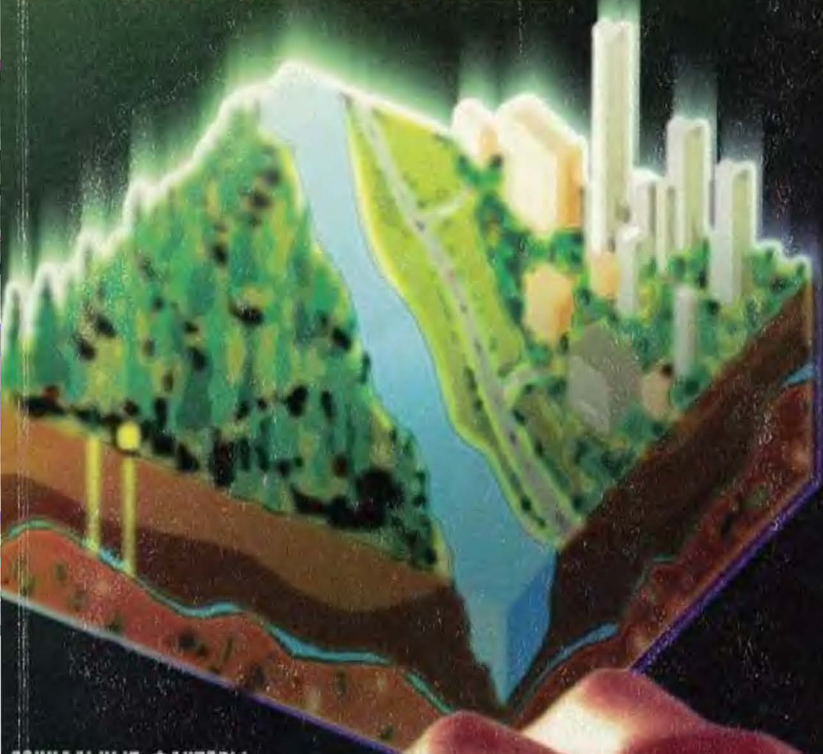


# Ю.К.КОРОЛЕВ



СОЦИАЛЬНЫЕ ФАКТОРЫ  
БИЗНЕС-ГЕОГРАФИЯ  
ЗЕМЛЕПОЛЬЗОВАНИЕ  
ИНЖЕНЕРНЫЕ ОБЪЕКТЫ  
ЭКОЛОГИЯ  
ПРИРОДНЫЕ РЕСУРСЫ



ЧАСТЬ I

ТЕОРЕТИЧЕСКАЯ  
ГЕОИНФОРМАТИКА  
ВЫПУСК 1



DATA+

**Ю.К. Королев**

# **Общая геоинформатика**

**Часть I**

**Теоретическая геоинформатика**

**Выпуск I**

Издательство ООО СП Дата+

МОСКВА

1998

редактор В.Гохман  
технический редактор В.Процкий

**ISBN 5-7312-0260-5**  
Издательство "DATA+"

Лицензия ЛРН№090157 от 25 февраля 1998 г.  
совместно с ДО"ГЛАГОЛЬ"

© "DATA+", 1998  
© Ю.Королев, 1998  
© А.Михайлов, оформление, 1998

# Содержание

|   |           |
|---|-----------|
| <b>Предисловие.</b>   | <b>6</b>  |
| <b>О структуре книги.</b>   | <b>11</b> |
| <b>Часть 1. Введение в понятие пространственных данных.</b>   | <b>21</b> |
| 1. Введение в пространственные информационные системы.  | 21        |
| 1.1. Организация пространственных данных.   | 21        |
| 1.2. Различные применения информационных систем для работы с пространственной информацией.                          | 32        |
| 1.2.1. Сферы применения ГИС.  | 33        |
| 1.2.2. Примеры различных требований к данным.   | 35        |
| 1.3. Некоторые компоненты ГИС.  | 39        |
| 1.3.1. Представление о ГИС как об инструментальном средстве.  | 41        |
| 1.3.1.1 Пространственные информационные системы и ГИС - соотношение этих понятий.                                   | 45        |
| 1.3.2. Физический компонент ГИС.  | 47        |
| 1.3.3. Программный компонент ГИС.   | 47        |
| 1.3.4. Данные как компонент ГИС.  | 48        |
| 2. Анализ потребностей. Задачи. Типы проблем, связанные с пространственной информацией.                             | 48        |
| 2.1. Проблемы и задачи при работе с пространственной информацией.   | 49        |
| 2.2. Информация положения и формы (пространственная в узком смысле слова) и информация атрибутивная (описательная). | 57        |



|  |           |
|--|-----------|
| 2.3. Измерения и пространственные взаимоотношения.   | 62        |
| 2.3.1. Свойства пространственных объектов.   | 63        |
| 2.3.2. Пространственные взаимоотношения. Различие между геометрической и топологической информацией (характеристиками). Взаимосвязь между топологическими и метрическими характеристиками. | 68        |
| 2.4. Категории проблем, возникающих при работе с пространственной информацией.   | 72        |
| 2.4.1 Типы пространственных проблем и типы организации пространственных данных.  | 72        |
| <b>Категории пространственных проблем.</b>   | <b>74</b> |
| 2.4.2 Другие аспекты.  | 80        |
| 2.5. Примеры комплексных задач.  | 83        |
| 2.5.1. Прогнозирование месторождений полезных ископаемых.  | 83        |
| 2.5.2. Моделирование инженерных сетей.   | 88        |
| 2.6. Основные категории средств и подходов, применяемых при решении пространственных проблем с помощью ГИС.  | 90        |
| 2.6.1 Классификация функций и инструментов ГИС.  | 91        |
| 2.6.2. Функции ввода данных.   | 91        |
| 2.6.3. Функции вывода и представления данных.  | 93        |
| 2.6.4. Функции обработки и анализа информации, администрирования данных.   | 94        |
| 2.6.5. Функции ГИС и типы применений ГИС.  | 94        |

|  |     |
|--|-----|
| 2.7 Элементы моделей пространственных данных, их взаимоотношения. Модели данных и их разновидности моделей данных. | 97  |
| 2.7.1 О моделях данных, структурах моделей данных и форматах файлов.   | 97  |
| 2.7.2. О представлении индивидуализированных объектов.   | 102 |
| 2.7.3. Об описательных характеристиках индивидуализированных объектов - атрибутах.                                 | 106 |
| 2.7.4. Формы представления непрерывных признаков - полей.  | 108 |
| 2.7.5. О некоторых требованиях к представлению пространственных данных.  | 109 |
| 2.7.6. Типы моделей пространственных данных - первый обзор.  | 110 |
| 2.7.7. Модели построения взаимоотношений между атрибутивной и пространственной графической информацией.            | 111 |
| 2.7.8. Модели организации пространственных объектов, организации связей между ними.                                | 112 |
| 2.7.9. Модели представления графической информации.  | 114 |
| 2.7.10 Подразделение моделей и форматов данных по их назначению.   | 115 |
| 2.7.11. Некоторые дополнительные моменты.  | 116 |
| 2.7.12. Краткое резюме и практические следствия.   | 117 |

## Предисловие.

Эта книга задумана как руководство по общей геоинформатике в первую очередь для широкого круга лиц, сталкивающихся сегодня с геоинформатикой в практическом плане, при освоении того или иного программного средства и его практическом применении для решения различных задач - своих собственных или заказчика. Из своего опыта общения с пользователями геоинформационных систем и близкого по назначению программного обеспечения, как с начинающими, так и с опытными, у автора сложилось определенное представление о том, каковы типичные проблемы, с которыми сталкиваются они на пути быстрого и эффективного практического использования достижений геоинформатики. И это впечатление заключается, в первую очередь, в том, что проблема не в качестве и полноте документации на конкретные программные средства, хотя это и очень важный аспект. Большинство современных программных средств, и западных, и отечественных, сегодня достаточно хорошо документированы. Конечно, трудно помочь тем, кто не желает знакомиться с документацией или пользуется нелегальными копиями программных продуктов, не о них сейчас речь. По-видимому, для многих является актуальной проблема, так сказать, "общеинформационной грамотности", как лет 10-12 назад проблема общекомпьютерной грамотности. Иначе говоря, часто не хватает общего взгляда на предмет, теоретической подготовки. Как минимум три оригинальные учебные пособия на русском языке уже существуют и достаточно широко известны\*. Они, однако, рассчитаны в основном на студентов ВУЗов и увязаны с учебными программами, в которых иногда отсутствуют, хотя бы просто в силу ограниченности объемов существующих курсов, определенные разделы, которые являются, на взгляд автора, необходимыми. Это касается как общих, теоретических разделов, так и еще в большей степени практических вопросов геоинформатики. Автор решил взять на себя смелость попытаться подготовить такой учебный курс, который отвечал бы его представлениям о потребностях сегодняшнего пользователя геоинформационных технологий. При определении его объема и содержания были поставлены следующие задачи. Во-первых, уделить значительное место теоретической части, так как именно этот аспект чаще всего оказывается за пределами документации на программные средства и курсов обучения для пользователей пакетов. Важность такой общей, не

---

\* 1. А.В. Кошкарев, В.С. Тикунов. Геоинформатика. Под ред. Д.В. Лисицкого. М., Картоцентр Геодезиздат, 1993, 213 стр., 2. Н.В. Коновалова, Е.Г. Капралов. Введение в ГИС. Изд. 2, М., ООО Библион, 1997, 160 стр., 3. Н.Н. Филатов. Применение ГИС при изучении окружающей среды. Петрозаводск, КГПУ, 104 стр.

привязанной к конкретным программным средствам подготовки подтверждается сегодняшней практикой. Во-вторых, эту теоретическую часть сделать максимально доступной и интересной одновременно различным группам лиц, интересующихся геоинформатикой. А это сегодня как представители самых различных естественнонаучных и инженерных специальностей, так и программисты, специалисты по информатике - крайне широкий круг лиц, резко отличающихся как по уровню знакомства с компьютерными технологиями и математикой, так и с дисциплинами географического цикла - геодезией, картографией и т.д.. Совместить в одной книге столь различные требования без ущерба для ее качества невозможно, поэтому неизбежны компромиссы, в первую очередь, в отношении строгости изложения. Кроме того, часть материала предполагается выделить в справочные приложения. (Правда, они будут выпущены только по завершении выпуска основных частей книги, см ниже). В некоторых примерах используется термины и понятия, специфичные для той или иной конкретной предметной области, которые могут быть незнакомы читателям, далеким от соответствующей проблематики. Однако автор стремился построить изложение так, чтобы можно было без существенных потерь для понимания основного содержания пропустить такие слишком конкретные, или даже, наоборот, слишком математизированные детали.

Автор, не являясь преподавателем ВУЗа, тем не менее часто встречался и обсуждал проблемы преподавания геоинформатики как с преподавателями, так и со студентами нескольких ВУЗов. Поэтому можно предположить, что книга будет также интересна студентам, встречающимся с геоинформатикой в различных курсах, как целиком посвященных этой дисциплине, так и затрагивающих ее в небольшой степени. Автор постарался организовать это издание так, чтобы уменьшить дублирование с известными ему учебными пособиями на русском языке. Однако следует отметить, что книга не приспособлена специально к учебной программе по геоинформатике в том или ином ВУЗе.

Книга отнюдь не является вполне оригинальной, значительное место занимают заимствования из различной, в основном англоязычной литературы, не переведившейся ранее на русский язык. Эти заимствования в основном носят характер вольного

перевода или пересказа. Но существенное место занимает и оригинальный авторский текст, иногда частично уже публиковавшийся в периодике или материалах конференций. Всегда, имея дело с такой работой, затрудняешься - считать ли ее переводом-компиляцией, или авторской работой. После некоторых колебаний автор все же принял решение подписаться под ней как автор и одновременно автор-составитель. Дело в том, что в нее даже заимствования вошли только те, где позиции автора и автора использованной литературы полностью совпадают.

Полный список использованной литературы будет приведен в конце издания.

Издание задумано выпускать в виде отдельных книг, выходящих по мере подготовки текста. Возможно, такая организация даже удобнее на практике.

О структуре книги. Она делится на две большие части - "Теоретическая геоинформатика" и "Практическая геоинформатика". Автор считает, что такое разделение полезно для изучения предмета. При этом общая теоретическая часть будет в таком выделенном виде полезнее и для тех, кто интересуется именно конкретными, практическими вопросами применения ГИС.

Что такое "Общая геоинформатика"? Автор считает, что общую геоинформатику следует отделять в процессе обучения от специальной геоинформатики. Точно так же полезно проводить определенную черту между теоретической и практической геоинформатикой. Раздел "Теоретическая геоинформатика" в данной книге задуман не как формальный курс, посвященный математическим основаниям геоинформационных технологий. Скорее, это попытка помочь пользователю этих технологий в расширении его эрудиции в сфере геоинформатики, скорее методология, чем конкретные методы. И те несколько более формальные, математические аспекты, которые все же рассматриваются в этом разделе, также имеют целью скорее проиллюстрировать некоторые общие положения, чем дать целостное представление о тех многочисленных разделах математики и информатики, которые составляют основу теоретической геоинформатики.

Практическая геоинформатика имеет дело с конкретными программными пакетами и практиками их применения. Она имеет дело с конкретными программными средствами - их возможностями, интерфейсом, особенностями реализации модели данных, ограничениями, встроенными языками и/или командами. Входит сюда также знание оптимальных областей применения для разных пакетов или модулей, понимание проблемы правильного выбора программных средств. Здесь же рассматриваются конкретные технологии - ввода данных, администрирования пространственных баз данных, технологии создания приложений ГИС, в том числе технологии программирования в среде ГИС. Под изучением практической геоинформатики не подразумеваются исключительно практические занятия за компьютером - здесь набирается достаточно полезного материала и для лекционного курса. Именно в таком ключе будет решена и раздел "Практическая геоинформатика" в данной книге. Это не будет просто пересказом стандартных руководств и собранием практических упражнений, а, скорее, попытка обобщения и осмысления практического опыта.

Автор полагает также, что при обучении полезно выделять (в разные курсы или в разные части одного курса) вопросы общей и специальной геоинформатики. Задачами общей геоинформатики как науки и как учебной дисциплины следует считать сферу установления и применения сущностных, общих закономерностей и особенностей работы с пространственной информацией, независящих от специфических особенностей конкретной области применения. Такое вычленение способствует лучшему пониманию принципиальных, базовых вопросов геоинформатики и дает лучшую основу для дальнейшего развития геоинформатики как науки. Последнее тем более представляется верным, что сегодня все более растет роль геоинформационных технологий как междисциплинарного, интегрирующего средства, которое позволяет эффективно объединять и комплексно изучать самые различные факты и знания на основании пространственного положения и взаиморасположения соответствующих объектов и явлений. Поэтому в сегодняшней геоинформатике все более растет роль понимания общих, принципиальных особенностей работы с пространственной информацией. Параллельно с этим растет и осознание того факта, что геоинформатика - это нечто большее, чем только средство работы с традиционными картами на компьютере.

Геоинформационные технологии внедряются во все большее число областей, и становится ясно, что обслуживание ею наук геолого-географического цикла - может быть, и вовсе не главное поле ее применения. Начинают возникать и очень специальные, нетрадиционные приложения геоинформационных технологий и может быть, с некоторым запозданием, обобщающие и теоретические работы, посвященные особенностям применения методов геоинформатики в тех или иных прикладных областях. То есть начинает развиваться геоинформатика специальная, точнее, множество специальных геоинформатик - геологическая геоинформатика, гидрологическая геоинформатика, геоинформатика транспорта, геоинформатика историческая, муниципальная, геоинформатика бизнеса и т.д.. Этот процесс пока не везде ярко проявлен, но тенденция ясна. Соответственно, на взгляд автора, должно быть организовано и обучение геоинформатике в ВУЗе. (Мы в данном случае рассматриваем модель обучения тех студентов, которые не специализируются собственно на геоинформатике). Курс общей геоинформатики базируется на том или ином курсе общей информатики или компьютерной грамотности и является основным и единым для всех или многих специальностей. Далее следуют курсы геоинформатики специальной, посвященной ее применению в интересах конкретной прикладной области. И желательно, конечно, чтобы все специальные дисциплины (к которым это имеет отношение), изучаемые впоследствии, использовали бы методы геоинформатики в практических работах. К сожалению, сейчас обычно картина не такова. После изучения курса геоинформатики и знакомства с современными программными средствами студент может больше не встретить ни ГИС, ни систему обработки данных дистанционного зондирования, ни GPS больше нигде до окончания ВУЗа. Итак, предлагается схема обучения геоинформатике в 4 этапа: первый - общекомпьютерная грамотность и общая информатика; 2 - общая геоинформатика; 3 - специальная геоинформатика; 4 - практическое применение геоинформационных технологий в курсах конкретных прикладных дисциплин как рядового уже освоенного инструментария. При этом практические занятия должны иметь место повсюду, а лекционная форма занятий - только в составе первых трех этапов. Данная работа посвящена только общей геоинформатике - ее теоретической и практической части.

И теоретические аспекты, и практические в книге рассматриваются по возможности так, чтобы обеспечить независимость основного

текста от особенностей тех или иных конкретных программных систем и воздержаться от их личных оценок. Тем не менее совершенно от них абстрагироваться невозможно, и автор, естественно, в примерах использует в основном те программные средства, которые лично ему лучше знакомы. Это в первую очередь программные продукты фирмы ESRI, Inc. (ARC/INFO, PC ARC/INFO, ArcView, ArcCAD, DAK, ArcExplorer, SDE, MapObjects и т.д.), и фирмы ERDAS, Inc. (ERDAS Imagine, MapSheets и т.д.). Личный опыт автора не исчерпывается знакомством только с этими семействами программных продуктов, и в их обзоре (который будет помещен в одном из последних выпусков) мы постараемся затронуть и многие другие системы из реально распространенных в России или представляющих принципиальный интерес. Но, конечно, данная работа не претендует на то, чтобы быть исчерпывающим справочником по существующим программным средствам.

## **О структуре книги.**

Как уже говорилось, работа будет выпускаться в виде отдельных небольших книг. Общее число и объем каждой из них сейчас еще уточняются.

Вся работа делится на четыре крупные части - теоретическая геоинформатика, практическая геоинформатика, дистанционное зондирование и приложения.

Приведем примерное содержание каждой части (в настоящее время еще возможны изменения, как в сторону изъятия, так и в сторону добавления разделов).

## **Часть 1 - теоретические вопросы геоинформатики**

Введение в понятие пространственных данных.

Организация пространственных данных.

Различные применения информационных систем для работы с пространственной информацией.

Сферы применения ГИС.



Примеры различных требований к данным.

Некоторые компоненты ГИС.

Представление о ГИС как об инструментальном средстве.

Пространственные информационные системы и ГИС - соотношение этих понятий.

Физический компонент ГИС.

Программный компонент ГИС.

Данные как компонент ГИС.

Формализованные и неформализованные знания как компонент ГИС.

Роль компьютерных технологий и автоматизации.

Анализ потребностей. Задачи. Типы проблем, связанные с пространственной информацией.

Информация положения и формы (пространственная в узком смысле слова) и информация атрибутивная (описательная).

Измерения и пространственные взаимоотношения.

Свойства пространственных объектов.

Пространственные взаимоотношения. Различие между геометрическими и топологическими характеристиками. Взаимосвязь между топологическими и метрическими характеристиками.

Категории проблем, возникающих при работе с пространственной информацией.

Типы пространственных проблем - и типы организации пространственных данных.

Категории пространственных проблем.

Основные категории средств и подходов, применяемых при решении пространственных проблем с помощью пространственных информационных систем

**(На этом заканчивается первая книга).**

Семантическая информация. Объекты. Поверхности. Данные.

Элементарные пространственные объекты.

Типы информации, связанные с пространственным объектом.

Метаданные.

Непространственные атрибуты.

Пространственные характеристики объектов.

Размерность пространственных объектов.

Геометрические элементы объектов - для двумерного и трехмерного пространства.

Непрерывно распределенные в пространстве признаки.

Пространственные и непространственные характеристики объектов в совместном рассмотрении.

## **Введение в способы организации пространственных данных в ГИС.**

Множественность способов представления пространственных и семантических данных. Динамичность способов представления - возможности преобразования одного способа представления в другой. Возможные при этом потери и получаемые преимущества.

Качество данных.

Временная компонента.

Позиционирование объекта в различных координатных системах

Непрерывные системы координат (непрерывное пространство).

Дискретные системы координат (дискретное пространство).

Системы привязки локальные и глобальные.

Картографические проекции.

Понятие расстояния.

Математические определения расстояний различных типов.

Способы представления линии. Списки координат и сплайны.

Упрощение (генерализация) линии.

Сглаживание линии.

Понятие сплайна.

Представление линий и полигонов - практические проблемы.

Явное и неявное задание объектов.

Понятие о фракталах и фрактальной геометрии.

Кривые заполнения пространства и размерность пространства.

Путь в пространстве. Связь с пространственной адресацией (системами привязки).

Топология. Графы. Площадные объекты. Порядок описания (ориентировка).

Сетевые структуры и графы.

Топология и выявление топологических некорректностей.

Классификация различных ситуаций на цифровых картах.

Некоторые процедуры, применяемые для выявления ошибок.

Типы площадных объектов.

Вложенность полигонов. Полигоны, содержащие в себе другие изолированные полигоны ("острова").

Соприкасающиеся полигоны. Регионы.

Формирование и хранение данных, описывающих топологические взаимоотношения.

Способы организации данных с делением пространства на регулярные и нерегулярные системы ячеек и иерархий.

Мозаики, двумерные разбиения (растры, сетки). Наборы точек.

Сетки (двумерные разбиения).

Системы точек. Их связь с сетками полигонов.

Масштаб и разрешение.

Геометрия регулярных сеток.

Способ кодирования данных.

Пространственные характеристики в растровой модели данных.

Моделирование поверхностей с помощью регулярных систем точек.

Некоторые варианты организации растровой модели данных.

Переменное пространственное разрешение: модель квадродерева.

Нерегулярные разбиения на основе треугольников.

Области близости. Полигоны Тиссена/Вороного. Домены Дирихле.

Триангуляция как метод разбиения поверхности.

Манипулирование пространственными данными.

Интерполяция и экстраполяция.

Базовые операции над линиями и точками.

Пересечение линий

Пересечение отрезков линий.

Процедура определения, находится ли точка внутри полигона.

Определение центраида (центра тяжести) полигона.

Пространственная статистика для точечных объектов.

Некоторые операции над полигонами.

Пересечение линий и полигона.

Объединение и пересечение полигонов.

Вычисление площадей.

Ситуации с атрибутами, возникающие при пересечении полигонов.

Количественные оценки формы полигона.

Вырезание одним полигоном другого.

Буферные зоны.

Процесс наложения полигонов и механика его выполнения.

Геометрическое преобразование пространственных данных. (трансформирование).

Трансформации сетка - пространственные объекты (растр - вектор).

Растрезация.

Векторизация.

Пространственный анализ

Базовые операции с растром.

Пространственное моделирование для регулярных данных.

Алгебра карт.

Некоторые операции над квадродеревом.

Операции над нерегулярными полигонами и графами.

Построение регионов.

Оверлейные операции (операции наложения карт).

Моделирование в векторной модели.

Интеграция и совместное использование разных представлений данных при решении реальных задач.

Построение структур пространственных баз данных.

Системы управления базами данных.

Стандартные методологические подходы к построению структур баз данных.

Концептуальная модель БД - подход "Объект - его взаимоотношения".

Логическая модель БД. Реляционная СУБД.

Перевод модели "Объект- его взаимоотношения" в реляционную модель.

Некоторые типы взаимоотношений между объектами.

Векторная нетопологическая и топологическая модели данных.

Представление отрезков прямых, полилиний и смешанных линий (включающих разные формы). "Спагетти".

Векторно-топологическая модель данных. систематическое описание.

Моделирование графа.

Моделирование рельефа.

Представление объема.

Регулярная модель данных. (Растр и квадродерево).

Объемы - моделирование через октадерево.

Выбор оптимального представления. Критерии выбора.  
Работа с комплексными, сложными, составными объектами.  
SQL и геоинформатика.  
Пространственные запросы.  
Точка в полигоне.  
Запросы полигоном.  
Запросы регионом.  
Запросы селектором (окружностью, прямоугольником).  
Запросы линией.  
Запросы буфером.  
Запросы многослойные (площадью).  
Запросы расстоянием.  
Запросы нахождения пути.  
Качество данных и вспомогательные моменты организации данных.  
Индексирование в базах данных.  
Пространственный индекс.  
Мультимедиа и ГИС.  
Концепция гиперкарты.  
Снимки и карта в совместном использовании.  
Imaging GIS. Роль растровой информации в ГИС  
Часть 2 - практические аспекты геоинформатики.  
Классификация программных средств геоинформатики  
ГИС и другие типы программного обеспечения.  
ГИС и САПР.  
ГИС и непространственные СУБД.  
ГИС и системы обработки информации (статанализ и др.).  
ГИС и системы моделирования.  
ГИС и графические редакторы.  
ГИС и издательские системы.

