

- Витрины данных синхронизированы и совместимы с корпоративным представлением. Имеется корпоративная модель данных. Существует возможность сравнительно лёгкого расширения хранилища и добавления новых витрин данных
- Гарантированная производительность

Недостатки:

- Существует избыточность данных, ведущая к росту требований на хранение данных
- Требуется согласованность с принятой архитектурой многих областей с потенциально различными требованиями (например, скорость внедрения иногда конкурирует с требованиями следовать архитектурному подходу)

Выше приведены основные варианты архитектур СППР. Выбор конкретного варианта зависит от условий, в которые поставлена проектная группа. На выбор архитектуры влияют ответы на такие вопросы как: нужен ли быстрый возврат от инвестиций; является ли проектная группа профессиональной; существует ли формализованная методология.

7.2. Методы поддержки принятия решений на основе информационных технологий

Для поддержки принятия решений с помощью информационных технологий, включая анализ и выработку альтернатив, в СППР используются следующие методы: 1) информационный поиск; 2) интеллектуальный анализ данных; 3) извлечение (поиск) знаний в базах данных; 4) рассуждение на основе прецедентов; 5) имитационное моделирование; 6) генетические алгоритмы; 7) искусственные нейронные сети; 8) методы искусственного интеллекта. Рассмотрим подробно каждый из них.

Информационный поиск (Information retrieval). Информационный поиск (ИП) (англ. *Information retrieval*) — процесс поиска *неструктурированной* документальной информации и наука об этом поиске²³. Термин «информационный поиск» был впервые введён Кельвином Муром в 1948 в его докторской диссертации, опубликован и употребляется в литературе с 1950.

Сначала системы автоматизированного информационного поиска, или информационно-поисковые системы (ИПС), использовались лишь для управления информационным взрывом в научной литературе. Многие университеты и публичные библиотеки стали использовать ИПС для обеспечения доступа к книгам, журналам и другим документам. Широкое распространение ИПС получили с появлением сети Интернет. У русскоязычных пользователей наибольшей популярностью пользуются поисковые системы Google, Яндекс и Рамблер.

Поиск информации представляет собой процесс выявления в некотором множестве документов (текстов) всех таких, которые посвящены

указанной теме (предмету), удовлетворяют заранее определенному условию поиска (запросу) или содержат необходимые (соответствующие информационной потребности) факты, сведения, данные.

Процесс поиска включает последовательность операций, направленных на сбор, обработку и предоставление необходимой информации заинтересованным лицам.

В общем случае поиск информации состоит из четырех этапов:

1. определение (уточнение) информационной потребности и формулировка информационного запроса;
2. определение совокупности возможных держателей информационных массивов (источников);
3. извлечение информации из выявленных информационных массивов;
4. ознакомление с полученной информацией и оценка результатов поиска.

Полнотекстовый поиск — поиск по всему содержимому документа. Пример полнотекстового поиска — любой интернет-поисковик, например www.yandex.ru, www.google.com. Как правило, полнотекстовый поиск для ускорения поиска использует предварительно построенные индексы. Наиболее распространенной технологией для индексов полнотекстового поиска являются инвертированные индексы.

Поиск по метаданным — это поиск по неким атрибутам документа, поддерживаемым системой — название документа, дата создания, размер, автор и т. д. Пример поиска по реквизитам — диалог поиска в файловой системе (например, MS Windows).

Поиск по изображению — поиск по содержанию изображения. Поисковая система распознает содержание фотографии (загружена пользователем или добавлен URL изображения). В результатах поиска пользователь получает похожие изображения. Так работают поисковые системы: Xcavator; Retrieve; PolarRose; Picollator Online by Recognition.

Методы информационного поиска. Адресный поиск - процесс поиска документов по чисто формальным признакам, указанным в запросе.

Для осуществления нужны следующие условия:

1. Наличие у документа точного адреса
2. Обеспечение строгого порядка расположения документов в запоминающем устройстве или в хранилище системы.

Адресами документов могут выступать адреса веб-серверов и веб-страниц и элементы библиографической записи, и адреса хранения документов в хранилище.

Семантический поиск - процесс поиска документов по их содержанию.

Условия: Перевод содержания документов и запросов с естественного языка на информационно-поисковый язык и составление поисковых образов документа и запроса. Составление поискового описания, в котором указывается дополнительное условие поиска.

²³ <http://ru.wikipedia.org/>

Принципиальная *разница* между адресным и семантическим поисками состоит в том, что при адресном поиске документ рассматривается как объект с точки зрения формы, а при семантическом поиске - с точки зрения содержания. При семантическом поиске находится множество документов без указания адресов. В этом принципиальное отличие каталогов и картотек. Библиотека - собрание библиографических записей без указания адресов.

Документальный поиск - процесс поиска в хранилище информационно-поисковой системы первичных документов или в базе данных вторичных документов, соответствующих запросу пользователя. Два вида документального поиска:

1. Библиотечный, направленный на нахождение первичных документов.
2. Библиографический, направленный на нахождение сведений о документах, представленных в виде библиографических записей.

Фактографический поиск - процесс поиска фактов, соответствующих информационному запросу.

К фактографическим данным относятся сведения, извлеченные из документов, как первичных, так и вторичных и получаемые непосредственно из источников их возникновения. Различают два вида поиска:

1. Документально-фактографический, заключается в поиске в документах фрагментов текста, содержащих факты.
2. Фактологический (описание фактов), предполагающий создание новых фактографических описаний в процессе поиска путем логической переработки найденной фактографической информации.

Информационный поиск как наука. Информационный поиск — большая междисциплинарная область науки, стоящая на пересечении когнитивной психологии, информатики, информационного дизайна, лингвистики, семиотики, и библиотечного дела.

