЭВМ». Адаптация к требованиям пользователя. М., Мир, 1984.
13. *Диалектический материализм* — методологическая основа теоретического естествознания. Киев, Наукова думка, 1976.
14. *Драгунов В. И., Айнемер А. И., Васильев В. И.* Основы анализа осадочных формаций. Л., Недра, 1974.
15. *Еганов Э. А.* О выделении объектов исследования в геологии. В кн.: Пути познания Земли. М., 1971, с. 263—272.
16. *Жуков Р. А., Груза В. В., Ткачев Ю. Р.* Проблемы выделения объектов исследования в геологии. — В кн.: Математические методы исследований в геологии. Л., 1975, вып. 3, с. 14—27. (Труды ВСЕГЕИ, Новая серия, т. 223).
17. *Кадастр* зарубежных стран, обладающих природными ресурсами нефти и газа, т. 1. Л., Недра, 1983. (Труды ВНИИЗарубежгеологии, вып. 40).
18. *Карогодин Ю. Н.* Седиментационная цикличность. М., Недра, 1980.
19. *Кондаков Н. И.* Логический словарь-справочник. М., Наука, 1975.
20. *Кренделев Ф. П., Кренделев С. Ф.* Эвристические методы в геологии. М., Наука, 1977.
21. *Круть И. В.* Исследование оснований теоретической геологии. М., Наука, 1973.
22. *Купцов В. И.* Роль философии в научном познании. М., Знание, 1976.
23. *Логвиненко Н. В.* Петрография осадочных пород. М., Высшая школа, 1974.
24. *Мартин Дж.* Организация баз данных в вычислительных системах. М., Мир, 1980.
25. *Месторождения* нефти и газа зарубежных стран. Справочник, М., Недра, 1976.
26. *Месторождения* нефти и газа СССР. Геология нефти и газа. Справочник, том 2, кн. 1. М., Недра, 1968.
27. *Методы* теоретической геологии, И. И. Абрамович, Ю. К. Бурков, В. В. Груза и др. Л., Недра, 1978.
28. *Мовшович Э. Б.* Системный подход к выделению нефтегазопосконых объектов. — В кн.: Системные исследования в геологии каустобиолитов. М., 1984, с. 107—135.
29. *Мовшович Э. Б.* Значение системно-структурных исследований для познания закономерностей размещения нефтегазопосконых объектов. — В кн.: Системные исследования в геологии каустобиолитов. М., 1984, с. 25—33.
30. *Мовшович Э. Б., Демчук В. М.* Моделирование выбора объектов поисково-разведочного бурения на нефть и газ. М., Изд. ВИЭМС, 1985. (Экспресс-информация. Сер.: Математические методы и автоматизированные системы в геологии, вып. 8).
31. *Оленин В. Б.* Нефтегеологическое районирование по генетическому принципу. М., Недра, 1977.
32. *Олле Т. В.* Предложения КОДАСИЛ по управлению базами данных. М., Финансы и статистика, 1981.
33. *Омелин В. М.* Диалоговая система как средство интерпретации геолого-геофизических данных: модель, принципы построения. — В кн.: Диалоговые системы и обработка графической информации в АСУ — Нефтегазразведка. Л., 1982, с. 5—25.
34. *Онопrienko В. И.* Функции и условия формализации геологических теорий. — В кн.: Применение математических методов и ЭВМ при поиске полезных ископаемых. Новосибирск, 1973, с. 27—39.
35. *Определение* достоверности прогнозной оценки ресурсов нефти и газа. Л., Недра, 1982.
36. *Применение* в геологии АФИС с интегрированной базой данных реляционного типа/Е. Н. Исаев, Э. Б. Мовшович, Л. Г. Олейниченко и др. М., Изд. ВИЭМС, 1985. (Обзор ВИЭМС. Сер.: Математические методы и автоматизированные системы в геологии).
37. *Принципы* выявления зон фашиального контроля нефтегазонакопления/Э. Б. Мовшович, М. Н. Кнепель, Л. И. Несмеянова, Л. А. Польстер. М., Недра, 1981.
38. *Система* эвристического прогнозирования геологических явлений в режиме диалога «Геолог — ЭВМ»/Г. И. Каратаев, Г. М. Беляков, В. В. Терлецкий, Л. И. Лобовкин. — В кн.: Формально-эвристические методы прогноза нефтегазонакопления. Минск, 1978, с. 21—36.
39. *Системный* подход в геологии (теоретические и прикладные аспекты). М., МИНХ и ГП, 1983.
40. *Системные* исследования в геологии каустобиолитов. М., Наука, 1984.
41. *Смирнова А. Я.* Построение автоматизированных фактографических информационно-поисковых систем в геологии. М., Недра, 1976.
42. *Смирнова А. Я.* Информационный анализ в геологии. М., Недра, 1985.
43. *Трофимчук А. А., Карогодин Ю. Н., Мовшович Э. Б.* Методологические вопросы геологии нефти и газа. Новосибирск, Наука, 1983.
44. *Философский* энциклопедический словарь. М., Советская энциклопедия, 1983.
45. *Чумаченко Б. А., Власов Е. П., Марченко В. В.* Системный анализ при геологической оценке перспектив рудоносности территорий. М., Недра, 1981.
46. *Шарапов И. П.* Логический анализ некоторых проблем геологии. М., Наука, 1977.
47. *Штофф В. А.* Проблемы методологии научного познания. М., Высшая школа, 1978.
48. *Шедровицкий Г. П.* Принципы и общая схема методологической организации системно-структурных исследований и разработок. — В кн.: Системные исследования. Методологические разработки. Ежегодник 1981. М., 1981, с. 193—207.