ГИС и современное картоиздание.

Электронное издание карт и пространственных БД.

ГИС и системы обработки изображения общего назначения.

ГИС и системы обработки данных дистанционного зондирования.

ГИС и системы управления в реальном времени.

ГИС и системы поддержки принятия решений.

ГИС и комплексные территориальные информационно-аналитические системы.

### **Программирование в ГИС.**

Прикладное и системное программирование.

Коммерческие и заказные программные системы.

Бесплатные и условно-бесплатные программные системы.

Программирование на алгоритмических языках в среде ОС и программирование в прикладной системе.

Компиляторы и интерпретаторы.

Процедурно-ориентированные и объектно-ориентированные языки программирования.

Обмен данными и динамическая связь между задачами.

Клиент-серверная технология.

Инструментальное средство MapObjects.

Концепция ODE - Open Development Environment

Источники данных и проблема ввода данных в ГИС.

Организация ввода пространственной информации

Пространственные базы данного коллективного пользования.

Архитектура клиент-сервер применительно к пространственным базам данных.

Удаленный доступ к пространственным базам данных с использованием телекоммуникаций. Глобальные компьютерные сети (Wide Area Networks). Internet и Intranet.

Картографические сервера для Internet/Intranet.

Каталоги пространственных данных на Internet.

Проблема администрирования пространственных баз данных.

Режим многопользовательского одновременного редактирования и проблема поддержки целостности коллективной пространственной базы данных.

## **Часть 3 - дистанционное зондирование**

## **Часть 4 - приложения.**

Справочная информация по некоторым программным средствам, источникам данных ДЗ, по источникам цифровой картографической информации, по некоторым форматам файлов, применяемых в геоинформатике.

Национальные и международные организации, общества, профессиональные союзы в области геоинформатики.

Мировые периодические издания в области геоинформатики.

Геоинформатика на Интернете.

Разумеется, книга не претендует на монографическую строгость и подробность изложения, и многие из упомянутых тем будут затронуты в ней лишь кратко, в той мере, в какой автор считает это необходимым для учебного курса. Ссылки на использованную литературу в тексте в основном будут опущены, как это нередко допускается в учебной литературе, чтобы сделать текст более читаемым. В основном они присутствуют только в практической части и в приложениях, там, где может потребоваться помощь в поиске дополнительной практически полезной информации.

Предполагается, что читатель знаком хотя бы в небольшой степени с компьютерными технологиями и имеет математическую подготовку на уровне общего курса высшей математики и курса элементарной статистики. Впрочем, более слабая подготовка по математике не является препятствием для понимания основного содержания книги. Кроме того, ряд справочных сведений по математике будет приведен в приложениях. Желательно также, чтобы читатель имел базовую подготовку по геодезии и картографии, причем, ее отсутствие также не будет большим



препятствием, и опять таки, планируется небольшой объем соответствующей информации поместить в приложениях.

Поскольку выход книг, представляющих данное издание, будет осуществляться поэтапно, автор с благодарностью рассмотрит все замечания и пожелания по более полному описанию отдельных разделов, а также предложения по расширению тематики книги. Ваши отзывы, замечания и предложения присылайте по электронной почте, по адресу [yuri@dataplus.msk.su](mailto:yuri@dataplus.msk.su)

# Часть 1. Введение в понятие пространственных данных.

## 1. Введение в пространственные информационные системы.

### 1.1. Организация пространственных данных.

*Информационные системы для работы с пространственной информацией. Понятие о пространственной информации. Информационные системы. Базы данных. Информация. Данные. Знания. Различные формы данных. Данные в реальном мире. Зафиксированные данные. Интерпретированные данные. Закодированные определенным образом. Структурированные или организованные различными способами. Послойный принцип организации пространственной информации. Объектно-ориентированный подход. Физическое представление данных. Карта как модель данных.*

По мере приближения к концу двадцатого столетия, мы вступаем в эпоху впечатляющих достижений в возможностях получения информации о природной и социальной сферах, а также в возможностях использования этих данных для решения научных и практических проблем. Можно сказать, что методы получения информации приобретают все более индустриальный характер. Объемы требуемой современной цивилизации и собираемой ею информации колоссально возрастают, и, естественно, требуют для обеспечения своего рационального использования привлечения современных, базирующихся на компьютерных технологиях средств - как для ее обработки и анализа, так и для организованного хранения, поиска нужной информации и другого манипулирования ею. В противном случае было бы неизбежно наступление информационного кризиса, связанного с утерей способности эффективно использовать имеющуюся информацию. Следствиями такого кризиса явились бы утрата, выпадение из научного и практического оборота части имевшейся информации, избыточное дублирование информации, увеличение расходов на повторный сбор уже имеющейся, но ненаходимой и недоступной информации, увеличение затрат на поиск и получение необходимой информации. В

целом развитие такого кризиса привело бы неизбежно к экономическим проблемам глобального характера, падению роли науки и образования, уменьшению роли развитых стран и сначала к замедлению научно-технического прогресса, а затем и к стагнации и деградации нашей цивилизации за счет уменьшения эффективности управления. Пользуясь образным выражением, можно сказать, что информационный кризис был бы равносителен склеротической закупорке главных сосудов нашей цивилизации. К счастью, те же достижения научно-технического прогресса, в первую очередь, современной электроники и математики, компьютерные технологии, которые сегодня многократно увеличили наши возможности по сбору информации, позволили создать и адекватные все возрастающим потребностям средства для ее хранения, поиска, переработки, распространения и анализа. Вся совокупность этих средств и методов обращения с информацией, с акцентом на те из них, которые имеют универсальный характер, а не определяются спецификой конкретной предметной области, называется информационными технологиями и является предметом рассмотрения информатики (общей информатики). В соответствии со специфическими потребностями конкретных предметных областей развиваются также базирующиеся на достижениях общей информатики ее специальные области, например, геологическая информатика, картографическая информатика, историческая информатика, информатика транспортной отрасли и т. п.. Программные и технические средства, реализующие информационные технологии на практике, очень многообразны. Те из них, которые предназначены для обеспечения доступа к информационным ресурсам - ввода информации, хранения ее, модификации, осуществления поиска необходимой информации и ее представления в нужном виде, называются информационными системами (информационно-поисковыми системами, ИПС). Понятие "информационная система" необязательно относится только к сфере компьютерных технологий. Существуют и аналоговые, некомпьютерные информационные системы, манипулирующие информацией в аналоговой, а не в цифровой форме. Например, библиотека - организованное хранилище книг, снабженное системой каталогов, инструкций и правил пользования, реализующее некоторые процедуры контроля состояния (инвентаризацию и реставрацию фонда) и обновления (пополнение фонда) - тоже информационная система. Аналоговая информационная система не

обязательно оперирует с человекочитаемой информацией на бумажных носителях. Например, организованный архив видеофильмов - кассет в формате VHS (Video Home System - стандарт для бытовых видеомагнитофонов) оперирует с информацией, электронным образом записанной на магнитной ленте, использование которой непосредственно человеком невозможно - требуются специальные технические средства - видеомагнитофон (видеоплеер) и телевизор. Но способ записи информации на ленту VHS видеомагнитофона - аналоговый, а не цифровой. Аналоговая информационная система может даже иметь определенный уровень автоматизации - например, та же фильмотека видеофильмов может иметь электромеханическую систему, снимающую со стеллажа нужную кассету и доставляющую ее транспортером в просмотровый зал. Однако сегодняшние информационные системы, как правило, являются цифровыми, то есть основаны на использовании компьютерной техники, и информация в них находится в цифровом виде. Хотя внутреннее для такой системы представление информации цифровое, в машиночитаемой форме, неприспособленной для прямого восприятия человеком, рядовой пользователь системы может никогда на практике не иметь дела с этим внутренним представлением и все время воспринимать информацию в удобной для него форме, используя мощные современные средства визуализации, в том числе, компьютерной графики. Современные информационные системы, как правило, создаются с использованием специального программного обеспечения, называемого системами управления базами данных (СУБД), а сами упорядоченные массивы данных, организованные с помощью СУБД, называются базами данных. Существуют специализированные пространственные информационные системы для работы с информацией об объектах и явлениях, которые имеют привязку к определенной позиции в пространстве, с информацией о тех объектах и явлениях, для которых важную роль играет их положение, форма, размеры, взаиморасположение по отношению к другим объектам и явлениям. Именно такие системы, в основном принадлежащие к классу геоинформационных систем, и являются одним из основных предметов рассмотрения в данной книге. Термин "пространственный", который мы употребили выше, имеет в данном контексте достаточно общий смысл. "Пространственные" объекты могут быть расположены в каком-то месте на земной поверхности, под ней или над ней, быть плоскими или объемными, могут вообще

не иметь никакой привязки именно к поверхности Земли с определенными для нее координатными системами. Важно, что объекты привязаны к некоторой координатной системе, возможно, местной и условной, и этот факт признается существенным и используется системой при организации данных и их использовании.

Специфический отдел информатики, имеющий дело с такой пространственно привязанной информацией, называется геоинформатикой.\* Соответственно выделяются и геоинформационные технологии как совокупность методов и приемов практического использования достижений геоинформатики для манипулирования пространственными данными, их представления и анализа. Как общая информатика имеет дело с общими свойствами информации и универсальными ее свойствами, а не со специфическими для конкретной предметной области, так и общая геоинформатика имеет дело с общими свойствами пространственной информации, независимо от конкретного ее содержания. И как для общей информатики существуют развивающиеся на ее пересечениях с конкретными предметными областями и научными дисциплинами ее специфические ветви, так и для геоинформатики также можно говорить о существовании или возможном появлении таких специфических ветвей - геологическая геоинформатика, геоинформатика в археологии, геоинформатика на железнодорожном транспорте. Таким образом, ни в коем случае нельзя трактовать геоинформатику в целом как информатику для геологии или географии или геодезии.

Поскольку даже сегодня, в информационный и компьютерный век, не вся информация хранится и обрабатывается с помощью компьютерных систем, мы продолжаем работать с книгами, папками и бумагами, фотографическими отпечатками и бумажными картами, то и предметом геоинформатики, строго говоря, являются не только компьютерные технологии. Но здесь методы геоинформатики, основанные на использовании общих свойств пространственной информации, как правило, неотделимы четко от традиционных методов и приемов, используемых в различных предметных областях, таких как картография, геология и т.п.

---

\* Это только одно из возможных определений геоинформатики, их сегодня существует множество. Автор не ставит здесь задачу дать их обзор и сопоставление. Интересующихся можно отослать например, к работам, А.В. Кошкарев, В.С. Тикунов. Геоинформатика Под ред. Д.В. Лисицкого. М., Картцентр-Геодезиздат, 1993, 213 стр., или А.М. Берлянт. Геоинформатика. М., Астрель, 1996, 208 стр. Не умаляя значение вопроса о четком определении понятия "геоинформатика", определении ее задач, предмета и метода исследований как науки, автор, тем не менее, считает, что это не является очень актуальной задачей на

Итак, в нашей книге мы по-преимуществу будем иметь дело с общей геоинформатикой - общей информатикой для пространственных объектов, пространственными информационными системами - базами данных пространственных объектов, созданными на основе идей геоинформатики, и с геоинформационными технологиями - вопросами практического применения геоинформатики для решения научных и практических задач. Мы не будем углубляться в специальные разделы геоинформатики, посвященные специфике тех или иных ее применений. В разделе "Практическая геоинформатика" мы, только для примера, кратко рассмотрим некоторые такие приложения.

Для перехода к дальнейшему нам важно разобраться в вопросе о том, что такое и как соотносятся друг с другом такие понятия как информация, данные и знания. Информация - это те данные, которые мы можем полезно использовать - например, увеличить наше знание о каком-либо явлении или предмете реального мира. Если мы имеем в нашей базе данных множество каких-то чисел и не вооружены абсолютно никаким, даже приблизительным знанием того, что это за числа, к чему относятся, какие величины представляют и для чего предназначены, то эти числа не являются для нас полезной информацией, как-то уменьшающей неопределенность нашего представления об окружающей действительности. Тем не менее эти непонятные цифры, хранимые в базе данных, тоже являются данными. Но информация извлекается из них только с применением нашего мышления или интуиции, базирующихся на некоторых знаниях о предмете.

Простой пример. Велосипед может быть графически представлен в виде кружков и нескольких линий, изображающих раму, педали и т.д. Такой геометрический образ велосипеда является собранием элементов данных. Каждый отдельный графический элемент можно рассматривать как данные. Но человеческое знание требуется для того, чтобы сделать вывод, что этот образ представляет в реальном мире предмет, который называется велосипедом и служит средством передвижения.

Другой пример. В базе данных имеются некоторые пары чисел. Только знание того, что эти числа являются картографическими координатами, понимание того, что такое вообще система

---

настоящем этапе и, тем более, для данной книги. Автор полностью согласен с И.К. Лурье ("Обучающие ГИС для наук о Земле", Информационный бюллетень ГИС ассоциации, №1(13), 1998, стр. 86), что "в понимании задач, предмета и метода геоинформатики нет однозначности, что естественно для быстро развивающейся отрасли знания".

прямоугольных координат, что такое картографические проекции и о какой картографической проекции идет речь в данном случае, как организована система координат в данной проекции - как направлены ее оси, где расположено начало координат, и, наконец, каковы масштабы по осям и в каких единицах представлены данные нам значения координат, позволяют придать смысл этим числам. Только зная все перечисленное, мы можем сказать, что данные, представленные этими числами, являются информацией о географическом положении каких-то объектов. Только имея какие-то “внешние”, дополнительные данные, можно сделать этот вывод. Для того же, чтобы понять или предположить, что же это за объекты, положение которых на Земле мы теперь знаем, требуется еще дополнительная информация. Таким образом, чтобы извлечь из отдельно взятых, не находящихся в каком-либо смысловом контексте, данных полезную информацию, необходима некоторая дополнительная информация особого, структурного плана, помогающая проинтерпретировать новые данные и ввести их в общую систему знаний об окружающем мире. Иначе говоря, данные, пока мы не имеем информации о том, как они связаны с системой знаний, в эту систему не входят и не являются полезной информацией. Как только мы привлекаем сведения о структуре их связи с существующей системой знаний, они оказываются связанными с этой системой, становятся ее частью и тем самым увеличивают объем этой системы знаний, становятся полезной информацией. Таким образом, знания предполагают системность, структурированность информации, обеспечивая возможность описания не только свойств, но и взаимосвязей между явлениями и объектами реального мира. Ценность данных возрастает на пути от разобщенных элементов данных к полезной информации и к систематическим знаниям.

С другой стороны, данные - это физическое представление информации. Наше познание мира, включающее такие процессы как отбор важнейшего, обобщение, анализ и синтез, обогащает нас информацией, которой возможно делиться с другими людьми, делая ее тем самым доступной не только для тех, кто ее создал. Такое накопление и обмен информацией возможны только с помощью данных - зафиксированных элементарных фактов.

Понятие “данные” имеет много аспектов. Какими могут быть данные?

1. Первичными фактами, объективно отвечающими в реальном мире предметам или условиям вне зависимости от способа фиксации и интерпретации человеком. Одно дерево выше другого, а река течет в направлении восхода Солнца независимо от того, наблюдает ли их человек, и если да, то как он это воспринимает и воспринимает ли вообще. В реальном мире есть нечто, что может быть зафиксировано нашими ощущениями непосредственно или с использованием технических средств и специальных приемов. На этом уровне тоже, вообще говоря, возможен обмен данными - привести кого-то за руку к реке и молча показать направление ее течения - но нетрудно понять, насколько редуцированы возможности такого обмена, да и является ли он, в сущности передачей информации - ведь без дополнительных сведений, хотя бы и сообщенных жестом, трудно догадаться, что речь идет именно о направлении течения реки.
2. Данные могут быть зафиксированы тем или иным способом, не включающим интерпретацию, например, записаны на магнитном носителе электронным датчиком, как сейсмические сигналы, или на фотопленку, как фотографии ландшафта.
3. Данные могут быть непосредственно в процессе их фиксации тем или иным способом проинтерпретированы человеком, например, полевая зарисовка ландшафта, сделанная географом, заполненная сотрудником социологической службы анкета-вопросник, словесное описание обнажения горных пород в полевом дневнике геолога.
4. Данные могут быть закодированы определенным образом, например, в виде бумажной карты с определенной легендой, в виде набора чисел - координат точек измерения глубин океана и самих измеренных глубин.
5. Данные могут быть структурированы или организованы определенным образом, например, в виде таблиц в результатах переписи населения, или в виде пространственной базы данных под управлением географической информационной системы.

Именно вопросы структурирования информации с помощью компьютерных систем занимают особо важное место в информатике



и в геоинформатике; естественно, они занимают важное место и в нашей книге. Нетрудно догадаться, что как существует масса вариантов расстановки мебели по квартире, то так же существует и еще большее число способов систематизировать и организовать информацию об окружающем нас мире. Многочисленные пакеты коммерческого программного обеспечения существуют для того чтобы практически выполнить такую организацию с помощью компьютерных технологий, в том числе и для работы с пространственной информацией. Для организации пространственной информации есть несколько преобладающих базовых форм такой организации, которые принято называть моделями данных. Постоянно растет и объем разработанной в геоинформатике теории, помогающей нам в процессе рациональной организации пространственной информации и данных. Эта теория включает сегодня как принципиальные, концептуальные основы такой организации, так и конкретные математические формализмы и разработанные типовые модели и структуры данных.

Одним из общепринятых принципов организации пространственной информации является последовательный принцип. Суть его заключается в том, что многообразная информация о какой-то территории организуется в виде серии тематических слоев, отвечающих конкретным потребностям. Каждый слой может содержать информацию, относящуюся только к одной или нескольким темам. Например, для целей изучения природных ресурсов такими темами могут выступать данные по геологии коренных пород, геологии четвертичных отложений, по почвам, типам землепользования, высотам рельефа местности, гидрографической сети, уклонам рельефа, экспозициям склонов, транспортной сети. Для задач планирования развития городской территории такой набор может включать данные по улицам, городским инженерным сетям, объектам транспортной инфраструктуры, различным типам зонирования городской территории, землевладениям и недвижимости. Такое подразделение информации на слои интуитивно понятно и привычно, и легко соотносится с общепринятыми принципами использования прозрачных калек-накладок при работе с бумажными картами. Последовательная организация данных предполагает, что слои в пространстве не имеют разрывов, и что везде мы имеем какую-то информацию - хотя бы "отсутствие объекта" или "нет данных о наличии или отсутствии объекта". Однако слои

необязательно должны в точности соответствовать тематическому делению. Например, несколько различных слоев может быть посвящено одной и той же, в принципе, теме, например, геологическому строению на разных уровнях срезов по глубине. Другой пример - отдельные слои для поэтажных планов каждого этажа зданий - тематика одна и та же, но во многих отношениях такое разделение очевидно удобно. Примеры чуть другого рода - представленные отдельными слоями данные могут относиться к разным точкам или разным периодам на временной шкале (однотипные данные переписи населения разных лет, карты средних температур января за разные годы.). (Рис. 1)

### Тематические слои      Временные срезы

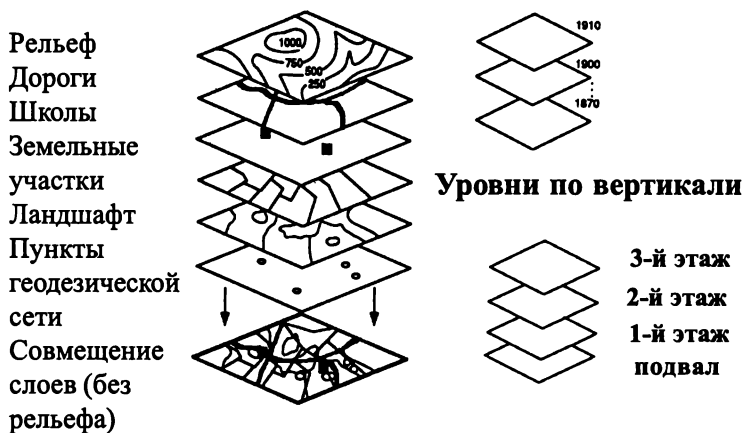
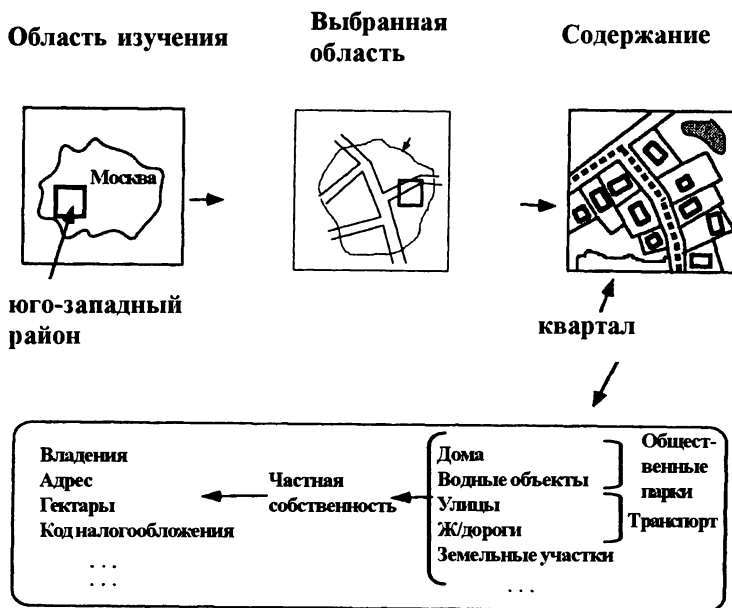


Рис. 1 Послойная организация данных

Аналогичное структурирование информации на темы, а также на отдельные листы по территории, широко распространено и при работе с аналоговой информацией, в первую очередь, с бумажными картами. Среди же компьютерных пространственных информационных систем такое послойное представление является сегодня доминирующим, явно отражая тот факт, что ГИС выросли на базе методов традиционного картографического анализа.

Другой принцип организации пространственной информации называется бесслоевым, или объектно-ориентированным подходом.

В нем отсутствует разбиение информации на тематические слои, точнее, отсутствует группировка объектов в тематические слои, с которыми можно манипулировать как с тематическими картами. Группировка объектов происходит более сложным образом, в соответствии с логическими взаимосвязями между ними, с построением иерархий, отвечающих их более общим и более частным свойствам. (Рис. 2) Такой подход, возможно, несколько ближе к структуре человеческого мышления и особенно эффективен тогда, когда требуется подчеркнуть индивидуальные логические взаимосвязи объектов, но менее эффективен, если необходимо вводить в рассмотрение также и непрерывно распределенные в пространстве признаки, такие как рельеф местности или данные о загрязнении почвы тяжелыми металлами. Впрочем, непреодолимой и резкой границы между этими двумя принципами организации пространственной информации нет. Представим себе такой вариант послойной организации, когда каждый индивидуальный объект вынесен в отдельный тематический слой - мы получим структуру, от которой уже недалеко до гибкости



**Рис. 2** Объектно-ориентированная организация данных

установления логических взаимосвязей, характерных для объектно-ориентированного подхода.

Для всех этих общих концепций организации пространственной информации существует и более детальный уровень представления, именуемый физическим уровнем представления. Например, если мы работаем с информацией о дорожной сети какого-то района, мы можем выразить ее в виде таблиц с номерами дорог по какому-то классификатору и массой числовых характеристик. Или мы можем представить эту же информацию в виде карты, где дороги отображены в виде линий, а те же связанные с ними числовые характеристики (ширина, число полос движения, тип покрытия) выражены в виде условных знаков (тип линии, цвет, толщина). Подразделение информации на физическом уровне может не в точности соответствовать ее подразделению на концептуальном уровне, например, физически сгруппированная в один слой информация может представлять несколько различных тематических слоев для пользователя. Такое раздельное рассмотрение разных уровней организации данных очень удобно на практике. Физический уровень организации данных, так же как и концептуальный (логический) является одним из предметов рассмотрения в этой книге.

Пользователи пространственной информации, еще задолго до появления компьютерных систем, уже работали с ней, используя какую-то определенную ее организацию, даже если и не выделяли четко проблему организации пространственной информации. Они в основном работали с бумажными картами. Карта, великолепное изобретение человечества, представляет собой, в сущности, картографическую модель действительности, один из способов организации пространственной информации. Она близка по идеологии, но не эквивалентна послыной организации информации в современных компьютерных системах. Карты очень разнообразны по принципам организации информации, но в целом более синтетичны и часто более эклектичны, чем результат следования строгим принципам организации информации в ГИС. Карта - это один из важнейших шагов на пути к современной геоинформатике, базирующейся на цифровых компьютерных системах. Долгий путь развития и применения бумажных карт нашел свое отражение в использовании и частичном заимствовании геоинформатикой

терминологии традиционной картографии, в воспроизведении с помощью компьютерных методов многих традиционных процедур работы с картами. Историческая связь геоинформатики и традиционных методов работы с картами очень тесная.