Информационный поиск рассматривает поиск информации в документах, поиск самих документов, извлечение метаданных из документов, поиск текста, изображений, видео и звука в локальных реляционных базах данных, в гипертекстовых базах данных таких, как Интернет и локальные интранет-системы. Существует некоторая путаница, связанная с понятиями поиска данных, поиска документов, информационного поиска и текстового поиска. Тем не менее, каждое из этих направлений исследования обладает собственными методиками, практическими наработками и литературой.

В настоящее время информационный поиск — это бурно развивающаяся область науки, популярность которой обусловлено экспоненциальным ростом объемов информации, в частности в сети Интернет. Информационному поиску посвящена обширная литература и множество конференций. Одной из наиболее известных является TREC, организованной в 1992 Министерством обороны США совместно с

Институтом Стандартов и Технологий (NIST) с целью консолидации исследовательского сообщества и развития методик оценки качества ИП.

Говоря о системах информационного поиска, употребляют термины запрос и объект запроса.

Запрос — это формализованный способ выражения информационных потребностей пользователем системы. Для выражения информационной потребности используется язык поисковых запросов, синтаксис варьируется от системы к системе. Кроме специального языка запросов, современные поисковые системы позволяют вводить запрос на естественном языке.

Объект запроса — это информационная сущность, которая хранится в базе автоматизированной системы поиска. Несмотря на то, что наиболее распространенным объектом запроса является текстовый документ, не существует никаких принципиальных ограничений. В частности, возможен поиск изображений, музыки и другой мультимедиа информации. Процесс занесения объектов поиска в ИПС называется индексацией. Далеко не всегда ИПС хранит точную копию объекта, нередко вместо неё хранится суррогат.

Задачи информационного поиска. Центральная задача ИП — помочь пользователю удовлетворить его информационную потребность. Так как описать информационные потребности пользователя технически непросто, они формулируются как некоторый запрос, представляющий из себя набор ключевых слов, характеризующий то, что ищет пользователь.

Классическая задача информационного поиска, с которой началось развитие этой области, — это поиск документов, удовлетворяющих запросу, в рамках некоторой статической коллекции документов. Но список задач ИП постоянно расширяется и теперь включает: вопросы моделирования; классификация документов; фильтрация документов; кластеризация документов; проектирование архитектур поисковых систем и пользовательских интерфейсов; извлечение информации, в частности аннотирования и реферирования документов; языки запросов и др.

Оценки эффективности информационного поиска. Существует много способов оценить насколько хорошо документы, найденные ИПС, соответствуют запросу. К сожалению, понятие степени соответствия запроса, или другими словами релевантности, является субъективным понятием, а степень соответствия зависит от конкретного человека, оценивающего результаты выполнения запроса.

Точность (Precision) - отношение числа релевантных документов, найденных ИПС, к общему числу найденных документов:

$$\text{Precision} = \frac{|D_{rel} \cap D_{retr}|}{|D_{retr}|},$$

где D_{rel} — это множество релевантных документов в базе, а D_{retr} — множество документов, найденных системой.

Полнота (recall) - отношение числа найденных релевантных документов, к общему числу релевантных документов в базе:

$$\text{Recall} = \frac{|D_{rel} \cap D_{retr}|}{|D_{retr}|},$$

где D_{rel} — это множество релевантных документов в базе, а D_{retr} — множество документов, найденных системой.

Выпадение (fall-out) - выпадение характеризует вероятность нахождения нерелевантного ресурса и определяется, как отношение числа найденных нерелевантных документов к общему числу нерелевантных документов в базе:

$$\text{Fall-out} = \frac{|D_{nrel} \cap D_{retr}|}{|D_{nrel}|},$$

где D_{nrel} — это множество нерелевантных документов в базе, а D_{retr} — множество документов, найденных системой.

F-мера (F-measure, мера Ван Ризбергена) - традиционно F-мера определяется, как гармоническое среднее точности и полноты:

$$F = 2 \times \text{Precision} \times \text{Recall} / (\text{Precision} + \text{Recall}).$$

Часто ее также называют F_1 мерой, потому что точность и полнота присутствуют в этой формуле с одинаковым весом.

Более общая формула для положительного вещественного α имеет вид:

$$F_\alpha = (1 + \alpha) \times \text{Precision} \times \text{Recall} / (\alpha \times \text{Precision} + \text{Recall}).$$

Интеллектуальный анализ данных (Data Mining). Интеллектуальный анализ данных (англ. *Data Mining*) — выявление скрытых закономерностей или взаимосвязей между переменными в больших массивах необработанных данных. Подразделяется на задачи классификации, моделирования и прогнозирования и другие. Термин «Data Mining» введен Григорием Пятецким-Шапиро в 1989 году.

Английский термин «Data Mining» не имеет однозначного перевода на русский язык (добыча данных, вскрытие данных, информационная проходка, извлечение данных/информации) поэтому в большинстве случаев используется в оригинале. Наиболее удачным непрямым переводом считается термин «интеллектуальный анализ данных».

Data Mining включает методы и модели статистического анализа и машинного обучения, дистанцируясь от них в сторону *автоматического* анализа данных. Инструменты Data Mining позволяют проводить анализ данных предметными специалистами (аналитиками), не владеющими соответствующими математическими знаниями.

Задачи, решаемые Data Mining:

1. Классификация — отнесение входного вектора (объекта, события, наблюдения) к одному из заранее известных классов.
2. Кластеризация — разделение множества входных векторов на группы (кластеры) по степени «похожести» друг на друга.
3. Сокращение описания — для визуализации данных, лаконизма моделей, упрощения счета и интерпретации, сжатия объемов собираемой и хранимой информации.
4. Ассоциация — поиск повторяющихся образцов. Например, поиск «устойчивых связей в корзине покупателя» (англ. market basket analysis) — вместе с пивом часто покупают орешки.
5. Прогнозирование
6. Анализ отклонений — Например, выявление нетипичной сетевой активности позволяет обнаружить вредоносные программы.
7. Визуализация

В литературе можно встретить еще ряд классов задач. Базовыми задачами являются первые три. Остальные задачи сводятся к ним тем или иным способом.