# ОГЛАВЛЕНИЕ

Предисловие	3
Глава 1 <b>Информация и информатика в структуре геологической деятельности</b>	8
1.1. Предпосылки математизации и компьютеризации геологической деятельности	8
1.2. Типы и процедуры преобразования фактографической информации в процессе геологической деятельности	13
1.3. Информатика в геологической деятельности: функциональный и организационный аспекты	30
Глава 2. <b>Методологические основы формализации геологической информации</b>	38
2.1. Идеализация и формализация. Их значение и функции в геоинформатике и геологическом познании	38
2.2. Организация понятийно-терминологической базы	44
2.2.1. Логические основы обеспечения сопоставимости геологических объектов и характеризующей их информации	47
2.2.2. Методологические основы построения геологических классификаций	61
2.2.3. Согласованность объектов и признаков	69
2.3. Содержательная постановка геологических задач	75
2.3.1. Соответствие предметов исследования решаемой задаче	75
2.3.2. Связность параметров	77
2.3.3. Согласованность данных по условиям измерения и масштабу	81
2.4. Согласование исходных данных с требованиями формального (математического) аппарата обработки	85
2.4.1. Обеспечение статистической однородности	85
2.4.2. Представительность и независимость исходных данных	87
2.5. Некоторые методические приемы организации и формализации исходных геологических данных	89
2.5.1. Информационный приоритет элементарных предметов геологии	90
2.5.2. Независимое поэлементное описание геологических объектов	91
2.5.3. Типизация геологических задач	99
Глава 3. <b>Организация геологических данных в автоматизированных системах</b>	108
3.1. Общие принципы организации обработки информации в системах	109
3.1.1. Типы систем по их назначению	109
3.1.2. Технологическая структура систем	111
3.1.3. Информационно-технологические аспекты общения пользователя с АС	117
3.2. Основы организации баз геологических данных	123
3.2.1. Концепция базы данных	125
3.2.2. Средства моделирования системы объектов предметной области в базе данных	127
3.2.3. Типы структур данных	130
3.2.4. Уровни организации данных в АФИС	139
3.3. Требования, предъявляемые АС к форме представления исходных данных	147
3.3.1. Представление информации в виде данных	147
3.3.2. Идентифицируемость объектов базы данных	157
3.4. Организация информационного обеспечения АФИС	159
3.4.1. Особенности геологической информации	160
3.4.2. Формирование массива данных	161
3.4.3. Формирование информационно-справочных запросов к АС	170
3.4.4. Информационные аспекты автоматизированного решения аналитических задач	178
Заключение	186
Список литературы	188

## ПРОИЗВОДСТВЕННОЕ ИЗДАНИЕ

Эдуард Борисович Мовшович

Марк Наумович Кнепель

Михаил Семенович Черкашин

## ФОРМАЛИЗАЦИЯ ГЕОЛОГИЧЕСКИХ ДАННЫХ ДЛЯ МАТЕМАТИЧЕСКОЙ ОБРАБОТКИ

Редактор издательства *А. М. Антокольская*  
Обложка художника *В. П. Христинина*  
Художественный редактор *Г. Н. Юрчевская*  
Технический редактор *Е. Л. Закашанская*  
Корректор *И. П. Розанова*

ИБ № 6799

Сдано в набор 08.12.86. Подписано в печать 30.01.87. Т-07630. Формат 60×90<sup>1/16</sup>.  
Бумага офсетная № 1. Гарнитура Литературная. Печать высокая. Усл. печ. л. 12,0.  
Усл. кр.-отт. 12,25. Уч.-изд. л. 13,93. Тираж 3280 экз. Заказ 1881/1036-2. Цена 80 коп.

Ордена «Знак Почета» издательство «Недра»,  
125047, Москва, пл. Белорусского вокзала, 3.

Московская типография № 6 Союзполиграфпрома при Государственном комитете СССР  
по делам издательств, полиграфии и книжной торговли.  
109088, Москва, Ж-88, Южнопортовая ул., 24.