

УДК 681.518(075.32)

О РАЗВИТИИ МЕТОДОЛОГИЧЕСКИХ ОСНОВ И КОНЦЕПЦИЙ ГЕОИНФОРМАТИКИ

С. К. Дулин, И. Н. Розенберг

В работе рассматривается история развития географической информатики как объединённой области, которая охватывает теорию географических информационных систем и методы анализа пространственных данных. Особое внимание уделено результатам математической формализации сформулированных гипотез, получивших подтверждение эмпирическим путём на соответствующих моделях. Описан подход к отображению семантики, заложенной в пространственной онтологии, в географические концептуальные схемы, представляющие информацию, хранимую в базах геоданных. Анализируются проблемы топологической группировки геоданных. Предлагается один из подходов систематизации и классификации геоданных, основанный на аппарате структурной согласованности.

1. Этапы развития геоинформатики

Географическая информационная система (ГИС) и анализ пространственных данных сформировались как две более или менее отдельные области научного исследования, однако последние десятилетия они демонстрируют заметное сближение друг к другу, так что сегодня уже можно утверждать, что эти две области включаются в географическую информатику, поддерживая и дополняя друг друга. Развитие географической информатики — GIScience [1, 2] было предопределено разработками в области ГИС и области анализа пространственных данных. Исследования в ГИС продвинули техническую способность обрабатывать пространственно распределённые данные, стимулируя реализацию отражения отношения между тем, что называется «географической действительностью», и формированием концептуально-го представления этой действительности в конечных цифровых формах, то есть в виде исчисляемого количества точек, линий, и областей в двумерном пространстве.

В конце 50-х и начале 60-х годов прошлого века, когда трудно было представить себе, что сложное графическое содержание карт подвластно числовой обработке компьютерами, вычислительные технологии сделали возможным возникновение автоматизированной картографии, которая фактически и привела к созданию ГИС. Создание сканеров и картографов в 60-х годах, наряду с быстрым усовершенствованием программного обеспечения, открыло поразительные возможности, даже на той, очень ранней стадии развития компьютеров [3, 4].

Тем не менее, только в 80-х годах произошло осознание появления новой парадигмы под названием ГИС и началось её бурное развитие. Возможно, самыми важными мотивациями развития ГИС были:

- практическая трудоёмкость получения точных размеров из карт и, соответственно, простота получения таких размеров из цифрового представления;
- потребность объединять данные (в том числе и на разных уровнях), характеризующие многие типы особенностей геоданных и отношений между ними в больших проектах;
- практическая проблема редактирования карт в картографическом промышленном процессе, которая стимулировала развитие автоматизированных редактирующих систем;
- проблемы управления большим количеством уточнений региональных данных при составлении проектов развития промышленных и социальных объектов.

Самая первая географическая информационная система — Канадская Географическая Информационная Система [5] — была разработана, чтобы автоматизировать обработку информации, собранной в форме карты Канадской кадастровой описи. Достаточно похожим образом американское бюро переписи населения создало простую ГИС для обслуживания подготовительного периода переписи 1970-го года. Следует заметить, что пространственный анализ в перспективе совершенствования методов обработки географической информации явно доминировал в ГИС с начала её возникновения.

Уже началу 70-х годов стали очевидными коммерческие выгоды создания интегрированного программного обеспечения для обработки географической информации. Первая коммерчески жизнеспособная ГИС появилась в начале 80-х. Этому способствовало появление персонального компьютера, позволившего приобрести достаточную вычислительную мощность в пределах

бюджета ведомства или фирмы. Кроме того, разработанное программное обеспечение управления реляционной базой данных устранило потребность самостоятельно создавать сложные программы обработки данных. В течение 70-х и 80-х годов ГИС конструировались и продавались как подчинённые процессоры или компьютерные приложения, которые могли выполнять операции на географических данных намного более рентабельно, чем люди. Собственно эксплуатацией ГИС как интегрированной программной среды для пространственного анализа и была заложена основа в форме подпрограмм для создания и редактирования электронных карт, и только впоследствии было добавлено большое количество специальных утилит.

Анализ пространственных данных является самым сложным из приложений ГИС, но он ни в коем случае не наиболее значителен с коммерческой точки зрения [6]. Наибольшая коммерческая ценность ГИС, как правило, заключается в её способности проследить местоположение и состояние необходимых объектов, вместо того чтобы проводить сложный анализ и моделирование.

По определению, анализ пространственных данных представляют собой коллекцию методик и моделей, которые явно используют пространственное соотношение элементов данных и взаимодействия между ними. К истории развития анализа пространственных данных имеет отношение большее число научных дисциплин, в основном, абстрактного математического характера: геометрия, топология, линейная алгебра, функциональный анализ, математическая статистика и др. В рамках анализа пространственных данных проводятся: исследование распределения данных, поиск глобальных и локальных выбросов, поиск глобальных трендов, проверки пространственной автокорреляции и выявление ковариаций между несколькими наборами данных. Но наиболее распространёнными являются исследования по выявлению подходящего метода интерполяции применительно к конкретным данным.

Методы интерполяции позволяют вычислить поверхность по точечным измерениям и получить (предсказать) значения переменной для каждой точки ландшафта. Все методы при построении поверхности основываются на предположении о подобии близлежащих точек опробования (элементов выборки). Методы интерполяции объединены в две группы: детерминированные и геостатистические. Методы детерминированной интерполяции

строят поверхности по точечным измерениям с учётом либо протяжённости сходства (например, метод обратного взвешенного расстояния), либо степени сглаженности (например, методы радиальных функций).

Геостатистические методы базируются и на детерминированных математических, и на статистических функциях, которые могут быть использованы для построения поверхностей и оценки точности (неопределённости) прогнозов. Методы геостатистической интерполяции (такие, как кригинг) учитывают статистические свойства точек измерений. Геостатистические методы количественно определяют пространственные автокорреляции между точками измерений и учитывают пространственное расположение точек опробования. Изменяющаяся степень корреляции между точечными данными играет важную роль в решении, какой вариант интерполяции лучше использовать. Например, если даным присущи локальные вариации, то целесообразно применить метод интерполяции с помощью детерминированных полиномиальных зависимостей, чтобы исключить крупномасштабные вариации из данных.

Уже в 50-е годы появились авангардные работы в области математики и статистики, которые оказали фундаментальными для развития анализа пространственных данных. Оригинальная работа [7] расширила авторегрессионные модели, наиболее значимые для анализа изменения временных рядов пространственных данных [8]. Это были практически первые пространственные авторегрессионные модели, пригодные для формального представления пространственных изменений, что позволило для некоторых типов данных выйти за пределы простого испытания на пространственную автокорреляцию.

Автокорреляция — это статистическая корреляция между случайно распределёнными в пространстве переменными одного и того же типа, атрибута, имени и т. д., когда корреляция зависит от расстояния и/или направления между местоположениями. С появлением работы [7] стало возможным определять формальное представление определённых типов пространственной структуры, тестировать значимость модели и качество доступа к данным.

Чуть позже, в 60-х годах, в интересах горнодобывающей промышленности был разработан метод «кригинг» — по имени южноафриканского геолога D. G. Krige [9]. Эта работа расширила теорию предсказания вероятностного процесса Вине-

ра-Колмогорова, распространив её на случай пространственных процессов, определённых на непрерывном географическом пространстве. Кригинг — это геостатистический метод интерполяции, использующий статистические параметры для более точного построения поверхностей. Он позволяет исследовать пространственные автокорреляции между данными, выполняя две группы задач: количественное определение пространственной структуры данных и создание прогноза.

Количественное представление (квантификация) пространственной структуры данных, известное как вариография (создание вариограмм), даёт возможность пользователям подобрать к данным модель пространственной зависимости. Для расчёта (прогноза) неизвестного значения переменной в заданном месте кригинг использует подходящую (подобранную) модель из вариографии, конфигурацию пространственных данных и значения в точках измерений вокруг данного местоположения. Поскольку в кригинге используются статистические модели, он обеспечивает построение разнообразных выходных карт поверхностей, включая прогнозные, а также прогноз стандартных ошибок, вероятности (правдоподобия), и квантильные.

Ещё одним подходом к исследованиям, который должен был стать одним из основных краеугольных камней пространственной статистики, была теория точечных процессов. К началу 50-х годов было разработано много статистических методов на основе квадратичного отклонения расстояния для проверки на пространственную рандомизацию в опорной точке. Работа [10] расширила квадратичную статистику, позволяя исследовать несколько масштабов карты одновременно.

Хотя термин *анализ пространственных данных* был уже введён в географическую литературу [11], длительное время не было признано существование области пространственной статистики как таковой. Можно упомянуть лишь обзорную монографию [12]. Были заложены теоретические основы анализа временных рядов [13, 14, 8], и анализа экономических временных рядов [15, 16]. Работа [17] обеспечила основание для введения нового класса моделей пространственных данных. Разработанные модели удовлетворяли двумерной версии свойства Маркова, представляющего временное изменение. В результате появления статьи [17] был определён общий класс пространственных моделей для дискретных и непрерывно оцениваемых переменных, определённых в двумерном пространстве.

В начале 80-х годов появилась монография [18], которая описывала усовершенствования, произошедшие в геостатистике. Эти результаты были включены в главу по пространственному сглаживанию и интерполяции. Другой существенной монографией в это время была книга по анализу образа точки [19]. В 1991 году публикация работы [20] обеспечила обзор методов и моделей анализа геостатистических данных.

ГИС и анализ пространственных данных непосредственно взаимодействуют на пространственной матрице данных. На концептуальном уровне, эта матрица состоит из строк и столбцов, где строки соответствуют экземплярам (элементам данных), а столбцы — признакам или атрибутам этих экземпляров. В самом простом случае используются два столбца, содержащие пару координат: широту и долготу, или x и y в некоторой специальной системе координат.

На практическом уровне, пространственная матрица данных — архив данных, собранных исследователем. В практических терминах структура и содержание этой матрицы — конечный продукт процессов концептуализации и представления, с помощью которых зафиксирована некоторая доля географической действительности. В некотором смысле — это выходные данные процесса цифровой фиксации мира. В другом смысле, матрица — опорная точка или ввод для анализа пространственных данных.

Необработанные данные чаще всего состоят из первоначальных измерений, сделанных с помощью инструментов. Термины *погрешность* и *точность* относятся к соответствию между результатами измерений и действительностью и уточнением повторных измерений, соответственно. Для многих инструментов эти параметры известны и могут использоваться для того, чтобы проанализировать воздействие ошибок на последующие исследования. Географическая информация, которая представлена на картах и в базах данных, редко составляется из первоначальных измерений. При этом основная часть географической информации, такая, как результаты компиляции, интерпретации, анализа, и вычисления, остаётся скрытой от пользователя.

На ранней стадии развития ГИС было принято измерять местоположения точек на поверхности Земли в абсолютных значениях относительно геодезической сетки земли, т. е. относительно Экватора и Гринвичского меридиана. С течением времени стало ясно, что точно знать абсолютное местоположение невозможно,

а относительные местоположения могут быть измерены намного более точно, чем абсолютные. Следовательно, более целесообразно проектировать базы геоданных (БГД) ГИС для хранения относительных местоположений и получать абсолютные местоположения, выводя их, когда это необходимо [21].

Ещё одной проблемой развития ГИС стала проблема неопределённости. Отображение реального мира не позволяет создать его совершенное представление. Независимо от используемого метода интерполяции, полученные данные всегда будут содержать ошибку. На её возникновение могут влиять разные факторы, такие, как неточность исходных данных, изменение физических свойств данных измерений, ошибки за счёт человеческого фактора, изменения, происходившие в процессе проведения измерений. Все эти факторы вносят свой вклад в величины ошибок измерений. Выбор модели ошибок измерений влияет на плавность результирующей карты поверхности и на величину стандартных ошибок. Моделирование ошибок измерений применяется в трёх методах кригинга — простом, обычном и универсальном. Ранние разработки абстрагировались от проблемы неопределённости, являющейся ахиллесовой пятой для ГИС [22]. В настоящее время этой проблеме уже уделяется определённое внимание, хотя очень небольшая часть коммерческих продуктов включает реализацию методов уменьшения неопределённости [23].

С самого начала развития ГИС в качестве основы для хранения геоданных выступали системы управления реляционными базами данных, которые позволили отделить специфическое программное обеспечение географической информационной системы от управления базой данных. Однако характерные для геоданных особенности должны были быть сохранены в отдельной, специальной базе данных, которая была неспособна использовать стандартное программное обеспечение управления базой данных, так как формы этих особенностей не могли быть определены в простой табличной структуре.

Только в конце 90-х годов развитие объектно-ориентированных систем управления базами данных позволило отчасти преодолеть эту проблему (хотя и оставив приоритет популярности реляционному подходу) и начинать доминировать в моделировании данных ГИС.

Первый принцип объектно-ориентированного подхода состоит в том, что каждая особенность на поверхности Земли является

примером класса, а второй — в том, что классы могут быть специализациями более общих классов. Специализированные классы наследуют все свойства более общих классов, и добавляют специальные собственные свойства. ГИС допускает обширный массив операций, основанных на этом подходе к представлению. Но и объектно-ориентированное моделирование также было не в состоянии адекватно представлять непрерывные формы геоданных. В то время, как понятие дискретных объектов является соответствующим для людей, транспортных средств, строений и промышленных объектов, оно несовместимо со многими явлениями в географическом мире, которые являются принципиально непрерывными: реки, дороги, или ландшафт.