Для задач классификации характерно «обучение с учителем», при котором построение (обучение) модели производится по выборке содержащей входные и выходные векторы.

Для задач кластеризации и ассоциации применяется «обучение без учителя», при котором построение модели производится по выборке, в которой нет выходного параметра. Значение выходного параметра («относится к кластеру ...», «похож на вектор ...») подбирается автоматически в процессе обучения.

Для задач сокращения описания характерно *отсутствие разделения на входные и выходные векторы*. Начиная с классических работ К. Пирсона по методу главных компонент, основное внимание здесь уделяется аппроксимации данных.

Можно выделить типичный ряд этапов решения задач методами Data Mining:

1. Формирование гипотезы;
2. Сбор данных;
3. Подготовка данных (фильтрация);
4. Выбор модели;
5. Подбор параметров модели и алгоритма обучения;
6. Обучение модели (автоматический поиск остальных параметров модели);
7. Анализ качества обучения, если неудовлетворительный переход на п. 5 или п. 4;
8. Анализ выявленных закономерностей, если неудовлетворительный переход на п. 1, 4 или 5.

Извлечение знаний в базах данных (Knowledge Discovery in Databases). Извлечение (поиск) знаний в базах данных (Knowledge Discovery in Databases – KDD) - процесс обнаружения полезных знаний в базах данных. Эти знания

могут быть представлены в виде закономерностей, правил, прогнозов, связей между элементами данных и др. Главным инструментом поиска знаний в процессе KDD являются аналитические технологии Data Mining, реализующие задачи классификации, кластеризации, регрессии, прогнозирования, предсказания и т.д.

Однако, в соответствии с концепцией KDD, эффективный процесс поиска знаний не ограничивается их анализом. KDD включает последовательность операций, необходимых для поддержки аналитического процесса. К ним относятся:

- Консолидация данных – процесс их извлечения из различных источников (OLTP-систем, СУБД, файлов отдельных пользователей, Интернета и т.д.) и загрузка в централизованное хранилище данных.
- Подготовка анализируемых выборок данных (в том числе, обучающих), загрузка их из хранилища данных или других источников в аналитическое приложение.
- Очистка данных от факторов, мешающих их корректному анализу, таких, как шумы и аномальные значения, дубликаты, противоречия, пропуски, фиктивные значения и т.д.
- Трансформация – оптимизация данных для решения определенной задачи. Обычно на данном этапе выполняется исключение незначимых факторов, снижения размерности входных данных, нормализация, обогащение и другие преобразования, позволяющие лучше «приспособить» данные к решению аналитической задачи.
- Анализ данных – применение методов и технологий Data Mining: построение и обучение моделей (нейронных сетей, деревьев решений, карт Кохонена и др.), решение задач классификации и регрессии, кластеризации, прогнозирования, поиска ассоциаций и т.д.
- Интерпретация и визуализация результатов анализа, их применение в бизнес-приложениях.

Knowledge Discovery in Databases не задает набор методов обработки или пригодные для анализа алгоритмы, он определяет последовательность действий, которую необходимо выполнить для того, чтобы из исходных данных получить знания. Данный подход универсальный и не зависит от предметной области, что является его несомненным достоинством.

Основоположниками концепции KDD считаются Григорий Пятецкий-Шапиро (Gregory Piatetsky-Shapiro) и Усама Файад (Usama Fayyad).

Рассуждение на основе прецедентов (Case-Based Reasoning). Прецедент - случай, имевший место ранее и служащий примером или оправданием для последующих случаев подобного рода. Вывод на основе прецедентов (CBR – Case-Based Reasoning) является подходом, позволяющим

решить новую задачу, используя или адаптируя решение уже известной задачи. Как правило, такие методы рассуждений включают в себя четыре основных этапа, образующие так называемый цикл рассуждения на основе прецедентов или CBR-цикл.

К преимуществам рассуждений на основе прецедентов можно отнести следующие аспекты:

- Возможность напрямую использовать опыт, накопленный системой без интенсивного привлечения эксперта в той или иной предметной области;
- Возможность сокращения времени поиска решения поставленной задачи за счет использования уже имеющегося решения для подобной задачи;
- Существует возможность исключить повторное получение ошибочного решения;
- Отсутствует необходимость полного и углубленного рассмотрения знаний о конкретной предметной области;
- Возможно применение эвристик, повышающих эффективность решения задач.
- К недостаткам рассуждений на основе прецедентов можно отнести следующее:
- При описании прецедентов обычно ограничиваются поверхностными знаниями о предметной области;
- Большое количество прецедентов может привести к снижению производительности системы;
- Проблематичным является определение критериев для индексации и сравнения прецедентов;
- Проблемы с отладкой алгоритмов определения подобных (аналогичных) прецедентов;
- Невозможность получения решения задач, для которых нет прецедентов или степень их сходства (подобия) меньше заданного порогового значения.

Основная цель использования аппарата прецедентов в рамках СППР и, в частности, систем экспертной диагностики сложных объектов, заключается в выдаче готового решения ЛПР для текущей ситуации на основе прецедентов, которые уже имели место в прошлом при управлении данным объектом или процессом.

На первом этапе CBR-цикла выполняется определение степени сходства текущей ситуации с прецедентами из библиотеки прецедентов системы и последующее их извлечение с целью разрешить новую проблемную ситуацию, сложившуюся на объекте.

Методы извлечения прецедентов. Для успешной реализации рассуждений на основе прецедентов необходимо обеспечить корректное извлечение прецедентов из библиотеки прецедентов (БП) системы

экспертной диагностики. Выбор метода извлечения прецедентов напрямую связан со способом представления прецедентов и соответственно со способом организации БП.

Основные способы представления прецедентов можно разделить на следующие группы:

- параметрические;
- объектно-ориентированные;
- специальные (в виде деревьев, графов, логических формул и т.д.).