## **1.2. Различные применения информационных систем для работы с пространственной информацией.**

Областей применения ГИС существует великое множество и, по всей видимости, число их и еще будет расти. В каждой из прикладных областей существуют свои специфические потребности и своя специфическая терминология, своя история внедрения геоинформационных технологий. На начальном этапе эти технологии развивались в значительной степени независимо в разных прикладных областях. Это является одной из причин того, что сегодня в мире насчитывается уже очень много коммерческих пакетов программ для работы с ГИС (может быть, даже слишком много), их функции часто в значительной степени совпадают, но при этом зачастую используется различная терминология для обозначения одинаковых или сходных функций и элементов моделей данных. Иногда напрямую терминология конкретной прикладной области проникает в терминологию пакета для обозначения тех функций и понятий, которые, в принципе, имеют отношение к общей геоинформатике, но в практике применения данного пакета имеют конкретное более узкое и специфическое назначение. Например, абсолютно одна и та же функция проведения изолиний равных значений числовой величины по регулярной сетке наблюдений в одном пакете может называться функцией построений изогипс (имея в виду рельеф местности), а в другом - функцией построения изолиний концентрации (имея в виду данные по геохимическим полям). Подобные терминологические проблемы, естественно, затрудняют понимание и обсуждение многих вопросов, связанных с геоинформатикой, равно как и сравнение и освоение пакетов программ. В этой книге мы будем стремиться использовать в первую очередь наиболее общую, не привязанную к конкретным приложениям, терминологию и только в примерах иногда приводить терминологию более специфическую. Правда, это заставит нас в большей степени, чем это обычно делается при описании

конкретного пакета, использовать абстрактную математическую терминологию.

### **1.2.1. Сферы применения ГИС.**

Из приведенного выше описания назначения геоинформатики ясно, что и построенные на ее базе геоинформационные технологии и геоинформационные системы, хотя они и называются часто географическими информационными системами, никак нельзя трактовать в общем случае как информационные технологии и информационные системы для географии (или геологии, геодезии). Они имеют значение и применение значительно более широкое, чем только в указанных дисциплинах. Приставка “гео” означает только использование “географического”, то есть пространственного принципа организации информации, только то, что это технологии и системы, предназначенные для работы с пространственной информацией. Поэтому области применения ГИС и геоинформатики сегодня находятся почти во всех сторонах человеческой деятельности. Можно сказать, что перечислить сферы применения ГИС не проще, чем перечислить сферы применения СУБД.

Сегодня можно назвать, оставляя в стороне сугубо научные приложения, следующие крупные области применения ГИС, причем этот список далеко не полный, и приводится просто для примера.

Управление земельными ресурсами, земельные кадастры.

Инвентаризация и учет объектов распределенной производственной инфраструктуры и управление ими.

Проектирование, инженерные изыскания и планирование в градостроительстве, архитектуре, промышленном и транспортном строительстве.

Тематическое картографирование практически в любых областях его применения.

Морская картография и навигация.

Аэронавигационное картографирование и управление воздушным движением.

Навигация и управление движением наземного транспорта.

Дистанционное зондирование.

Управление природными ресурсами (водными, лесными и т.д.).

Представление и анализ рельефа местности.

Моделирование процессов в природной среде, управление природоохранными мероприятиями.

Мониторинг состояния окружающей среды.

Реагирование на чрезвычайные и кризисные ситуации.

Геология, минерально-сырьевые ресурсы и горнодобывающая промышленность.

Планирование и оперативное управление перевозками.

Планирование развития транспортных и телекоммуникационных сетей.

Маркетинг, анализ рынка.

Археология.

Комплексное управление и планирование развития территории, города.

Безопасность, военное дело и разведка.

Общее и специальное образование.

Сельское хозяйство.

Никакого особенного порядка в этом списке нет. Отметим только, что в нем присутствуют как термины, относящиеся к предмету активности (земельные ресурсы), так и относящиеся к задачам или целям (планирование развития, управление перевозками), а также и к методам и средствам (дистанционное зондирование).

В этой большой области приложений можно выделить несколько основных типов. Одни связаны с задачами учетно-инвентаризационного типа, акцент делается на данных и измерениях (например, задачи земельного кадастра или управления распределенной производственной инфраструктурой большого предприятия). Другие связаны с задачами управления и принятия решений. В третьих акцент делается на моделировании и сложном анализе данных. Первый тип имеет наиболее важное значение, хотя бы в силу того, что на этот тип задач приходится максимальное число реализованных и находящихся в режиме эксплуатации систем,

в том числе крупнейшие по числу пользователей и объемам собранных данных. (Для приложений этого типа вообще характерна работа с большим числом объектов и высокая детальность изучения территории.) Использование ГИС в качестве информационно-справочных систем начинает также широко использоваться в обучении. Отметим также, что вне зависимости от того, используются ли при работе с данными мощные аналитические процедуры и сложные запросы, ГИС очень часто используются как средства поддержки принятия решений. Достигаемая здесь эффективность, даже при использовании минимальных средств, доступных геоинформатике, часто очень высока за счет высокой наглядности картографической визуализации информации и удобства доступа к информации.

Знание конкретных потребностей той или области применения, той или иной задачи, является важным для рационального выбора программных средств и построения эффективной структуры базы данных. Например, для одних применений наиболее эффективной является послойная организация данных, для других - объектная "бесслойная" модель данных. Часто оптимальным решением оказывается комбинирование этих двух подходов. Более подробно мы этот вопрос разберем в главах 2 и 3.

### **1.2.2. Примеры различных требований к данным.**

Поиск пути - одна из основных человеческих потребностей. Знание того, как сориентироваться в системе улиц города, найти оптимальный маршрут от своего дома до работы, как объяснить дорогу к метро заблудившемуся прохожему - это только некоторые из проблем нахождения маршрута, в которые мы вовлечены ежедневно. В последнее время, однако, идет много разговоров о внедрении индивидуальных электронных средств навигации для автомобиля, которые способны автоматизировать этот процесс, и даже созданы и поступили в продажу первые образцы таких устройств. В общем, такой программно-аппаратный комплекс должен выполнять следующие функции:

1. Каким-то способом определить текущее положение автомобиля.
2. Получить "задание" - положение или характеристики места назначения.



3. Получить карту требуемого участка, включающего и текущее положение, и место назначения. Карта, возможно, будет использована при анализе ситуации и при представлении результата.
4. Используя какой-то метод, установить оптимальный путь к месту назначения.
5. Полученную информацию надо представить человеку.

Некоторые из действующих образцов подобных устройств имеют только алфавитно-цифровой дисплей и при общении с человеком оперирует только названиями улиц и направлениями движения (“по улице X 750 м к северо-востоку, затем поворот направо на улицу Y и по ней еще 1000 м к юго-востоку”). Однако уже появились системы с графическими мониторами, способными отображать карту улиц. Такая система способна непрерывно отслеживать положение автомобиля (например, используя систему спутниковой навигации) и показывать его на карте по ходу движения. При этом карта смещается по мере перемещения на новые участки.

Функциональные возможности подобной системы могут быть очень широки - с ней можно не только решать задачи поиска маршрута, но она способна давать подсказки при движении с учетом текущего положения машины, например, предупреждая водителя звуковым сигналом о том, что на следующем повороте надо сворачивать... Возможно организовать также оперативное внесение изменений в дорожную ситуацию - например, сообщить системе расчета маршрута, что какая-то улица в данный момент недоступна для проезда из-за дорожных работ или автомобильной пробки. Эти изменения могут вноситься вручную водителем, также возможно их автоматическое получение по радио.

Однако давайте перейдем от рассмотрения возможностей системы (они быстро развиваются и совершенствуются) к вопросу о данных, необходимых для ее функционирования. Нам необходима координатная информация, чтобы отобразить в виде карты форму и взаиморасположение дорог. Нам нужны, вероятно, местоположения каждого почтового адреса, чтобы их можно было использовать в диалоге с системой, а также названия конкретных мест и районов города, чтобы они могли быть использованы в качестве ориентиров. Естественно, необходимы названия улиц. Если мы хотим, чтобы система была достаточно универсальной, она должна быть способна

обрабатывать различные особые ситуации, например, не терять работоспособности в случаях, когда одна и та же улица имеет несколько названий (разные названия для правой и левой стороны, как это имеет место в некоторых странах), когда почтовые адреса распределены вдоль улицы не в строгом порядке и т.п. Некоторая дополнительная информация может быть необходима при поиске оптимального маршрута, например, места и время проведения дорожных работ, среднее время проезда по определенным улицам с учетом интенсивности движения (возможно, различной в разные часы и дни недели и неодинакового в разных направлениях). Также очень важно знать, какие улицы соединены в смысле возможности сделать поворот, каков характер такого соединения (из какого направления движения возможен поворот, в какое время суток и недели, каково среднее время выполнения поворота в данном направлении с учетом интенсивности движения). В общем можно сделать вывод, что очень и очень много разнообразных данных, в том числе статистических, оценочных (среднее время движения) и весьма оперативных (сведения о перекрытии улицы из-за аварии или автомобильной пробки) необходимы для правильного и практически полезного функционирования такой системы.

Рассмотрим следующий пример, лежащий скорее в области научных исследований. Стоит задача проанализировать влияние типов землепользования в определенном ландшафте на количество необходимых для питания растений химических веществ, вымываемых из почвы в речную сеть. Как каждый случай выпадения атмосферных осадков, так и рассматриваемый в долгосрочном плане режим осадков создают определенные условия существования для площадного и локализованного поверхностного стоков, просачивания воды в почву и подземного стока. Количество и скорость вымывания из почвы питательных веществ зависит от многих факторов, в числе которых уклон рельефа, режим грунтовых вод, проницаемость почв, тип покрывающей растительности. Поэтому важно иметь всю эту (и, возможно, многую другую) информацию, достаточно детальную, для всей изучаемой территории. Детальное знание рельефа местности, крутизны и экспозиции склонов, расположение (в определенной последовательности в пространстве) участков с различным типом землепользования - это только небольшая часть информации, очевидно необходимой для решения проблемы. Менее очевидно,

Скан  
страницы  
отсутствует

Скан  
страницы  
отсутствует

функции ГИС, для работы которого необходимо некоторое компьютерное оборудование. И, разумеется, требуется человек со своими знаниями и навыками, с пониманием потребностей в решении конкретных задач и способный взаимодействовать с системой. То есть, требуется некоторая организация различных ресурсов. Здесь мы должны упомянуть о том, что в самом понятии ГИС заключена определенная двойственность. Этим словом называют как собственно сам тот программный инструментарий, который позволяет организовать базу данных пространственной информации и производить над ней различные манипуляции, так и организованные с его помощью данные, а еще чаще понимают под ГИС весь комплекс программных и аппаратных средств, организованных с их помощью данных, конкретных процедур и технологий и человеческих интеллектуальных ресурсов, обеспечивающих его функционирование. Поэтому можно одновременно встретить как выражения “ГИС ARC/INFO версия 7.1” или “ГИС для персонального компьютера”, так и “ГИС по редким видам растений Приморского края” или “Муниципальная ГИС города Таганрога”. Такого же типа двойственность характерна и для использования словосочетания “база данных”. В разговорном языке под этим также иногда понимают как программное обеспечение для управления базами данных (СУБД), так и конкретные наборы данных, организованные с помощью такого программного обеспечения, а также и сочетание организованного набора данных и программного обеспечения, настроенного для работы с этими данными и решения на их основе конкретных задач. Кстати, такая нечеткость в использовании понятий “ГИС” и “база данных” встречается как в русском, так и в английском языке. Конечно, это несколько осложняет жизнь, и мы будем стремиться в этой книге использовать слова “ГИС” таким образом, чтобы всегда было ясно, о чем именно в данном случае идет речь и, во всяком случае, не называть “ГИС” сами содержащиеся в системе данные безотносительно к их организации с помощью той или иной программной системы. Глубинной причиной появления такой двойственности в языке является отражение того факта, что ГИС как разновидность программного обеспечения (как и СУБД) является только инструментальным средством, оболочкой без содержания, неспособной решать никакие практические задачи без наполнения конкретными данными. (Или, возможно, получения к ним удаленного доступа, например, через Интернет).

### 1.3.1. Представление о ГИС как об инструментальном средстве.

По типу выполняемых функций и по режиму использования геоинформационные системы могут быть очень различны. Например, от системы, используемой для поддержки принятия оперативных решений, может потребоваться практически непрерывно и быстро отвечать на запросы на получение информации. В этом случае мы имеем систему запросного типа, типа информационно-поисковой системы. Например, в городском управлении внутренних дел с помощью электронной карты отображается текущее положение всех патрульных машин. При поступлении сигнала о происшествии требуется срочно определить, какая из машин находится сейчас ближе к месту происшествия и найти для нее оптимальный маршрут следования. Здесь функция системы заключается в быстрых ответах на множественные частные вопросы, причем в данном примере видно, что это могут быть не просто запросы типа обычных хорошо определенных запросов к базе данных, а запросы, ответ на которые требует выполнения определенного пространственного анализа (нахождение кратчайшего пути, например). Во всяком случае, здесь конечное в определенном смысле решение находится прямо в процессе функционирования системы.

Или же система может быть ориентирована на подготовку и оформление в заданном виде информации для ее последующего использования уже вне системы. Например, система может готовить информацию по определенному запросу в виде таблиц или карт. Эти выходные документы могут готовиться на любом носителе - в бумажной или цифровой форме. Независимо от этого, такие формы выдачи результатов анализа или запроса, будучи предназначены для последующего использования, содержат обычно гораздо больше информации, чем в первом случае, и более сложны по форме. Режим же поддержки принятия оперативных решений требует быстрой реакции на запросы, которые должны быть по возможности сформулированы очень четко и конкретно, а объем выдаваемой информации обычно очень небольшой.

Два рассмотренных существенно различных режима использования ГИС не исчерпывают всего спектра возможностей, но к ним или к их

комбинации можно свести подавляющее большинство реальных ситуаций. Для обеспечения любого из этих режимов система в целом так или иначе должна быть способна выполнять несколько групп функций. Рассмотрим их в самом первом приближении.

1. Система должна обеспечивать средства для формирования цифрового представления пространственных объектов и явлений, иначе говоря, должны обеспечиваться функции сбора, кодирования и ввода информации.
2. Для поддержания массива данных в актуальном состоянии должны обеспечиваться средства редактирования, обновления, эффективного хранения данных, а также реорганизации данных и преобразования их в разные формы, средства для контроля правильности и качества данных.
3. Требуется обеспечивать средства для получения информации, как в ее первичном виде, так и обобщенной в различных смыслах (генерализованной, суммарной, усредненной), а также средства для анализа, моделирования ситуаций и процессов и интегрирования разнородной информации.
4. Система должна обеспечивать выполнение сложных запросов на получение информации. Выполнение таких сложных запросов само по себе также является методом анализа.
5. Наконец, система должна обладать способностью представлять результаты работы в виде человекочитаемых наглядных документов - таблиц, карт, диаграмм и т.п..

В некотором смысле, любая работающая ГИС, разумеется, является частью гораздо более общей инфраструктуры информационного обеспечения потребностей человечества. При этом относительная и абсолютная роль ГИС (или, скажем иначе, - информационных систем для работы с пространственной информацией) для человечества уже сейчас значительна и продолжает возрастать. Отметим, что с точки зрения информатики в окружающем нас мире всегда можно выделить “физический”, “вещественный”, реально наблюдаемый компонент - например, в городе это здания, улицы, рельеф местности со своими условиями стока поверхностных вод, инженерные сети, транспортные коммуникации и т.п. - назовем это “реальной инфраструктурой” (или сочетанием “реальных инфраструктур”). Выделим также компонент

наших знаний обо всем этом, который может выражаться как в некоторой упорядоченной системе понятий, так и во вполне вещественных предметах, таких как бумажные планы и карты, документы и, например, компьютерные информационные системы. (Все это вместе можно обозначить как информационную инфраструктуру). Наряду с этим можно отдельно выделить также инфраструктуру управления, которая в своей реальной деятельности использует информационную инфраструктуру, описывающую объекты, отношения и процессы реального мира для рационального управления объектами реального мира.

В каком-то смысле аналогию с этой ситуацией можно провести и рассматривая саму информационную систему. В ней также можно выделить “физический” компонент (или, как чаще говорят, физический уровень) - это аппаратное, техническое обеспечение и способ кодирования данных в их внутримашинное представление, и сами эти данные, записанные тем или иным способом на машинные носители, а также логический уровень организации информации - способ ее организации в те или иные взаимосвязанные структуры (который в конечном счете всегда выражен также в каких-то элементах данных на физическом уровне, но эти данные, определяющие организацию данных, могут быть “напрямую” недоступны рядовому пользователю).

Практически любые программные средства являются инструментом, средством для решения каких-либо конкретных прикладных задач. Исключения разве что лишь в сфере общесистемного программного обеспечения (например, операционные системы), да в сфере компьютерных игр - ПО, специально созданного для развлечения. Однако можно подразделить прикладные программные средства на те, которые предназначены для решения содержательных задач конечного пользователя практически сразу, “из коробки” (например, текстовый редактор - установил на компьютер, и сразу можешь начинать работать - вводить и редактировать текст), и на те, которые сами по себе являются только полуфабрикатом, инструментом для разработки приложений для конечного пользователя. (это, например, системы разработки программ на алгоритмических языках - компилятор и средства отладки для программирования на языке С, например.) Естественно, для программиста - разработчика



приложений эти инструментальные средства являются никаким не полуфабрикатом, а законченным продуктом, обеспечивающим выполнение необходимых ему функций также сразу, “из коробки”. Отметим также, что очень многие прикладные программы, готовые к немедленному использованию конечными пользователями-непрограммистами, сегодня имеют также и встроенные средства программирования разного уровня, начиная от возможностей написания простых макрокоманд, обеспечивающих удобство работы со стандартными последовательностями часто повторяющихся операций, до серьезных языковых средств, позволяющих приспособить прикладной пакет к каким-то специальным потребностям или написать сложные приложения. В силу этих и других причин различие между готовыми к использованию прикладными программами и программами - инструментальными средствами разработчика не такое четкое и однозначное. Часто на одну и ту же программную систему можно взглянуть с разных точек зрения. Программное обеспечение для информационных систем (системы управления базами данных) и пространственных информационных систем добавляет еще один аспект к этой проблеме. Помимо того, что представители этих классов программных систем также могут иметь свои средства программирования для разработки приложений, важную роль играет также вопрос о данных. Как обычную СУБД, так и программное обеспечение для пространственной информационной системы (например, ГИС) в качестве готового к использованию средства может рассматривать не только программист, но и специалист, формирующий структуру базы данных и/или наполняющий эту базу информацией. (Эти процессы могут не требовать программирования.) Для него эта система готова к использованию - созданию конкретной базы данных и, возможно, наполнению ее информацией, немедленно. Для другой же категории пользователей программное обеспечение само по себе, пока оно не сопровождается той или иной подготовленной базой данных, вещь, к использованию не готовая, и прикладной системой не является, а представляет из себя только еще инструментарий, пустую оболочку, которую еще надо наполнить содержанием. Здесь опять проявляется условность деления программного обеспечения на прикладные системы и на инструментальные средства, а также ярко проявляется и важность данных как полноправного компонента информационной системы. В информационной системе они часто составляют ее

наиболее важную часть, стоимость организованных в базу данных может многократно превышать стоимость всех остальных компонентов (программного и аппаратного обеспечения) вместе взятых. В особенности это верно именно для пространственных информационных систем, поскольку процесс организации и ввода пространственных данных, включающих графический компонент, часто особенно трудоемок и дорог.

### **1.3.1.1 Пространственные информационные системы и ГИС - соотношение этих понятий.**

Мы до сих пор использовали эти два понятия почти как синонимы, не останавливаясь на различиях между ними. И преимущественно мы использовали выражение “пространственные информационные системы”, хотя оно в профессиональном языке специалистов по геоинформатике и в литературе встречается гораздо реже, чем ГИС. Пространственные информационные системы - понятие в функциональном плане более узкое, чем ГИС. Каждая ГИС является в своей основе пространственной информационной системой, поскольку организует, хранит пространственную информацию и обеспечивает к ней доступ - выполняет непосредственно информационные запросы пользователя или обслуживает информацией работу других функций - своих собственных или функционирующих в рамках других программных средств. В тот набор функций, которые должна выполнять пространственная информационная система, вообще говоря, не входят те очень многочисленные и разнообразные функции аналитического и моделирующего характера, которыми сегодня обладают развитые ГИС и благодаря которым, собственно, они и находят сегодня такое многообразное и эффективное применение. То есть, любая ГИС как свое функциональное подмножество обязательно содержит пространственную информационную систему.

С другой стороны, сегодня уже имеются такие программные продукты, которые, являясь пространственными информационными системами, выполняют только ее узкие функции, не являясь ГИС в общепринятом смысле слова. Такие системы обеспечивают ввод, хранение и выдачу по запросу пространственной информации, оставляя все функции по ее визуализации, редактированию, проведению анализа за отдельными программными пакетами. Речь

идет о так называемых “серверах пространственных данных”. Их клиентами, то есть системами, обеспечивающими выполнение остальных функций, могут быть самые разные пакеты программ, в том числе - разнообразные ГИС. Но не только. Это могут быть и более простые пакеты, не дотягивающие до уровня полнофункциональной ГИС, но обеспечивающие формирование запросов и визуализацию полученной от сервера в ответ на запрос пространственной информации. Это даже могут быть вообще пакеты не из разряда традиционных средств геоинформатики, вообще имеющие мало отношения к каким-либо ГИС, но которым требуются для работы (постоянно или эпизодически) какие-либо пространственные данные. Например, это могут быть пакеты САПР\*, статистического анализа данных. В качестве примера сервера пространственных данных, оформленного в виде отдельного от ГИС программного продукта, можно назвать Spatial Database Engine (SDE) фирмы ESRI.

Сделав такое пояснение, мы, однако, в дальнейшем по-прежнему будем использовать термины “ГИС” и “пространственная информационная система”, применяя последний в тех случаях, когда мы акцентируем внимание именно на соответствующих базовых функциях. В вопросах же практического применения мы чаще будем использовать более привычный и компактный термин “ГИС”.

Естественно, что приступая к изучению основ геоинформатики, мы вначале будем в первую очередь интересоваться более общими, базовыми понятиями и функциями, то есть вначале нас будут больше интересовать функции пространственной информационной системы, независимо от того, является ли она подсистемой в ГИС или самостоятельным продуктом. То есть в первую очередь нас интересует вопрос об организации данных, методах ввода и модификации данных, организации запроса и вообще доступа к данным, и уже затем - те многочисленные более аналитические функции, которые делают полнофункциональную ГИС более, чем информационной системой. Хотя надо ясно понимать, что четкой границы здесь нет. Элементарные функции запроса, визуализации информации часто являются уже и функциями ее анализа.

---

\* САПР - системы автоматизированного проектирования (англ. CAD - Computer Aided Design). Известным примером САПР широкого применения является пакет AutoCAD фирмы Autodesk.

### **1.3.2. Физический компонент ГИС.**

В качестве “физического”, “материального” компонента пространственной информационной системы (или ГИС) мы рассматриваем компьютерное (аппаратное) обеспечение, понимая его в широком смысле. Это не только сами компьютеры тех или иных типов, но также и разнообразные, как принято говорить, периферийные устройства, обеспечивающие хранение данных на внешних, возможно, съемных, носителях информации, ввод и вывод информации в различной форме, а также оборудование компьютерных сетей и средств коммуникации, позволяющее различным, в том числе иногда и весьма удаленным компьютерам\* взаимодействовать друг с другом и обмениваться информацией.