Индустрия ГИС недавно приняла компонентно-ориентированные подходы к программному обеспечению, разрушая ранее принятые традиции монолитных пакетов соединением частей компонентов многократного использования. Это может привести к значительным успехам в интеграции ГИС с другими формами программного обеспечения, которые используют те же самые стандарты, особенно пакеты для статистического анализа за [24].

Сегодня ГИС может быть определена как *вычислительное приложение*, способное к созданию, хранению, управлению, визуализации и анализу географической информации. Это определяет её самые сильные прикладные возможности в управлении ресурсами, управлении полезными ископаемыми, передаче данных, городском и региональном планировании, маршрутизации транспортных средств и доставке посылок, и во всех науках, которые имеют дело с поверхностью Земли.

2. Характерные особенности и общие свойства географической информации и пространственно-распределённых данных

Термин «географический» относится к особенностям и явлениям, происходящим на или около поверхности Земли. Следовательно, географическая информация — это информация о таких особенностях и явлениях. Этот термин также подразумевает определённый диапазон пространственного разрешения, который исключает любые квантовые или релятивистские эффекты и, таким образом, находится жёстко в рамках теории Ньютона.

Географическая информация могла бы быть определена более формально как состоящая из элементарных пар формы $\langle x, z \rangle$, где x — местоположение в пространстве-времени, и z — ряд свойств этого местоположения [25]; или же как информация, которая является приводимой к таким парам.

Термин «географический» является специализацией термина «пространственный», который с помощью расширения соотносится с любой пространственно-временной структурой и любым пространственным разрешением. Пространственная структура может содержать географическую структуру, как в случае вселенной, но географическая структура может также содержать перемещаемые в её пределах пространственные структуры. По этой причине термин «геопрограммный» существенно избыточен и идентичен «географическому».

Термин «географический» наследует многие свойства «пространственного», добавляя новые свойства и, таким образом, специализируя определение «пространственного». Если считать, что «пространственный является особенным», как в [26, 6], то «географический» должен быть ещё более особенным, и теория географической информации должна быть отличной от теории пространственной информации, наследуя всю общность последней, но добавляя свои собственные специфические особенности. Специфическая природа географической информации накладывает ограничения, сужая множество вариантов, рассматриваемых в общем пространственном случае.

Если вполне допустимо представлять пространственную информацию с помощью абстрактных методов, предполагая возможность установления равенства положения точек, то для географической информации практически невозможно определить это равенство. Невозможно определить, находится ли точка точно на линии, или совпадают ли две линии точно, основываясь только на знании местоположения. Таким образом, существующие утилиты «точка-в-многоугольнике» должны определять только одну альтернативу — внутри или вне, а подпрограммы наложения многоугольников могут вывести равенство положения, используя только определённую пользователем позиционную толерантность, но никак не точное сравнение.

Вообще, неблагоприятно вычислить топологию из геометрии, а лучше позволить независимо определённой топологии отвергать результаты геометрии, когда эти два подхода находятся в противоречии. Теория географической информации часто мо-

жет позволить себе не принимать во внимание возможность определения точного равенства положения объектов, потому, что это предполагает нереалистичный уровень задания точности позиционирования объектов.

Можно утверждать, что переменная обладает пространственной зависимостью, если существуют корреляции между её значениями в различных точках. Степень подобия между значениями в двух точках, как правило, увеличивается по мере приближения их друг другу, подразумеваемая степень непрерывности и гладкости. Географическая информация обладает этим типом пространственной зависимости, которой можно придать статус закона [27] и сформулировать следующим образом: «Все элементы поверхности подобны, но соседние элементы более подобны, чем отдалённые.»

Свойство, выраженное в законе, подтверждается статистикой пространственной автокорреляции. Очевидно, что этот закон имеет отношение только к географическим пространствам, и не применим ко всем пространствам в общем случае.

Существует весьма важный вопрос географической науки: должна ли цель исследования состоять в том, чтобы обнаруживать общие истины, или же в том, чтобы документировать определённые факты? Эти два альтернативных подхода называются *дедуктивным* и *индуктивным*, соответственно. В географическом контексте они выражены как различные стратегии изучения поверхности Земли.

В *дедуктивной стратегии* исследование успешно, если оно раскрывает принципы, которые являются верными всюду, в то время, как *индуктивные стратегии* поддерживают детальное исследование уникальных характеристик мест, которые возможно, приведут к обобщениям. Ясно, что дедуктивная стратегия требует некоторой степени однородности области, скорее даже не по форме, но, вероятно, в процессах, которые изменяют и формируют эту область. Поиск таких процессов доминирует при дедуктивном подходе. С другой стороны, индуктивный подход не требует никакой однородности на любом уровне.

Поверхность Земли обладает поразительной разнообразностью, и часто необходимо сканировать большую долю поверхности, чтобы столкнуться со всем её разнообразием, так как маленькая область поверхности типично представляет только маленькую долю полного изменения любой переменной. Из этого следует, что результаты анализа почти всегда явно зависят от географиче-

ческих границ региона исследования и что изменение границ приведёт к другим результатам. Маленькая область не обеспечивает репрезентативную пробу поверхности Земли. Так же, как и пространственная зависимость, пространственная разнородность является определяющей характеристикой определения географического пространства.

Гетерогенная природа географического мира неизбежно приводит к предпочтительному *исследованию формы*. Геоморфологи, например, долго утверждали, что исследование процесса имеет большее значение и ценность, чем исследование формы, и понимание того, как мир функционирует, более важно, чем понимание того, как это *выглядит*.

Интерес к процессам естественен, имея в виду человеческие процессы, которые изменяют пейзаж, а также физические процессы типа эрозии и тектонической активности. Поэтому полнота географической информации невозможна без исследованной преднамеренных, нормативных модификаций пейзажа человеческим действием [28]. Однако, на сегодняшний момент информация о географическом мире является преимущественно статической, и большая часть геоинформации хранится в форме снимков, сделанных в определённые моменты времени. Недостаточное внимание к временным изменениям в ГИС не позволяет пока избежать статического представления мира, что остаётся одним из самых важных препятствий на пути развития ГИС.

Динамические модели процессов на основе начального состояния системы должны вычислять будущие состояния, основанные на соответствующих функциях и параметрах. В этом смысле динамические модели процессов подобны операциям ГИС, которые принимают на вход географическую информацию и производят новую географическую информацию на выходе. С точки зрения объектно-ориентированного подхода, динамические модели процессов являются родственными методами, которые могут быть *инкапсулированы* в объектные классы.

В основе ГИС лежит обработка картографического отображения, и метафора карты естественным образом определяет архитектуру и функционирование ГИС. Но возникло много приложений, где геоанные не вписываются в традиционное понятие карты, и в таких случаях метафора карты имеет тенденцию ограничивать исследование.

ГИС работает хорошо со статическими данными и значительнее хуже, когда она призвана анализировать временные ряды, де-

тализованные данные в движении, или транзакциях. В настоящее время у ГИС не возникает принципиальных проблем при обработке двумерных данных, но существует много ограничений, когда важно третье пространственное измерение. Кроме того, технология ГИС не приспособлена для того, чтобы иметь дело со случаями, где отсутствуют существенные количества данных, от которых зависит принятие решения.

ГИС оперирует абсолютными позициями геоанных и их точной геометрией. Тем не менее, большинство операций и моделируемых процессов в пространстве являются инвариантными масштабированию, отражению или вращению, но не изменению относительной позиции, т.е. большинство пространственных преобразований требует только знания матрицы пространственных взаимодействий между геоанными W , а не их точных координат. В пространственном случае элементы этой матрицы могут быть установлены равными уменьшающейся функции расстояния, или длины общей границы, или же равными 1, если геоанные совместно используют общую границу, и 0 — в противном случае. Матрица W по сути своей похожа на матрицу инцидентности, используемой в теории графов.

Как выражение пространственных отношений, матрица W компактна и удобна. ГИС с её значительно более сложным аппаратом для представления пространственных свойств является полезной для того, чтобы вычислить элементы W от расстояний или смежности, и чтобы затем отобразить результаты анализа в форме карты. Однако, для фактического анализа необходимо использовать не только квадратную матрицу взаимодействий W , но и прямоугольную пространственную матрицу данных, упомянутую выше.

По сути, ГИС и матрица W отражают два различных мировоззрения: одно — обеспечение полного, *непрерывного представления* пространственного изменения, и другое — обеспечение намного более абстрактного, *дискретного представления*. Во втором мировоззрении стремление к точному представлению абсолютной позиции, которая до сих пор характеризовала развитие ГИС, важно только для подготовки данных и отображения результатов.

ГИС сегодня представляет устойчивое основание для анализа за пространственных данных. Предлагаемые коммерческие ГИС обеспечивают диапазон методик для анализа и визуализации в дополнение к существенным вспомогательным функциям пре-

образования, изменения проектирования и шага дискретизации. Используются различные стратегии интеграции, чтобы связать ГИС с более специализированными приложениями.

Интернет-взрыв конца 90-х годов произвёл существенное изменение перспектив развития ГИС и переориентировал большую часть исследований в эту область. С 1995 года разработчики ГИС поддерживают совместное использование географической информации в сети Интернет. На сегодня имеется достаточно бесплатной географической информации, доступной широкому кругу пользователей и размещённой на многочисленных web-сайтах (например, <http://www.alexandria.ucsb.edu>, <http://www.gisa.ru>, <http://www.datarplus.ru>, <http://www.gismp.com>).

Были разработаны также и новые технологии, позволяющие пользователям ГИС взаимодействовать с отдалёнными web-сайтами непосредственно через программное обеспечение клиента. В результате ГИС представляет на сегодняшний день многобillionную долларовую индустрию с миллиардами, расходуемыми ежегодно на сбор данных и их распространение, разработку программного обеспечения и разнообразных приложений. ГИС проникли фактически во все дисциплины, которые в той или иной мере имеют дело с поверхностью Земли, от атмосферной науки до океанографии, к криминологии и хронологии. Миллионы людей всё больше используют GPS и другие устройства, чтобы определять своё местоположение и направление движения.

До сих пор пространственно-временные данные были не очень доступны. Однако широкое внедрение систем слежения на основе GPS и развёртывание отображающих систем с высоким разрешением на спутниках позволяет получить изображения поверхности Земли с разрешением более 1 м., открывая диапазон новых возможностей для анализа социальной инфраструктуры и построенной формы ландшафта.

До 90-х годов о ГИС преобладало представление как о некоем скрупулёзном средстве, выполняющем слишком утомительные, сложные или отнимающие много времени задачи. Предполагалось, что ГИС станет всё более и более мощной, осуществляя всё большую и большую пропорцию известных методов анализа пространственных данных. Появление Интернета вызвало фундаментальный пересмотр этого представления и в то же самое время увеличило список потенциально возможных исследований.

ГИС в этом новом контексте представляется как среда и средство коммуникации данных о поверхности Земли. В этом представлении анализ пространственных данных обеспечивается системой локальных и удалённых служб. Местоположение геоданных приобрело функциональность: местоположение, которое является предметом анализа, местоположение пользователя, местоположение хранения данных и местоположение, обеспечивающее аналитическое обслуживание.

Развитие и внедрение ГИС привели к массовой популяризации пространственных методов и возрастанию значения старых географических дисциплин. Мощные инструментальные средства ГИС оказались в руках сотен тысяч пользователей. При этом многие из них не имеют никакого понятия о теории анализа пространственных данных. В результате этого участились случаи неправильного употребления и неверного истолкования геоданных и конфликты между использованием ГИС и научным осмыслением полученных выводов.

К сожалению, усовершенствование ГИС не оказало влияние на решение проблемы неопределённости. Не представляет труда вывести геоданные двойной точности, учитывая 14 десятичных знаков. Но в действительности по-прежнему нет такой фактической точности геоданных или вычислений, и результаты представляются с числовой точностью, определённой их погрешностью измерений.

ГИС, будучи коммерческим программным продуктом, является, тем не менее, полигоном для обширных научных исследований, находясь на грани между объективным научным знанием и прагматичными подходами, определяемыми субъективным миром человеческого рассуждения. Это приносит значительную прибыль за счёт привлечения как широких слоёв пользователей, так и лиц, принимающих решение, но в то же самое время представляет риски неосмысленного использования получаемых результатов.

3. Соотношение образного и вербального в геоинформатике

Общепринятое кодирование геоинформации в цифровой растровой и/или векторной форме с ограниченным использованием знаковых представлений влечёт за собой ограничения в части

организации семантического поиска геоданных, обусловленные, с одной стороны, отсутствием унификации описания единой модели для вербальной и невербальной информации [29], а, с другой стороны, сложностью структуризации системы геоязыков, используемых для формирования геоинформационных ресурсов. Знания о пространственных объектах не могут быть достаточно адекватно переданы только в лингвистической и/или формально-логической формах, о чём свидетельствует использование широкого спектра геоописаний разной геоязыковой принадлежности — невербальных средств описания научных знаний.

На стыке картографии и семиотики — лингвистической науки, исследующей свойства знаков и знаковых систем, сформировался особый раздел — *картографическая семиотика* (картосемиотика), в рамках которой разрабатываются общая теория систем картографических знаков. В ней изучаются проблемы, касающиеся происхождения, классификации, свойств и функций картографических знаков и способов картографического изображения.

