

Проектирование приложений информационных системах

Лекция №2

Тема: Постановка задачи обработки информации. Основные виды, алгоритмы и процедуры обработки информации

Обработка информации состоит в получении одних «информационных объектов» из других «информационных объектов» путем выполнения некоторых алгоритмов и является одной из основных операций, осуществляемых над информацией, и главным средством увеличения ее объема и разнообразия.

На самом верхнем уровне можно выделить числовую и нечисловую обработку. В указанные виды обработки вкладывается различная трактовка содержания понятия «данные». При числовой обработке используются такие объекты, как переменные, векторы, матрицы, многомерные массивы, константы и т.д. При нечисловой обработке объектами могут быть файлы, записи, поля, иерархии, сети, отношения и т.д. Другое отличие заключается в том, что при числовой обработке содержание данных не имеет большого значения, в то время как при нечисловой обработке нас интересуют непосредственные сведения об объектах, а не их совокупность в целом.

С точки зрения реализации на основе современных достижений вычислительной техники выделяют следующие виды обработки информации:

- последовательная обработка, применяемая в традиционной фоннеймановской архитектуре ЭВМ, располагающей одним процессором;
- параллельная обработка, применяемая при наличии нескольких процессоров в ЭВМ;
- конвейерная обработка, связанная с использованием в архитектуре ЭВМ одних и тех же ресурсов для решения разных задач. Причем если эти задачи аналогичны, но не тождественны, то это последовательный конвейер, если задачи одинаковые – векторный конвейер.

Принято относить существующие архитектуры ЭВМ с точки зрения обработки информации к одному из следующих классов.

1. Архитектуры с одиночным потоком команд и данных (SISD). К этому классу относятся традиционные фоннеймановские однопроцессорные системы, где имеется центральный процессор, работающий с парами «атрибут - значение».

2. Архитектуры с одиночными потоками команд и данных (SIMD). Особенностью данного класса является наличие одного (центрального) контроллера, управляющего рядом одинаковых процессоров. В зависимости от возможностей контроллера и процессорных элементов, числа процессоров, организации режима поиска и характеристик маршрутных и выравнивающих сетей выделяют:

- матричные процессоры, используемые для решения векторных и матричных задач;
- ассоциативные процессоры, применяемые для решения нечисловых задач и использующие память, в которой можно обращаться непосредственно к информации, хранящейся в ней;
- процессорные ансамбли, применяемые для числовой и нечисловой обработки;
- конвейерные и векторные процессоры.

3. Архитектуры с множественным потоком команд и одиночным потоком данных (MISD). К этому классу могут быть отнесены конвейерные процессоры.

4. Архитектуры с множественным потоком команд и множественным потоком данных (MIMD). К этому классу могут быть отнесены следующие конфигурации:

- мультипроцессорные системы;
- системы с мульти обработкой;
- вычислительные системы из многих машин;
- вычислительные сети.

Основные процедуры обработки данных представлены на рисунке 4.?????

Создание данных, как процесс обработки, предусматривает их образование в результате выполнения некоторого алгоритма и дальнейшее использование для преобразований на более высоком уровне.

Модификация данных связана с отображением изменений в реальной предметной области, осуществляемых путем включения новых данных и удаления ненужных.

Контроль, безопасность и целостность направлены на адекватное отображение реального состояния предметной области в информационной модели и обеспечивают защиту информации от несанкционированного доступа (безопасность) и от сбоев и повреждений технических и программных средств.

Поиск информации, хранимой в памяти компьютера, осуществляется как самостоятельное действие при выполнении ответов на различные запросы и как вспомогательная операция при обработке информации.

Поддержка принятия решения является наиболее важным действием, выполняемым при обработке информации. Широкая альтернатива принимаемых решений приводит к необходимости использования разнообразных математических моделей.

Создание документов, сводок, отчетов заключается в преобразовании информации в формы, пригодные для чтения как человеком, так и компьютером. С этим действием связаны и такие операции, как обработка, считывание, сканирование и сортировка документов.

При преобразовании информации осуществляется ее перевод из одной формы представления или существования в другую, что определяется потребностями, возникающими в процессе реализации информационных технологий.

Реализация всех действий, выполняемых в процессе обработки информации, осуществляется с помощью разнообразных программных средств.

Наиболее распространенной областью применения технологической операции обработки информации является принятие решений.