В большинстве случаев для представления прецедентов достаточно простого параметрического представления, т.е. представления прецедента в виде набора параметров с конкретными значениями и решения (диагноз по проблемной ситуации и рекомендации ЛППР):

$CASE(x_1, x_2, \dots, x_n, R)$,

где $x_1 \dots x_n$ – параметры ситуации, описывающей данный прецедент ($x_1 \in X_1, x_2 \in X_2, \dots, x_n \in X_n$), R – диагноз и рекомендации ЛППР, n – количество параметров прецедента, а X_1, \dots, X_n – области допустимых значений соответствующих параметров прецедента.

Существует целый ряд методов извлечения прецедентов и их модификаций:

1) *Метод ближайшего соседа (NN – Nearest Neighbor)*. Это самый распространенный метод сравнения и извлечения прецедентов. Он позволяет довольно легко вычислить степень сходства текущей проблемной ситуации и прецедентов из БП системы. С целью определения степени сходства на множестве параметров, используемых для описания прецедентов и текущей ситуации, вводится определенная метрика. Далее в соответствии с выбранной метрикой определяется расстояние от целевой точки, соответствующей текущей проблемной ситуации, до точек, представляющих прецеденты из БП, и выбирается ближайшая к целевой точка.

Метод определения ближайшего соседа (ближайших соседей) также широко применяется для решения задач классификации, кластеризации, регрессии и распознавания образов [Башмаков и др., 2005].

Основными преимуществами данного метода являются простота реализации и универсальность в смысле независимости от специфики конкретной проблемной области. К существенным недостаткам метода можно отнести сложность выбора метрики для определения степени сходства и прямую зависимость требуемых вычислительных ресурсов от размера БП, а также неэффективность при работе с неполными и зашумленными исходными данными.

На практике применяются различные модификации указанного метода [Варшавский и др., 2006]. Обычно решение выбирается на основе нескольких ближайших точек (соседей), а не одной (метод k ближайших соседей). Возможно использование метода ближайшего соседа, основанного на знаниях о предметной области (определенных зависимостях между параметрами объекта).

2) *Метод извлечения прецедентов на основе деревьев решений*. Этот метод предполагает нахождение требуемых прецедентов путем разрешения вершин дерева решений. Каждая вершина дерева указывает, по какой ее ветви следует осуществлять дальнейший поиск решения. Выбор ветви осуществляется на основе информации о текущей проблемной ситуации. Таким образом, необходимо добраться до концевой вершины, которая соответствует одному или нескольким прецедентам. Если концевая вершина связана с некоторым подмножеством прецедентов, то тогда для выбора наиболее подходящего из них может использоваться метод ближайшего соседа. Такой подход рекомендуется применять для больших БП, т.к. основная часть работы по извлечению прецедентов выполняется заранее на этапе построения дерева решений, что значительно сокращает время поиска решения.

3) *Метод извлечения прецедентов на основе знаний*. В отличие от методов, описанных выше, данный метод позволяет учесть знания экспертов (ЛППР) по конкретной предметной области (коэффициенты важности параметров, выявленные зависимости и т.д.) при извлечении. Метод может успешно применяться совместно с другими методами извлечения прецедентов, особенно когда БП имеет большие размеры и предметная область является открытой и динамической.

4) *Метод извлечения с учетом применимости прецедентов*. В большинстве систем, использующих механизмы рассуждений на основе прецедентов, предполагается, что наиболее схожие с текущей проблемной ситуацией прецеденты являются наиболее применимыми в этой ситуации. Однако это не всегда так. В основе понятия извлечения на основе применимости (адаптируемости) лежит то, что извлечение прецедентов базируется не только на их сходстве с текущей проблемной ситуацией, но и на том, насколько хорошую для желаемого результата модель они собой представляют. Т.е. на выбор извлекаемых прецедентов влияет возможность их применения в конкретной ситуации. В некоторых системах эта проблема решается путем сохранения прецедентов вместе с комментариями по их применению. Использование указанного подхода позволяет сделать поиск решения более эффективным, заранее отбрасывая часть заведомо неперспективных прецедентов.

Помимо рассмотренных методов для извлечения прецедентов могут успешно применяться и другие методы (например, аппарат искусственных нейронных сетей).

Имитационное моделирование — это метод, позволяющий строить модели, описывающие процессы так, как они проходили бы в действительности. Такую модель можно «проиграть» во времени как для одного испытания, так и заданного их множества. При этом результаты будут определяться случайным характером процессов. По этим данным можно получить достаточно устойчивую статистику.

Имитационное моделирование — это метод исследования, при котором изучаемая система заменяется моделью с достаточной точностью

описывающей реальную систему и с ней проводятся эксперименты с целью получения информации об этой системе. Экспериментирование с моделью называют имитацией (имитация — это постижение сути явления, не прибегая к экспериментам на реальном объекте).

Имитационное моделирование — это частный случай математического моделирования. Существует класс объектов, для которых по различным причинам не разработаны аналитические модели, либо не разработаны методы решения полученной модели. В этом случае математическая модель заменяется имитатором или имитационной моделью.

Имитационная модель — логико-математическое описание объекта, которое может быть использовано для экспериментирования на компьютере в целях проектирования, анализа и оценки функционирования объекта.

К имитационному моделированию прибегают, когда:

- дорого или невозможно экспериментировать на реальном объекте;
- невозможно построить аналитическую модель: в системе есть время, причинные связи, последствие, нелинейности, стохастические (случайные) переменные;
- необходимо симитировать поведение системы во времени.

Цель имитационного моделирования состоит в воспроизведении поведения исследуемой системы на основе результатов анализа наиболее существенных взаимосвязей между ее элементами или другими словами — разработке симулятора (английский термин — *simulation modeling*) исследуемой предметной области для проведения различных экспериментов.