Картографическая семиотика включает три основных раздела:

- 1) *картографическая синтактика* — изучает правила построения и употребления знаковых систем, их структурные свойства, грамматику языка карты;
- 2) *картографическая семантика* — исследует соотношения условных знаков с самими отображаемыми объектами и явлениями;
- 3) *картографическая прагматика* — изучает информационную ценность знаков как средства передачи информации, и особенности их восприятия читателями карты.

Основная задача картографической семиотики — разработка языка карты, т.е. знаковой системы, включающей условные обозначения, способы изображения, правила их построения, употребление и чтения при создании и использовании карт. Именно язык карты определил современные особенности картографической визуализации. Наибольший вклад в развитие теории картографической семиотики сделал А.А. Лютий, считавший, что язык карты составляет главный предмет исследования картографии. Он впервые выдвинул идею о существовании в языке карты двух подязыков [30]: один обеспечивает отображение

содержательной (сущностной, субстанциональной) определённости картографируемых объектов, другой — позволяет передать информацию о взаимном размещении объектов, их форме и ориентации в пространстве.

Язык карты — это объектный язык картографии. Его главными функциями (как и картографии вообще) являются коммуникативная функция, т.е. передача некоторого объёма информации от создателя карты к пользователю, и познавательная функция — получение новых знаний о картографируемом объекте. С развитием новых технологий, методов моделирования и с появлением множества новых геоизображений возникает проблема расширения понятий картографической семиотики на всю систему геоизображений.

Согласно теории А.А. Лютого [31], геоинформационные описания должны основываться на трёх основных сферах представления геознаний: невербальные знания, которые не могут быть представлены в лингвистической форме, вербальные знания, которые не могут быть адекватно переведены в невербальную форму, и та часть знаний, которая может быть представлена и в вербальной, и в невербальной формах. В комплексе проводимых исследований карты, аэрофотоснимки, литологические и стратиграфические колонки, картоподобные диаграммы, палеотектонические схемы, геохимические диаграммы, схемы корреляции и так далее, должны рассматриваться как соответствующим образом оформленные *геотексты*.

В результате исследований А.А. Лютого возникло новое научное направление — *геосемиотика*, занимающаяся разработкой вопросов знакового представления геоинформации. С точки зрения разрешения проблемы образного и вербального, А.А. Лютий близко подошёл к *геоинформационно-семантическому моделированию геосферы*. В одной из последних своих работ он с авторами высказал мнение о необходимости создания «единой семиотической системы Электронного образа Земли» [32].

Основываясь на подходе А.А. Лютого, геоинформационный портал можно рассматривать как электронную библиотеку, которая включает геотексты на разных геоязыках. Кроме геотекстов, библиотека, конечно же, включает вербальную информацию, позиционированную во времени и пространстве, то есть пространственно-определённые вербальные тексты на разных естественных языках.

Таким образом, геоинформационный портал представляется как семиотически неоднородный по языковой принадлежности его текстов. Наряду с моноязычными вербальными текстами в нём могут присутствовать поли- и кроссязычные тексты. Аналогично и геотексты в нём могут быть моноязычными, поли- и кроссязычными.

В настоящее время отечественные исследовательские проекты по созданию геоинформационного портала находятся на начальных стадиях выполнения. С концептуальной точки зрения, эти проекты могут быть разделены на два основных направления:

— разработка принципов интеграции пространственной информации на растровой и/или векторной основе с присоединёнными базами фактографических и метаданных, а также пространственно-определённой вербальной информации;

— разработка принципов и подходов к многоуровневому семантическому моделированию и интеграции пространственной и пространственно-определённой вербальной информации с одновременным использованием электронных знаковых форм представления геотекстов, а также растровой и/или векторной форм их представления.

ГИС по сути своей визуальна, и большая часть её потенциала направлена на отображение именно визуальной информации. Визуальное взаимодействие очень эффективно для человеческого восприятия, так как оно позволяет очень быстро просматривать большое количество информации. Визуальные навыки человека изучены достаточно хорошо. Они обеспечивают доступность визуальной технологии даже для детей. Использование человеком вербальных средств общения, напротив, постоянно требует изучения и совершенствования, оставаясь неким средством неполного и неточного отображения визуальной информации о предметной среде.

Однако, числовые методы исследований явно преобладают, поэтому существует очень сильное смещение исследований в пользу числового анализа, предоставляющего конкретные, хорошо структурированные данные, что можно наблюдать в обычно выражаемом презрении к приложениям ГИС, которые «только делают карты».

Современные исследования в области анализа пространственных данных основываются на активном сотрудничестве образного, числового и вербального, что позволяет обеспечивать прин-

ципально новыми методиками создаваемые коммерческие продукты.

4. Отображение пространственной онтологии в географическую концептуальную схему

Значение интеграции географической информации увеличилось из-за новых возможностей, являющихся результатом углублённого познания многогранных связей окружающего мира и обусловленных увеличивающимися возможностями представления географической информации. Онтология в информатике, согласно современным толкованиям, является «точной спецификацией концептуализации предметной области», но с определёнными ограничениями в зависимости от области интересов, и должна включать словарь терминов и некоторые спецификации их значений.

Использование онтологий способствует созданию адекватных концептуальных моделей, обеспечивая качественное, контролируемое информационное интегрирование. Одна из центральных современных задач геоинформатики — найти способ связи формального представления семантики, заложенной в онтологии, с концептуальными схемами, описывающими информацию, сохранённую в базах геоданных (БГД). Главный ожидаемый результат — формальная структура, которая реализует отображение пространственной онтологии в географическую концептуальную схему.

Отображение онтологий в концептуальные схемы необходимо производить, используя три различных уровня абстракции: формальный, уровень доменов и прикладной уровень. На формальном уровне, с высоким уровнем абстракции, выделяют концепции, чтобы представить схему и определить онтологию. На уровне доменов, схема выступает в роли характерной логической модели данных, пригодной для сопоставления с наполненным онтологией. На прикладном уровне необходимо сосредоточить внимание на специфике географических приложений.

Влияние онтологий в географических методологиях разработки систем отличается от методологий разработки традиционных систем, причём, коренные отличия прослеживаются именно на концептуальном этапе конструирования.

При традиционном подходе к моделированию систем разработчику требуется зафиксировать взгляд пользователя на реальный мир и отразить его в формальной концептуальной модели. При этом разработчик модели следует парадигме объектно-ориентированного подхода или подхода сущность-связи, в зависимости от имеющейся в его распоряжении среды программирования. В любом случае выбранный подход заставляет разработчика модели мысленно отображать концепции пользователя в абстракции, доступные в его парадигме выбора. Это отображение делается недостаточно строго, с учётом специфики данного конкретного случая, привнося, тем самым, противоречия и погрешности, которые неизбежно ведут к конфликтам между концепциями пользователя и абстракциями, зафиксированными в концептуальной модели.

Основная причина этих конфликтов — отсутствие начального соглашения между пользователем и разработчиком модели на уровне концепций реального мира. Такое соглашение могло бы быть установлено посредством использования онтологии, которая является общедоступной спецификацией прикладной области. Если онтология, основанная на взгляде пользователя на мир, была бы предварительно стенирована и формализована так, чтобы это могло оказывать влияние на процесс разработки модели, такие конфликты были бы менее вероятны. С другой стороны, консолидация концепций и знаний, представленных концептуальной схемой, могла бы быть полезной на начальных шагах конструирования онтологий.

Масштабное выполнение отображений между онтологиями и концептуальными схемами может ускорить разработку онтологий, так как концептуальные схемы широко распространены как модели документирования самого широкого спектра информации для функционирующих информационных систем. С другой стороны, эти отображения могут сокращать цикл разработки прикладных программ, получая из онтологии более точный набор концепций, на основе которых концептуальная схема может быть разработана более корректно.

Автоматическая связь между концепциями онтологии и приятивами концептуальной схемы позволила бы пользователю обращаться к информации в базе данных, используя высокоуровневые концепции [33, 34]. В [35] предлагается организовать общие концептуальные схемы для прикладных информационных хранилищ, чтобы отображать разнородные концептуальные схе-

мы в виде общей онтологии верхнего уровня. Работы [36, 37] используют интеграцию онтологий, чтобы улучшить результаты интеграции схемы. Их система вывода обучается на процессе интеграции онтологий для того, чтобы вывести глобальные схемы из локальных. Этот подход сохраняет онтологии и схемы на различных уровнях, но хотелось бы получить инструмент прямого отображения онтологий на концептуальные схемы.

Следует заметить, что специфика географического мира в достаточной мере определяет параметры отображения онтологий в концептуальные схемы [38, 39]. Чтобы адекватно представлять географический мир, необходимо иметь компьютерные представления географических знаний, которые способны фиксировать не только описательные атрибуты пользовательских концепций, но также и описывать геометрические и позиционные компоненты этих концепций. Эти представления также должны фиксировать пространственные и временные зависимости между экземплярами этих концепций. В отличие от случая обычных информационных систем, большинство пространственных и временных зависимостей не представлены в ГИС и чаще всего могут просто выводиться путём использования различных географических функций. Поэтому обязательно должна быть привнесена дополнительная семантика в концептуальную схему географического приложения, семантические спецификации которой, являющиеся частью онтологии этого приложения, зафиксированы в разработчиком модели.

Исследования, связанные с интеграцией данных, берут начало с середины 80-х годов, и сегодня это направление широко распространено среди исследователей ГИС [40–47]. Новое поколение информационных систем должно обладать способностью обрабатывать семантическую разнородность, возникающую в результате использования разнородных источников информации, предоставляемых Интернет и средствами распределённого вычисления [48].

Семантика информационной интеграции оказывается в центре внимания всё большего количества исследователей [41, 43–45, 47–51]. Использование множественных онтологий становится основной особенностью современных информационных систем, если в них предполагается поддержка семантики при интеграции информации. Онтологии могут фиксировать семантику информации, могут быть представлены в формальном языке и могут также использоваться, чтобы хранить связанные метаданные,

допуская таким образом семантический подход к информационному интегрированию.

Сложность и богатство географической информации и трудности её моделирования усложняют решение вопросов интероперабельности ГИС, таких, как интеграция различных моделей географических сущностей и различного компьютерного представления этих объектов (то есть, растрового и векторного). Ранее, когда онтологий впервые обратили на себя внимание исследователей ГИС [52–62], у многих из них возник вопрос: были ли онтологии фактически известными концептуальными методами моделирования данных, скрытыми под новым названием [63]? В работе [64] полагают, что онтологии могут использоваться как строительные блоки концептуальных схем. Можно считать, что главное различие между онтологией и концептуальной схемой в том, что они сформированы с различными целями [65]. В то время, как онтология детально описывает специфический домен, концептуальная схема создана для того, чтобы описать содержание базы данных.

Другими словами, онтология — внешнее описание по отношению к информационным системам и является спецификацией возможных миров, в то время, как концептуальная схема — внутреннее средство информационной системы и является спецификацией одного возможного мира [66].

Онтологии семантически богаче, чем концептуальные схемы баз данных, и, тем самым, ближе к когнитивной модели пользователя. Концептуальные схемы представляют модель данных, которые должны быть сохранены в базе данных. Онтология представляет концепции реального мира. Несмотря на то, что суть некоторой концепции, выраженной в онтологии, может быть только одна, концептуальная схема будет по-разному фиксировать эту концепцию в зависимости от особенности необходимого представления географических данных.

Необходимо также указать, что существует чёткое различие между концептуальной моделью данных и концептуальной схемой. Модель данных относится к способу моделирования любой базы данных, включая её систему обозначений. Модель сущность-связь (ER) [67], объектная методика моделирования (OMT) [68] и унифицированный язык моделирования (UML) [69] — примеры концептуальных моделей данных. Концептуальные схемы, с другой стороны, относятся к результату моделирования, то есть, они представляют собой набор диаграмм,

использующих данную модель данных как язык, чтобы выразить структуры данных разрабатываемого приложения.

Практически, концептуальные схемы ограничены логическими средствами, используемыми на имеющейся вычислительной технике. Кроме того, концептуальная схема для представления пространственных данных зависит от неявного предположения, что её компоненты являются измеримыми. Поэтому, концептуальные схемы приписывают для каждого из их компонентов одно или несколько соответствующих компьютерных представлений. Тем самым концептуальная схема требует тесной связи с набором компьютерных представлений, в то время, как онтология не требует такого связывания.

Фактически, использование онтологий развивается с более широкими возможностями обеспечения консолидации знаний, что, несомненно, выходит за рамки возможностей, которыми обладают текущие инструментальных средства и методы моделирования данных.

В традиционном концептуальном моделировании онтологии или игнорируются или испытывают недостатки в формальных методах [57], но информация о структурах данных, классов, и доменов, которые составляют концептуальную схему, может быть адаптирована и использована для формирования классов онтологий. С подобной проблемой сталкиваются и программисты, у которых способ кодирования варьируется от их индивидуальных предпочтений [70]. И программист, и разработчик модели имеют их собственные онтологии, заданные в явном или неявном виде.

В [35] впервые введён термин «*информационные системы, управляемые онтологиями*» для систем, которые используют формально определённые онтологии, а в [71] предложена структура для разработки географических приложений, использующих онтологии. Структура использует онтологии в качестве основы для интеграции географической информации. На основе объединения онтологий, которые связаны с источниками географической информации, был разработан механизм, позволяющий интегрировать географическую информацию преимущественно на основе её значения. Так как интегрирование может происходить на различных уровнях, были также разработаны основные механизмы для изменения уровня детализации. Использование онтологии, переведённой в разряд элемента информационный

системы, является базисом *географических информационных систем, управляемых онтологией* (ODGIS).