### **1.3.3. Программный компонент ГИС.**

В качестве программного компонента конкретной работающей ГИС в первую очередь надо рассмотреть, конечно, программное обеспечение, специально предназначенное для работы с пространственной информацией, то есть сами геоинформационные системы разных классов, уровней и типов. Но не только. Хотя это и не является основным предметом рассмотрения в этой книге, не следует забывать, что реальная действующая ГИС всегда использует и другое программное обеспечение, как минимум, операционную систему компьютера. Нередко используется и дополнительное программное обеспечение для организации компьютерной сети, доступа в глобальную сеть Интернет, организации дополнительной защиты информации от несанкционированного доступа. В отдельных случаях вместе с ГИС, во взаимодействии с ней, используется и дополнительное программное обеспечение для решения специальных задач, как, например, пакеты САПР или углубленного статистического анализа данных. Если ГИС находится в процессе развития и совершенствования выполняемых ею функций, вместе с ней могут использоваться те или иные инструментальные средства программирования, такие как компиляторы и отладчики. ГИС может тесно взаимодействовать с общеофисным программным обеспечением, таким как текстовые редакторы, электронные таблицы, системы поддержки документооборота. Наконец, особо важную роль из других типов программных пакетов могут играть

системы обработки данных дистанционного зондирования и СУБД. Относительно обоих этих типов ПО сегодня часто бывает трудно определить, являются ли они внешними по отношению к ГИС, или являются ее частью, настолько высока в отдельных случаях степень их интеграции с ГИС. Взаимоотношения ГИС с другими типами программного обеспечения мы будем подробнее рассматривать позднее, а пока только хотели отметить многообразие и сложность их взаимоотношений.

#### **1.3.4. Данные как компонент ГИС.**

Мы уже упоминали, что данные являются важнейшим компонентом любой функционирующей информационной системы, составляя нередко ее наиболее ценную, в том числе и в прямом стоимостном смысле, часть. В ГИС, как в пространственных информационных системах, роль и стоимость компоненты данных в типичном случае многократно возрастает.

## **2. Анализ потребностей. Задачи. Типы проблем, связанные с пространственной информацией.**

В этой главе мы попытаемся разобраться в том, какие, собственно, типы проблем и задач вызвали к жизни появление геоинформационных технологий. В чем состоит та часто упоминаемая специфика пространственной информации, которая требует особого отношения к такой информации? Какие типы проблем возникают при работе с пространственной информацией, независимо от ее конкретного содержания и области применения? Что такое пространственный запрос, и какие виды пространственных запросов существуют? Следует отметить, что эта область, несмотря на наличие ряда публикаций, в целом еще, возможно, не получила достаточного теоретического осмысления, применяемые подходы к классификации и проблем, и запросов разнятся. Вероятно, данная область является одним из перспективных “направлений роста” геоинформатики.

## 2.1. Проблемы и задачи при работе с пространственной информацией.

*Ограничения. Пространственные измерения. Пространственные взаимоотношения. Решения, выводы принимаемые на основе пространственной информации. Моделирование на основе пространственных данных. Моделирование процессов. Предсказательная сила пространственных моделей.*

Осмысленный подход к созданию, приобретению, оценке и использованию пространственной информации должен, конечно, включать рассмотрение вопроса о том, какие, собственно, конкретные проблемы научного или практического характера с помощью этой информации требуется решить. Предпочтительнее, чтобы проблема была бы определена максимально четко. Любые задачи и проблемы можно обычно формулировать на различных уровнях, с разной степенью детальности. В крупных организациях может существовать много подразделений и отдельных пользователей с различными потребностями по работе с пространственной информацией. Их задачи на содержательном уровне могут различаться радикально. При этом они могут использовать одну и ту же или совершенно различную информацию. Например, и анализ преступности в городе, и планирование работ по ремонту дорожного покрытия, и планирование новых маршрутов городского транспорта, и инвентаризация и мониторинг состояния уличного освещения, и выбор оптимальных мест для размещения уличной рекламы - решение всех этих задач в значительной степени опирается на использование информации о городских улицах. Естественно, и один пользователь в рамках решения одной содержательной задачи может для одного и того же множества данных формулировать различные пространственные запросы. Что еще более для нас сейчас интересно, это то, что очень разные вопросы, с которыми различные пользователи при решении своих специфических прикладных задач обращаются к пространственной информации, могут быть сведены к небольшому числу типов, отличающихся с точки зрения общей геоинформатики. Эти типы следует уметь выделять из практически бесконечного разнообразия запросов к пространственной информационной системе, возникающих в различных прикладных областях. Практическое значение такой типизации весьма велико. Только умея видеть общее

за частным, распознавая различные типы запросов не с точки зрения их содержательного смысла, а с точки зрения общих особенностей работы с пространственной информацией, можно правильно подойти к определению необходимой функциональности пространственной информационной системы, выбору способа организации информации в ней, то есть в конечном счете - к правильному выбору программного обеспечения и построению оптимальной структуры базы данных.

Вот пример нескольких пространственных запросов, сформулированных на содержательном уровне в рамках решения одной задачи. Пример относится к области управления национальными парками в США, анализируются рекреационные свойства ландшафта и расположение в нем объектов инфраструктуры парка - троп для перемещения посетителей и смотровых вышек для наблюдения красот природы.

Какова высота смотровых вышек (может быть, различная для разных вышек)?

Где расположены смотровые вышки?

Какие из них соединены пешеходной тропой?

Как пройти к каждой из смотровых вышек?

Какой из возможных путей их обхода будет самым коротким?

Какая часть территории видна с каждой из смотровых вышек?

Какая растительность в пределах этих наблюдаемых зон?

Какую площадь занимают различные виды растительности?

Какую площадь занимают различные типы древесной растительности в пределах заданных интервалов абсолютных высот местности?

Какую площадь занимают древесные насаждения различных типов в пределах зон в 500 метров вокруг каждой из пешеходных троп?

Нетрудно придумать примеры и из других прикладных областей, касающиеся параметров пространственного положения и взаиморасположения объектов. Пространственные запросы можно определенным образом систематизировать. Простейшее, что приходит в голову - это выделить две группы, отличающиеся

**направлением** нашего запроса - “Что это?” и “Где это находится?”, или, иначе говоря, “Что мы имеем в конкретном месте?” и “Где находится какой-либо объект (объекты) с заданными характеристиками?” В первом случае мы имеем дело с поиском объекта в пространственной базе данных по его положению, во втором случае ищется положение для уже выбранных объектов. Первый случай можно назвать “картографическим интерфейсом запроса к базе данных”, “запросом информации “через” карту”. Второй случай можно сравнить с выдачей отчета, распечатки с найденными данными при использовании обычной СУБД, и назвать его можно “созданием картографического отчета”. В последнем случае результат поиска выдается на карту, или карта строится по результатам информационного поиска.

На оба типа пространственных запросов могут накладываться **дополнительные ограничения**. Эти ограничения могут быть различного характера, например:

1. Основанные на характеристиках объекта, например, при запросе “Что находится в данном месте?” мы можем производить поиск только среди информации, относящейся к лесному хозяйству, игнорируя ту ее часть, которая описывает геологическое строение.
2. Основанные на числе объектов, например, “Найти любые три жилых дома в пределах данного микрорайона” для проведения выборочного опроса населения.
3. Основанные на расстоянии от заданного объекта, например “Найти все аэродромы в радиусе 300 км от Москвы”.
4. Основанные на требовании обязательного попадания в пределы заданной площади или заданного типа площадных объектов, например, “Найти все землевладения, расположенные хотя бы частично на луговых глеевых оподзоленных почвах”, или “Найти все острова на Онежском озере”.

Как мы видим, дополнительные ограничения могут носить как **непространственный** (1,2), так и **пространственный** (3,4) характер. Такие пространственные ограничения могут использовать понятия **расстояния** или **включения** (попадания искомого объекта в пределы заданного), а также и другие варианты взаимоотношений объектов.



Некоторые варианты пространственных запросов, включающих пространственные ограничения, продемонстрированы на рисунке (Рис.3).

“Что мы имеем в этой точке, вдоль этой линии, в пределах этого многоугольника?”

“Что мы имеем вдоль этой линии или в пределах этой площади?”

“Что мы имеем в пределах этой буферной зоны?” (Зона на заданном расстоянии от заданного объекта)

“Что мы имеем в этой точке?”

“Что мы имеем в этом круге?”

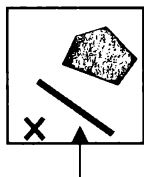
“Что мы имеем в этой зоне?”

“Где в пределах этого прямоугольного окна свободные территории?”

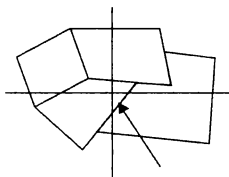
Как видно, могут существовать как многообразные типы искомым пространственных объектов, так и многообразие пространственных объектов, используемых для наложения дополнительного пространственного ограничения на поиск.

Кроме пространственных запросов на поиск информации или объектов, может потребоваться провести измерения, или определить пространственные взаимоотношения объектов. Пространственные измерения могут включать, например, измерения расстояний или площадей (“Расстояние от административного центра”, “Размеры озера”, “Расстояние до ближайшей железной дороги”), а определения пространственных взаимоотношений, например, могут быть анализом соседства - “Какие земельные участки граничат с данным?”, а также иметь вид “Пересекаются ли где-нибудь Рижская железная дорога и Дмитровское шоссе в Москве?” Отметим, что предпоследний пример одновременно является и примером анализа пространственных взаимоотношений и примером пространственного запроса на поиск объектов.

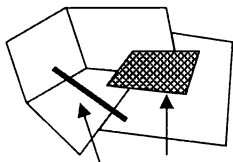
Нетрудно представить себе еще и такой тип пространственных проблем, как нахождение оптимального места для размещения какого-либо объекта. Такая задача представляет собой, в сущности, один из вариантов поддержки принятия пространственных решений.



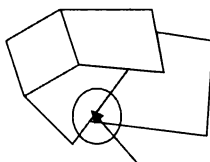
Что у нас на этой  
площади или на этой  
линии?



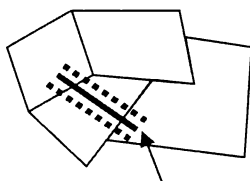
Что у нас в этой  
точке?



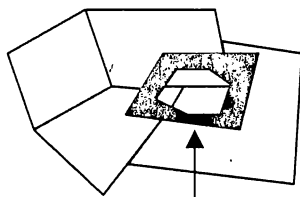
Что у нас на этой  
площади или на этой  
линии?



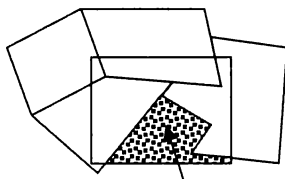
Что у нас в пределах этой  
окружности?



Что у нас в этой  
буферной зоне?



Что у нас в этой зоне?

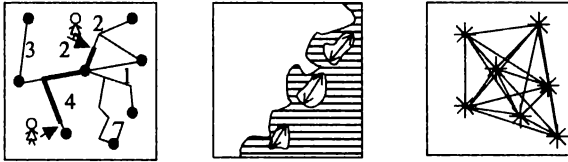


Какие участки в пределах  
этого окна не заняты  
объектами?

Рис. 3 Некоторые типы пространственных запросов

Она относится к числу наиболее важных для практической геоинформатики и может применяться в различных вариантах - определение наилучшего места для размещения нового ресторана или автозаправочной станции, оптимальное распределение учащихся по окрестным школам, определение оптимальной трассы для нового магистрального газопровода, нахождение оптимального варианта распределения зон обслуживания для станций скорой помощи в городе. Системы поддержки принятия пространственных решений могут иметь дело с выводами в отношении объектов, реально ненаблюдаемых, на основе изучения наблюдаемых объектов. Таковы, например, практически все задачи прогнозирования месторождений полезных ископаемых в геологии. Мы наблюдаем выходы тех или иных горных пород, измеряем геофизические и геохимические поля, и на основе их изучения и установления характера взаимоотношения между ними делаем выводы относительно возможности обнаружения месторождений тех или иных полезных ископаемых или в целом на заданной территории или даже в конкретном месте, иногда даже с вполне конкретным выводом "Бурить скважину именно здесь!" Итак, мы получаем фактически еще один вид пространственных проблем - проблему предсказания, пространственного прогноза. (Рис. 4)

Последняя категория проблем включает и задачи пространственного моделирования, которое может иметь, а может и не иметь динамический характер. Например, если мы строим модель распространения паводковой волны по долине реки и прогнозируем величину поднятия уровня воды в определенном месте в определенное время, то это, конечно, использование в прогнозе динамической модели процесса. Если же мы строим прогноз значений гравитационного поля в промежутках между точками наблюдений, мы имеем дело со статической моделью. Однако в обоих случаях результат - предсказание на основе косвенной или неполной информации. Поддержка принятия пространственного решения может быть связана с использованием статической или динамической пространственной модели, а может быть связана и с использованием модели (динамической или статической), не имеющей дела с пространственной информацией. Например, задача определения оптимального положения для новых ресторанов может включать в себя использование прогноза вероятного спроса на их услуги, выполняемого с использованием корреляционного анализа по косвенным данным - спросу на определенные виды товаров.

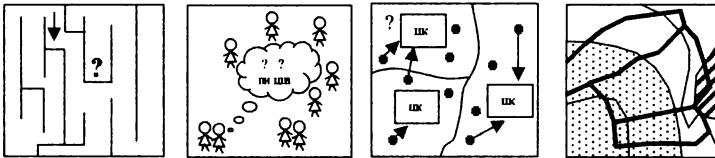


- а) Расстояние по сети      Размер барьерных островов      Удаленность административных центров друг от друга



- б) Схема для ориентировки      Схема стран-соседей

- а) Примеры измеренных характеристик для различных типов объектов.  
 б) Примеры схематических карт, демонстрирующих пространственные взаимоотношения.



- Проблема выбора маршрута      Проблема выбора объекта по его положению      Проблема распределения учеников по школам и построение зон обслуживания для каждой школы      Проблема размещения

Рис. 4 Некоторые типы пространственных проблем.

Подводя краткий итог, приведем таблицу - сводку основных типов активности, связанной с использованием пространственных данных.

| <b>Дано:</b>                             | <b>Требуется найти</b>   |
|--|--|
| Имя объекта                              | Местоположение объекта   |
| Характеристики объекта                   | Имя объекта  |
| Местоположение объекта                   | Имена объектов   |
|  | Характеристики объектов  |
| <b>Категория действия</b>                | <b>Конкретизация действия</b>  |
| Провести измерения                       | площади, расстояния, формы, и т.д.   |
| Выявить пространственные взаимоотношения | включения, направления, прилегания, связности, и т.д.  |
| Принять решение                          | выбрать место размещения, выбрать маршрут или путь, распределить ресурсы.  |
| Сделать прогноз                          | на основе прямых геометрических построений, используя статическое моделирование, используя динамическое моделирование. |

Здесь мы только в общих чертах коснулись пространственных проблем и их типизации. Подробнее мы познакомимся с ними чуть позже.

## 2.2. Информация положения и формы (пространственная в узком смысле слова) и информация атрибутивная (описательная).

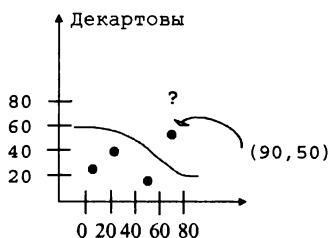
*Системы координат - полярная, прямоугольная. Пространственная ссылка. - - через конкретные прямоугольные или азимутальные координаты. - через указание соседних объектов. - через линейные координаты вдоль известного объекта. - через описанный прямоугольник (через значения минимальных и максимальных координат). - по имени объекта или его коду. - по нахождению в заданном сегменте делимости, регулярной или нет (на известном листе карты или в заданной стране).*

Два упомянутых в предыдущем параграфе понятия всегда в явной или неявной форме присутствуют в наших рассуждениях при обсуждении любой пространственной проблемы - это положение и характеристики, свойства. Они как бы определяют два возможных взгляда на предмет, по разному направляя мыслительный процесс. Хотя на практике два этих взгляда и нелегко разделить. И более того, этими такими, казалось бы, элементарными понятиями часто бывает трудно воспользоваться. Представим себе ситуацию, когда вертолет привозит человека с завязанными глазами куда-то в середину большого лесного массива. Естественно, первый вопрос, “Где я?”. Но что может послужить в этой ситуации системой координат, системой привязки, точкой отсчета? Как узнать свое местонахождение, передвигаясь по лесу? Это представляет собой трудную проблему. (Конечно, если в кармане нет GPS приемника и листа нужной топографической карты). Другой пример. “Опишите природу пустыни Сахара”. Но что такое, в точности, Сахара? Что определяет ее границы? Они достаточно неопределенны.

Человечество за тысячелетия своей истории выработало несколько вариантов таких систем привязки или систем координат. Они могут быть частными, используемыми разово в одном единственном случае (“расстояние в шагах и направление по отношению к солнцу от места посадки вертолета”) или использоваться во многих случаях как стандартные (общеземная сетка параллелей и меридианов). Системы координат могут быть полярными, основанными на направлении на конкретную точку и расстоянии до нее. (Такие системы называют также азимутальными). Другой

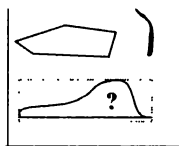
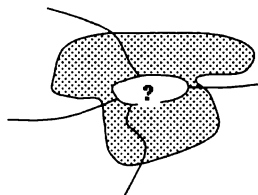
общеупотребительный вариант - это **прямоугольные системы координат**, использующие также два числовых значения, представляющие собой два расстояния во взаимно перпендикулярных направлениях. (Прямоугольные системы координат называют еще **картезианскими**, или **Декартовыми** по имени математика Декарта). (Рис. 5)

а) По координатам:



б) Указанием соседей

с) Указанием координат  
е) Указанием вмещающего прямоугольника листа карты (единицы регулярной делимости пространства)

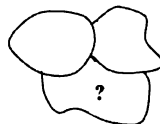
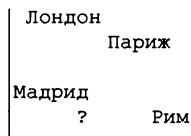
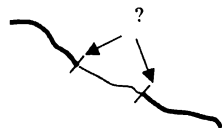


|   |     |
|---|-----|
| 1 | 2   |
| 3 | ? 4 |

б) Указанием соседей

д) По названиям мест

е) Указанием административных округов (единицы нерегулярной делимости пространства)



**Рис. 5 Способы указания положения**

Системы координат могут иметь в качестве начала отсчета какой-то реальный объект, например, железнодорожную станцию, или использовать воображаемый объект, не существующий в действительности, как точка пересечения экватора и гринвичского меридиана. Системы координат и точки начала отсчета в них могут иметь универсальный характер, и быть общепризнанными стандартами для многочисленных применений, или они могут быть местными, условными, не привязанными к какому-либо стандарту или вообще к какой-либо другой системе отсчета. В качестве примера можно привести сетку условных квадратов с буквенно-цифровыми обозначениями, на которую разбит туристский план города. Кроме того, системы отсчета развиваются со временем. Древнейшие карты строились, исходя из представления о плоской Земле, современные, естественно, принимают во внимание сфероподобную форму нашей планеты.

Сегодняшние ГИС всегда работают с прямоугольной системой координат, обычно построенной на плоскости. При этом обычно имеется возможность связи с географическими (градусными) координатами (которые представляют собой вариант азимутальных). Эта связь реализуется или в возможности преобразования координат из одной системы в другую для всех или избранных объектов в пространственной базе данных переписыванием их значений в файлах, или в возможности пересчета “на лету” координат текущего положения курсора, или даже в преобразовании “на лету” всего картографического представления из одной системы в другую, как это делается в ArcView GIS. Базой для определения абсолютного положения и измерений расстояния является в них евклидова геометрия. Хотя в ряде случаев для этой цели используются другие подходы. Кроме того, при работе с ГИС могут быть важными иногда и какие-то языковые, описательные параметры, имеющие отношение к системе координат и привязки, в ряде случаев для того, чтобы организовать и определить какую-то частную систему координат, как в примере с вертолетом, частью для того, чтобы организовать удобное взаимодействие человека и машины.

Хотя мы многого не знаем о том, как пространственная информация обрабатывается в мозгу человека, некоторые простые, но фундаментальные положения можно считать уже достаточно надежно установленными.



Во-первых, разумеется, сознание человека, а соответственно, и он сам в своей обыденной деятельности не пользуется обязательно какой-то единой прямоугольной системой координат. Люди часто для описания положения какого-либо объекта или пути используют разные для каждого случая системы отсчета, например, по отношению к знакомому им объекту, например, их дому, центру города. Или могут быть использованы азимутальные системы привязки. Например, коренные жители Гавайских островов использовали расстояние от/до центра острова в сочетании с положением напротив определенной точки периметра острова. В сущности, подобную систему отсчета и сейчас используют иногда моряки в прибрежном плавании. (“На траверзе острова Кильдин, расстояние 1 миля.”)

Во-вторых, характерные объекты ландшафта, природные или созданные человеком, и их имена собственные могут использоваться, иногда во множественном числе, для создания системы привязок. В частности, в старых городах Европы обычна ситуация, когда разные участки одной и той же, в сущности, улицы имеют разные наименования, производные от расположенных на этих участках известных исторических объектов. Наоборот, в США в городах обычно меньше таких характерных ориентиров, и доминирует чисто цифровая и формальная система наименования улиц.

В-третьих, сами языковые конструкции, используемые для описания положения и направления, часто достаточно неопределенны по значению или могут использоваться в различных обстоятельствах в различных целях. Например, слово “Север” в различных обстоятельствах может обозначать и точный азимут на магнитный Север по компасу из данной точки, и обобщенное, примерное направление, и множество направлений на Север из всех точек какого-то объекта, и неопределенно большую область на Земном шаре или неопределенную область в пределах какой-то страны.

Обычно пространственные объекты и явления рассматриваются в какой-то из систем привязки или координат. Два эти понятия - система привязки и система координат не являются полными синонимами. Системой координат нам представляется разумней

называть только системы числовых координат. Понятие система привязки является более общим термином, включающим в себя и понятие система координат. Два другие термина - позиция и местоположение, место, тоже не являются синонимами, хотя в обыденной жизни различие между ними часто игнорируется. Термин позиция лучше употреблять по отношению к положению существующего объекта, а место - по отношению к определенной точке в системе привязки, безразлично к тому, находится в ней какой-то объект, или не находится. Предположив на минуту, что объект реального мира имеет четкие и определенные границы, нам необходим какой-то метод доступа к носителю информации, чтобы зафиксировать его положение. Например, нам требуется пометить его на листе бумаги или произвести соответствующие действия в ГИС. Иначе говоря, необходима такая операция, как привязка объекта, которая может быть выполнена как абсолютная, то есть как фиксация его положения в некоторой универсальной и общепринятой системе координат, или как относительная, по отношению к какому-либо другому известному объекту. Далее мы будем широко использовать термин локатор, или характеристика положения, для именованя тех средств или характеристик объекта, которые мы используем для определения его положения.

Пространственная привязка, то есть фиксация положения в пространстве, может быть выполнена в различных формах. (Рис.5)

1. В виде конкретной системы числовых координат, как прямоугольной, так и азимутальной (полярной).
2. Указанием объектов-соседей, то есть линейных или площадных объектов, которые касаются искомого.
3. Заданием линейной привязки, то есть линии и расстояния, измеренного вдоль нее, как пикетаж по профилю нивелирования или километровые столбы вдоль дороги.
4. С помощью указания минимального описанного многоугольника, включающего искомый объект (то есть, фактически, указания минимальных и максимальных прямоугольных координат по осям X и Y).
5. С помощью имени объекта или его числового кода-номера.
6. Указанием того участка в системе разбиения территории (или объема), который содержит искомый объект. При этом

разбиение может быть на участки правильной или неправильной формы, одинакового или разного размера. Например, указанием номенклатуры листа масштаба 1:25000, в пределах которого расположен объект. Или указанием страны, в пределах которой он расположен (“Город Грац расположен в Австрии”).

Подробно все эти формы пространственной привязки мы рассмотрим позднее, а пока только отметим, что некоторые формы могут определять положение независимо от того, существует ли в данном месте объект (1, 6, 4 варианты), а некоторые подразумевают фактическое наличие объекта в данном месте (2, 5 варианты). Вариант 3 подразумевает наличие специального референсного (ссылочного) объекта - линии, но не обязательно самого искомого объекта. Кроме того, ясно, что некоторые системы привязки в состоянии определять положение объекта однозначно, а другие нет. Например, это относится к использованию имен. Мало того, что может существовать несколько однотипных объектов с одним названием, иногда и объекты разных типов имеют одинаковые названия (Нью-Йорк - название и штата и города, Вашингтон - также название и штата и города, да еще и расположенных в разных концах США.

## **2.3. Измерения и пространственные взаимоотношения.**

Приняв, что рассматриваемые объекты реального мира могут быть четко отграничены от их окружения (это не всегда так, но эти более сложные ситуации мы пока не рассматриваем), и их положение в пространстве может быть задано в той или иной системе привязки, выполнение различных задач с использованием информации об этих объектах может потребовать определения пространственных характеристик этих объектов и наличия средств измерения этих характеристик.