Имитационное моделирование позволяет имитировать поведение системы, во времени. Причём плюсом является то, что временем в модели можно управлять: замедлять в случае с быстропротекающими процессами и ускорять для моделирования систем с медленной изменчивостью. Можно имитировать поведение тех объектов, реальные эксперименты с которыми дороги, невозможны или опасны.

Имитация, как метод решения нетривиальных задач, получила начальное развитие в связи с созданием ЭВМ в 1950х — 1960х годах.

Можно выделить две разновидности имитации:

- Метод Монте-Карло (метод статистических испытаний);
- Метод имитационного моделирования (статистическое моделирование).

Виды имитационного моделирования. Агентное моделирование — относительно новое (1990е-2000е гг.) направление в имитационном моделировании, которое используется для исследования децентрализованных систем, динамика функционирования которых определяется не глобальными правилами и законами (как в других парадигмах моделирования), а наоборот. Когда эти глобальные правила и законы являются результатом индивидуальной активности членов группы. Цель агентных моделей —

получить представление об этих глобальных правилах, общем поведении системы, исходя из предположений об индивидуальном, частном поведении ее отдельных активных объектов и взаимодействиях этих объектов в системе. Агент — некая сущность, обладающая активностью, автономным поведением, может принимать решения в соответствии с некоторым набором правил, взаимодействовать с окружением, а также самостоятельно изменяться.

Дискретно-событийное моделирование — подход к моделированию, предлагающий абстрагироваться от непрерывной природы событий и рассматривать только основные *события* моделируемой системы, такие как: «ожидание», «обработка заказа», «движение с грузом», «разгрузка» и другие. Дискретно-событийное моделирование наиболее развито и имеет огромную сферу приложений — от логистики и систем массового обслуживания до транспортных и производственных систем. Этот вид моделирования наиболее подходит для моделирования производственных процессов. Основан Джеффри Гордоном в 1960х годах.

Системная динамика — парадигма моделирования, где для исследуемой системы строятся графические диаграммы причинных связей и глобальных влияний одних параметров на другие во времени, а затем созданная на основе этих диаграмм модель имитируется на компьютере. По сути, такой вид *моделирования* более всех других парадигм помогает понять суть происходящего выявления причинно-следственных связей между объектами и явлениями. С помощью системной динамики строят модели бизнес-процессов, развития города, модели производства, динамики популяции, экологии и развития эпидемии. Метод основан Форрестером в 1950 годах.

Популярные системы имитационного моделирования: AnyLogic; Arena; eM-Plant; Powersim; GPSS.

Генетические алгоритмы (genetic algorithm). Генетический алгоритм (англ. *genetic algorithm*) — это эвристический алгоритм поиска, используемый для решения задач оптимизации и моделирования путем последовательного подбора, комбинирования и вариации искомым параметров с использованием механизмов, напоминающих биологическую эволюцию. Является разновидностью эволюционных вычислений (англ. *evolutionary computation*). Отличительной особенностью генетического алгоритма является акцент на использование оператора «скрещивания», который производит операцию рекомбинации решений-кандидатов, роль которой аналогична роли скрещивания в живой природе. «Отцом-основателем» генетических алгоритмов считается Джон Холланд (англ. *John Holland*), книга которого «Адаптация в естественных и искусственных системах» (англ. *Adaptation in Natural and Artificial Systems*) является основополагающим трудом в этой области исследований.

Описание алгоритма. Задача кодируется таким образом, чтобы её решение могло быть представлено в виде вектора («хромосома»). Случайным образом создаётся некоторое количество начальных векторов («начальная популяция»). Они оцениваются с использованием «функции

приспособленности», в результате чего каждому вектору присваивается определенное значение («приспособленность»), которое определяет вероятность выживания организма, представленного данным вектором. После этого с использованием полученных значений приспособленности выбираются вектора (*селекция*), допущенные к «скрещиванию». К этим векторам применяются «генетические операторы» (в большинстве случаев «скрещивание» - crossover и «мутация» - mutation), создавая таким образом следующее «поколение». Особи следующего поколения также оцениваются, затем производится селекция, применяются генетические операторы и т. д. Так моделируется «эволюционный процесс», продолжающийся несколько жизненных циклов (*поколений*), пока не будет выполнен критерий остановки алгоритма. Таким критерием может быть:

- нахождение глобального, либо субоптимального решения;
- исчерпание числа поколений, отпущенных на эволюцию;
- исчерпание времени, отпущенного на эволюцию.

Генетические алгоритмы служат, главным образом, для поиска решений в очень больших, сложных пространствах поиска.

Таким образом, можно выделить следующие этапы генетического алгоритма:

1. Создание начальной популяции
2. Определение (задание) функций приспособленности для особей популяции (оценивание)
(Начало цикла)
 1. Выбор индивидов из текущей популяции (селекция)
 2. Скрещивание и/или мутация
 3. Вычисление функций приспособленности для всех особей
 4. Формирование нового поколения
 5. Если выполняются условия останова, то (конец цикла), иначе (начало цикла).

Создание начальной популяции. Перед первым шагом нужно случайным образом создать некую начальную популяцию; даже если она окажется совершенно неконкурентоспособной, генетический алгоритм все равно достаточно быстро переведет ее в жизнеспособную популяцию. Таким образом, на первом шаге можно особенно не стараться сделать слишком уж приспособленных особей, достаточно, чтобы они соответствовали формату особей популяции, и на них можно было подсчитать функцию приспособленности (Fitness). Итогом первого шага является популяция N , состоящая из N особей.

Отбор. На этапе отбора нужно из всей популяции выбрать определенную ее долю, которая останется "в живых" на этом этапе эволюции. Есть разные способы проводить отбор. Вероятность выживания особи h должна зависеть от значения функции приспособленности $Fitness(h)$. Сама доля выживших s обычно является параметром генетического алгоритма, и ее просто задают заранее. По итогам отбора из N особей

популяции N должны остаться sN особей, которые войдут в итоговую популяцию N' . Остальные особи погибают.