Наиболее распространённая в геоинформатике концептуальная модель основывается на том, что географическая действительность представлена либо как совокупность полностью определимых объектов, либо как гладкие, непрерывные пространства — поля. Объектная модель представляет мир как поверхность, заполненную дискретными, идентифицируемыми сущностями с геометрическим представлением и описательными атрибутами. Непрерывная модель рассматривает географическую действительность как набор пространственных распределений в географическом пространстве.

Некоторые авторы считают [72], что объектная и непрерывная модели имеют основное смысловое значение при размещении их в декартовой системе координат, являющейся априорной каркасной системой для расположения пространственных явлений. При таком подходе декартова система координат — просто нейтральный контейнер, в пределах которого происходят все материальные процессы. Каждый объект пространства связан к одной или нескольким точкам расположений на Земле, благодаря чему фиксируются абсолютные положения объектов в пространстве.

Альтернативой абсолютному местоположению является понятие *относительного пространства* [73], составленное на основе относительных положений в пространстве существующих географических объектов. Пространственные модели взаимодействия и модели «положение–локализация», используемые на транспорте, — примеры приложений, которые используют понятие относительного пространства.

Общепринятая технология ГИС, как правило, реализует подход на основе абсолютного местоположения, так как большинство геометрических представлений, используемых в GIS, такие, как решётки, TIN-сети, плоские векторные карты, основаны на понятии *геосвязанного расположения*. Это не удивительно, ведь понятия объектов и областей могут быть обобщены в едином формальном определении.

В формальном представлении географическое пространство обычно сводится к области, представляющей интерес для решения поставленных задач. Географическое поле определяется соотношением $f = (R, V, \lambda)$, где R — географическая область, V — набор атрибутов и $\lambda: R \rightarrow V$ — отображение между точками

в R и значениями в V (в OpenGIS [74] λ обозначает функцию покрытия).

Геообъекты представляют индивидуальные сущности географической области. Пусть дан набор географических областей $R_1, \dots, R_n$ и набор атрибутов $A_1, \dots, A_n$, с доменами $D(A_1), \dots, D(A_n)$, тогда географический объект определяется соотношением $(a_1, \dots, a_n, S_1, \dots, S_m)$, где a_i — его описательные атрибуты ($a_i \in D(A_i)$) и S_i — его географические расположения ($S_i \subseteq R_i$).

Онтология играет существенную роль в создании ГИС, так как она позволяет устанавливать соответствия и взаимосвязи среди различных доменов пространственных сущностей и отношений [53]. Использование онтологий будет способствовать улучшению функционирования информационных систем, благодаря тому, что удастся избежать проблем, вызванных противоречиями между онтологиями, неявно внедрёнными в ГИС, конфликтами между онтологическими концепциями и реализацией, и конфликтами между онтологиями «здорового смысла» пользователя и математическими концепциями программного обеспечения [57].

Однако следует помнить о важности внесения фундаментальных семантических описаний в самом начале проектирования [43]. Важность роли онтологий в моделировании пространственной неопределённости отмечена в [60]. В [75] обсуждается необходимость разрабатывать пространственные информационные теории, которые ориентированы на проблемы пользователей ГИС, а не на проблемы реализации.

Другой семантический подход к интеграции географической информации представлен в GeoCosm [76] — доступный через сеть ГИС-протокол для интеграции автономных распределённых разнородных геоданных. Здесь предлагается каноническая модель, которая интегрирует концептуальные схемы, при этом онтология используется для разрешения конфликтов среди информационных источников.

Ниже онтологии классифицированы согласно их зависимости от специфики задачи или точки зрения пользователя [77]:

— *онтологии верхнего уровня* описывают очень обобщённые понятия. В OGDIS онтология верхнего уровня описывает обобщённое понятие пространства. Например, теория, описывающая части и целое и их отношение к топологии, названная мереотопологией [78], как раз относится к этому уровню.

— *онтологии* домена описывают словарь, связанный с групповым доменом типа дистанционного восприятия или городской среды.

— *онтологии задачи* описывают задачу или действие типа интерпретации изображения или оценки вредного шума.

— *прикладные онтологии* описывают концепции, зависящие как от домена так и от задачи, и обычно являются их специализацией. Они представляют потребности пользователя относительно спецификации приложения.

Традиционный подход к созданию информационной системы обычно начинается с концептуального этапа проектирования, в результате которого, устраняя несущественные детали и сосредотачиваясь на отображении существенной относительно поставленной задачи информации, разработчик создаёт концептуальную схему базы данных. За этим следует логический этап проектирования, на котором высокоуровневая концептуальная схема преобразуется во внутреннюю схему, или схему выполнения. Эта схема поддерживается логическими средствами и функциями системы управления базой данных (СУБД) и на её основе загружаются фактически хранимые данные [79]. Наконец, этап физического проектирования определяет потребности системы в терминах значений структур данных, соответствующих концепциям концептуальной схемы.

Концептуальные схемы независимы от СУБД, в то время как конкретные схемы экземпляров данных строго подчинены СУБД.

В случае географических приложений имеются следующие различия. Сначала, на этапе концептуального моделирования, разработчик должен решить, как будет представлена каждая концепция реального мира: будет ли это единственный объект, или полностью определённый класс сущностей, или гладкая, непрерывная пространственная сущность в виде поля, раstra и т. д. Таким образом, ему следует выбрать геометрические представления, которые будут использоваться для отражения каждой концепции.

Следует заметить, что одна и та же самая концепция может быть представлена различными геометрическими представлениями. Например, почва может быть представлена TIN, регулярной сеткой или набором изолиний, а река может быть представлена как линия или как площадной объект, охватывающий пространство между его берегами.

Однако, концептуальный уровень и уровень выбора представления должны быть фактически объединены в проекте, так как разработчик модели определяет выбор вариантов представления после того, как решение относительно объектов или полей принято, потому, что это решение требуется прежде, чем на этой стадии моделирования может быть возобновлена разработка спецификаций пространственных расположений. Поэтому эта стадия моделирования названа *концептуальным уровнем представлений* [80].

На концептуальном уровне представлений оправданы любые полезные преобразования между представлениями, чтобы избежать избыточности. Пусть, например, данные относительно почвы представлены в виде TIN. Однако, необходимо также представление почвы в виде набора изолиний. Изолинии могут быть получены с помощью некоторого алгоритма конвертирования TIN, и эта операция может быть определена на уровне представлений. При этом концепции почвы или её основных географических характеристик расположения не изменяются в процессе выбора варианта представления или когда производятся преобразования между представлениями.

Позже должны быть определены возможные варианты каждого представления [81]. Так как эти варианты существенно зависят от типа представления, которое было выбрано, но независимы от проектируемой ГИС, этой фазе проектирования соответствует отдельный уровень, называемый *уровнем способов изображения представлений*. На этом уровне чёткость или лёгкость визуальной интерпретации данных в выбранной среде способов изображения представления (экран, бумага) определяется графическими параметрами типа очертания символа, цвета, типа линии или текстыры заполнения площадного объекта.

С уже разработанными представлениями и их способами изображения процесс перемещается на уровень реализации, в котором пространственные структуры данных обычно заполняются необходимыми данными путём использования инструментальных средств и языков, которыми располагает ГИС или СУБД, с возможностью хранения пространственных структур данных. Так как географические приложения требуют некоторой картографической основы, типа основной карты, существует также потребность в стадии преобразования данных.

Имеются три различных уровня абстракций, на которых находятся как онтологии, так и концептуальные схемы: формальный, доменный и прикладной уровни.

Первый уровень — *формальный*, в котором большее количество абстрактных концепций включено в конструкцию онтологий и концептуальных схем. В случае онтологий на этом уровне мы имеем абстракции формальных особенностей сущностей [54], такие, как понятия времени и пространства. В случае концептуальных схем, на этом уровне мы находим основные идеи, взятые из концептуальных моделей данных, то есть, понятия, которые широко используются в концептуальном моделировании данных: объекты, поля и связи.

Когда мы сопоставляем содержание формального уровня с географическими приложениями, мы получаем второй уровень абстракции, или *уровень доменов*. На этом уровне, онтологии описывают словарь, который используется, чтобы представить действительное содержание домена знания, типа географии и географической информации.

Онтология географических понятий [53], которая описывает географическое пространство, географические объекты, и явления географического пространства, явно присутствует на этом уровне. Когда подобное перемещение к верхнему уровню абстрагирования делается при работе с концептуальными моделями, следует обратить внимание на то, что здесь, как и в случае онтологий, необходим словарь, чтобы выразить абстрактные понятия формального уровня. Концептуальные модели данных, типа OMT-G, обеспечивают эту потребность, определяя графическую систему обозначений, основанную на элементарных понятиях. Эти элементарные понятия соответствуют семантическому содержанию формального уровня: Прямоугольник означает объект или поле, пунктирная линия означает пространственную связь, треугольник обозначает спецификацию, и так далее. Поэтому, примитивы, составляющие концептуальную модель данных, используются как форма языка, для определения данных, необходимых конкретному приложению на следующем уровне абстракций.

На *прикладном уровне*, онтологии более конкретны, являясь уточнением специализаций уровня доменов, и формируются в соответствии с концепциями поля знания, заданного приложением. На этом уровне, онтология — это то, что в [54] названо Е-онтологией, — способ осмысления набора концепций, позволяющего

совместно использовать его группой пользователей, чтобы согласованно с ним взаимодействовать. В концептуальном моделировании данных на прикладном уровне примитивы концептуальной модели объединены, чтобы образовать удобочитаемые диаграммы, из которых можно получить детальные потребности приложения по организации данных. Более определённо, используются примитивы модели данных, чтобы определить концептуальную схему приложения, перемещаясь от уровня доменов до прикладного уровня.

Следует обратить внимание на то, что конечный продукт концептуального моделирования данных, который является концептуальной схемой, использует ряд концепций от предыдущих уровней, внедрённых в схему и изобразённых с помощью специального графического «языка». В обычном процессе моделирования пользователю предлагается ознакомиться с этим языком, чтобы квалифицированно проверить обоснованность схемы. Аналогично, если нужно автоматически извлечь онтологию из схемы, следует формализовать трансляцию системы обозначений модели в онтологические концепции, и затем использовать эти концепции как основу для поиска специфических для приложения понятий и определений.

В некотором смысле, это эквивалентно обучению компьютерной программы чтению и пониманию концептуальных схем, заполняя общий справочный фрейм набором онтологических концепций формального уровня и уровня доменов.

В заключение, приведём набор определений, которые могут использоваться в качестве базиса для алгебраической формулировки отображения между пространственной онтологией и географической концептуальной схемой. Определения, данные ниже, составлены согласно концепциям, которые были введены для определения OMT-G примитивов [80].

Определение 1. Терм — это тройка $\tau = [\eta, \delta, A]$, $\tau \in T$, где η — строка символов, содержащих название терма, δ — строка символов, содержащих его определение, и A — набор атрибутивных доменов $A_1, A_2, \dots, A_n$, каждый связанный с множеством значений V_i .

Определение 2. Отношение $\varphi: T \rightarrow T$, $\varphi \in \Phi$ — есть функция из T в T такая, что для каждого терма $\tau_1 \in T$, существует терм $\tau_2 = \varphi(\tau_1)$, $\tau_2 \in T$.

Определение 3. Семантическое отношение σ между двумя терминами — это отношение, которое принадлежит набору семантических отношений $\Sigma = \{\text{Нурегушту, Нурупушту (является), Метеопушту (является частью), Синонимия}\}$, $\Sigma \subseteq \Phi$.

Определение 4. Пространственное отношение ρ между двумя терминами — это отношение, которое принадлежит набору пространственных отношений $P = \{\text{смежность, пространственная совместимость, близость, связность}\}$, $P \subseteq \Phi$.

Определение 5. Пояснительное отношение k между двумя терминами — это отношение, которое принадлежит набору пояснительных отношений K . К специально определён для каждой онтологии и $K \subseteq \Phi$.

Определение 6. Онтология — это пара $\Theta = [T, \Phi]$, где $T = \{\tau_1, \tau_2, \dots, \tau_n\}$ — набор термов, и $\Phi = \{\varphi_1, \varphi_2, \dots, \varphi_n\}$, и $\exists \varphi_i \in (\Sigma \cup K)$.

Определение 7. Пространственная онтология — это пара $\Theta_s = [T_s, \Phi_s]$, где $T_s = \{\tau_{s1}, \tau_{s2}, \dots, \tau_{sn}\}$ — набор термов, и $\Phi_s = \{\varphi_{s1}, \varphi_{s2}, \dots, \varphi_{sn}\}$, $\Phi_s \supseteq (\Sigma \cup K \cup P)$, и $\exists \varphi_{si} \in P$.

Определение 8. Сущность — это кортеж $e = [n, A]$, где n — строка символов, указывающих имя объекта, A — набор атрибутов, каждый из которых связывается с заданным доменом. Каждая сущность e принадлежит E , набору всех сущностей, т. е. $e \in E$.