### 2.3.1. Свойства пространственных объектов.

*Свойства, определяемые для индивидуального объекта и только для группы объектов. Характер расположения - компактный и рассеянный. Характер размещения - равномерный и сгруппированный. Характер окружения - число соседей или Характер чередования - наличие закономерностей в последовательности расположения объектов.*

Для начала рассмотрим примеры пространственных характеристик, определенные для индивидуальных объектов. (Рис. 6)

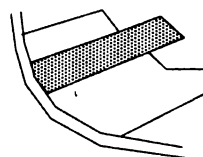
1. **Длина**, например, длина реки или длина береговой линии озера.
2. **Площадь** поверхности, например, площадь озера, острова или конкретного земельного участка.
3. **Объем**, например, количество земли, которое необходимо удалить для постройки дорожной выемки.



Длина реки



Объем нефти в месторождении



Периметр и площадь землевладения



Форма административных округов



Врезание береговой линии в сушу



Уклон откоса



Центры объектов и группы объектов

**Рис. 6** Некоторые основные пространственные характеристики.

4. Форма, например, округлая или вытянутая, прямолинейная или изогнутая.
5. Регулярность или неправильность формы, например, угловатая, состоящая из прямолинейных отрезков граница штата Юта в США, или резко нерегулярная форма морского побережья в Далмации на Адриатике.
6. Ориентировка, например, направление длинной оси друмлина.
7. Объект может иметь центр, как, например, середина дороги или центр города.
8. Уклон, крутизна склона - это пространственное свойство, например, горы.

(Примеры можно продолжить, конечно.)

Как обстоит дело с измерением этих свойств? Все замечательно просто, если мы уже имеем отображение этих объектов на карте или в геоинформационной системе. Или если мы в состоянии посетить объекты на местности и произвести прямые наблюдения и измерения. Однако мы получим различные ответы, например, при измерении длины, в зависимости от точности используемых нами измерительных устройств, а также в зависимости от масштаба, в котором мы изначально зафиксировали объекты. (То есть при разной детальности рассмотрения). Длина береговой линии даже такого относительно простого по конфигурации побережья как Мурманский берег Кольского полуострова будет значительно отличаться, если мы произведем ее измерения по картам масштаба 1:1000000 и 1:50000. В случае более изрезанного побережья различия будут еще значительнее. Поэтому обычно измерения пространственных характеристик объектов реального мира включают и рассмотрение вопроса о необходимой и достаточной точности измерения, допустимой величине ошибки, источниках данных, методах и инструментах проведения измерений. В одних случаях требуется высокая инструментальная точность, как в описании землевладения в земельном кадастре, в других - достаточно визуальной качественной оценки, как это часто бывает при описании формы объекта.

Более того, некоторые пространственные характеристики являются величинами многомерными. Например, форма может пониматься

просто как соответствие по форме той или иной правильной геометрической фигуре (например, кругу или прямоугольнику), а может учитываться еще и каким-то образом измеренная количественно степень отклонения от этой правильной формы. Например, степень кривизны линии. Для описания склона часто недостаточно простого указания его уклона, он может быть различен в различных точках, склон может быть вогнутым или выпуклым. Часто пространственную характеристику можно измерить не одним единственным способом, так что указание ее значения без описания способа измерения может быть лишено смысла. Часто требуется указывать еще и точное место измерения характеристики в пределах объекта. В ряде случаев для одного объекта может потребоваться несколько измерений одной характеристики, в особенности если она меняется со временем. Например, и площадь, и форма некоторых объектов, как техногенных, так и природных может меняться значительно. Например, некоторые соленые и пересыхающие озера и водотоки в аридных областях, положение главного русла реки в пределах широкой горной долины, береговая линия водохранилища, зависящая от режима его использования. Или, например, площадь и форма территории, в пределах которой на спрос сильно влияют покупки, производимые жителями города Москвы, сильно различаются для зимнего и летнего (дачного) сезонов. (И, кстати еще - "сильно влияют". Как сильно? - требуется точно определить ту степень влияния, начиная с которой нас начинает интересовать эта проблема.) В таких случаях иногда прибегают к описанию пространственных характеристик с использованием вероятностного или статистического подхода.

Для более-менее однородных по тематике типов объектов, например, для озер, могут рассматриваться их групповые пространственные характеристики. Например, определенные типичные формы бывают характерны для озер, связанных своим происхождением с тектоническими разломами на древних щитах, другие формы характерны для озер в местах активного развития многолетнемерзлых грунтов.

Как для описания изменчивых во времени пространственных характеристик одного объекта, так и для описания обобщенных, групповых характеристик нескольких однородных объектов, часто

используется один из вариантов такого обобщенного, генерализованного описания.

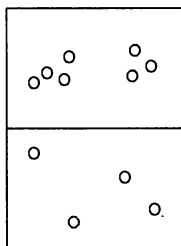
1. Минимальное и максимальное значения характеристики или диапазон возможных значений.
2. Среднее значение характеристики и характеристика степени ее изменчивости (дисперсия, среднеквадратическое отклонение и т.п.).

Некоторые пространственные характеристики имеют смысл и определены только для некоторых множеств объектов, а не для единичных объектов. Кроме того, хотя площадь и форма и могут быть определены для единичного озера, они имеют смысл и для группы озер, например, как средние величины для группы или как суммарная площадь всех озер в пределах данной территории.

Примеры пространственных характеристик для групп объектов:  
(Рис. 7)

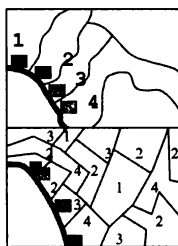
1. Рисунок распределения в пространстве несвязанных между собой отдельных объектов может отличаться (объекты могут располагаться сгруппированно или рассеянно).
2. Объекты могут располагаться компактно или порознь (например, все земельные участки одного владельца могут располагаться компактно вблизи его домовладения, или быть разбросаны на большой территории, перемежаясь чужими участками).
3. Объекты могут граничить с различным числом других объектов того же типа. В том числе, и с нулевым их числом, как, например, изолированные вырубki в лесу. (Например, различные государства граничат с различным числом соседних государств.).
4. Объекты могут быть разным образом связаны, соединены друг с другом, как, например, транспортные сети различных авиакомпаний связаны друг с другом наличием общих аэропортов как пунктов пересадки.
5. Группу объектов могут характеризовать определенные направления, как, например, система преимущественных путей и направлений сезонных миграций диких животных.

Сгруппированное



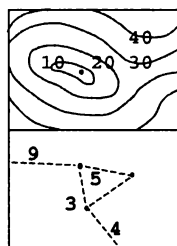
Рассеянное  
Распределения

Компактное



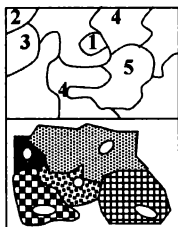
Разбросанное  
Размещение

Аккумулирующие



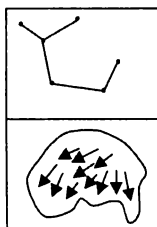
Отдельные  
Расстояния

Соседство



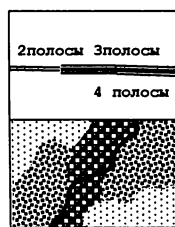
Поляны в лесу  
Изолированность  
или соседство с  
себе подобными

Маршруты авиалиний



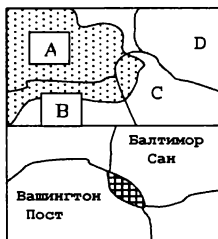
Миграционные  
потоки  
Связи и потоки

Шоссе



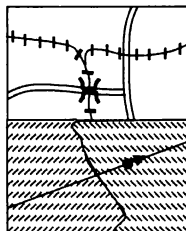
Типы ландшафтов  
Порядок в  
последовательности  
расположения

Почвы и производство пшеницы



Распространение газет  
Частичное перекрытие

Шоссе и железные дороги



Авиалиния и два  
государства  
Пересечения

**Рис. 7** Пространственные характеристики для группы объектов



6. Группы пространственных объектов могут характеризоваться наличием или отсутствием определенной последовательности, закономерности в порядке их расположения (зоны высотной поясности растительности в горах обычно сменяют друга последовательно, а вблизи крупного города сначала обычно располагаются плотно населенные пригороды, а не малолюдные сельскохозяйственные территории).

В этой части книги мы не будем рассматривать конкретные методы проведения измерений, вопросы точности измерений, допустимых ошибок, а сосредоточимся больше на принципиальных вопросах, на типизации свойств пространственных объектов.

Обратим внимание, что ряд характеристик, упомянутых выше (которые можно рассматривать как пространственные), относящихся к группам объектов, фактически определяется характеристиками положения объекта и его атрибутивными (описательными) характеристиками. Например, это касается высотной поясности растительности или типов территории вблизи города. Вообще, когда рассматриваются особенности взаиморасположения объектов, упорядоченности в их расположении, последовательности расположения, мы обычно не можем обойтись одними геометрическими или топологическими характеристиками, а учитываем атрибутивные характеристики объектов, задающие деление их на классы, ранжирование, степень подобия по значениям каких-либо непространственных параметров. Хотя из этого правила и есть исключения - постепенное расширение русла реки вниз по ее течению.

### **2.3.2. Пространственные взаимоотношения. Различие между геометрической и топологической информацией (характеристиками). Взаимосвязь между топологическими и метрическими характеристиками.**

*Геометрическая информация (метрическая) - координаты центра полигона, расстояние между объектами. Неметрическая (топологическая) информация - связность, ориентировка линии, направление обхода контура, прилегание, что в чем содержится (отношения включения), перекрытие, пересечение. Пример - расстояние между городами по железной дороге может быть*

*измерено как в километрах, так и в числе перегонов между станциями. Характеристики, существующие для единичного объекта и существующие только для пар объектов. Атрибутивная информация и пространственные характеристики объектов. Сложная взаимосвязь атрибутивной и пространственной информации.*

В приведенных в предыдущем параграфе примерах неявно проявились самые фундаментальные различия между пространственными характеристиками разных типов. Во-первых, следует различать те из них, которые требуют измерений координат объектов, то есть являются **геометрическими, или просто метрическими**, характеристиками, основанными, например, на расстоянии между объектами, и те свойства, которые основаны на **неметрических** свойствах, как, например, наличие связи между объектами. Последние характеристики или свойства называются **топологическими**. К топологическим относят такие свойства как:

**связность** (наличие связи между объектами);

**ориентация, направленность** (от объекта А к объекту В, а не ориентация в смысле азимута линии);

**примыкание** (наличие общей границы, точек);

**включение, содержание внутри себя**.

Связность мы проиллюстрировали примером авиалиний (см. стр. 66). Пример направленности - река, имеющая направление течения. Примыкание относится к примеру со странами-соседями. Включение касается, например, вырубок в лесу в пределах лесного квартала или островов на озере.

Более подробно топологические характеристики мы рассмотрим впоследствии. Пока отметим, что они обычно не имеют смысла при наличии единственного объекта, а только при наличии нескольких объектов, системы объектов, так как они характеризуют **взаимоотношения** объектов. Поэтому эти характеристики часто называют **топологическими отношениями**.

Некоторые пространственные характеристики можно оценить как измерением (метрически), так и определением топологических отношений. Например, расстояние между двумя железнодорожными

станциями можно описать как расстояние в километрах вдоль железной дороги, так и числом перегонов, промежуточных станций. Близость может рассматриваться и с точки зрения евклидовой геометрии, и с точки зрения примыкания, касания объектов (касаются они друг друга или нет).

Можно заметить, что определение или измерение ряда пространственных характеристик требует наличия хотя бы пары объектов, или, как иногда говорят, диады. Диаду можно рассматривать как особый тип объектов (уже неэлементарный, так как состоит из двух объектов). Например, таблица, содержащая время или стоимость проезда между разными городами, в которой и по столбцам, и по колонкам расположены названия городов, в качестве основного элемента содержит ячейки, определяемые двумя городами. То есть стоимость билета есть функция диады городов. Диаду не всегда можно (и нужно) рассматривать в виде отдельного типа объектов, но тем не менее в некоторых случаях это полезно. Например, при изучении миграции населения, в некоторых транспортных задачах.

Наконец, если мы станем рассматривать пространственные объекты, относящиеся к нескольким различным на содержательном уровне темам, мы можем заметить появление еще одного типа пространственной характеристики, связанной с сонахождением или согласованной изменчивостью различных объектов. Это чрезвычайно важный момент, значение его для анализа пространственных, а через них часто и причинно-следственных взаимосвязей между объектами и явлениями трудно переоценить. Сравнивая особенности совместного распределения в пространстве объектов с различными характеристиками и разных типов, мы приходим к изучению совмещения или наложения (полного или неполного) объектов в пространстве (например, границ земельных участков и типов почв), пересечения объектов (пересечения железной и шоссеной дорог). Важно, что при изучении такого рода пространственных взаимоотношений нам часто тоже нельзя обойтись без рассмотрения значений атрибутов, например, задающих деление объектов на классы по темам. Так, линейные объекты могут пересекаться в реальном мире, если это две шоссеные дороги, если это шоссеная и железная дороги, но очевидно не имеют реального пересечения, если это водопровод и линия электропередачи.

Пространственные взаимосвязи объектов могут рассматриваться, конечно, и для случая трехмерного пространства, хотя мы и приводим все примеры для пространства двумерного. Для анализа пространственных взаимосвязей и изучения особенностей распределения значений какого-либо свойства (выраженного, например, значениями атрибута однотипных объектов) часто привлекается специальный математический аппарат, называемый пространственной статистикой ( spatial statistics).

Наконец, говоря об изучении совместного распределения или размещения объектов, относящихся к различным темам, полезно вспомнить о послойном принципе организации пространственной информации (см. параграф 1.1). В ГИС, организованной в соответствии с этим принципом, операции, требующиеся для изучения и выявления таких взаимоотношений, реализуются как операции между тематическими слоями, как операции взаимодействия этих слоев. При этом часто они являются операциями пространственного наложения (spatial overlay) двух или нескольких слоев в пределах всей области изучения или какой-то ее части, аналогично тому, как это традиционно производилось при работе с бумажными картами с помощью прозрачных калек и светостола. В результате выполнения этих операций выявленные взаимоотношения обычно фиксируются в виде некоторого нового слоя, содержащего, например, в случае наложения двух слоев с площадными объектами, новые площадные объекты, являющиеся всеми имеющимися комбинациями двух исходных. Несмотря на известную громоздкость этой операции, она обладает большим аналитическим потенциалом, а та общность подхода, которая здесь достигается, является одним из сильных аргументов в пользу систем с послойной организацией информации.

В завершение параграфа еще несколько слов о связи и различиях между геометрическими (метрическими) и топологическими (неметрическими) характеристиками. Из самого названия ясно, что метрические характеристики требуют наличия некоторой метрики, то есть, на практике, какой-то системы координат, единицы измерения и указания как мерить (по прямой или вдоль некоторой линии, например). Выражаются измерения обычно в численных значениях, для которых употребимы понятия “больше-меньше”, и результаты измерений зависят от принятой метрики (например, от положения

начала координат и выбора единицы измерения). При любых ее изменениях, также как при любых изменениях формы объектов, результаты измерений могут меняться. (Не все и не всегда, например, при повороте или сдвиге координатных осей площадь объекта не изменится). А вот топологические свойства, не являясь метрическими, не зависят от выбора системы координат, и не меняются даже при многих очень значительных изменениях формы объектов. Точнее, они не изменяются при таких преобразованиях формы объектов, где не происходит их разрывов или склеек, то есть соединений того, что ранее не соединялось. Топологические свойства и взаимоотношения объектов можно вывести из метрических (геометрических), а вот обратное неверно. Если у нас имеются координаты всех точек или уравнения двух отрезков, мы можем сказать, пересекаются они или нет, а из одного факта пересечения линий ничего нельзя сказать об их форме, положении и размерах. В целом можно сказать, что топологические свойства носят более общий, фундаментальный характер.

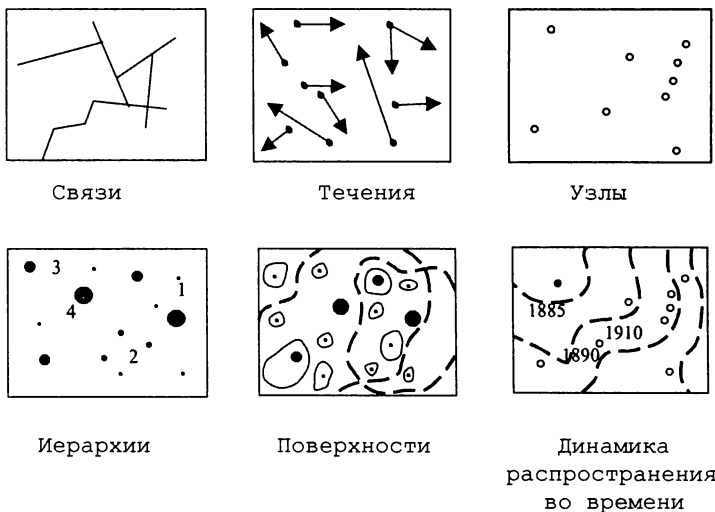
## **2.4. Категории проблем, возникающих при работе с пространственной информацией.**

Теперь мы можем подробнее познакомиться с категориями пространственных проблем, кратко упомянутых ранее, в параграфе 2.1.

### **2.4.1 Типы пространственных проблем и типы организации пространственных данных.**

**Вектора. Сети. Узлы. Иерархии. Поверхности.  
Непрерывные числовые поля.**

Типы пространственных проблем, решением которых мы занимаемся, имеют отношение к пространственной организации соответствующих объектов. Точнее говоря, и различные способы организации пространственной информации в ГИС дают нам различные возможности по решению тех или иных пространственных проблем, и различные типы организации пространственных объектов в реальном мире порождают различные типы проблем, с ними связанных. (Рис.8) Так, например, разбирая типизацию пространственных проблем, есть смысл отдельно рассмотреть такие типы пространственной организации:



**Рис. 8 Категории пространственных организаций.**

1. Множество направлений перемещения или течения, одним из вариантов представления которых являются отдельные вектора или векторные поля (перемещения воздушных масс или направления миграций населения).
2. Пространственные сети, такие как системы инженерных коммуникаций в городе или система железных дорог страны.
3. Узлы - особого рода точечные объекты, являющиеся пересечениями линейных объектов, например, перекрестки дорог.
4. Иерархии объектов, или их ранжированные ряды (главные города административных единиц разного порядка или множество городов, разбитое на классы по численности населения.)
5. Поля или поверхности, описывающие распределение в пространстве какого-то непрерывного признака (геофизические, геохимические поля, рельеф местности).
6. Поля, описывающие динамику распространения в пространстве с течением времени какого-либо явления или величины

(распространение эпидемического заболевания, фильтрация химического загрязнения в подземных водах).

Из приведенного выше списка примеров видно, что решение пространственных проблем может включать анализ таких категорий, как движение и перемещение, развитие и организационная структура, помимо таких собственно пространственных категорий, как расположение, взаиморасположение, расстояние и конфигурация.

## Категории пространственных проблем.

*Категории пространственных проблем по: - типам рассматриваемых объектов- типам пространственных характеристик, принимаемых в рассмотрение. - тому, какие элементы заданы, и какие требуется найти. - содержательным целям.*

Более детальный подход к выделению типичных пространственных проблем должен принимать во внимание:

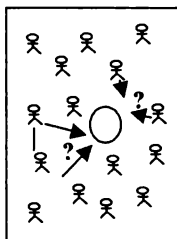
1. Типы объектов, которые рассматриваются и типы их комбинаций.
2. Типы пространственных свойств этих объектов.
3. Элементы, которые даны и элементы, которые требуется найти.
4. Цели.

Например, нахождение пути через лабиринт - это, в сущности, проблема организации последовательности линий - так собрать в определенной последовательности отрезки лабиринта, чтобы гарантировать выход из него. Она может иметь и усложненные варианты, если потребовать найти оптимальный путь из нескольких имеющихся альтернатив. В случае поиска оптимального места для размещения новой закусочной, пространственные объекты могут быть точкой (местоположение здания ресторана), могут вовлекаться в рассмотрение также площадные объекты (административные округа), линейные объекты (дорожная сеть), другие точечные объекты (существующие закусочные). (Рис.9)

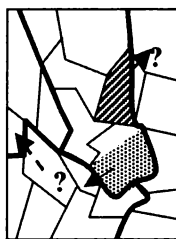
Итак, первый тип пространственных проблем - проблема нахождения оптимального местоположения объекта, задача



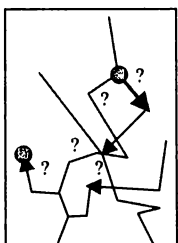
Размещение



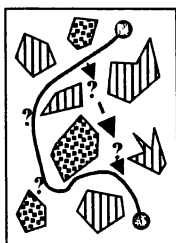
Распределение



Районирование



Маршрут



Путь

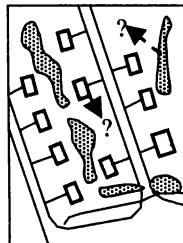
Взаиморазмещение  
группы объектов

Рис. 9 Несколько типов проблемы размещения в двух измерениях .

**размещения (location problem).** В ней искомым неизвестным является одно - оптимальное в смысле некоторого заданного критерия положение нового объекта. В условиях могут быть заданы пространственное распределение возможных потребителей ресурсов (услуг), существующая транспортная сеть, определяющая возможности их доступа к новому объекту, возможно, также, и дополнительные пространственные ограничения (например, набор свободных земельных участков, на которых возможно строительство). Должен быть также задан критерий, который служит для проверки оптимальности положения. Например, если мы ищем наилучшее положение для новой дешевой придорожной закусочной, то мы можем решать задачу, пытаясь выявить положения, где максимальным будет поток возможных посетителей, не заботясь об учете уровня их покупательной способности, то есть не принимая во внимание никакие особенности пространственного распределения



жителей с разным уровнем дохода. Но если мы ищем оптимальное положение для дорогого ресторана, то такой учет, вероятно, уже не годится, и мы должны оптимизировать не количество посетителей, а среднюю выручку за день или за неделю (с учетом того, что в дешевой закусочной стоимость всех заказов примерно одинакова, а в ресторане может отличаться значительно). Кроме того, в качестве дополнительных ограничений могут выступать положения других объектов, конкурирующих за ресурсы (посетителей), например, закусочные конкурирующих фирм. Задача усложняется, но остается принципиально сходной, если требуется найти оптимальное положение не одного, а нескольких объектов, и требуется оптимизировать общий критерий для всех объектов (например, общую выручку всех новых закусочных). Наконец, как стоимость создания новой, так и стоимость эксплуатации каждой закусочной может отличаться в различных местах, например, из-за различия в стоимости земли, местных налогов, стоимости подвоза продовольствия. Таким образом данный тип задачи может усложняться многократно, оставаясь в то же время задачей выбора местоположения.

Задачу более сложного и принципиально другого типа мы имеем, если требуется не только определить оптимальное положение объектов, как правило, нескольких, но и произвести оптимальное распределение связанных с ними ресурсов. Например, требуется не только определить оптимальное размещение поликлиник в городе (возможно, с разными максимальными возможностями по числу ежедневно принимаемых людей), но и оптимально распределить, приписать жителей к этим поликлиникам. Эта комбинированная задача называется задачей размещения-распределения (location-allocation problem). Она тоже может выступать во многих различных усложненных вариантах. Например, помимо неравной пропускной способности поликлиник, пространственного распределения жителей и реалий доступности по транспортной сети, может потребоваться учет таких требований, как различный демографический состав населения на разных участках территории и, соответственно, разная вероятность заболеваний и разная частота обращений в поликлинику. Наконец, может потребоваться выполнение определенного баланса между требованием немедленно (или в течении заданного времени) обслужить каждого посетителя (минимизация очередей) и минимизацией времени работы поликлиники “в недогрузку”, то есть

когда врачи ждут больных. Это же положение может быть выражено в экономических категориях. В определенных “пиковых” ситуациях может быть возможно или невозможно перераспределение ресурсов заново (например, направление части посетителей из перегруженной поликлиники в другую менее загруженную).

Третий тип задач вытекает из второго и является в чистом виде задачей оптимального распределения ресурсов, если положение всех объектов уже задано (allocation problem). В практике применения ГИС это одна из самых важных задач. Отметим, что во всех рассмотренных проблемах, связанных с ресурсами (понимаемыми обобщенно), существуют два типа объектов: обладающие ресурсом (или производящие его) - источники ресурсов и потребляющие ресурс (потребители ресурсов). И потребители, и источники ресурсов могут быть объектами различных геометрических типов, не обязательно только точечными. Например, в задачах организации сбора мусора или снега на улицах города источником ресурсов являются линейные объекты, а потребителем - точечные. Распределение ресурсов может быть неодинаковым по разным объектам. (Один магазин может продавать за день в 10 раз больше определенного товара, чем другой, а снабжаться они могут с одного оптового склада.) Потребители ресурсов могут иметь меняющиеся во времени потребности, в том числе с наложением случайных флуктуаций, которые могут быть описаны только статистически. И вся задача распределения ресурсов, обеспечения ресурсами может тогда приобретать статистический характер (“При заданном уровне знаний о спросе, заданном расположении магазинов и заданном расположении складов и их емкости, заданной конфигурации и пропускной способности транспортной сети обеспечить отсутствие перебоев в торговле данным товаром с вероятностью 95 процентов”. Или “С вероятностью 99 процентов в течение 10 минут после сигнала пожарная команда должна прибыть к месту пожара в пределах своей зоны обслуживания. При этом для ряда объектов должно одновременно выполняться требование с вероятностью 99.9 процентов доехать до них за 5 минут”).