Размножение. Размножение в генетических алгоритмах обычно половое - чтобы произвести потомка, нужны несколько родителей; обычно, конечно, нужны ровно два. Размножение в разных алгоритмах определяется по-разному - оно, конечно, зависит от представления данных. Главное требование к размножению - чтобы потомок или потомки имели возможность унаследовать черты обоих родителей, "смешав" их каким-либо достаточно разумным способом. Вообще говоря, для того чтобы провести операцию размножения, нужно выбрать $(1-s)p/2$ пар гипотез из N и провести с ними размножение, получив по два потомка от каждой пары (если размножение определено так, чтобы давать одного потомка, нужно выбрать $(1-s)p$ пар), и добавить этих потомков в N' . В результате N' будет состоять из N особей. Почему особи для размножения обычно выбираются из всей популяции N , а не из выживших на первом шаге элементов N_0 (хотя последний вариант тоже имеет право на существование)? Дело в том, что главный бич многих генетических алгоритмов - недостаток разнообразия (diversity) в особях. Достаточно быстро выделяется один-единственный генотип, который представляет собой локальный максимум, а затем все элементы популяции проигрывают ему отбор, и вся популяция "забывается" копиями этой особи. Есть разные способы борьбы с таким нежелательным эффектом; один из них - выбор для размножения не самых приспособленных, но вообще всех особей.

Мутации. К мутациям относится все то же самое, что и к размножению: есть некоторая доля мутантов m , являющаяся параметром генетического алгоритма, и на шаге мутаций нужно выбрать mN особей, а затем изменить их в соответствии с заранее определенными операциями мутации.

Искусственные нейронные сети (ИНС) — математические модели, а также их программные или аппаратные реализации, построенные по принципу организации и функционирования биологических нейронных сетей — сетей нервных клеток живого организма. Это понятие возникло при изучении процессов, протекающих в мозге при мышлении, и при попытке смоделировать эти процессы. Первой такой моделью мозга был перцептрон. Впоследствии эти модели стали использовать в практических целях, как правило, в задачах прогнозирования.

С точки зрения *машинного обучения*, нейронная сеть представляет собой частный случай методов распознавания образов, дискриминантного анализа, методов кластеризации и т. п. С *математической* точки зрения обучение нейронных сетей, это многопараметрическая задача нелинейной оптимизации. С точки зрения *кибернетики*, нейронная сеть используется в задачах адаптивного управления и как алгоритмы для робототехники. С точки зрения *развития вычислительной техники и программирования*, нейронная сеть — способ решения проблемы эффективного параллелизма. А

с точки зрения *искусственного интеллекта*, ИНС является основой философского течения коннективизма и основным направлением в структурном подходе по изучению возможности построения (моделирования) естественного интеллекта с помощью компьютерных алгоритмов.

Задача принятия решений и управление близка к задаче классификации. Классификации подлежат ситуации, характеристики которых поступают на вход нейронной сети. На выходе сети при этом должен появиться признак решения, которое она приняла. При этом в качестве входных сигналов используются различные критерии описания состояния управляемой системы.

Сами ИНС представляют собой систему соединённых и взаимодействующих между собой простых процессоров (искусственных нейронов). Такие процессоры обычно довольно просты, особенно в сравнении с процессорами, используемыми в персональных компьютерах. Каждый процессор подобной сети имеет дело только с сигналами, которые он периодически получает, и сигналами, которые он периодически посылает другим процессорам. И тем не менее, будучи соединёнными в достаточно большую сеть с управляемым взаимодействием, такие локально простые процессоры вместе способны выполнять довольно сложные задачи.

Нейронные сети не программируются в привычном смысле этого слова, они *обучаются*. Возможность обучения — одно из главных преимуществ нейронных сетей перед традиционными алгоритмами. Технически обучение заключается в нахождении коэффициентов связей между нейронами. В процессе обучения нейронная сеть способна выявлять сложные зависимости между входными данными и выходными, а также выполнять обобщение. Это значит, что, в случае успешного обучения, сеть сможет вернуть верный результат на основании данных, которые отсутствовали в обучающей выборке.

Этапы решения задач:

1. сбор данных для обучения;
2. подготовка и нормализация данных;
3. выбор топологии сети;
4. экспериментальный подбор характеристик сети;
5. экспериментальный подбор параметров обучения;
6. собственно обучение;
7. проверка адекватности обучения;
8. корректировка параметров, окончательное обучение;
9. вербализация сети с целью дальнейшего использования.

Следует рассмотреть подробнее некоторые из этих этапов.

Классификация ИНС. Классификация по *типу входной информации*: аналоговые нейронные сети (используют информацию в форме действительных чисел); двоичные нейронные сети (оперируют с информацией, представленной в двоичном виде).

Классификация по *характеру обучения*: обучение с учителем — выходное пространство решений нейронной сети известно; обучение без учителя — нейронная сеть формирует выходное пространство решений только на основе входных воздействий (такие сети называют самоорганизующимися); обучение с подкреплением — система назначения штрафов и поощрений от среды.

Классификация по *характеру настройки синапсов*: сети с фиксированными связями (весовые коэффициенты нейронной сети выбираются сразу, исходя из условий задачи, при этом: $dW/dt = 0$, где W — весовые коэффициенты сети); сети с динамическими связями (для них в процессе обучения происходит настройка синаптических связей, то есть $dW/dt \neq 0$, где W — весовые коэффициенты сети).

Классификация по *времени передачи сигнала*. В ряде нейронных сетей активирующая функция может зависеть не только от весовых коэффициентов связей w_{ij} , но и от времени передачи импульса (сигнала) по каналам связи τ_{ij} . По этому в общем виде активирующая (передающая) функция связи c_{ij} от элемента u_i к элементу u_j имеет вид: $c_{ij}^* = f[w_{ij}(t), u_i^*(t - \tau_{ij})]$. Тогда *синхронной* сетью называют такую сеть, у которой время передачи τ_{ij} каждой связи равно либо нулю, либо фиксированной постоянной τ . *Асинхронной* называют такую сеть, у которой время передачи τ_{ij} для каждой связи между элементами u_i и u_j свое, но тоже постоянное.