Определение 9. Географическая сущность — это тройка $g = [n, A, \text{гер}]$, где n — строка символов, указывающих имя класса, A — набор атрибутов, каждый из которых связывается с заданным доменом, и гер — вариант представления, выбранный для класса (гер $\in \text{REP}$, где REP — точка, линия, многоугольник, вершина (узел), однонаправленная дуга, двунаправленная дуга, изолиния, выборка, TIN, плоский подраздел, двумерные соты). Каждый пространственный объект принадлежит G , набору всех пространственных объектов, то есть, $g \in G$, $G \subseteq E$.

Определение 10. Отношение r : $E \rightarrow E$, $r \in R$ является функцией от набора сущностей E к самому себе, так что для каждой сущности $e_1 \in E$ существует сущность $e_2 = r(e_1)$ такая, что $e_2 \in E$. Каждое отношение принадлежит R , — набору всех отношений, то есть, $r \in R$.

Определение 11. Семантическое отношение m между двумя сущностями является отношением, которое принадлежит на-

бору пространственных отношений $M = \{\text{обобщение, специализация, агрегирование}\}$, и $M \subseteq R$.

Определение 12. Пространственное отношение s между двумя сущностями — это отношение, которое принадлежит набору пространственных отношений $S = \{\text{пространственная взаимосвязь, пространственное агрегирование, сеть}\}$, и $S \subseteq R$.

Определение 13. Пояснительное отношение x между двумя терминами — это отношение, которое принадлежит набору пространственных отношений $X = \{\text{простая ассоциация}\}$, и $X \subseteq R$.

Определение 14. Концептуальная схема — это пара $C = [E, R]$, и $R \supseteq (\text{MUX})$.

Определение 15. Концептуальная схема для географической информации — это пара $C_g = [E_g, R_g]$, где $E_g \subseteq E$ и $\exists e_{gi} \in G$, $R_g \subseteq R$, и $\exists r_{gi} \in (\text{MUX})$, и $\exists r_{gj} \in S$.

На основании предыдущих определений может быть формально представлено отображение между онтологиями и концептуальными моделями.

Определение 16. Отображение $\psi_{\tau-e}(\tau, e)$ между термом τ в пространственной онтологии Θ_s и сущностью e в концептуальной схеме для географической информации C_g таково, что $\forall (\tau_i) \in \Theta_s$, $\exists (e_j) \in C_g \mid \psi_{\tau-e}(\tau_i, e_j) \in \Psi_{\tau-e}$.

Отображение $\psi_{\tau-e}(\tau, e)$ определяется с помощью отображений между элементами термина в онтологии и элементами сущности в концептуальной схеме:

$$\psi_{\tau-e}(\tau, e) = \psi_1(\eta, n) + \psi_2(\delta, \emptyset) + \psi_3(A, A) + \psi_4(\emptyset, \text{гер}),$$

где $\emptyset$ представляет невозможность отображения.

Определение 17. Отображение $\psi_{\varphi-r}(\varphi, r)$ между соотношением в пространственной онтологии Θ_s и отношением в концептуальной схеме для географической информации C_g — таково, что $\forall (\varphi_i) \in \Theta_s$, $\exists (r_j) \in C_g \mid \psi_{\varphi-r}(\varphi_i, r_j) \in \Psi_{\varphi-r}$.

Отображение $\psi_{\varphi-r}(\varphi, r)$ определяется с помощью отображений между тремя видами отношений в онтологии (семантическими, пространственными, и пояснительными) и соответствующими соотношениями в концептуальной схеме: $\psi_{\varphi-r}(\varphi, r) = \psi_5(\sigma, m) + \psi_6(\rho, s) + \psi_7(k, x)$.

Определение 18. Отображение $\psi(\Theta_s, C_g)$ между пространственной онтологией Θ_s и концептуальной схемой для географической информации C_g таково, что $\forall (\tau_i, \varphi_i) \in \Theta_s$, $\exists (e_j, r_j) \in C_g \mid \psi_{\tau-e}(\varphi_i, r_j) \in \Psi_{\varphi-r} \mid \psi_{\varphi-r-e}(\varphi_i(\tau_i), r_j(e_j)) \in \Psi_{\varphi\tau-r_e}$.

5. Согласованное формирование и сопровождение интегрированных геоинформационных ресурсов

При размещении геоданных, представленных средствами разнотипной геоязыковой принадлежности, особый интерес вызывает проблема группировки данных в наборы, классы, категории, объекты и другие классификации, соответствующие однородности (достаточной близости) геоданных, дающей основание для помещения их в одну и ту же таблицу. Набор «похожих» пространственных объектов, таких как участки или реки, хранится в таблице и называется *классом объектов*.

Наборы связанных классов объектов, имеющих одинаковую пространственную привязку, могут быть организованы в более крупную структуру, называемую *набором классов объектов*. Каждый пространственный объект в базе геоданных (БГД) содержит форму (геометрию) и может участвовать в топологических отношениях. Способность хранения геометрии объекта является одним из обязательных условий существования БГД, поскольку каждый пространственный объект всегда должен быть доступен для отображения и анализа.

Что же касается отношений, то помимо типичных (общих) для реляционной БД отношений типа «род-вид» или «часть-целое», необходимо позаботиться об отображении топологических (соединение, узел пересечения и др.) и пространственных (касание, внутри, снаружи и др.) отношений. Можно заметить, что в действительности пространственные отношения могут быть представлены в рамках реализации топологических отношений заданием соответствующей топологии.

Обеспечивающие функционирование ГИС объекты в БГД хранятся в связанных реляционных таблицах. Часть из этих таблиц содержит атрибутивные данные, другие — представляют собой группировки пространственных объектов. Специальные таблицы задают отношения между пространственными объектами, правила проверки корректности и домены атрибутов.

Большинство реальных географических объектов, имеющих четкие очертания, целесообразно моделировать с помощью векторного представления. Векторные данные описывают форму пространственных объектов в виде упорядоченного набора координат с ассоциированными атрибутами. Такое представление поддерживает некоторые геометрические операции, среди которых расчёт длины или площади, идентификация перекрестий и пересечений с другими объектами, нахождения смежных или находящихся в близости пространственных объектов.

Векторные объекты принято классифицировать по их размерности следующим образом [82].

— Точки являются фигурами без измерений, они представляют географические объекты, которые слишком малы, так что можно пренебречь их размерами. Каждая точка хранится как одна пара координат (x, y) и набор атрибутов;

— Линии являются фигурами с одним измерением, они представляют географические объекты, размеры которых существуют только в одном измерении. Каждая линия хранится как упорядоченный набор пар координат (x, y) и набор атрибутов. Сегменты линий могут быть прямыми, округлыми, эллиптическими или сплайновыми.

— Полигоны являются фигурами с двумя измерениями, они представляют площадные географические объекты. Каждый полигон хранится как набор сегментов, образующих его границу.

Есть ещё один тип, который условно можно отнести к векторным данным, — это надпись (аннотация). Надписи являются описаниями, которые связаны с пространственными объектами и отображают названия или атрибуты.

Векторные данные в базе геоданных обладают структурой, которая управляет хранением пространственных объектов в соответствии с их размерностью и отношениями.

Фоновые изображения, объекты с размытыми очертаниями и оригиналы результатов съёмки, напротив, целесообразно подерживать в растровой (сеточной) форме в виде пиксельных значений на двухмерной сетке, называемой растром.

При необходимости отображения физических рельефов с помощью им трёхмерностью, используется нерегулярная триангуляционная сеть (TIN), являющаяся, по сути, моделью поверхности. База геоданных хранит TIN в виде интегрированного набора узлов с их значениями высоты и треугольников со сторонами, соединяющими узлы. Высоту (или z -значение) можно интерпо-

лизовать для любой точки в пределах географического экстенда (области простираения) TIN. Модель данных TIN позволяет выполнять различные виды анализа поверхности, оценку видимости произвольных точек поверхности из заданной точки наблюдения др.

Из вышеперечисленного только векторные объекты могут использоваться как самостоятельные субъекты, участвующие в топологических отношениях. Растры и TIN, подобно рисункам или другим OLE-объектам, участвуют в организации БД как введённые и обрабатываются специальными приложениями.

Очень важным и достаточно принципиальным моментом является выбор одной из трёх вышеперечисленных категорий, к которой следует причислять *векторный объект*.

Во-первых, векторный объект обязан своим появлением процедуре векторизации некоторой части растрового изображения. Соответственно, его характеристики напрямую зависят от выбранных параметров этой процедуры.

Во-вторых, вопрос о возможности пренебрежения размерами векторного объекта решается в зависимости от условий поставленной задачи. Так, имеющий все основания считаться площадным объект «станция» при укрупнении детализации рассмотрения может перейти в категорию точечного объекта, а при дальнейшем агрегировании параметров задачи просто исчезнуть из рассмотрения как несущественный фактор, находящийся внутри диапазона выбранной точности. Это не означает, что он исключается из БГД, но в рамках данного рассмотрения в категории точечных объектов его нет. Таким образом, мы подходим к необходимости динамического формирования категорий векторных объектов в зависимости от условий взаимодействия с проблемной средой и поставленной задачи.

От принадлежности к определённой категории зависит участие векторного объекта в топологических отношениях. Топология, устанавливающая топологические отношения между географическими объектами, делает возможным проведение расширенного пространственного анализа и играет фундаментальную роль в обеспечении качества БГД.

В настоящий момент можно указать следующие базовые типы топологий [1, 82]:

— *линейно-узловая топология* (линейные объекты могут иметь общие конечные точки);

- *полигональная топология* (площадные объекты могут иметь общие границы);
- *топология маршрутов* (линейные объекты могут иметь общие сегменты);
- *топология регионов* (площадные объекты могут перекрываться);
- *узловая топология* (линейные объекты могут иметь общие конечные вершины);
- *точечные события* (вершины линейных объектов могут совпадать с точечными объектами).

В одних и тех же топологических взаимоотношениях могут участвовать несколько классов полигональных, точечных и линейных объектов. В современных ГИС пользователи могут устанавливать, какие топологии являются подходящими для соответствующих слоёв данных. Топология обычно реализуется в виде набора правил целостности, определяющих поведение пространственно взаимосвязанных геоданных и объектных классов. Топологические правила могут быть заданы для объектов внутри объектного класса или для объектов, принадлежащих двум или более объектным классам. Они позволяют моделировать пространственные отношения связности, смежности, близости и совпадения, а также контролировать целостность совпадающей геометрии у различных классов объектов. Для этого в ГИС вводят инструментальные средства, интерпретирующие пространственные условия и обеспечивающие поиск и исправление любых нарушений этих правил.

Все процедуры установления топологических отношений и группировки геоданных в существующих ГИС предназначены для использования специалистом в процессе настройки ГИС на решение определённого класса задач. Для реализации в ГИС автоматической адаптации к изменению параметров взаимодействия пользователя с предметной областью нужен принципиально новый инструмент поддержки согласованности структуры топологических отношений и группировки в динамическом режиме. В следующем разделе описывается аппарат *теории структурной согласованности*, подход на основе которого позволяет приблизиться к динамической реализации указанных процессов.

6. Аппарат теории структурной согласованности

Не конкретизируя состав совокупности объектов, подлежащих группировке в некие наборы, рассмотрим эту совокупность как множество взаимосвязанных абстрактных элементов. В самом общем случае объекты заданной совокупности имеют формализованную структуру, т.е. могут быть описаны некоторым конечным набором параметров (реквизитов).

Таким образом, пусть имеется совокупность однородных объектов любой природы $O = \{o_i\}$ ($i = 1, \dots, m$), где однородность понимается в смысле описания каждого объекта из рассматриваемой совокупности в виде одинакового упорядоченного набора (вектора) из m характеристик — реквизитов: $o_i = (p_1^i, \dots, p_m^i)$. В информационном моделировании, говоря о сравнении любых двух объектов из этой совокупности, понимают оценку их сходства на основе интегрированного сходства соответствующих характеристик этих двух объектов [83]. При этом, в зависимости от характера и условий решаемой задачи, в сравнении могут участвовать либо все m реквизитов, либо некоторое их подмножество из $k \leq m$ элементов. Естественно, что сами реквизиты могут иметь различную природу и тип, главное, что для каждого p_j имеется некоторая функция, позволяющая дать численную оценку для каждой пары характеристик из этой группы. В реальных приложениях не все объекты из рассматриваемой совокупности могут обладать полным набором признаков: для некоторых из них значения отдельных признаков могут быть неизвестными или неприменимыми. В таких ситуациях однородность всей совокупности объектов может быть восстановлена, если присвоить отсутствующим признакам нулевые значения, восстановив, таким образом, вектор характеристик $\{p^i\}$ до размерности k .

Введём функцию F сходства объектов по совокупности k выбранных реквизитов, нормированную на максимальный диапазон значения реквизита на множестве из n объектов [83]. Для двух объектов o_i и o_j , сходство которых устанавливается с точностью до k реквизитов $\{p_i\}_{mk}$, функция может быть записана как:

$$F(o_i, o_j) = \frac{1}{k} \sum_{i=1}^k w_{ml} \frac{|p_{ml}^i - p_{ml}^j|}{\max |p_{ml}^i - p_{ml}^j|}.$$

Здесь $0 \leq w_{ml} \leq 1$ — вес ml -го реквизита, а $\max |p_{ml}^i - p_{ml}^j|$ задаёт диапазон варьирования ml -го реквизита.