Легко понять, что проблемы, связанные с распределением ресурсов и выбором местоположения тесно связаны с задачами нахождения кратчайшего или в ином смысле оптимального пути.

Полезно заметить, что две близкие по содержательному смыслу задачи - выбор оптимального маршрута при наличии predetermined transport (например, дорожной) сети (route finding problem), и выбор оптимального маршрута передвижения по бездорожью (path finding problem или cross-country movement problem) - существенно различаются в геоинформатике. За двумя этими задачами стоит совершенно различная организация данных и несколько разный математический аппарат. Первые задачи относятся к группе задач анализа пространственных сетей, и опираются на векторное представление данных, вторые используют растровую модель данных и методы минимизации на стоимостной поверхности. И в том, и в другом случае задачи могут различаться по сложности, например, включать также **барьеры (barriers)** - участки, линии или точки на линиях (дорогах) или полностью непроходимые, или затрудняющие передвижение в определенной степени. Для задачи нахождения оптимального маршрута по заданной дорожной сети решение может проводиться с учетом только расстояния (то есть длин линий, составляющих дорожную сеть), или с учетом каких-то других их характеристик (среднее время движения). Эти характеристики, влияющие на скорость перемещения, или, иначе говоря, какое-то сопротивление движению, распределенное вдоль линейных объектов (или связанные с другими типами объектов, например, узлами), получили общее наименование **импеданса**, то есть обобщенного сопротивления какому-то перемещению. В общем случае это могут быть уже не транспортные задачи, связанные с перемещением объекта, а например, задачи анализа электрических сетей, трубопроводных систем. Тем не менее для столь различных задач в рамках общей геоинформатики предлагается эффективный общий подход, основанный на общих свойствах пространственных сетей. Для задач поиска оптимального пути в условиях отсутствия predetermined дорожной сети также может иметься много усложняющих вариантов. Например, может приниматься во внимание различное, иногда очень большое, число факторов, влияющих на передвижение по местности. Эти факторы могут иметь различную значимость как в целом, то есть обладать тем или иным **весом фактора**, так и различные **значения каждого фактора** могут обладать различными **весами**. Комбинируя эти весовые коэффициенты можно построить очень сложную модель проходимости местности, даже привязанную к характеристикам конкретной модели транспортного средства. (Конечно, при наличии

всей необходимой информации о местности). К важнейшей информации о местности, которая требуется для подобных моделей, относятся в первую очередь некоторые цифровые поля, то есть информация о некоторых свойствах, непрерывно распределенных в пространстве: рельеф местности и особенно его производная характеристика - уклон рельефа (крутизна склона), характеристики почвы, растительного покрова. Какой-то фактор, например, уклон до определенных значений почти не оказывает влияния на передвижение (уклоны в 0.5 - 1 градус, например), а далее его влияние возрастает вплоть до абсолютной невозможности перемещения (по английски - "no go"). При дальнейшем усложнении модели может потребоваться учесть тот, например, факт, что какое-то транспортное средство определенный уклон может преодолеть влоб, но не под углом 60 и более градусов к направлению движения. Или оно в состоянии двигаться вверх по уклону 40 градусов, но не более чем на протяжении 100 метров и только при предварительном разгоне до определенной скорости. Может появиться необходимость решать и комбинированные задачи, когда ищется вариант передвижения с частичным использованием имеющейся сети дорог. Другие пространственные сети (речная сеть) также могут влиять и накладывать ограничения на перемещение. Можно представить себе и ситуацию, когда перемещение по абсолютно одинаковой местности в одном месте может быть возможным, а в другом - неприемлемым. Например, по причине невозможности или высокой цены доставки дополнительного горючего в определенные участки территории.

Проблема классификации территории по комплексу признаков (проблема районирования) - одна из самых обычных и самых важных пространственных проблем. В некоторых случаях она решается с применением методов многомерных классификаций, где объектом классификации является какой-то небольшой участок местности, элементарная единица заданной схемы делимости территории. Нередко проблема решается просто комбинированием информации, относящейся к разным тематическим слоям. Обычным решением здесь являются так называемые оверлейные операции, или операции наложения. Но иногда возникают задачи районирования, требующие объединения различных типов пространственных проблем. Например, если мы хотим получить оптимальный вариант зонирования городской территории, учитывающий, помимо других

показателей, степень доступности территории с помощью определенных средств общественного транспорта. Последняя подзадача является разновидностью задачи анализа пространственных сетей, и часто рассматривается и как самостоятельный тип пространственных проблем.

Можно выделить как самостоятельный тип проблем также проблему определения оптимального взаиморасположения различных объектов, может целого большого их комплекса (layout problem). Фактически это задача взаиморазмещения уже очень близка конкретным прикладным задачам проектировщика, градостроителя, ландшафтного архитектора.

#### **2.4.2 Другие аспекты.**

*Например, некоторые пространственные проблемы требуют включения в рассмотрение фактора времени. Проблема выявления произошедших с течением времени изменений в ситуации. О динамических моделях природных и техногенных процессов.*

Дальнейшая классификация пространственных проблем может, например, производиться по типу того пространственного свойства, которое должно быть оптимизировано или установлено в процессе решения проблемы. Например, в конкретной задаче может потребоваться минимизировать суммарное расстояние, которое жителям данного городского района требуется, чтобы посещать свои поликлиники. Но для некоторых ситуаций оптимизация только одного такого критерия может оказаться недостаточной. Например, может быть дополнительно поставлено условие, чтобы ни для одного из жителей это расстояние не превышало бы некоторого порога, или этот порог не превышался бы для 95 % населения. В некоторых случаях может потребоваться совместное нахождение как оптимального значения пространственной характеристики (положения объекта, например), так и атрибутивной характеристики (например, в том же примере с поликлиникой определить еще и необходимую пропускную способность (число принимаемых в день пациентов). Последнее может быть величиной многомерной (имея в виду, что для врачей различных специальностей эта величина может быть различной).

Некоторые пространственные проблемы включают в рассмотрение время. Часто это только время перемещения из пункта в пункт, как в рассмотренных выше задачах транспортного типа (route finding и path finding). Фактор времени возникает также в задачах динамического моделирования, причем это может быть время самого разного порядка - от минут в моделях распространения загрязнений в воздушной среде до миллионов лет, если мы говорим о процессе формирования рельефа под воздействием геологических процессов. В определенном смысле факторы времени и пространства тесно связаны - классическая формула физики дает эту связь через скорость. Поэтому так удобны и широко применяются в геоинформатике, например, способы визуализации областей равной доступности через изохроны - линии, ограничивающие области, которые могут быть достигнуты за определенный промежуток времени объектом,двигающимся из заданного центра с заданной скоростью. Такой анализ может проводиться как при наличии заданной дорожной сети, так и при ее отсутствии или для комбинированных ситуаций. Этот метод отображения является одним из главных при решении задач выделения зон обслуживания.

Говоря о классификации пространственных проблем, нельзя не упомянуть и об их различиях с точки зрения возможностей практического разрешения. Некоторые из типов проблем решаются легко для определенных моделей организации пространственных данных, и решаются с трудом или не разрешимы вовсе при другой организации пространственных данных. В дальнейшем мы подробнее разберем этот вопрос. Сейчас только еще раз подчеркнем важность оптимального выбора модели данных с учетом не только затрат на организацию и ввод информации, но и всех возможных проблем, которые предстоит решать. Например, многие проблемы с легкостью решаются в растровых моделях (на регулярных сетках), но с трудом - в векторных, и наоборот. Для решения некоторых пространственных проблем важно, чтобы в модели данных поддерживались определенные типы пространственных объектов, например, маршруты (транспорта, например), наложенные на существующую сеть линейных объектов (например, улиц). Или может потребоваться возможность трактовать пространственные сети линейных объектов как ориентированный граф (этот подход мы подробно рассмотрим позднее). В некоторых системах узлы (точки пересечений линейных объектов) не могут быть представлены как

отдельный тип объектов. Это также может вызвать затруднения при решении определенного круга задач.

Помимо поддержки определенного типа объектов может потребоваться поддерживать в модели данных определенные типы взаимоотношений между ними. Например, при решении многих задач может потребоваться, чтобы система была способна в явном виде использовать такие отношения между объектами, как связность или прилегание, соседство, то есть явно поддерживать топологические отношения между объектами. Без этого ряд проблем может не иметь удобных для практики методов решения, или они могут оказаться нерационально трудоемкими в вычислительном отношении. Вообще, вопрос практической реализации методов решения той или иной пространственной проблемы часто упирается в возможность или невозможность произвести все необходимые вычисления за практически приемлемое время. Несмотря на колоссально возросшую вычислительную мощь компьютеров и хорошие перспективы ее дальнейшего увеличения, все равно некоторые алгоритмы, имеющие отношение к решению пространственных проблем, требуют при больших объемах данных такого времени вычислений, что их практическое применение в ряде случаев оказывается невозможным. Особенно это относится к алгоритмам переборного характера, когда методом решения является просто последовательный перебор всех возможных вариантов (например, вариантов размещения объекта) и сравнение их друг с другом. Это значит, что могут встретиться такие пространственные проблемы, которые с легкостью решаются для относительно небольших объемов исходных данных, но по мере увеличения их объема требующееся для вычисления время возрастает очень резко и нелинейно, то есть существует практический предел на, например, число объектов, включаемых в рассмотрение при анализе. Кроме этого, могут встретиться и другие практические ограничения, связанные с необходимостью получения всех данных, необходимых для анализа. Принципиально или на отдельных специально подобранных примерах проблема может оказаться решаемой, но в реальной ситуации подбор всех необходимых данных может оказаться невозможен или слишком дорог.

Итак, на том мы завершаем наше краткое рассмотрение пространственных проблем тем, что немаловажным оказывается

критерий практических ограничений, возникающих при реализации того или иного способа решения проблемы, или, иначе говоря, область применимости метода ее решения.

## **2.5. Примеры комплексных задач.**

Многие пространственные проблемы характеризуются значительной комплексностью, многофакторностью и неопределенностью, как в постановке, так и в методе решения. В особенности это относится к проблемам, касающимся в том или ином смысле природной среды. Это понятно - чем сложнее объект, тем сложнее проблемы. Но и многие инженерные приложения ГИС сегодня ставят задачи, требующие привлечения сложных методов и интенсивных вычислений. В сложности очень многих реальных задач, нелегко сводящихся к типовым пространственным проблемам, рассмотренным выше, и заключается одна из серьезных трудностей практического использования ГИС. Знание не только конкретных функций конкретной ГИС, но и понимание общих принципов и подходов геоинформатики, как и разработка ее специальных прикладных разделов, сможет, думается помочь в этом.

Приведем пару примеров для иллюстрации реальной сложности некоторых задач на практике.

### **2.5.1. Прогнозирование месторождений полезных ископаемых.**

Сегодня - это одна из очень типичных задач, которые ставят перед ГИС. Для нее характерна работа с огромным, но тем не менее обычно недостаточным объемом исходной информации. Исходная информация обычно представляет собой комбинацию трех ее типов. Во-первых, информация о непрерывно распределенных в пространстве числовых признаках, то есть числовые поля. (А еще чаще - только некоторые ограниченные данные для восстановления этого поля, обычно замеры его величин по некоторой нерегулярной сети точек). Во-вторых, масса точечных объектов, имеющих обычно очень большой набор атрибутов, как измеренных в сильных (количественных) шкалах, так и качественных. В третьих - это масса площадных и линейных объектов, также имеющих большой набор



разнообразных характеристик, но в первую очередь - качественных. Существует обширная литература, посвященная попыткам систематизации, типизации и формализации геологических задач. Мы здесь дадим только очень упрощенную и неполную схему, при этом касающуюся только двумерных (плоских) задач, в то время как геологические задачи, по крайней мере те из них, что имеют дело с крупным масштабом рассмотрения объектов, высокой детальностью работ, всегда существенно трехмерны, имеют дело не с картами только, а с объемами - телами в трехмерном пространстве. Мы же про это здесь пока забудем.

Первая задача связана с построением достаточно полного представления распределения в пространстве (у нас - двумерном, то есть на карте) значений какого-либо признака или нескольких признаков, которые реально определимы только в некоторых точках или участках. Речь может идти как о качественном, так и о количественном признаке, или о какой-то их комбинации. Например, нас интересует распределение по площади содержания в горных породах какого-либо химического элемента, или распределение пород, обладающих какой-либо определенной характеристикой, например, возрастом образования. Эта задача обычно включает и интерполяцию, и экстраполяцию, то есть определение искомым значений как в точках пространства, лежащих между точками (областями) с известными значениями, так и за пределами всей области наблюдений на некотором отдалении от нее. Эта фактически обычная задача картирования распределения какого-либо признака при неполноте наблюдений. Неполнота наблюдений - самая обычная и повседневная проблема в геологии. Получить прямым наблюдением или измерением достаточно полную информацию о каждой точке пространства, даже двумерного, просто невозможно - из-за отсутствия физической возможности произвести наблюдение (например, из-за несплошной обнаженности данного участка территории из-под растительного и почвенного покрова), а также и из экономической невозможности произвести наблюдения с требуемой детальностью.

Вторая проблема встает при необходимости опять-таки восстановить на какой-то площади значения определенного показателя, но уже ненаблюдаемого и неизмеряемого непосредственно, а только на основании каких-либо косвенных,

непрямых данных. Это первая задача со всеми ее сложностями, но с дополнительными проблемами оценки взаимосвязи между наблюдаемыми и искомыми признаками, определением информативности различных признаков и т.д.. Именно к этой задаче обычно сводятся многие подходы при поисках месторождений полезных ископаемых.

Наконец, третий тип задач - это задачи районирования территории по комплексу признаков, выявления участков, более и менее похожих друг на друга по комбинации значений признаков.

Нетрудно заметить, что все эти задачи взаимосвязаны и на практике постоянно переходят одна в другую. Отметим, что мы упомянули только статические задачи, не интересуясь типами природных процессов, которые сформировали наблюдаемую геологическую ситуацию, ни, тем более, динамикой этих процессов, то есть относительным или абсолютным временем, когда происходили какие-то природные явления. Набор методов, который может применяться при решении этих задач может быть весьма разнообразен. В числе прочих, часто применяются методы математической статистики и особый ее раздел - пространственная статистика. Но базовые, основные подходы даются геоинформатикой, хотя, начиная применять ее методы, специалисты-геологи даже и не подозревали о существовании такой науки.

Рассмотрим один очень упрощенный и абстрактный пример (Рис.10). Мы не будем здесь разъяснять и комментировать специальную геологическую терминологию, используемую в примере. Ее понимание здесь не имеет существенного значения.

Имеются следующие карты распределения в пространстве наблюдаемых или измеряемых признаков:

Содержание магния в кальците, содержание железа в сфалерите, удлинение оолитов, степень сжатия складок, материал выполнения жил, содержание натрия в мусковите, плотность кливажа, сульфидная минерализация, размеры кристаллических зерен, зоны дробления. Среди этих признаков есть количественные и качественные. Для всех сначала решается задача типа 1 (не представлено на рисунке) - построение карт распределения значений признака по всей



Рис. 10

территории изучения. Пока это не выполнено, мы сплошь и рядом будем иметь ситуацию, когда в одном месте нам известен один набор признаков, в другом - другой, а это крайне осложняет анализ, может быть, даже делая его невозможным. В связи с этим заметим, что выполнение этой задачи - построение карт распределения признаков по отдельным наблюдениям - может потребовать преобразования модели данных, способа организации информации в пространственной базе данных, часто это будет преобразование из векторного в растровое представление.

Далее выполняется комбинирование первичных (наблюдаемых) признаков и формирование по ним карт распределения рудоконтролирующих факторов - карта палеотемператур, карта степени деформаций, карта проницаемости горных пород. Это еще более сложный и неоднозначный процесс, он опирается на наше знание причинно-следственных связей между различными признаками и характером и параметрами геологических процессов. Тем самым он является еще более специфичным для данной конкретной прикладной области (геологии) и, соответственно, здесь меньше роль методов и приемов общей геоинформатики.

Комбинирование пространственных атрибутов дает возможность выделить участки, более перспективные для обнаружения месторождения руд определенного типа - или из общих представлений об условиях их образования, или по аналогии с известным их местонахождением. Такое комбинирование может выполняться с использованием разных моделей данных. Иногда при этом роль стандартных методов общей геоинформатики оказывается очень высока.

В случае трехмерной задачи все значительно усложняется за счет меньшей доступности точек для наблюдения и измерения первичных признаков прямыми методами - это могут быть, кроме поверхности, только отдельные линии скважин и горных выработок. В остальных случаях приходится даже для измерения первичных признаков привлекать косвенные геофизические методы. Во вторых, сами объемы вычислений и сложность математических процедур для работы с трехмерной моделью многократно возрастают.

При решении таких сложных и специфических задач разделение функций между ГИС и другим программным обеспечением может

быть проведено различным образом. Например, нередко необходимая (в данном примере - особенно на первой и второй стадиях) статистическая обработка данных выполняется в среде какого-то специализированного пакета. Те же операции, в которых выполняется собственно пространственный анализ данных и визуализация, обычно сегодня выполняются в среде ГИС. Но еще недавно здесь вместо ГИС общего назначения использовались специализированные пакеты, предназначенные для анализа именно геологических данных и прогнозирования месторождений. (Например, отечественные системы типа "Регион", "АСОД-Прогноз", "Лидер" и др.) Сегодня наблюдаются две тенденции. С одной стороны растет понимание универсальности многих подходов в работе с пространственными данными, и многое переносится в среду ГИС общего назначения. С другой стороны, упрощается интеграция различных программных средств для решения одной задачи. И сегодня трудно сказать, по какому пути пойдет, например, расширение применения статистических методов в геоинформатике - путем включения все большего числа функций статистического анализа непосредственно в среду пакета ГИС (Например, в ARC/INFO 7 мы уже имеем немалый набор возможностей), или будут развиваться средства межпакетной интеграции (чему опять таки есть примеры в среде ARC/INFO 7 - интеграция этой системы и статпакета S-Plus).

### **2.5.2. Моделирование инженерных сетей.**

Не менее сложные задачи возникают в сфере работы с инженерными сетями (например, инженерными коммуникациями города). Имеются в виду, например, системы водо и газоснабжения, электрические сети, канализация бытовая и ливневая, телефонная сеть. До некоторой степени возникающие здесь задачи подобны задачам транспортного типа, рассмотренным ранее, поскольку в обоих случаях основа - это анализ пространственных сетей. Отсюда естественно проистекает требование к топологической корректности всех пространственных данных. Естественно также, что при работе с инженерными сетями важное значение имеют функции ГИС как информационно-справочной системы - инженерное хозяйство даже небольшого города очень велико, и задачи инвентаризации, организации множества документов как по сетям в целом, так и по их отдельным элементам и устройствам очень важны сами по себе.

Но все еще гораздо сложнее. В этих сетях постоянно протекают или могут протекать многочисленные процессы, зачастую физически достаточно сложные. Поэтому могут возникнуть требования динамически отслеживать реальное состояние объектов, например, нагрузку электрических сетей или степень заполненности ливневой канализации, а также требования промоделировать процессы в сетях, чтобы, например, убедиться, что в своем сегодняшнем состоянии сеть ливневой канализации справится с ливневыми осадками определенной интенсивности. Здесь мало иметь топологически корректную пространственную информацию и достаточный набор атрибутов (сечения и уклоны труб, например). Нужна еще более или менее полная математическая модель течения воды по безнапорным трубопроводам. Опять-таки, эта модель может быть реализована в рамках отдельного специального пакета, а ГИС будет выступать в роли только пространственной базы данных, средства визуализации и анализа пространственных сетей. Или же эта гидродинамическая модель реализуется прямо в ГИС. Сегодня снова есть примеры того и другого подхода. В данном случае, однако, больше шансов за то, что будет сильнее развиваться именно встраивание расчетной модели в ГИС. Почему? Эта модель оказывается более тесно связанной с задачами пространственного анализа, чем это кажется с первого взгляда. Во-первых, эффект для одной и той же системы ливневой канализации может оказаться различным при выпадении одного и того же количества осадков одной и той же интенсивности в разных местах города и при разной форме и направлении движения пятна осадков. Это зависит и от рельефа местности в городе, и от конфигурации сети. Например, ясно, что эффект будет различен в случаях движения зоны выпадения ливневых осадков в направлении основного стока системы канализации или в противоположном направлении. В первом случае может возникнуть эффект дополнительной аккумуляции и система может переполниться. Важно также иметь возможность оценить, учитывая характеристики городского ландшафта, где могут образовываться при ливне потоки воды повышенной мутности (например, за счет смыва глины с территорий строительных площадок) и где, соответственно, велика опасность заиливания и засорения системы. Не менее, если не более, сложная специфика имеется при работе с другими типами инженерных сетей.

## **2.6. Основные категории средств и подходов, применяемых при решении пространственных проблем с помощью ГИС.**

*Средства сбора, получения, ввода данных. Средства управления данными, поддержания целостности данных, организации запросов. Средства визуализации, вывода, оформления данных и результатов. - Собственно средства обработки и анализа данных.*

Набор инструментов, с помощью которых ГИС или пространственная информационная система выполняют свои функции, можно рассматривать в различных аспектах. Во-первых, их можно разбить на несколько больших групп - группа функций, связанная с вводом, изменением и организацией данных, с организацией запросов и выводом данных, с анализом данных. К ним добавляются часто недооцениваемые по важности функции администрирования пространственной базы данных, поддержания ее целостности в процессе модификации, обеспечения необходимого режима разграничения прав доступа к информации. Во-вторых, все эти инструменты можно рассматривать на уже упоминавшемся физическом ("внутреннем") уровне - включая детали организации данных на конкретных носителях, функционирование компьютера, сети и периферийных устройств. Можно рассматривать их на уровне конечного пользователя, который не обязан знать, "как все это устроено внутри", но которому желательно знать и понимать логику выполняемых операций и то, как практически ими можно воспользоваться. Можно вообще не вникать в существо выполняемых системой операций даже на логическом и концептуальном уровне, а оценивать только соответствие получаемых результатов поставленной задаче.

Еще один аспект заключается в том, как эти средства задокументированы и как организована работа с ними пользователя - можно оценивать качество пользовательского интерфейса, его удобность, логичность организации, насколько легко они осваиваются пользователем.

Можно говорить - и это часто служит предметом обсуждения - о большей или меньшей функциональной полноте возможностей того или иного пакета. Этот аспект часто является главным при

классификации и сравнении программных средств и даже является определяющим при отнесении программного средства вообще к ГИС или пространственной информационной системе. Существует даже выражение “полнофункциональная ГИС”, которое, правда, сегодня вытесняется термином “профессиональная ГИС” - именно потому, что представление о том, что считать “полным” набором функций для ГИС меняется с течением времени и вряд ли может быть закреплено навсегда каким-либо списком.

### **2.6.1 Классификация функций и инструментов ГИС.**

Сами же отдельные функции и инструментальные средства ГИС можно классифицировать, например, по такой схеме:

1. С каким типом представления данных работает данная функция - графическим (чертежом, картой или картинкой), алфавитно-цифровым, звуковым.
2. С пространственной или непространственной информацией мы имеем дело (или, может быть, и с той и с другой).
3. Отношение к вводу-выводу информации из системы - имеет ли место пополнение ее новыми данными, вывод данных или это операция, не производящая ввода-вывода данных.
4. Является ли функция одношаговой - неразделимой на уровне пользователя на более элементарные операции, или она включает много отдельных шагов.
5. Требуется ли данная операция диалога с пользователем, или после запуска целиком может выполняться без его участия.

Это, конечно, только очень грубая и примерная схема. В нашем дальнейшем изложении (в первой части книги) нас не будет особенно интересовать качество и форма организации пользовательского интерфейса, а больше существо выполняемых функций.