Классификация по *характеру связей*:

1. Сети прямого распространения (Feedforward).

Все связи направлены строго от входных нейронов к выходным (примерами таких сетей являются перцептрон Розенблатта, многослойный перцептрон, сети Ворда).

2. Рекуррентные нейронные сети

Сигнал с выходных нейронов или нейронов скрытого слоя частично передается обратно на входы нейронов входного слоя (обратная связь). Рекуррентная сеть Хопфилда «фильтрует» входные данные, возвращаясь к устойчивому состоянию и, таким образом, позволяет решать задачи компрессии данных и построения ассоциативной памяти. Частным случаем рекуррентных сетей является двунаправленные сети. В таких сетях между слоями существуют связи как в направлении от входного слоя к выходному, так и в обратном. Классическим примером является Нейронная сеть Коско.

3. Радиально-базисные функции

Искусственные нейронные сети использующие в качестве активационных функций радиально-базисные (такие сети сокращенно называются RBF-сетями). Общий вид радиально-базисной функции:

$$f(x) = \phi\left(\frac{x^2}{\sigma^2}\right), \text{ например, } f(x) = e^{-\frac{x^2}{\sigma^2}},$$

где x — вектор входных сигналов нейрона, σ — ширина окна функции, $\phi(y)$ — убывающая функция (чаще всего, равная нулю вне некоторого отрезка).

Радиально-базисная сеть характеризуется тремя особенностями: 1. Единственный скрытый слой; 2. Только нейроны скрытого слоя имеют нелинейную активационную функцию; 3. Синаптические веса связей входного и скрытого слоев равны единице.

4. Самоорганизующиеся карты

Такие сети представляют собой, соревновательную нейронную сеть с обучением без учителя, выполняющую задачу визуализации и кластеризации. Эти сети являются методом проецирования многомерного пространства в пространство с более низкой размерностью (чаще всего, двумерное), применяются также для решения задач моделирования, прогнозирования и др. Они являются одной из версий нейронных сетей Кохонена. Самоорганизующиеся карты Кохонена служат, в первую очередь, для визуализации и первоначального («разведывательного») анализа данных. Сигнал в сеть Кохонена поступает сразу на все нейроны, веса соответствующих синапсов интерпретируются как координаты положения узла, и выходной сигнал формируется по принципу «победитель забирает все» — то есть ненулевой выходной сигнал имеет нейрон, ближайший (в смысле весов синапсов) к подаваемому на вход объекту. В процессе обучения веса синапсов настраиваются таким образом, чтобы узлы решетки «располагались» в местах локальных сгущений данных, то есть описывали кластерную структуру облака данных, с другой стороны, связи между нейронами соответствуют отношениям соседства между соответствующими кластерами в пространстве признаков.

Удобно рассматривать такие карты как двумерные сетки узлов, размещенных в многомерном пространстве. Изначально самоорганизующаяся карта представляет из себя сетку из узлов, соединенный между собой связями. Кохонен рассматривал два варианта соединения узлов — в прямоугольную и гексагональную сетку — отличие состоит в том, что в прямоугольной сетке каждый узел соединен с 4-мя соседними, а в гексагональной — с 6-ю ближайшими узлами. Для двух таких сеток процесс построения сети Кохонена отличается лишь в том месте, где перебираются ближайшие к данному узлу соседи.

Известные типы сетей: Перцептрон Розенблатта; Многослойный перцептрон; Сеть Джордана; Сеть Элмана; Сеть Хэмминга; Сеть Ворда; Сеть Хопфилда; Сеть Кохонена; Когнитрон; Неокогнитрон; Хаотическая нейронная сеть; Осцилляторная нейронная сеть; Сеть встречного распространения; Сеть радиальных базисных функций (RBF-сеть); Сеть обобщенной регрессии; Вероятностная сеть; Сиамская нейронная сеть; Сети адаптивного резонанса.

Методы искусственного интеллекта (Artificial intelligence). Искусственный интеллект (ИИ) (англ. *Artificial intelligence, AI*) — это наука о создании интеллектуальных машин и систем, особенно интеллектуальных компьютерных программ, направленная на то, чтобы понять человеческий интеллект. При этом применяемые методы не обязательно должны быть

биологически правдоподобны. Проблема состоит в том, что неизвестно какие вычислительные процедуры мы хотим называть интеллектуальными. А так как мы понимаем только некоторые механизмы интеллекта, то под интеллектом в пределах этой науки мы понимаем только вычислительную часть способности достигнуть цели в мире.

Различные виды и степени интеллекта существуют у многих людей, животных и некоторых машин, интеллектуальных информационных систем и различных моделях экспертных систем с различными базами знаний. При этом как видим такое определение интеллекта не связано с пониманием интеллекта у человека — это разные вещи. Более того, эта наука моделирует человеческий интеллект, так как с одной стороны, можно изучить кое-что о том, как заставить машины решить проблемы, наблюдая других людей, а с другой стороны, большинство работ в ИИ вовлекают изучение проблем, которые требуется решать человечеству в промышленном и технологическом смысле. Поэтому исследователи ИИ свободны использовать методы, которые не наблюдаются у людей, если это необходимо для решения конкретных проблем.

Именно в таком смысле термин ввел Джон Маккарти в 1956 году на конференции в Дартмутском университете, и до сих пор, несмотря на критику тех, кто считает, что интеллект — это только биологический феномен, в научной среде термин сохранил свой первоначальный смысл, несмотря на явные противоречия с точки зрения человеческого интеллекта.