При таком определении функции F она принимает значения из отрезка $[0, 1]$, причём единица означает «полное сходство» элементов o_i и o_j , а ноль — их «полное различие» (отсутствие сходства). Промежуточные значения могут быть интерпретированы как интегральная степень сходства двух объектов по k выбранным реквизитам, что даёт возможность рассматривать это множество в качестве полного неориентированного графа с нагруженными нормированными связями.

Если на основе значений $F(o_i, o_j) = f_{ij}$ устанавливать знак связи, то он уже будет интерпретироваться как результат объективной оценки сходства объектов, а не как субъективная оценка эксперта.

Одной из важных задач формирования структуры знаковых связей для последующего распределения n объектов по N классам на основании этой структуры является выбор порогового значения α функции F , до которого (т.е., когда $0 \leq F(o_i, o_j) \leq \alpha$) объекты o_i и o_j считаются несходными по k признакам и после которого ($\alpha < F(o_i, o_j) \leq 1$) — сходными. Если присвоить знак «минус» всем связям с $0 \leq F(o_i, o_j) \leq \alpha$ и знак «плюс» — в противном случае, то полученная знаковая структура будет представлять собой дискретную знаковую модель рассматриваемого множества.

Более общий случай возникает, если задать два пороговых значения α и β и присваивать знак «минус», если $0 \leq F(o_i, o_j) \leq \alpha$, значение 0, если $\alpha \leq F(o_i, o_j) \leq \beta$, и знак «плюс», если $\beta \leq F(o_i, o_j) \leq 1$. Этим вводится диапазон индифферентности (α, β), значение связи внутри которого говорит об индифферентности к сходству по k выбранным признакам. При равенстве $\alpha = \beta$ мы возвращаемся к предыдущему случаю.

После расстановки полученных знаков связей структура множества примет вид знакового графа с соответствующей матрицей связности. Теперь можно перейти к формулированию задачи *структурной согласованности множества* с возможностью автоматической адаптации к изменению параметров топологических отношений и группировки в динамическом режиме.

Задача поиска согласованной структуры знаний всегда связана с организацией этих знаний адекватно решаемым задачам. Такая организация, являясь одновременно и инструментом поиска знаний, релевантных поставленной задаче, должна обладать определённой гибкостью, что обусловлено необходимостью пополнения её новыми знаниями и динамизмом изменения пара-

метров задачи. Изменения состава базы знаний влекут за собой переустановление взаимосвязей между отдельными элементами в базе знаний и последующие целенаправленные структурные преобразования, направленные на согласование её отдельных компонентов в смысле взаимоотношений и взаимосоответствий ряда параметров.

Наиболее распространённый подход к установлению согласованности основывается на оценке парных взаимосвязей между элементами рассматриваемой совокупности объектов. Однако для учёта взаимного влияния связей следует анализировать тернарные отношения, введя правило, позволяющее каждое тернарное отношение относить к одному из двух типов: согласованному или рассогласованному. В ряде работ [84, 85] эти типы получили названия соответственно *консонансного* и *диссонансного состояния*. Можно представить всю совокупность имеющих объектов в виде множества тернарных отношений (треугольников), что даёт возможность построить некоторую систему классификации внутреннего состояния множества взаимосвязанных однородных объектов, позволяющую определить его структурную согласованность.

Одной из возможностей исследования тернарных связей является критерий Хайдера [84], согласно которому треугольник является консонансным, если положительная взаимосвязь между любой парой его вершин соответствует идентичности связей этих двух вершин с третьей, иначе — диссонансным.

Основным результатом исследований свойств консонансного множества явилась теорема [84], согласно которой любое консонансное множество M_K (состоящее только из консонансных треугольников) можно представить в виде пары подмножеств M_1 и M_2 : $M_K = M_1 \cup M_2$, таких, что любые два объекта o_i и o_j , принадлежащие одному и тому же подмножеству, связаны между собой положительной связью, а принадлежащие разным подмножествам — отрицательной. Одно из рассмотренных в теореме подмножеств M_1 или M_2 может оказаться пустым.

В результате произвольных изменений знаков связей консонансное множество «портится» и преобразуется в так называемое *ассонансное* (часть треугольников переходит в диссонанс). Обратные преобразования переводят полученное ассонансное множество обратно в консонанс. Понятно, что, обладая свободой преобразования связей между объектами заданного множества, его можно перевести в любое другое состояние. Однако если

учесть, что нечётное число изменений знака любой связи эквивалентно её однократному изменению [84], а чётное не изменяет знака этой связи, то можно говорить о наборе связей, однократное изменение знаков которых переводит любое множество в заданное состояние. Такие связи получили название *сильных диссонансных связей* исходного множества, а задача поиска согласованного состояния этого множества сводится таким образом к задаче поиска его сильных диссонансных связей.

Консонансное множество из N объектов не единственное: существует $[N/2] + 1$ различных видов таких множеств. Это определяет неоднозначность существующих наборов сильных диссонансных связей, приводящих произвольное ассонансное множество к консонансу: каждый набор приводит к своему типу консонанса. Следовательно, можно говорить о существовании такого минимального набора, который приводит заданное ассонансное множество к консонансному типу. При этом консонансное множество, в которое переводится заданное ассонансное множество с помощью изменения знаков минимального набора сильных диссонансных связей, называется *консонансным прообразом* заданного ассонансного множества [84].

Консонансный прообраз соответствует согласованному состоянию произвольного ассонансного множества, поэтому задача поиска ближайшего по типу консонансного множества связана с согласованием по структурным признакам исходной совокупности объектов.

Так как консонансное множество представимо в виде совокупности двух подмножеств, таких, что объекты внутри каждого из этих подмножеств связаны положительной связью, а из разных подмножеств — отрицательной, то задача приведения произвольного ассонансного множества в консонанс в этом случае заключается в разбиении его на два класса объектов, однотипных внутри каждого из них.

Необходимость определения согласованного множества, которое может состоять из более чем двух подмножеств, приводит к определению *поликонсонанса* и *консонанса степени n* [86]. Под поликонсонансом степени n будем понимать согласованное состояние множества M , при котором это множество состоит не более, чем из n подмножеств, таких, что объекты внутри каждого из них связаны только положительными связями, а объекты из разных подмножеств — только отрицательными связями. Под

консонансом степени n будем понимать поликонсонанс степени n , состоящий в точности из n классов (рис. 1).

o_1	+	-	-	-	-	-
+	o_2	-	-	-	-	-
-	-	o_3	+	-	-	-
-	-	+	o_4	-	-	-
-	-	-	-	o_5	+	o_6
-	-	-	-	-	+	+
-	-	-	-	-	+	o_6

Консонанс 3-й степени

Рис. 1

Важное следствие такого определения консонанса касается диссонансного множества, которое теперь можно рассматривать как консонанс степени N , где N — количество элементов в данной совокупности. Таким образом, диссонансное множество можно рассматривать как согласованное по структуре связей его объектов, и задача приведения к согласованному состоянию теперь касается исключительно ассонансного множества.

Заметим также, что задача приведения некоторого произвольного ассонансного множества к поликонсонансу степени n соответствует задаче разбиения данного множества на не более, чем n классов, объединяющих «похожие» или «близкие» по структуре связей объекты, в то время, как задача приведения к консонансу степени n соответствует задаче разбиения данного множества ровно на n таких классов.

В реальных условиях часто приходится иметь дело с разреженными графами, в которых некоторые связи могут быть неопределёнными или несуществующими. Для возможного рассмотрения такого случая введём понятие *индифферентной связи*, соответствующей отсутствию информации о связи между парой выбранных объектов. Индифферентная связь не является ни положительной, ни отрицательной и её естественно пометить символом 0 и считать, что её присутствие не влияет на определение вида состояния множества.

Одним из подходов, который может привести к определённому уменьшению перебора в процессе поиска поликонсонанса, может служить использование различного рода эвристики, например, предложенных в работах [85–87]. Эвристические методы способны давать существенные результаты в некоторых

практических применениях и представляют собой достаточно сложный метод снижения перебора при поиске консонанса. Заметим, однако, что весьма целесообразным оказывается и использование более простых и тривиальных методов. Например, если учесть, что разбиение множества на классы в соответствии с поликонсонансом, соответствует объединению отдельных его элементов в группы по структурному признаку, то становится очевидным, что два объекта из рассматриваемой совокупности, имеющие тождественные связи с остальными объектами из этого множества и имеющие положительную взаимосвязь между собой, окажутся в результате такой реструктуризации в одном классе. А раз так, то можно изначально выявить элементы с одинаковыми структурами связей и условно представить их как единый (интегрированный) объект из рассматриваемой совокупности, что приводит к снижению общего количества объекта и, следовательно, — к снижению перебора. Сложность задачи выявления объектов с тождественными связями соответствует задаче сортировки множества объектов, для которой существует множество достаточно эффективных и практически реализуемых алгоритмов.

С введением понятия консонанса степени $n > 2$ изменяется и представление о множестве, названном выше ассонансным. Таким множеством названа совокупность объектов, которая не является ни консонансной, ни диссонансной.

Рассмотрим, теперь множество, матрица связности которого представлена на рис. 2.

o_1	+	-	-	-	-	-
+	o_2	-	-	-	-	-
-	-	o_3	+	-	-	-
-	-	-	+	o_4	-	-
-	-	-	-	-	o_5	+
-	-	-	-	-	-	o_6

Рис. 2

Очевидно, что оно является поликонсонансом третьей степени, однако, если рассматривать консонанс степени 2, то отнесительно такого консонанса исходное множество оказывается ассонансным. Следовательно, с понижением степени консонанса, консонансное множество может преобразоваться в ассонансное.

Аналогичные рассуждения, проведённые в обратном порядке, позволяют сделать вывод, что с повышением степени консонанса, ассонансное множество может преобразоваться в консонансное. Таким образом, одной из задач анализа структурной согласованности множества является задача, связанная с выявлением типа состояния, в котором находится изучаемое множество. Способы решения этой задачи изложены в [85].

Другая задача, о которой мы уже говорили при изучении свойств консонанса степеней 2, — это устранение рассогласованности приведением множества в согласованное состояние минимально возможными изменениями начальной структуры множества. Решение этой задачи связывалось с поиском сильных диссонансных связей.

Рассмотрим теперь случай консонанса степеней $n > 2$. Анализ структуры матрицы связности консонансного множества показывает, что объекты o_i и o_j , принадлежащие одному и тому же подмножеству имеют одинаковую структуру связей, т.е. они неразличимы по связям с другими объектами данного множества. Если объекты o_i и o_j принадлежат разным подмножествам, то связи с объектами из этих же двух подмножеств у них противоположны по знаку, а с объектами из других подмножеств — одинаковы. Это свойство консонанса позволяет сформулировать правило, по которому любой из его объектов можно перевести из одного подмножества в любое другое: достаточно у этого объекта изменить знаки всех его связей с объектами из этих двух подмножеств.

Процедура перевода объекта из одного подмножества в другое может интерпретироваться как перевод вершины адекватного системе знакового графа из одной компоненты в другую, поэтому эту операцию можно назвать *поверхинным изменением*. В [85] показано, что любая последовательность поверхинных изменений не выводит множество из вида состояния. Под видом состояния множества понимается одно из трёх: консонанс, диссонанс или ассонанс.

Основная идея создания алгоритма оценки состояния произвольного ассонансного множества заключается в его сравнении с некоторым, быть может, произвольно взятым консонансным прообразом и последующим «синхронным» преобразованием этого прообраза и исходного множества с использованием операции поверхинного изменения связей с тем, чтобы новый консонансный прообраз оказался «ближе» к ассонансному множеству.

Естественно, главной задачей теперь становится выработка критерия оценки «близости» множеств.

Рассмотрим два произвольных множества M_1 и M_2 , состоящих из одних и тех же N объектов и различающихся только взаимосвязями между этими объектами. Для любой пары объектов o_i и o_j из этих двух множеств определим число r_{ij} по следующему правилу:

- а) $r_{ij} = 0$, если $i = j$;
- б) $r_{ij} = 0$, если $i \neq j$ и связи между o_i и o_j в этих множествах различны;
- с) $r_{ij} = 1$, если $i \neq j$ и связи между o_i и o_j в этих множествах совпадают.

В качестве характеристики различия множеств M_1 и M_2 введём набор из N чисел $v_1, v_2, \dots, v_N$, где $v_i = \sum_{j=1}^N r_{ij}$ для $i = 1, \dots, N$ и построим из этих компонентов вектор $V = (v_1, v_2, \dots, v_N)$, выражающий пообъектное различие в знаках у двух рассматриваемых множеств. Если рассматриваемые множества представить в виде графов, где в качестве вершин выступают объекты этих множеств, то любую компоненту вектора V можно интерпретировать как сумму различий в связях соответствующей вершины в обоих графах. Такая интерпретация позволяет рассматривать вектор V как вектор поверхинных различий двух множеств. Сумма компонентов этого вектора

$$S_v = \sum_{i=1}^N v_i$$

называется суммой поверхинных различий для этих множеств.

Понятно, что $0 \leq v_i \leq N - 1$ для любого i . Отсюда следует, что сумма поверхинных различий S_v не превышает $N(N - 1)$ и всегда чётна.

Заметим, что задание множества посредством вектора поверхинных различий от некоторого другого множества в общем случае неоднозначно, что можно продемонстрировать на следующем примере (рис. 3).