### **2.6.2. Функции ввода данных.**

Вопросы ввода данных занимают одно из важнейших мест среди функций ГИС. Если же говорить о важности этих функций на данном



конкретном этапе внедрения геоинформационных технологий, то это, может быть, и центральное место. Следует различать функции по первичному вводу информации в пространственную базу данных, так называемый массовый, или первичный ввод или первичная загрузка базы данных, и функции вторичного ввода, который производится уже в процессе эксплуатации системы. В некоторых случаях эти функции могут не различаться - возможны ситуации, когда заполненная однажды данными ГИС функционирует без изменения набора данных в ней. Но это достаточно редкий случай. Чаще бывает, что система начинает функционировать и выполнять конкретные задачи при еще незавершенном ее первичном заполнении. Тогда функции первичного и вторичного ввода четко неразделимы. Для чего вообще стоит останавливаться на этом моменте? Дело в том, что и задачи массового первичного ввода и вторичного ввода имеют определенную специфику, и нередко разные технологические схемы, разные функции и даже разное программное обеспечение будут оптимальным для использования в том и другом случае. Первичный ввод часто выполняется сразу в больших объемах, нередко путем переноса данных из другой компьютерной системы (импорт данных) или путем оцифровки большого количества часто однотипных бумажных документов. Для этого иногда создаются временные группы специалистов или вообще эта задача поручается другой специализированной организации. Большие объемы однотипных данных дают возможность затратить значительное время на разработку узкоспециализированной, но эффективной именно для данной конкретной ситуации технологии ввода (например, подобрать оптимальный режим сканирования карт, оптимальное ПО для векторизации и тонко настроить параметры его работы.). Вторичный ввод данных может быть различным по характеру. Это может быть модификация, редактирование существующих данных - изменение значений атрибутов объектов, редактирование объектов с изменением их формы или/и положения. Это может быть дополнение существующих тематических слоев новыми объектами, как в пределах той территории, где уже имеются данные в базе, так и с наращиванием, расширением территории. Это может быть дополнение базы данных новыми тематическими слоями, ранее в ней отсутствовавшими, и объектами в них. В целом это могут быть более разнообразные процессы, чем при первичном вводе, использующие более разнообразные технологии и функции системы.

В целом для них всегда характерна необходимость согласования вновь вводимых данных с уже имеющейся в пространственной базе данных информацией (например, может потребоваться необходимость точного согласования - обеспечения точного совпадения общих границ прежнего и нового объектов). Вторичный ввод информации может потребовать модификации структуры пространственной базы данных, например, добавления нового тематического слоя.

В некоторых случаях ввод данных в пространственную базу данных может быть сопряжен с процессом первичного сбора данных. Такие технологии находят все больше применения. Например, данные могут поступать с цифровых геодезических приборов, приемников глобальной спутниковой системы позиционирования непосредственно в ГИС, эксплуатирующуюся в полевых условиях на мобильном компьютере. И атрибутивные, описательные характеристики объектов могут вводиться прямо в процессе их получения - с клавиатуры человеком, работающим в поле или непосредственно с полевого прибора или стационарно установленного датчика. Подобные технологии, равно как и технологии с использованием данных дистанционного зондирования - космических и аэросъемок, если их дешифрирование - выделение на них объектов производится на компьютере, позволяют совершенно миновать бумажную стадию при подготовке данных для ГИС. Бумажные карты и отдешифрированные отпечатки снимков устраниаются из процесса ввода. Карта становится только средством визуализации той или иной информации из ГИС и рассматривается в таком качестве уже не в разделе, посвященном вводу данных, а вместе с функциями вывода данных.

### **2.6.3. Функции вывода и представления данных.**

Функции вывода могут отличаться у систем, ориентированных на выполнение потока отдельных информационных запросов и систем, ориентированных на выполнение серьезных аналитических операций. Последние чаще выдают вещественные документы - твердые копии карт, таблиц, диаграмм, отчетов. Запросного типа системы выдают преимущественно карты, краткие таблицы и диаграммы только на экране.

#### **2.6.4. Функции обработки и анализа информации, администрирования данных.**

Функции обработки и анализа информации - наверное самые важные. Некоторые потребности по анализу мы уже кратко охарактеризовали ранее, в основном на примерах. Здесь только отметим, что функции обработки могут быть связаны только с пространственной компонентой информации, только с описательной (атрибутивной) компонентой или с обеими.

Разумеется, важна группа функций, обеспечивающих - администрирование базы данных, контроль ее состояния, управление доступом.

#### **2.6.5. Функции ГИС и типы применений ГИС.**

По требованиям к наличию или относительной развитости тех или иных функций в программном обеспечении ГИС, можно выделить 12 основных направлений, типов применения ГИС. Редко сегодня встречаются программные продукты, специально предназначенные только для выполнения задач какого-либо одного из этих направлений, но и все направления также не всегда полностью покрываются какой-либо одной системой. Мы можем рассматривать также группировку функций ГИС в соответствии с этими типами применений.

1. Простое копирование бумажных карт в электронной форме с помощью сканера или дигитайзера (automated mapping). Основная цель - возможность просматривать карты на экране и получать копии карт на плоттере с легкостью внесения необходимых изменений в графический компонент. Это одна из самых простых и примитивных функций ГИС. Здесь не предусматривается ни активной работы с описательными характеристиками объектов, ни какие-либо аналитические функции. Система визуализации (условных знаков и других элементов оформления) может быть развита, но она статична. В типичном случае не предусматривается возможность задания условного знака через значения атрибутивных характеристик. Подобным образом поставленная задача в основном интересовала производителей бумажных карт (тех из них, кто издавал карты путем копирования, видоизменения или пересоставления старых, не имея дело с

реальными новыми данными или с быстро меняющейся ситуацией), а также организации (например, владеющие инженерными сетями), которые были заинтересованы в ведении графического документооборота для целей обзора и инвентаризации своего распределенного по территории хозяйства.

2. Компьютерный подход к тематическому картографированию - множественность атрибутов, связанных с одними и теми же объектами дает возможность для одной и той же контуровки легко получать множество карт разного тематического содержания, с разными схемами классификации для одного и того же признака, с использованием разных картографических методов представления информации (карты плотностей точек, цветовые легенды, вписанные диаграммы, масштабируемые значки...). (Thematic mapping). Более развитые функции, в основном за счет управления визуализацией через значения атрибутов. Классическая функция сегодняшних настольных ГИС и ГИС-вьюеров. В отличие от предыдущей группы хорошо обеспечивается визуализация площадных объектов.

3. Комплексирование информации методом наложения карт (оверлейные операции) (Map overlay). В примитивном виде (как просто включение и отключение тематических слоев и просмотр различных их комбинаций на экране) обеспечивается практически везде, в том числе в группах 1 и 2. Однако как полноценные операции, где в результате формируются новые слои, и в их формировании участвуют атрибуты объектов и при этом происходит перегруппировка или объединение атрибутов для вновь формирующихся объектов, обеспечивается далеко не всеми системами. Одна из очень серьезных и важных на практике аналитических функций ГИС - задачи районирования территории по комплексу признаков, выбора оптимального месторасположения нового объекта и др. Практический интерес представляют в особенности функции наложения с участием площадных объектов.

4. Анализ пространственных данных, включающий рассмотрение их описательных характеристик, таких, например, как средний размер населенных пунктов в пределах определенной территории или урожай зерновых в пределах определенной климатической зоны.

5. Различные виды статистики по атрибутам (непространственной) по пространственным объектам (+ картографическое представление результатов).

6. Пространственная статистика, включающая характеристики размещения объектов, но не включающая рассмотрение значений их атрибутов (например, характеристика степени равномерности (случайности) распределения объектов с помощью критерия хи-квадрат).

7. Пространственный анализ (использующий статистические методы или нет) формы объектов или других характеристик, требующих проведения пространственных измерений на самих объектах.

8. Пространственная статистика, включающая в рассмотрение как расположение, так и атрибутивные характеристики объектов (например, построение карт трендов, пространственная корреляция между удалением от дорожной сети и степенью нарушенности первичного биоценоза).

9. Пространственный анализ, использующий моделирование (“что если”, моделирование процессов с построением их математических моделей...).

10. Интерактивный просмотр информации в пространственной базе данных с включением неформальных элементов анализа ее человеком (традиционный метод работы, где все решения и анализ - за человеком, а информационная система выполняет только техническую работу, обеспечивая удобство и наглядность работы с информацией за счет “количественного” фактора - увеличения скорости доступа к информации - обеспечивается качественный скачок.

11. Решение задачи методом организации запроса к пространственной базе данных - формулирование условий и ограничений для поиска объектов и визуализация найденных объектов.

12. Комбинирование логического и пространственного анализа - системы поддержки принятия решений и экспертные системы, учитывающие пространственную информацию.

Как видно, эта классификация типов применения ГИС или групп функций не является делением на какие-то конкретные прикладные

области применения, хотя определенные связи с конкретными задачами, конечно, просматриваются. Здесь внимание концентрируется на типах информации, которые должны быть задействованы в работе, на типах функций, которые работают с этой информацией. При всей условности данной классификации, она удобна как независимая от конкретных областей применения, и в дальнейшем в нашей книге мы будем неоднократно на нее ссылаться.

## **2.7 Элементы моделей пространственных данных, их взаимоотношения. Модели данных и их разновидности.**

Все предыдущее рассмотрение типов потребностей, измерений и пространственных взаимоотношений, пространственных проблем, хотя и было несколько абстрактным, имело основной целью показать, взаимосвязь этих понятий, показать, насколько это возможно на данном уровне рассмотрения, что очень многое определяется разумной организацией данных в пространственной информационной системе. То есть, иначе говоря, очень важную роль играет правильный выбор модели пространственных данных.

### **2.7.1 О моделях данных, структурах моделей данных и форматах файлов.**

Модель данных - это концептуальный уровень организации данных, логический уровень. Если говорить о компьютерных моделях данных, а нас только они здесь и интересуют, то это всегда цифровое представление данных. Не только числовые характеристики, но и информация положения, и пространственная геометрическая (положения и формы), и пространственная топологическая (взаиморасположение и связи объектов), и описательная словесная (неколичественная) информация всегда в компьютерных моделях данных присутствуют в числовой форме. Термины типа “полигон”, “полилиния”, “дуга”, “идентификатор”, “таблица” как раз относятся к этому уровню, равно как и понятие “слой”, “тема”, “способ индексирования”. Это более уровень пользователя и администратора базы данных, а также и разработчика систем, и касается все это не столько программирования, сколько некоторых разделов

математики, с одной стороны, и соотношения элементов модели данных с действительными объектами реального мира, с другой. Основным решаемым на этом уровне вопросом - это вопрос степени адекватности модели данных решаемой задаче. Достаточно ли точно, с сохранением всех ли нужных связей выбранная модель данных позволит смоделировать в компьютере реальную ситуацию? И к рассмотрению проблем на этом уровне обязательно следует привлекать пользователя - специалиста по конкретной задаче, конкретному применению ГИС.

Более детальный уровень рассмотрения организации данных часто называется структурой модели данных. Здесь уже фигурируют и математические и программистские термины, такие, как матрицы, списки, системы ссылок, указатели, механизмы сжатия информации, а основные вопросы, рассматриваемые на этом уровне - это вопросы эффективности в смысле экономии места и быстродействия, то есть уже чисто внутренние компьютерные вопросы эффективности.

На следующем по детальности уровне организации данных мы уже имеем дело со структурой файлов данных и их конкретными форматами. Это уже вопросы размещения информации в файле - что вынесено в его заголовок, и как он организован, что в какой последовательности и какими формами представления чисел записано в этих файлах. Это уровень уже чисто программистский, рядовому конечному непрограммирующему пользователю вникать в этот уровень обычно уже нет необходимости совсем. Другое дело, что назначение конкретных файлов в определенных структурах данных, возможности и ограничения по использованию конкретных форматов файлов знать полезно.

Ну, и наконец, следует сказать об уровне организации конкретной структуры базы данных ГИС, которая уже может быть уникальна для каждого конкретного проекта - какие объекты вносятся в базу данных, как они распределены между слоями (если мы используем принцип послойной организации данных), какие используются классификаторы и т.д. Этот уровень сегодня не является предметом нашего рассмотрения. Разумеется, данная схема весьма условна и не претендует на строгость изложения.

Выбор того или иного способа организации данных в ГИС определяет очень многое, почти все. Определяет гораздо больше, чем выбор того или иного конкретного программного пакета. Выбор модели данных напрямую определяет многие функциональные возможности создаваемой ГИС, так как некоторые функции по работе с пространственной информацией просто невозможно реализовать для определенных типов организации данных, или эти функции будут обеспечиваться только путем чрезмерно сложных манипуляций. Способ организации данных в ГИС, то есть выбранная модель данных, напрямую определяет также и применимость тех или иных технологий ввода данных. От выбора модели данных в не меньшей степени зависит также достижимая пространственная точность представления геометрической информации, возможность накопления в пространственной базе качественного, кондиционного и внутренне непротиворечивого материала. Возможности по организации тех или иных процедур контроля качества и непротиворечивости пространственных данных существенно зависят от выбранной модели данных. Возможность организовать работу с большими объемами данных или с точными данными по большим территориям также связана не только с особенностями и ограничениями конкретного программного пакета, но в еще большей мере - с типом и особенностями выбранной модели данных. Такие важные для практики аспекты, как удобство редактирования и обновления данных, возможность организации многопользовательской работы с пространственной базой данных в режиме редактирования - тоже связаны в первую очередь с моделью организации данных, и уже во вторую - с выбором конкретного программного обеспечения.

Хотя модели данных, структуры данных - это достаточно абстрактная категория и, может быть, самый сложный и математизированный раздел теоретической геоинформатики, определенное представление об этом должны иметь, на взгляд автора, и конечные пользователи ГИС, а не только системные администраторы и программисты. Тем более это относится к руководителям проекта, технологам, специалистам, занятым контролем качества и сертификацией цифровых карт. Наконец, эта тема важна даже для руководителей более высокого уровня, участвующих в решении принципиальных и стратегических вопросов внедрения и развития ГИС-технологий.



Дело в том, повторим, что выбор той или иной модели организации пространственных данных - вопрос гораздо более принципиальный, чем выбор конкретного программного пакета в рамках одного типа модели и тем более конкретной компьютерной платформы. Ошибки в решении этого вопроса могут проявиться решающим образом в самой возможности выполнения ГИС требуемых функций, в возможности расширения списка этих функций в будущем, в успешности или неуспешности проекта с экономической точки зрения и, что очень важно - могут определить ценность накапливаемых пространственных баз данных в долговременной перспективе. Будут ли Ваши данные совместимы с другими, будут ли они нужны другим, наконец, не придется ли Вам на каком-то этапе эксплуатации и развития системы бросить все и переходить на принципиально другую систему, поддерживающую другую модель данных, и придется ли Вам при этом отвергнуть весь уже накопленный материал и вводить его заново, или нет - это во многом зависит от правильного ответа на вопрос о выборе способа организации пространственных данных, модели данных. Как известно, обмен данными между двумя разными ГИС и даже полная смена программного пакета - обычно не очень большая проблема, если используемые в них модели данных близки. В других же случаях может оказаться так, что объем труда, необходимого для конвертирования существующих данных сопоставим с затратами на повторный ввод информации или даже превышает их. И это все несмотря на то, что проблемы формальной конвертации форматов файлов обычно не существует. Так что понятия “защита инвестиций, вкладываемых в создание ГИС” и “модель данных” не столь далеки друг от друга, как это кажется на первый взгляд. Такие сугубо практические моменты мы затрагиваем здесь, в разделе, посвященном теоретической геоинформатике, не случайно. Мы хотим подчеркнуть практическую важность этого раздела, касающегося моделей данных. К сожалению, иногда его значение недооценивается. Мы и в следующих разделах книги еще неоднократно будем обращаться к теме организации данных в ГИС и пространственных базах данных, рассматривая этот вопрос и более детально, и более формально, и более практически. Пока же в заключение данного выпуска мы рассмотрим вопрос об организации данных в ГИС в самом общем виде.

Вспомним, что информация о реальном мире частично относится к индивидуализированным объектам, а частично к непрерывно

распределенным в пространстве свойствам. Последние можно рассматривать как пространственные поля каких-либо характеристик, числовых или нечисловых. Например, индивидуализированные объекты - дома. Они находятся на некоторой местности, рельеф которой удобно представить себе в виде числового поля - поля абсолютных высот рельефа. Независимо от способа организации информации в нашей ГИС, независимо от того, насколько детально и правильно мы в ней представили рельеф местности, у нас не возникает сомнений, что эта характеристика носит принципиально непрерывный характер - нигде не может быть участка с "никаким" рельефом местности, с отсутствием этой характеристики. Могут быть участки, где рельеф скрыт от непосредственного наблюдения (под зданием), где значение этой характеристики неоднозначно (вертикальные или нависающие обрывы), но сама непрерывность ее не вызывает сомнений (если, конечно, считать за ту же характеристику и рельеф суши, и рельеф дна водоемов - или водной поверхности). И тем самым просматривается ее принципиальное отличие от индивидуализированных объектов. На самом деле взаимоотношения здесь более сложные и не столь антагонистичные, но эти детали мы опускаем.

Индивидуализированные объекты, в свою очередь, могут иметь четкие, определенные границы или положение, достаточно хорошо описываемые, например, линией их контура (для площадного объекта), или быть размытыми с нечеткими границами (с зоной постепенного перехода на границе, с границей, определенной статистически или с границей, переменной во времени.) Как всегда, с нечеткими объектами все гораздо сложнее, мы их здесь пока подробно рассматривать не будем.

Пространственные поля могут быть полями характеристик, измеряемых в разных шкалах (количественных и качественных). Они могут быть одномерными и многомерными, то есть представляющими одну или одновременно несколько характеристик. Они могут быть скалярными и векторными, то есть характеризоваться числом без направления, как, например, поле температур, так и числом с направлением в пространстве (например, поле ветра, характеризующееся величиной и направлением). Существуют и более сложные их типы, но нам достаточно и этого.

## 2.7.2. О представлении индивидуализированных объектов.

Для начала рассмотрим индивидуализированные объекты.

Индивидуальные объекты могут иметь различную размерность в геометрическом смысле - быть точечными (нулевая длина и ширина, размерность 0), линейными (нулевая ширина при ненулевой длине, размерность 1) и площадными (ненулевая длина и ширина, размерность 2). Следует отделять от понятия размерность объекта понятие размерности пространства его описания. Линейное пространство (размерности 1) - это, например, система координат вдоль линии, такая как километровые столбы вдоль шоссе; двумерное плоское пространство (размерности 2) - это, например, карта с системой картографических координат; трехмерное пространство (размерности 3) - это наше обычное "реальное" пространство. Встречается даже термин 2.5-мерное пространство - он относится к 3-х мерному пространству, в котором определена некоторая неплоская в общем случае поверхность с системой двумерных координат на ней. Непрерывного объема нет - ничего нельзя позиционировать в свободном трехмерном пространстве, но каждая точка имеет координаты  $X, Y, Z$ . Однако каждая из точек с одними  $X$  и  $Y$  имеет только одно  $Z$ . Такого типа пространство удобно для описания пространственных данных об объектах местности вместе с рельефом местности, на которой они находятся, для работы с перспективным изображением местности, для создания, как сейчас принято говорить, "виртуальных" отображений действительного или смоделированного ландшафта - высокореалистичного и динамичного его отображения. Также и многие существенно трехмерные, например, некоторые геологические задачи, могут быть успешно смоделированы с помощью 2.5-мерного пространства, достаточно только расширить его до возможности оперировать не с одной, а с несколькими поверхностями.

Легко заметить, что размерность пространства описания связана определенным образом с размерностью объекта, а именно: максимальная размерность объекта равна размерности пространства описания. На линии можно расположить точечные и линейные объекты, на карте - точечные, линейные и площадные, в трехмерном пространстве - точечные, линейные, площадные и объемные.

Упомянутое подразделение объектов на элементарные геометрические типы характерно для картографии. Точечные объекты отображаются на картах немасштабными условными знаками (маркерами), линейные объекты - линейными условными знаками, площадные - площадными.\* Нетрудно, однако, заметить, что такое деление условно и зависит от масштаба рассмотрения. Для геоинформатики этот факт имеет большее значение, чем для традиционной картографии, так как карта в ГИС - объект динамичный, с меняющимся по ходу работы масштабом рассмотрения. К тому же иногда единая пространственная база данных создается по бумажным картам нескольких масштабов или вообще без участия бумажной карты прямо по полевым наблюдениям или в результате дешифрирования данных дистанционного зондирования.

Объекты могут относиться к разным категориям и находиться между собой в сложных соотношениях, например, образовывать иерархические структуры соподчиненности. К примеру, линейные объекты могут быть реками (элемент гидрографии), железными дорогами, автомобильными дорогами, центральными линиями улиц (элементы транспортной сети), линиями газопроводов, линиями нефтепроводов (элементы трубопроводной сети). Автомобильные дороги, например, могут быть далее подразделены на множество различных классов по разным их признакам и сочетаниям признаков (значение, ширина, число полос движения, тип дорожного покрытия, его состояние, время постройки, время и содержание последнего ремонта или инспекции, принадлежность к той или иной организации обслуживания, интенсивность движения - может быть, разная в разное время суток, день недели, сезон - дальше можно продолжать очень долго). И из этих признаков можно построить не единственную систему классификации дорог - в зависимости от того, какие признаки считать более существенными, а какие - относительно менее существенными, им подчиненными. Ясно, что такие смысловые подразделения и группировки объектов в разные категории не могут быть абсолютными на все случаи жизни - выбор той или иной из них зависит от стоящих перед нами задач. Это не исключает, однако, того, что для конкретных массовых применений такие группировки не могут быть заданы как стандартные или рекомендуемые - то есть имеют право на существование и

---

\*Кстати, заметим, что в традиционной картографии, даже на топографических картах, и, тем более, на картах более сложных по содержанию, например, геологических, уделено определенное внимание и площадным объектам с нечеткими границами - заболоченные территории на топокартах, края зон наложенных изменений типа ороговикования на геологических картах. Заметим также, что в современных ГИС недостаточно развиты

необходимы стандартные классификаторы объектов, например, имеющие иерархическую структуру. Однако ясно, что служить основанием для организации пространственных данных в ГИС или пространственной базе данных сколько-нибудь общего назначения такие классификаторы объектов не могут - получающиеся структуры данных оказываются негибкими и не могут обеспечить решение сколько-нибудь широкого и меняющегося спектра задач. Конкретно иерархические структуры к тому же могут иметь проблемы в организации удобных связей с внешними СУБД, которые сегодня в основном строятся на базе реляционной модели (см. в последующих выпусках).

С индивидуальным объектом, если мы хотим трактовать его как индивидуальный, должен быть связан его уникальный идентификатор - например, какой-то номер, формально присваиваемый ему программой в процессе ввода или имеющий содержательный смысл и вне пространственной базы данных, например, его номер по какому-то перечню или кадастру. Это может быть, в принципе, и какое-то уникальное нечисловое имя - например, уникальное название, неповторяющееся более нигде в пределах области изучения. (Или только той ее части, в пределах которой требуется сохранять уникальность идентификаторов. В последнем случае мы имеем как бы двухуровневое деление, и имя (код) этой части может рассматриваться как составная часть идентификатора (префикс идентификатора) объекта.) В общем, идентификатор необходим, в противном случае объект не является уникальным и его нельзя трактовать как в полной мере индивидуальный, самостоятельный. Идентификатор объекта, как в виде номера, так и в виде имени и в реальной жизни, и в компьютерной системе может меняться - он необязательно сопровождает объект на протяжении всей истории его существования. Например, объект в реальном мире может быть переименован - Сталинград - Волгоград. В компьютерной системе изменение идентификатора объекта часто бывает необходимо при реорганизации структуры базы данных - например, при слиянии двух прежде отдельных ее частей. Но изменение идентификатора объекта и в том, и в другом случае - операция совершенно особого рода, требующая особого внимания и контроля и часто выполняющаяся с помощью специальных средств и приемов. У идентификаторов объектов могут быть дополнительные имена-синонимы для удобства использования. Но это случай нетипичный для организации данных в компьютерных системах. Использование

---

средства представления и работы с такими объектами, особенно если говорить и об их участии в анализе, и об их наглядной визуализации. Здесь также видится одно из возможных интересных направлений развития будущей геоинформатики.

не номеров, а имен собственных в качестве идентификаторов объектов хотя и возможно в принципе, но не очень удобно и надежно - в России есть несколько (много!) рек с названием Быстрая, на земном шаре несколько городов с названием Paris, и даже на одном листе топографической карты 1:100000 нередко можно встретить различные географические объекты с одними и теми же названиями. Индивидуальный объект имеет определенное положение, позицию. Информация о его положении, однако, тоже неудобна для использования в качестве его идентификатора. Объект, как природный (река), так и техногенный (мост) может со временем изменить свое положение. Кроме того, позиционная информация, как мы уже видели, являясь метрической, может отличаться по точности (инструмент, метод измерения, тот или иной способ представления численных значений координат в данном формате файла), а также может отличаться из-за использования той или иной системы координат, той или иной картографической проекции, той или иной модели эллипсоида.