В философии не решён вопрос о природе и статусе человеческого интеллекта. Нет и точного критерия достижения компьютерами «разумности», хотя на заре искусственного интеллекта был предложен ряд гипотез, например, тест Тьюринга или гипотеза Ньюэлла — Саймона. Поэтому, несмотря на наличие множества подходов, как к пониманию задач ИИ, так и созданию интеллектуальных информационных систем можно выделить два основных подхода к разработке ИИ:

- нисходящий, семиотический — создание экспертных систем, баз знаний и систем логического вывода, имитирующие высокоуровневые психические процессы: мышление, рассуждение, речь, эмоции, творчество и т. д.;
- восходящий, биологический — изучение нейронных сетей и эволюционных вычислений, моделирующих интеллектуальное поведение на основе более мелких «неинтеллектуальных» элементов.

Причем последний подход, как правило, является критикой первого подхода, а сам скорее не относится к науке о ИИ в смысле данном Джоном Маккарти — их объединяет только общая конечная цель.

Подходы к определению искусственного интеллекта. Единого ответа на вопрос чем занимается искусственный интеллект, не существует. Почти каждый автор, пишущий книгу об ИИ, отталкивается в ней от какого-либо определения, рассматривая в его свете достижения этой науки. Обычно эти

определения сводятся к следующим: тест Тьюринга; когнитивное моделирование; логический подход; агентно-ориентированный подход.

Кроме перечисленных существуют еще интуитивные подходы:

- Самый общий подход предполагает, что ИИ будет способен проявлять поведение, не отличающееся от человеческого, причём, в нормальных ситуациях. Эта идея является обобщением подхода теста Тьюринга, который утверждает, что машина станет разумной тогда, когда будет способна поддерживать разговор с обычным человеком, и тот не сможет понять, что говорит с машиной (разговор идёт по переписке).
- Писатели-фантасты часто предлагают ещё один подход: ИИ возникнет тогда, когда машина будет способна чувствовать и творить. Так, хозяин Эндрю Мартина из «Двухсотлетнего человека» начинает относиться к нему как к человеку, когда тот создаёт игрушку по собственному проекту. А Дейта из Звёздного пути, будучи способным к коммуникации и научению, мечтает обрести эмоции и интуицию.

Если в основе работы СППР лежат методы искусственного интеллекта, то говорят об *интеллектуальной СППР*, или ИСППР.

Проблемы, решаемые с помощью методов ИИ:

1. Информационное обеспечение технологии смысловой обработки данных. Включает разработку компонент: информационного обеспечения сетевого взаимодействия; хранения и использования баз данных и знаний; интерпретации семантики текстов сообщений и документов; содержательного поиска текстовых документов и сообщений по запросам на естественном языке; осмысленного сетевого обмена.

2. Коннект-анализ потоков электронных сообщений корпоративных пользователей. Включает средства: быстрого составления картины связей между участниками обмена электронными сообщениями; выявления взаимосвязи между деятельностью субъектов электронных сообщений и конкретными ситуациями (экономическими, политическими, социальными, технологическими); распознавания устойчивых системно-организованных структур и деловых созвездий и оценивания роли отдельных субъектов в их деятельности.

3. Взаимодействие «человек-машина» на естественном языке. Включает: построение концептуальных моделей понятийной области пользователя, с применением семантических сетей и методов онтологий; создание интерфейса и интерактивной технологии взаимодействия экспертов и топ-менеджеров.

4. Контент-анализ предметной области. Включает применение инструментария онтологий и семантических сетей, обеспечивающих: формальное и концептуальное описание терминов предметной области и отношений между ними; построение онтологий в сетевых структурах; определение общего и персонального словарей участников информационного

обмена; совместное использование людьми и программными интеллектуальными агентами общего понимания структуры информации; повторное использование знаний в предметной области; автоматизированный анализ знаний в предметной области.

5. Представление и управление знаниями. Включает: возможность ИС аккумулировать в логическом виде знания экспертов в различных предметных областях; автоматизированное извлечение знаний из нормативных документов, учебников, справочников, отчетов исследований; автоматизированный синтез советов; выполнение классификации; представление знаний в логической форме для автоматической обработки; объяснение пользователю решения задачи.

7.3. Управление на базе ситуационных центров

Чтобы дать определение «ситуационный центр», необходимо предварительно разобраться с понятием *ситуация*. Само это слово используется повседневно в самых разных смыслах, порой неотделимых от таких понятий, как состояние, событие, процесс, положение и т. д.

Ситуация - совокупность обстоятельств, положение, обстановка²⁴

Ситуация — одноактность и неповторимость наступления множества событий, стечения всех жизненных обстоятельств и положений, открывающихся восприятию и деятельности человека²⁵

Ситуация есть принуждение к принятию решения, свобода же состоит в выборе решения²⁶

Ситуация — это то, что создает систему и предшествует ей, а также то, что определяет состояние системы и наличествует в ней, кроме того, это то, что приводит к распаду системы или ее преобразованию. *Ситуации* — "кокон", который "обволакивает" системы²⁷

Проблемная ситуация (от греч. *problema* — задача, задание и лат. *situatio* — положение) — 1) содержащее противоречие и не имеющее однозначного решения соотношение обстоятельств и условий, в которых разворачивается деятельность индивида или группы; 2) психологическая модель условий порождения мышления на основе ситуативно возникающей познавательной потребности, форма связи субъекта с объектом познания. Проблемная ситуация характеризует взаимодействие субъекта и его окружения, а также психическое состояние познающей личности, включенной в объективную и противоречивую по своему содержанию среду²⁸.

Проблемная ситуация — осознание, возникающее при выполнении практического или теоретического задания, того, что ранее усвоенных

²⁴ из словаря «Словарь Ожегова»

²⁵ <http://ru.wikipedia.org/wiki/Ситуация>

²⁶ Жан Поль Сартр. «Человек обречен на свободу»

²⁷ Солодухо Н.М. Манифест ситуационного движения // Ситуационные исследования. Выпуск 1: Ситуационный подход. Казань, 2005.

²⁸ Психологический словарь <http://psychology.net.ru/dictionaries/psy.html?word=712>