Множество, представленное знаковой матрицей связности B , имеет вектор поверхинных различий V , состоящий из всех единиц, относительно тривиального консонансного множества, представленного матрицей A . При этом такой же вектор повершин-

A					B					C					V				
α_1	+	+	+	+	α_1	+	–	+	–	α_1	+	–	–	–	1				
+	+	α_2	+	+	+	α_2	+	–	–	+	α_2	–	–	–	1				
+	+	+	α_3	+	–	+	α_3	+	–	–	–	α_3	+	+	1				
+	+	+	+	α_4	+	–	+	+	–	–	–	+	+	α_4	1				

Рис. 3

ных отличий множество B имеет и с консонансным множеством, представленным матрицей C .

Множеству, минимально удалённому по сумме изменяемых связей от множества M , соответствует минимальная сумма компонентов вектора поверхностных различий с M . Укажем условие, при котором поверхностное преобразование приводит к уменьшению суммы поверхностных различий.

Пусть дано множество из N объектов. Построим произвольное (например, тривиальное) консонансное множество из этих же объектов и определим вектор поверхностных различий для этих двух множеств:

$$V = (v_1, v_2, \dots, v_N), \quad \text{где} \quad v_i = \sum_{j=1}^N r_{ij}.$$

Сумма поверхностных различий определяется как сумма компонентов этого вектора:

$$S_v = \sum_{i=1}^N v_i = \sum_{i=1}^N \sum_{j=1}^N r_{ij}.$$

Предположим, что существует объект, для которого сумма его «плохих» связей больше половины всех связей этого объекта. Для определённости будем считать, что это объект α_1 : $v_1 > (N-1)/2$. Поверхшинное изменение связей этого объекта приводит к реорганизации вектора поверхностных различий так, что новая сумма будет определена как

$$S'_v = \sum_{i=1}^N v'_i = v'_1 + \sum_{i=2}^N \sum_{j=1}^N r'_{ij} = v'_1 + \sum_{i=2}^N r'_{i1} + \sum_{i=2}^N \sum_{j=2}^N r'_{ij}.$$

Поскольку при поверхностном изменении все связи объекта изменяются на противоположные, то $v'_1 = N-1-v_1$. В силу

симметричности связей $r_{ij} = r_{ji}$, следовательно,

$$\sum_{i=2}^N r'_{i1} = \sum_{j=2}^N r'_{1j} = v'_1 = N-1-v_1.$$

При поверхностном изменении связей объекта α_1 связи между α_i и α_j ($i > 1$ и $j > 1$) остаются неизменными, отсюда

$$\sum_{i=2}^N \sum_{j=2}^N r'_{ij} = \sum_{i=2}^N \sum_{j=2}^N r_{ij}.$$

Таким образом, получаем

$$S'_v = (N-1-v_1) + (N-1-v_1) + (S_v - 2v_1).$$

Далее,

$$S'_v - S_v = 2(N-1-v_1) - 2v_1.$$

Таким образом, если мы хотим добиться того, чтобы поверхностное изменение связей некоторого объекта привело к уменьшению суммы поверхностных различий, нам необходимо выбирать для этого объекты, для которых справедливо соотношение:

$$v_i > \frac{N-1}{2}.$$

Однако, можно показать, что алгоритм, основанный на последовательном изменении связей у одного из элементов множества, в общем случае, не даёт желаемого результата и полученное в результате множество не будет минимально удалённым консонансным преобразованием.

Проиллюстрируем это на следующем примере, выяснив, одновременно, причину подобной ситуации (рис. 4).

На рисунке представлена знаковая матрица связности ассоциансного множества из девяти элементов. Вектор V_1 представляет собой вектор поверхностных различий данного множества от консонансного множества $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5\}$; $\{\alpha_6, \alpha_7, \alpha_8, \alpha_9\}$, состоящего из тех же объектов. Все компоненты этого вектора не превосходят половины всех связей соответствующих объектов и, следовательно, любое поверхностное изменение любого из объектов данного множества приведёт только к увеличению суммы поверхностных различий. Однако, приведённое консонансное множество нельзя считать минимально удалённым от данного ассоциансного, поскольку можно построить множество $\{\alpha_5\}$:

тем — для объекта o_i . Учитывая, что при однократной операции поверхностных переборов связи объекта изменяются на противоположные, приходим к выводу, что связи между объектами из выбранной группы останутся также неизменными:

$$\sum_{i=1}^k \sum_{j=1}^k r'_{ij} = \sum_{i=1}^k \sum_{j=1}^k r_{ij}.$$

По причине симметричности связей между объектами, оставшиеся два слагаемых оказываются равными

$$\sum_{i=k+1}^N \sum_{j=1}^k r'_{ij} = \sum_{i=1}^k \sum_{j=k+1}^N r'_{ij}.$$

Учитывая все эти соотношения, можно записать:

$$S'_v - S_v = 2 \sum_{i=1}^k \sum_{j=k+1}^N r'_{ij} - 2 \sum_{i=1}^k \sum_{j=k+1}^N r_{ij}.$$

Введём $x_i = \sum_{j=k+1}^N r_{ij}$ — количество «плохих» связей объекта o_i , входящего в группу из k объектов, с объектами, не входящими в эту группу. Поскольку эти связи меняются на противоположные, то имеем: $x'_i = (N - k - x_i)$. Тогда условия уменьшения общей суммы поверхностных различий будут выглядеть следующим образом:

$$\begin{aligned} S'_v - S_v &= 2 \sum_{i=1}^N x'_i - 2 \sum_{i=1}^N x_i = 2 \sum_{i=1}^N (N - k - x_i) - 2 \sum_{i=1}^N x_i = \\ &= 2k(N - k) - 4 \sum_{i=1}^N x_i < 0, \end{aligned}$$

что выполнимо при условии

$$\sum_{i=1}^N x_i > \frac{k(N) - k}{2}.$$

Заметим, что в случае группы из одного объекта ($k = 1$), мы приходим к сформулированному ранее правилу $v_i > (N - 1)/2$.

Важным фактором в деле практической реализации предложенного алгоритма поиска согласованного состояния произвольно взятого множества является эффективность его работы,

определяемая в первую очередь необходимыми временными затратами на поиск ближайшего по структуре связей консонансного прообраза. К сожалению, с ростом N — количества объектов в множестве — трудоёмкость алгоритма возрастает как 2^N [88]. Если исходное множество является консонансным или диссонансным, то эта проблема решается предложенным в начале этой работы методом определения типа состояния множества на основе оценки строк в матрице связности множества. В случае ассонансного множества немаловажное значение приобретает изначально задаваемый консонансный прообраз данного множества: чем «ближе» он оказывается к искомому, тем меньше количество шагов нам потребуется для решения задачи.

Однако, выбор первоначального прообраза — это отдельная задача, на решение которой оказывают существенное влияние дополнительные факторы, в том числе связанные с субъективной оценкой экспертом условий решаемой им задачи, параметров окружающей обстановки и пр. Хотя можно предложить и чисто математические методы определения исходного прообраза, нельзя не отметить и другую совокупность факторов, оказывающих существенное влияние на эффективность алгоритма [88, 89]. Эти факторы связаны с применением дополнительных результатов, направленных на уменьшение количества необходимых переборов при выполнении тех или иных операций в рамках выполнения всего алгоритма.

Использование этих результатов в практической реализации алгоритма уменьшения согласованности множества взаимосвязанных объектов позволило в значительной степени добиться повышения эффективности работы программных средств, что проявилось, прежде всего, во временных характеристиках симетры.

7. Заключение

Знание о пространственных объектах является предметом геоинформатики, при этом геоинформация первична по отношению к таким знаниям, ибо она обеспечивает всё необходимое для их пополнения. Знание о пространственных объектах — это проверенный практикой результат познания окружающей действительности, адекватное отражение её в мышлении человека или на каком-либо носителе, — это представление фактов, утверждений о фактах или правила получения утверждений о фактах

из имеющихся фактов. Благодаря геоинформатике становятся реальностью уникальные возможности представления знания об окружающем мире на уровне описываемых визуальных концепций ментального мира человека, обеспечивая единообразие представления информации практически во всех сферах деятельности, связанных с использованием рельефа земной поверхности.

История развития геоинформатики показывает, что получение геоинформации первоначально напрямую связывалось с анализом пространственных данных, что, в свою очередь, стимулировало развитие соответствующих методов. Возникновение компьютерной обработки геоданных повлекло за собой бурный рост количества автоматизированных систем обработки геоинформации, названных ГИС. В связи с этим последние десятилетия характеризовались разработкой ГИС, максимизирующих коммерческую выгоду предоставлением широкого диапазона методик визуализации пространственных объектов за счёт интеллектуализации пользовательского интерфейса. Справедливости ради стоит отметить, что развитие ГИС привело и к массовой популяризации пространственных методов, однако смещение акцента в сторону коммерциализации визуальных представлений, очевидно. В связи с этим следует заметить, что ГИС, являясь исследовательским полигоном идей геоинформатики, в последнее время обнаруживают серьёзный дефицит внедрения методов анализа и обработки пространственных данных, который не позволяет в достаточной мере отражать информационные процессы в природе и обществе на уровне современных требований.

Проведённый в работе анализ публикаций позволяет сделать вывод, что наиболее актуальными перспективными направлениями в развитии геоинформатики в ближайшем будущем представляются исследования, связанные с разработкой эффективного аппарата унификации и стандартизации представления, обработки, анализа и интерпретации различных видов геоинформации. Это, в частности, находит своё подтверждение в идее глобального проекта «Электронная Земля: научные информационные ресурсы и информационно-коммуникационные технологии», представляющего собой совокупность проектов, направленных на решение актуальной задачи создания ИТ-инфраструктуры как распределённой среды сбора, обработки и анализа разнородной информации в области наук о Земле [32]. Кроме этого, весьма востребованными в ближайшее время окажутся исследования по представлению мобильной геоинформации — оперативной ин-

формации о местоположении и его характеристиках, получаемой мобильными наземными средствами со спутниковых навигационных систем.

Основной задачей представленной работы была попытка объединить взгляды известных мировых учёных на перспективы развития геоинформатики. Основываясь на мнениях авторов, здесь следует считать работы М. Goodchild, F. Folpessa, G. Samaga, В. Smith, А. А. Лютого, А. М. Берлянта. Всё обсуждение, включая и ретроспективу, построено по принципу оценки взаимовлияния теории географических информационных систем и методов анализа пространственных данных. В последнем разделе для анализа топологической группировки геоданных авторами предложен оригинальный подход систематизации и классификации геоданных на основе аппарата теории структурной согласованности.

Список литературы

1. Goodchild M.F. Geographical information science // Int. J. of Geographical Information Systems. 1992. V. 6. P. 31–45.
2. Duckham M., Goodchild M.F., Worboys M.F. Fundamentals of geographic information science. — N. Y.: Taylor and Francis, 2003.
3. Foresman T. W. The history of GIS: Perspectives from the pioneers. — N. J.: Prentice Hall PTR, Upper Saddle River, 1998.
4. Geographical information systems: Principles and applications / Ed. by Maguire D.J., Goodchild M.F., Rhind D. W.. — Harlow: Longman Scientific and Technical, 1991.
5. Tomlinson R.F., Calkins H. W., Marble D.F. Computer handling of geographical data: An examination of selected geographic information systems. — Paris: UNESCO Press, 1976.
6. Longley P.A., Goodchild M.F., Maguire D.J., Rhind D. W. Geographic information systems and science. — N. Y.: Wiley, 2001.
7. Whittle P. On stationary processes in the plane // Biometrika. 1954. V. 41. P. 434–449.
8. Kendall M.G. Time series. — London: Griffin, 1976.
9. Matheron G. Principles of geostatistics // Economic Geology. 1963. V. 58. P. 1246–1266.
10. Greig-Smith P. The use of random and contiguous quadrats in the study of the structure of plant communities // Annals of Botany. 1952. V. 16. P. 293–316.

11. *Berry B.J.L., Marble D.F.* Spatial analysis. — N. J.: Prentice Hall, Englewood Cliffs, 1968.
12. *Matern B.* Spatial variation (2nd ed). — N. Y.: Springer, Berlin Heidelberg, 1986.
13. *Anderson T. W.* The statistical analysis of time series. — N. Y.: Wiley, 1971.
14. *Grenander U., Rosenblatt M.* Statistical analysis of stationary time series. — N. Y.: Wiley, 1957.
15. *Johnston J.* Econometric methods. — N. Y.: McGraw Hill, 1963.
16. *Malinvaud E.* Statistical methods of econometrics. — Amsterdam: North Holland Publishing Company, 1970.
17. *Besag J.E.* Spatial interaction and the statistical analysis of lattice systems // Journal of the Royal Statistical Society. 1974. V. B36. P. 192–225.
18. *Ripley B.D.* Spatial statistics. — N. Y.: Wiley, 1981.
19. *Diggle P.* Statistical analysis of spatial point patterns. — London: Academic Press, 1983.
20. *Cressie N.* Statistics for spatial data. — N. Y.: Wiley, 1991.
21. *Goodchild M.F.* Measurement-based GIS. In: Shi W., Fisher P.F., Goodchild M.F. (eds) Spatial data quality. — N. Y.: Taylor and Francis, 2002. P. 5–17.
22. *Goodchild M.F.* Uncertainty: The Achilles heel of GIS? // Geo Info Systems. 1998. V. 11. P. 50–52.
23. *Zhang L., Goodchild M.F.* Uncertainty in geographical information. — N. Y.: Taylor and Francis, 2002.
24. *Unger M.J., Goodchild M.F.* Integrating spatial data analysis and GIS: A new implementation using the Component Object Model (COM) // Int. J. of Geographical Information Science. 2002. V. 16. P. 41–54.
25. *Goodchild M.F., Egenhofer M.J., Kemp K.K., Mark D.M., Sheppard E.* Introduction to the Varenus Project // Int. J. Geogr. Info. Sci. 1999. V. 13. P. 731–745.
26. *Anselin L.* What is Special About Spatial Data? // Alternative Perspectives on Spatial Data Analysis. Technical Report 89-4. — Santa Barbara, Calif.: National Center for Geographic Information and Analysis, 1989.
27. *Tobler W.R.* A Computer Movie: Simulation of Population Change in the Detroit Region // Economic Geography. 1970. V. 46. P. 234–240.
28. *McHarg I.L.* Design with Nature. — Garden City, N.Y.: The Natural History Press, 1969.