Тем не менее, естественно, что информация о положении (а также форме, размерах) объекта крайне важна в ГИС. Обычно такую информацию, как бы она ни была выражена - числовыми координатами, или как-то еще, называют информацией положения (локатором) и отделяют ее от информации идентификации (идентификатора). Вся остальная информация об объекте может рассматриваться как его атрибуты - набор характеристик. Атрибуты можно подразделить на пространственные и непространственные. Пространственные - это, например, периметр и площадь площадного объекта, длина линейного. Непространственные атрибуты могут быть самыми различными - числовыми, текстовыми значениями каких-то величин, описывающих объект. Пространственные атрибуты часто являются функциями параметров положения - например, периметр площадного объекта может рассчитываться из координат слагающих его контур точек.

Можно говорить об объектах элементарных (обычно это точка, линия и полигон) и об объектах неэлементарных (группировках), представляющих объединения (постоянные или временные группировки) элементарных объектов. Если такая группа в свою очередь имеет уникальный идентификатор, то она тоже может рассматриваться как индивидуальный объект. Такая группировка может быть организована на базе как однотипных, так и разнотипных

объектов. В последнем случае назовем такие объекты (группировки) комплексными объектами. Более того, в нее могут входить также и объекты неэлементарные, которые уже, в свою очередь, являются группировками.

### 2.7.3. Об описательных характеристиках индивидуализированных объектов - атрибутах.

Чуть подробнее коснемся атрибутов объектов. Для них важнейшей характеристикой является тип использованной шкалы измерений. Общепринято деление шкал (и, соответственно, данных в этих шкалах) на “качественные” и “количественные”. К “качественным” относят шкалы номинальную (наименований) и ординальную (порядковую, ранговую). К “количественным” относят интервальную (интервалов) и рациональную (отношений) шкалы.

Отметим, что это деление не имеет ничего общего с формой записи или кодированием значений - и данные в номинальной шкале могут быть представлены (и всегда в конечном счете представляются в компьютере) числом. Но это число как бы и не численное значение, это просто код класса, число здесь выступает просто как заменитель названия. В случае номинальной и порядковой шкал для этого числа не имеют смысла некоторые арифметические операции (для порядковой шкалы - имеется только операция упорядочения и сравнения “больше-меньше-равно”, а для номинальной - только сравнение “равно-не равно”).

### Проиллюстрируем различие между разными шкалами.

|                 | Качественные   |   | Количественные  |  |
|-----------------|--|---|---|--|
| <b>Шкалы</b>    | Номинальная  | Порядковая  | Интервальная  | Рациональная                                       |
| <b>Свойства</b> | Наличие/отсутствие<br>Качественное различие<br>Отнесение к одному и тому же или другому классу | Место в определенной последовательности<br>Относительная позиция по сравнению с другими<br>Больше, меньше или равно | Различие<br>Условное начало отсчета (условный нуль)<br>Вычитание, сложение - на сколько | Отношение - во сколько раз<br>Умножение<br>Деление |

| Порождающий процесс | Классификация        | Ранжирование                | Измерения                              |  |
|---------------------|----------------------|-----------------------------|--|--|
| Примеры             | Название города      | Большой город               | Температура в градусах Цельсия         | Число жителей в городе                   |
|                     | Тип землепользования | Самая богатая гумусом почва | (Не Кельвина!)                         | Площадь земельного участка               |
|                     | Название шоссе       | Шоссейная дорога            | Азимут в градусах                      | Расстояние между двумя горными вершинами |
|                     |                      | местного значения           | Прирост числа жителей города за 10 лет |  |

Кроме этого, атрибуты могут быть первичными (измеренными, введенными) и вторичными, расчетными, полученными расчетом из значений других атрибутов. Частный случай таких вторичных рассчитываемых - это атрибуты (обычно пространственные), которые рассчитываются исходя из позиционных параметров объектов (например, периметр).

И, конечно, сегодня, в эпоху мультимедиа, вполне естественным является расширение понятия атрибутов объекта на другую связанную с ним информацию - растровую графику - фотографию объекта или отсканированную схему каких-то документов, с ним связанную, видеofilm или компьютерную анимацию, звуковую информацию. Поэтому можно говорить о классических (алфавитно-цифровых) атрибутах объекта, а также о его расширенных атрибутах (мультимедийных, в частности). В современных ГИС возможно иногда связывание с объектом некоторого действия, например, запуск какой-то программы или макро при активизации объекта указанием на него. Возможно также при активизации объекта (например, площадного объекта "Московская область" на карте России) перейти к другой карте - более подробной карте Московской области с районным делением или к карте Москвы - столице области. Таким образом другую карту, организованную в рамках той же (например, векторно-топологической) или другой (например, растровой) модели данных или какой-то индивидуальный объект в них можно тоже рассматривать в качестве атрибута первого объекта. При этом возможна организация связи между ними различным образом, в том числе с установлением иерархических связей соподчиненности между как индивидуальными объектами, так и их множествами (картами, тематическими слоями). Все это позволяет создавать конкретные структуры баз данных очень



сложные и изощренные. Но это и усложняет теоретическое рассмотрение моделей данных. С одной стороны, теряется четкость понятия “атрибут” - атрибутом одного объекта могут быть другие объекты, в том числе их позиционная информация, их графический образ. С другой стороны, появляются “атрибуты-действия”, и усложняется проведение границы между тем, что относится обычно к категории “данные” (атрибут) и тем, что является компьютерной программой и рассматривается обычно в категории “знания” или “инструменты работы с данными”. Все это вызывает необходимость введения понятия “обобщенные атрибуты” и ярко демонстрирует условность и относительность деления информации об объекте на атрибутивную и неатрибутивную (графическую, например). Этот пример также хорошо демонстрирует сложность взаимоотношений между способами организации данных и способами организации и воплощения наших знаний.

#### **2.7.4. Формы представления непрерывных признаков - полей.**

Некоторое свойство, непрерывно распределенное в пространстве, например, на земной поверхности, удобно в математическом смысле рассматривать в качестве поля. (Вообще говоря, разного вида - скалярного, векторного, тензорного, трехмерного или двумерного, определенного только на поверхности земли или с ней не связанного, а только проектируемого на нее или пересекаемого ею, но всех сложностей мы сейчас касаться не будем). Типичные формы представления подобных непрерывных свойств таковы:

Нерегулярная сеть точек - произвольно расположенные точечные объекты, в качестве атрибутов имеющие значение поля в данной точке. С помощью такого способа представления, если не иметь очень густо расположенных по сравнению с пространственной изменчивостью поля точек, трудно гарантировать его адекватное представление. Сеть может быть слишком редкая, или точки, выбранные случайно, не попадают на характерные представительные места/значения, или, наоборот, точки выбраны неслучайно и тоже не являются представительными (например, все значения кислотности почвы измерялись только там, где можно было легко передвигаться по местности, то есть на голых водоразделах, а не в заболоченных речных долинах).

Способ представления изолиниями - наиболее распространен в традиционной картографии. Обычно нет никакой информации о поведении поля между изолиниями, и вторая проблема состоит в том, что по одним и тем же исходным данным (обычно это точки нерегулярной сети) интерполяция и последующее проведение изолиний может быть сделана не единственным способом. Способ представления очень привычный, но не самый удобный для анализа.

Наиболее удобен для многих случаев вариант, когда поля задаются регулярно расположенными в пространстве точками достаточной густоты (регулярная модель), особенно когда это точки не интерполированные из нерегулярных, а измерения, проведенные по регулярной сети. Из них легко перейти к любой другой форме представления.

Несколько особняком стоит модель данных TIN (Triangulated Irregular Network), специально предназначенная для представления поверхностей значений, полей (например и в первую очередь - поверхности рельефа местности). Это также нерегулярная сеть точек, но точек связанных, соединенных особым образом выбранной сетью прямых отрезков, при этом эти отрезки (называемые часто ребрами - edges) образуют множество треугольников, как бы грани (фасеты). Наличие таких связей между точками дает некоторое представление (линейное приближение) о поведении поля (или форме поверхности) на данном участке в промежутке между точками. Поэтому модель данных типа TIN часто позволяет получать более качественное и более экономное представление поверхностей (полей). В особенности удачным оказывается применение этой модели данных для ситуации, когда изменчивость поля (например, рельефа местности) очень разная в разных частях территории, например, если район изучения захватывает и горный район, и плоскую предгорную равнину. К сожалению, многие типы аналитических задач трудно выполнимы на этом типе модели данных.

### **2.7.5. О некоторых требованиях к представлению пространственных данных.**

Отметим, что карты (это касается как бумажных карт, так и их компьютерного представления в ГИС) могут быть геометрически

точные (рассчитаны на то, чтобы правильно с точностью до искажений картографической проекции передавать форму, размеры и абсолютное положение объектов; по таким картам можно производить те или иные геометрические измерения) и могут быть только топологически точны (например, схема железных дорог, на которой ни формы линий не соблюдаются, ни масштаб не одинаков в разных местах. Но если города А и В соединены железной дорогой, и она проходит именно через город С, то этот факт точно отражен в графике). Такие “карты”, по которым принципиально нельзя мерить расстояния и координаты объектов, но которые сохраняют все топологические свойства нужного множества объектов, например, дорожной сети (что с чем связано, или что с чем граничит, например) - чаще называются не картами, а схемами, а их компьютерное представление часто может рассматриваться как граф. Конкретная карта может быть одновременно и геометрически, и топологически точна, но могут быть и карты, точные геометрически, но не топологически, или наоборот.

Карту, равно как и упомянутый граф дорожной сети, можно рассматривать как разновидности моделей пространственных данных, может быть, компьютерные, может быть, нет.

### **2.7.6. Типы моделей пространственных данных - первый обзор.**

Итак, какие же бывают модели пространственных данных. Как часто случается в жизни, их нельзя расклассифицировать по одной оси, они различаются как бы “в разные стороны”.

Несколько схематизируя ситуацию, заметим, что в ГИС о модели данных можно говорить в нескольких смыслах. Во-первых, сами пространственные данные, а это в значительной части графическая, позиционная компонента, могут быть организованы различно по своему внутреннему устройству, в соответствии с разными моделями. Во-вторых, можно говорить и о разных моделях организации атрибутивных данных внутри себя, безотносительно к графической компоненте. Наконец, и это уже момент, специфичный именно для ГИС, можно говорить о разных моделях отношений между пространственной и атрибутивной информацией, или, упрощая, между графической и описательной. Такое раздельное рассмотрение

названных аспектов полезно не всегда, иногда модель данных весьма тесно интегрирует пространственную и атрибутивную информацию.

### **2.7.7. Модели построения взаимоотношений между атрибутивной и пространственной графической информацией.**

В целом существует три подхода к организации совместной работы с пространственной и атрибутивной информацией, три модели такого взаимодействия. Иногда этот аспект именуют принципами взаимодействия ГИС с базой данных. Во-первых, наиболее известная и распространенная на сегодня модель этих взаимоотношений - геореляционная, называемая также иногда гибридной или смешанной. Однако последние два термина лучше оставить за гибридными растрово-векторными моделями, и мы будем использовать термин “геореляционная”. В этом случае пространственная компонента организована по-своему, а атрибутивная - по своему, между ними просто устанавливаются и поддерживаются связи через идентификатор объекта. Пространственная информация, метрическая, а в некоторых системах также и топологическая, хранится совершенно отдельно от атрибутивных в своих файлах или системах файлов. Атрибутивная информация организована в таблицы, которые управляются с помощью реляционной СУБД. Эта СУБД может быть как встроенной в ПО ГИС как его функциональная подсистема или может быть внешней по отношению к ГИС. Иногда, как в ARC/INFO или ArcView, реализуются одновременно оба подхода - есть и простая встроенная подсистема, и возможно использование внешних СУБД для хранения атрибутивной информации. Этот подход исходит из того, что трудно добиться одновременной оптимизации хранения и графических, и атрибутивных данных. Кроме того, во многих применениях ГИС сегодня существует тенденция как бы достраивать ГИС к существующим под управлением СУБД большим базам данных, добавления в них пространственных компонент и соответствующих средств работы с ним.

Второй вариант - интегрированный. Здесь предусматривается использование средств реляционных СУБД для хранения как графической, так и атрибутивной компоненты. В этом случае ГИС выступает как бы в качестве некоторой надстройки над СУБД. Этот

подход обладает рядом преимуществ, особенно для крупных хранилищ информации, работающих в режиме активного многопользовательского использования, когда существенной проблемой становится обеспечение целостности данных. Однако современные традиционные реляционные СУБД мало подходят для работы с пространственными объектами общего вида (отличными от точечных). Примеры такого подхода в коммерческом ПО известны - например, GFIS фирмы IBM, но широкого распространения они не получили.

Наконец третий подход - объектный. Обладает очень многими привлекательными сторонами, в особенности в части возможности описания в нем очень сложных структур данных, взаимоотношений между объектами и возможностями решать многочисленные задачи моделирования. Однако в чистом виде для общего вида пространственных задач применим с трудом, гораздо более интересен появляющийся сегодня союз реляционных СУБД и объектного подхода в виде объектно-реляционной модели данных.

### **2.7.8. Модели организации пространственных объектов, организации связей между ними.**

Два общих принципа, два подхода к организации пространственных данных нами уже упоминались ранее (см. стр. 28). Это различные принципы группировки объектов в логически связанные структуры более высокого порядка. (Они могут в той или иной мере сочетаться с разными только что описанными моделями взаимоотношений пространственных объектов и их атрибутов.) Один - это принцип послойной организации информации (его часто называют классическим), второй - опять же основан на объектно-ориентированном подходе. Надо сразу отметить, что объектно-ориентированный подход здесь понимается не обязательно в том смысле, что в объектно-ориентированном программировании (и только что упомянут выше), и что между двумя упомянутыми принципами организации информации "антагонистического противоречия", конечно, нет. Послойный принцип организации информации очень нагляден и хорошо соотносится с приемами традиционной картографии. Он заключается в том, что находится некоторое деление объектов на тематические слои; и объекты, отнесенные к одному слою, образуют некоторую логически (а часто

и физически) отдельную единицу данных - например, они собираются в один файл или в одну директорию, они имеют единую и отдельную от других слоев систему идентификаторов, к ним можно обращаться как к некоторому множеству. Например, мы выносим в один слой все объекты гидрографии, или все шоссейные дороги, или все, относящееся к растительному покрову. Чаще всего при этом организуется также и деление одного тематического слоя по горизонтали - по аналогии с отдельными листами карт. Это делается в основном из удобства администрирования баз данных и чтобы избежать работы с чрезмерно большими файлами. Для случая векторно-топологических моделей данных (то есть тех, которые позволяют строить и хранить в самой модели данных топологические отношения между объектами) обычно существуют некоторые дополнительные ограничения - в один лист одного тематического слоя можно поместить объекты не всех геометрических типов одновременно. Векторные нетопологические модели данных в этом смысле предоставляют больше свободы, но все равно часто и в них в один слой помещаются только объекты одного геометрического типа. Число слоев при послойной организации данных может быть ограничено, может быть практически не ограничено в зависимости от конкретной реализации. При послойной организации данных очень удобно манипулировать большими группами объектов, представленных слоями, как единым целым, например, включая или выключая слои для визуализации, определять операции, основанные на взаимодействии слоев. В целом можно сказать, что послойная организация данных имеет большой аналитический потенциал. Она часто используется при организации как векторно-топологических, так и векторно-нетопологических моделей пространственных данных для ГИС. (Так же как нередко и в САПР. Там, впрочем, отдельные слои могут обычно включать все типы объектов и само это деление на слои носит более простой и формальный характер - они часто используются только для управления визуализацией и как удобный способ задания умолчаний для чертежа). И абсолютно преобладает послойный принцип организации в растровых моделях данных для представления непрерывно распределенных признаков.

Объектно-ориентированный принцип организации данных в ГИС акцентирует внимание не столько на общих свойствах объектов (моделируемых через деление на слои в предыдущем подходе),

сколько на их положении в какой-либо сложной иерархической схеме классификации, на взаимоотношениях между объектами. В силу этого удобно отображаются различные родственные и генетические отношения между объектами, отношения соподчиненности, функциональные связи между объектами. В целом этот подход в чистом виде менее распространен, чем послойный, вероятно, отчасти из-за больших трудностей практической организации всей системы взаимосвязей объектов. Трудозатраты на подготовительный этап формирования конкретной структуры базы данных могут быть очень велики. Могут быть и определенные сложности с изменением сформированной структуры базы данных и приспособлении ее к новым задачам. Для информации о природных объектах вообще применяется реже, чем послойный подход. Трудно также представить себе растровую ГИС, эффективно реализующую объектный подход.

Сегодня в моделях данных в ГИС преобладает послойный принцип организации информации. Два этих подхода, (если отвести крайности) вообще не противоречат друг другу. Существуют возможности использования обеих подходов совместно (имеются такие примеры). В будущем можно ожидать более тесного их взаимодействия, эти тенденции уже отчетливо видны. Однако сегодня могут существовать проблемы обмена информацией между системами, использующими два этих разных подхода. Иногда высказывается мнение, что объектно-ориентированный подход более эффективен при работе с ГИС только по техногенным объектам, а послойный - при работе с природными объектами и для случая, когда имеешь дело и с теми и с другими. Я не уверен, что это положение носит общий характер, скорее оно касается (если вообще справедливо) только каких-то конкретных систем.

### **2.7.9. Модели представления графической информации.**

Не менее важным является деление моделей данных на векторные, оперирующие непосредственно с координатами объектов и слагающих его точек и модели данных с делением пространства, где положение объекта или слагающего его элемента задается принадлежностью к некоторому дискрету, элементу делимости пространства. Наиболее распространенным видом моделей с

делением пространства являются растровые модели данных, то есть использующие регулярную сетку точек или, что эквивалентно, разбиение пространства на регулярные ячейки прямоугольной (чаще квадратной) формы. Однако хочется подчеркнуть, что хотя их практическое значение и распространенность гораздо больше, чем у других моделей данных с делением пространства, вся эта группа очень разнообразна и представляет значительный теоретический интерес. Отдельные на первый взгляд экзотические модели данных из этой группы находят довольно неожиданное применение в комбинации с векторными и растровыми данными.

### **2.7.10 Подразделение моделей и форматов данных по их назначению.**

Разумеется, можно также проводить деление моделей и форматов данных по их назначению. Для использования в ГИС, для использования в САПР, для использования в простых чертежно-рисовальных системах. ГИСовские модели и форматы обычно четко отделяются на сегодняшний день от всех прочих за счет необходимости организации и активного использования связи между информацией положения и атрибутивной информацией объекта. Серьезные ГИС сегодня также должны обладать возможностями работать с непрерывно распределенными данными, что часто заставляет в одной системе использовать комплексные (часто говорят, гибридные) модели данных (то есть фактически несколько разных моделей данных). Все это делает ГИСовские модели данных и форматы в среднем более сложными, чем все остальные. Можно думать, что и передний край теоретического осмысления общих вопросов о моделях и форматах данных также сегодня проходит в области геоинформатики.

Если вернуться только к области ГИС, то все равно сохраняется подразделение моделей и форматов данных по типу использования. Я говорю о моделях и форматах внутренних (реально используемых системой при ее основном функционировании) и моделях и форматах обменных, предназначенных для обмена с другими (ГИСовскими и неГИСовским пакетами).



### 2.7.11. Некоторые дополнительные моменты.

Здесь можно сделать несколько замечаний. Во-первых, относительно деления позиционных параметров на параметры положения и параметры, определяющие форму объекта. Только для точечных объектов это всегда полные синонимы - точка формы не имеет, или ее положение - это и есть ее форма. Для линейных же и площадных объектов ситуация иная. В дальнейшем мы еще будем подробнее рассматривать способы описания линейных и площадных объектов в различных моделях данных, а сейчас только заметим, что они могут описываться как координатами каких-то их точек, так и каким-то описанием их формы, например, заданием функций, график которой дает форму требуемого объекта или просто словесным описанием формы - "квадрат со стороной 2 метра". Нетрудно заметить, что в последнем случае (а иногда и в предпоследнем) информацию положения можно отделить от информации, описывающей форму, поместив ее просто в атрибуты объекта. Для однозначности надо только определить, в каком месте объекта расположена та его точка, положение которой задается локатором. Назовем эту точку привязочной точкой объекта, определив ее положение в собственной системе координат объекта, а точку в пространстве, которая задается для такого объекта локатором - точкой вставки объекта. Это подход традиционно используемый в отношении внемасштабных условных знаков для точечных объектов. Но он может быть обобщен, если трактовать графическое представление условного знака как пространственный объект, и вообще использоваться в других ситуациях, если нам удастся описать форму, размер и ориентировку объекта через его атрибуты, например. Правда, многие задачи, требующие взаимодействия пространственных объектов, анализа их взаимоотношений, могут затруднены при таком представлении. Эти рассуждения показывают нам условность подразделения информации об объекте на атрибутивную и информацию о его форме, а также возможность иногда рассматривать отдельно информацию о положении, форме, размере и ориентировке (угле поворота) объекта.

## 2.7.12. Краткое резюме и практические следствия.

Мы познакомились с некоторыми общими принципами построения моделей данных в пространственных информационных системах, затронули кратко вопросы о разновидностях моделей данных. Были приведены некоторые отдельные соображения, показывающие сложность вопроса - как только мы выходим за рамки единственной конкретной программной системы с ее моделью данных и ее ограничениями, многое, даже определения и классификация элементов, слагающих модель данных, становятся непростой задачей, многое оказывается дискуссионным.

Мы затронули также принципиальную важность вопроса о выборе модели данных - часто это вопрос номер 1, когда мы задумываемся о работе с пространственными данными и только начинаем планировать ее. Это относится как к разработчикам оригинальных программных систем - новых инструментальных средств для геоинформатики, так и к разработчикам приложений и к пользователям, которые стоят перед проблемой выбора программного продукта для реализации проекта.

Геоинформационные технологии принесли компьютерные методы в практику многих и многих областей знания, где до этого использование компьютеров было ограниченным и вспомогательным, усадили за компьютер много специалистов, которые до знакомства с ГИС практически не имели дела с какими-либо информационными технологиями. Это порождает определенные проблемы как в обучении геоинформатике, так и в практическом освоении технологий. Может быть, сейчас стоит назвать некоторые из этих проблем, поскольку они в значительной мере связаны с предыдущим текстом.

1. Многие пользователи не знают, не могут четко сформулировать, чего они хотят. Процесс постановки задачи и выбора модели данных, программного средства и конкретной технологии связаны поэтому очень тесно. На этих этапах очень желательно тесное взаимодействие специалиста-прикладника, способного формулировать задачу на содержательном уровне и специалиста по общей геоинформатике, выступающего в роли системного аналитика. Весьма вероятно, что этот процесс должен

быть итеративным - идти методом последовательных приближений.

2. Пользователи могут думать, что применение какого-то программного обеспечения само по себе решит их проблемы. Однако никакие самые совершенные средства обработки и анализа не состояниии решить никаких проблем без достаточного обеспечения качественными данными и продуманной постановки задачи.

3. Некоторые операции, возможные теоретически, на практике оказываются слишком дорогими за счет необходимости сбора дополнительной информации, за счет того, что алгоритмы их реализации требуют слишком больших вычислений даже для современных компьютерных систем. Уровень реализации некоторых методов анализа в современных программных средствах может быть недостаточен для их применения в определенных ситуациях (ограничения на число объектов...).

**Ю.К. Королев**

Общая геоинформатика

Часть I. Теоретическая геоинформатика

Выпуск 1

СП ООО Дата+, 1998 г., 118 стр.

Книга предназначена для широкого круга специалистов, начинающих и продолжающих знакомиться с геоинформатикой в практическом плане, осваивая те или иные программные средства ГИС и обработки данных дистанционного зондирования.

Планируется, что книга будет состоять из четырех частей: теоретическая геоинформатика, практические аспекты геоинформатики, дистанционное зондирование, справочные приложения. Издание будет осуществляться по отдельным выпускам.

Первый выпуск посвящен общим принципам организации пространственной информации.

Издание рекомендуется широкому кругу специалистов в области информационных технологий, кадастра, окружающей среды, а также других областей, применяющих или собирающимся применять геоинформационные технологии.

Издание будет также полезно студентам, преподавателям ВУЗов.



## КОРОЛЕВ ЮРИЙ КОНСТАНТИНОВИЧ

Родился в 1950 г. в Москве. По специальности - геолог-съемщик.  
В 1973-1991 г.г. работал в объединении "Аэрогеология", затем  
ТОО "Геоцентр". В настоящее время - главный эксперт СП Дата+