29. *Зацман И.М., Лютий А.А.* Семиясфера электронного образа Земли и знаковое представление геотекстов // Системы и средства информатики. — М.: Наука, 2001. — Вып. 11. — С. 132–148.
30. *Лютий А.А.* О сущности языка карты // География и природные ресурсы. 1985. № 3. С. 42–151.
31. *Лютий А.А.* Язык карты: сущность, система, функции. — М.: ГЕОС, 2002. — 327 с.
32. *Лютий А.А., Зацман И.М., Куренков С.А.* Семиясфера Электронного образа Земли: основные структурные составляющие // Картография XXI века: теория, методы, практика: Тез. докл. 2-й Всеросс. конф. по картографии. Т. 1. — М., 2001. — С. 154–161.
33. *Moulton A., Madnick S.E., Siegel M.D.* Knowledge representation architecture for context interchange mediation: Fixed income securities investment examples // W01: WEBN — First International Workshop on Electronic Business Hubs: XML, Metadata, Ontologies, and Business Knowledge on the Web. — Munich, 2001.
34. *Moulton A., Madnick S.E., Siegel M.D.* Cross organizational data quality and semantic integrity: Learning and reasoning about data semantics with context interchange mediation // MIT Sloan School of Management Working Paper 108, 2001.
35. *Guarino N.* Formal ontology and information systems, in Guarino N. (Ed.) Formal Ontology in Information Systems. — Amsterdam: IOS Press, 1998. — P. 3–15.
36. *Hakimpour F., Geppert A.* Global schema generation using formal ontologies // Conceptual Modeling — ER 2002, 21st International Conference on Conceptual Modeling. — Tampere, 2002.
37. *Hakimpour F., Timpf S.* Using ontologies for resolution of semantic heterogeneity in GIS // 4th AGILE Conference on Geographic Information Science. — Brno, 2001.
38. *Anselin L.* What is special about spatial data? Alternative perspectives on spatial data analysis. — Santa Barbara, CA: NCGIA, 1989.
39. *Egenhofer M.* What's special about spatial? Database requirements for vehicle navigation in geographic space // Sigmod Record. 1993. V. 22. P. 398–402.
40. *Goodchild M., Egenhofer M., Fegeas R., Kottman C.* Interoperating Geographic Information Systems. — Norwell, MA: Kluwer Academic, 1999.
41. *Bishr Y.* Semantic aspects of interoperable GIS. — Netherlands: Wageningen Agricultural University, 1997.
42. *Bishr Y.* Overcoming the semantic and other barriers to GIS interoperability // Int. J. of Geographical Information Science. 1998. V. 12. P. 299–314.

43. *Harvey F.* Designing for interoperability: Overcoming semantic differences, in Goodchild M., Egenhofer M., Fegeas R., Kottman C. (Eds.) *Interoperating Geographic Information Systems*. — Norwell, MA: Kluwer Academic, 1999. P. 85–98.
44. *Gahegan M.* Characterizing the semantic content of geographic data, models, and systems, in Goodchild M., Egenhofer M., Fegeas R., Kottman C. (Eds.) *Interoperating Geographic Information Systems*. — Norwell, MA: Kluwer Academic, 1999. P. 71–84.
45. *Kashyap V., Sheth A.* Semantic heterogeneity in global information system: The role of metadata, context and ontologies // *Cooperative Information Systems Current Trends and Directions* / Ed. by M. Papazoglou, G. Schlageter. — London: Academic Press, 1996. — P. 139–178.
46. *Mena E., Kashyap V., Illarramendi A., Sheth A.* Domain specific ontologies for semantic information brokering on the global information infrastructure // *Formal Ontology in Information Systems* / Ed. by N. Guarino. — Amsterdam: IOS Press, 1998. P. 269–283.
47. *Worboys M., Deen S.* Semantic heterogeneity in geographic databases // *SigmodRecord*. 1991. V. 20. P. 30–34.
48. *Sheth A.* Changing focus on interoperability in information systems: from system, syntax, structure to semantics // *Interoperating Geographic Information Systems* / Ed. by M. Goodchild, M. Egenhofer, R. Fegeas, C. Kottman. — Norwell, MA: Kluwer Academic, 1999. P. 5–29.
49. *Kuhn W.* Defining semantics for spatial data transfer // *Sixth International Symposium on Spatial Data Handling*. — Edinburgh, Scotland, 1994.
50. *Camara G., Souza R., Freitas U., Monteiro A.* Interoperability in practice: Problems in semantic conversion from current technology to OpenGIS // *Lecture Notes in Computer Science*. 1999. V. 1580. P. 129–138. (Interoperating Geographic Information Systems — Second International Conference, INTEROP'99 / Ed. by A. Vckovski, K. Brassel, H.-J. Schek).
51. *Rodriguez A.* Assessing semantic similarity among spatial entity classes // *Spatial Information Science and Engineering*. — Orono, ME: University of Maine. 2000. P. 182.
52. *Smith B., Mark D.* Geographical categories: An ontological investigation // *Int. J. of Geographical Information Science*. 2001. V. 15. P. 591–612.
53. *Smith B., Mark D.* Ontology and geographic kinds // *International Symposium on Spatial Data Handling*. — Vancouver, BC, Canada, 1998.

54. *Smith B.* An introduction to ontology, in Pequet D., Smith B., Brogaard B. (Eds.) *The Ontology of Fields*. — Santa Barbara, CA: National Center for Geographic Information and Analysis. 1998. P. 10–14.
55. *Smith B., Mark D.* Ontology with human subjects testing: An empirical investigation of geographic categories // *The American Journal of Economics and Sociology*. 1999. V. 58. P. 245–272.
56. *Mark D.* Toward a theoretical framework for geographic entity types // *Lecture Notes in Computer Science*. 1993. V. 716. P. 270–283 (Spatial Information Theory / Ed. by A. Frank, I. Campari).
57. *Frank A.* Spatial ontology // *Spatial and Temporal Reasoning* / Ed. by O. Stock. — Dordrecht, The Netherlands: Kluwer Academic, 1997. — P. 135–153.
58. *Frank A.* Tiers of ontology and consistency constraints in geographical information systems // *Int. J. of Geographical Information Science*. 2001. V. 15. P. 667–678.
59. *Fonseca F., Egenhofer M.* Ontology-driven geographic information systems // 7th ACM Symposium on Advances in Geographic Information Systems. — Kansas City, MO, 1999.
60. *Bittner T., Winter S.* On ontology in image analysis in integrated spatial databases // *Lecture Notes in Computer Science*. 1999. V. 1737. P. 168–191 (Integrated Spatial Databases: Digital Images and GIS / Ed. by P. Agouris, A. Stefanidis).
61. *Camara G., Monteiro A., Paiva J., Souza R.* Action-driven ontologies of the geographical space: Beyond the field-object debate // *GIScience 2000 — First International Conference on Geographic Information Science*. — Savannah, GA, 2000.
62. *Rodriguez A., Egenhofer M., Rugg R.* Assessing semantic similarity among geospatial feature class definitions // *Lecture Notes in Computer Science*. 1999. V. 1580. P. 1–16 (Interoperating Geographic Information Systems — Second International Conference, INTEROP'99 / Ed. by A. Vckovski, K. Brassel, H.-J. Schek).
63. *Winter S.* Ontology: Buzzword or paradigm shift in GI science? // *Int. J. of Geographical Information Science*. 2001. V. 15. P. 587–590.
64. *Fikes R., Farquhar A.* Distributed repositories of highly expressive reusable ontologies // *IEEE Intelligent Systems*. 1999. V. 14. P. 73–79.
65. *Cui Z., Jones D., O'Brien P.* Semantic B2B integration: Issues in ontology-based applications // *Sigmod Record Web Edition*. 2002. V. 31. P. 920–926.

66. *Bishr Y.A., Kuhn W.* Ontology-based modeling of geospatial information // 3rd. AGILE Conference on Geographic Information Science, Helsinki, Finland, 2000.
67. *Chen P.S.S.* The entity-relationship model: Towards a unified view of data // ACM Transactions on Database Systems. 1976. V. 1. P. 9–36.
68. *Rumbaugh J., Blaha M., Premierani W., Eddy F., Lorenzen W.* Object-Oriented Modeling and Design. — Englewood Cliffs, NJ: Prentice-Hall, 1991.
69. Rational Software Corporation. The unified language: Notation guide, version 1.1 // 1.1 ed: Rational Software Corporation, 1997.
70. *Sowa J.* Knowledge Representation: Logical, Philosophical, and Computational Foundations. — Pacific Grove, CA: Brook/Cole, a division of Thomson Learning, 2000.
71. *Fonseca F.* Ontology-driven geographic information systems // Spatial Information Science and Engineering. — Orono: University of Maine, 2001. P. 118.
72. *Couclelis H.* People manipulate objects (but cultivate fields): Beyond the raster-vector debate in GIS // Lecture Notes in Computer Science. 1992. V. 639. P. 65–77 (Theories and Methods of Spatio-Temporal Reasoning in Geographic Space / Ed. by A.U. Frank, I. Campari, U. Formentini).
73. *Couclelis H.* From cellular automata to urban models: New principles for model development and implementation // Environment and Planning B: Planning and Design. 1997. V. 24. P. 165–174.
74. OpenGIS. The OpenGIS® Guide-Introduction to Interoperable Geoprocessing and the OpenGIS Specification. — Wayland, MA: Open GIS Consortium, 1996.
75. *Kuhn W.* Metaphors create theories for users // Lectures Notes in Computer Science. 1993. V. 716. P. 366–376 (Spatial Information Theory / Ed. by A. Frank, I. Campari).
76. *Ram S., Khatri V., Zhang L., Zeng D.D.* GeoCosm: A semantics-based approach for information integration of geospatial data // Conceptual Modeling — ER 2001, 21st International Conference on Conceptual Modeling, Yokohama, Japan, 2001.
77. *Guarino N.* Semantic matching: Formal ontological distinctions for information organization, extraction, and integration // Information Extraction: A Multidisciplinary Approach to an Emerging Information Technology, International Summer School, SCIE-97, Frascati, Italy, 1997.
78. *Smith B.* On drawing lines on a map // Lecture Notes in Computer Science. 1995. V. 988. P. 475–484 (Spatial Information Theory — A

- Theoretical Basis for GIS, International Conference COSIT'95 / Ed. by A. Frank and W. Kuhn).
79. *Elmasri R., Navathe S.* Fundamentals of database systems (3rd ed). — Reading, MA: Addison-Wesley, 2000.
80. *Borges K., Davis C., Laender A.* OMT-G: An object-oriented data model for geographic applications // Geoinformatica. 2001. V. 5. P. 221–260.
81. *Davis C., Laender A.* Multiple representations in GIS: Materialization through map generalization, geometric and spatial analysis operations // 7th ACM Symposium on Advances in Geographic Information Systems. — Kansas City, MO, 1999.
82. *Зейлер М.* Моделирование нашего мира (руководство ESRI по проектированию базы геоданных). — California: ESRI Press, 1999. — 254 с.
83. *Дулин С.К.* Знаковая интерпретация сходства для согласованности множества объектов // Изв. АН СССР. Техн. кибернетика. 1990. № 5. С. 22–25.
84. *Дулин С.К.* Об одной процедуре уменьшения структурной согласованности системы знаний // Изв. АН СССР. Техн. кибернетика. 1987. № 2. С. 22–34.
85. *Дулин С.К., Киселёв И.А.* Управление структурной согласованностью в базе знаний // Изв. АН СССР. Техн. Кибернетика. 1991. № 5. С. 29–39.
86. *Дулин С.К.* Согласование структур в условиях расширенного понятия консонанса // Изв. АН СССР. Техн. кибернетика. 1989. № 5. С. 86–93.
87. *Dulin S.K.* The Approach to Structural Consistency of Situations' Models in Active Knowledge Base // Proceedings of the 1995 ISIC Workshop of 10th IEEE International Symposium on Intelligent Control, AdRem, Inc., Monterey, USA, 1995. P. 153–158.
88. *Дулин С.К.* Разработка средств контроля согласованности компонентов системы знаний // Интеллектуальные системы, АТН РФ. Т. 1, вып. 1–4. 1996. С. 157–164.
89. *Дулин С.К., Киселёв И.А.* Интеллектуальная система организации многоуровневой согласованной базы знаний // Программные продукты и системы, № 4. 2002. С. 19–22.