К зависимости от степени информированности о состоянии управляемого процесса, полноты и точности моделей объекта и системы управления, взаимодействия с окружающей средой, процесс принятия решения протекает в различных условиях:

Принятие решений в условиях определенности. В этой задаче модели объекта и системы управления считаются заданными, а влияние внешней среды - несущественным. Поэтому между выбранной стратегией использования ресурсов и конечным результатом существует однозначная связь, откуда следует, что в условиях определенности достаточно использовать решающее правило для оценки полезности вариантов решений, принимая в качестве оптимального то, которое приводит к наибольшему эффекту. Если таких стратегий несколько, то все они считаются эквивалентными. Для поиска решений в условиях определенности используют методы математического программирования.

Принятие решений в условиях риска. В отличие от предыдущего случая для принятия решений в условиях риска необходимо учитывать влияние внешней среды, которое не поддается точному прогнозу, а известно только вероятностное распределение ее состояний. В этих условиях использование одной и той же стратегии может привести к различным исходам, вероятности появления которых считаются заданными или могут быть определены. Оценку и выбор стратегий проводят с помощью решающего правила, учитывающего вероятность достижения конечного результата.

Принятие решений в условиях неопределенности. Как и в предыдущей задаче между выбором стратегии и конечным результатом отсутствует однозначная связь. Кроме того, неизвестны также значения вероятностей появления конечных результатов, которые либо не могут быть определены, либо не имеют в контексте содержательного смысла. Каждой паре «стратегия - конечный результат» соответствует некоторая внешняя оценка в виде выигрыша. Наиболее распространенным является использование критерия получения максимального гарантированного выигрыша.

Принятие решений в условиях многокритериальности. В любой из перечисленных выше задач многокритериальность возникает в случае наличия нескольких самостоятельных, не сводимых одна к другой целей. Наличие большого числа решений

усложняет оценку и выбор оптимальной стратегии. Одним из возможных путей решения является использование методов моделирования.

Решение задач с помощью искусственного интеллекта заключается в сокращении перебора вариантов при поиске решения, при этом программы реализуют те же принципы, которыми пользуется в процессе мышления человек.

Экспертная система пользуется знаниями, которыми она обладает в своей узкой области, чтобы ограничить поиск на пути к решению задачи путем постепенного сужения круга вариантов.

Для решения задач в экспертных системах используют:

- метод логического вывода, основанный на технике доказательств, называемой резолюцией и использующей опровержение отрицания (доказательство «от противного»);
- метод структурной индукции, основанный на построении дерева принятия решений для определения объектов из большого числа данных на входе;
- метод эвристических правил, основанных на использовании опыта экспертов, а не на абстрактных правилах формальной логики;
- метод машинной аналогии, основанный на представлении информации о сравниваемых объектах в удобном виде, например, в виде структур данных, называемых фреймами.

Источники «интеллекта», проявляющегося при решении задачи, могут оказаться бесполезными либо полезными или экономичными в зависимости от определенных свойств области, в которой поставлена задача. Исходя из этого, может быть осуществлен выбор метода построения экспертной системы или использования готового программного продукта.

Процесс выработки решения на основе первичных данных, можно разбить на два этапа:

- выработка допустимых вариантов решений путем математической формализации с использованием разнообразных моделей;
- выбор оптимального решения на основе субъективных факторов.

Информационные потребности лиц, принимающих решение, во многих случаях ориентированы на интегральные технико-экономические показатели, которые могут быть получены в результате обработки первичных данных, отражающих текущую деятельность предприятия. Анализируя функциональные взаимосвязи между итоговыми и первичными данными, можно построить так называемую информационную схему, которая отражает процессы агрегирования информации. Первичные данные, как правило, чрезвычайно разнообразны, интенсивность их поступления высока, а общий объем на интересующем интервале велик. С другой стороны состав интегральных показателей относительно мал, а требуемый период их актуализации может быть значительно короче периода изменения первичных данных - аргументов.

Для поддержки принятия решений обязательным является наличие следующих компонент:

- обобщающего анализа;
- прогнозирования;
- ситуационного моделирования.

В настоящее время принято выделять два типа информационных систем поддержки принятия решений.

Системы поддержки принятия решений DSS (Decision Support System) осуществляют отбор и анализ данных по различным характеристикам и включают средства:

- доступа к базам данных;
- извлечения данных из разнородных источников;
- моделирования правил и стратегии деловой деятельности;
- деловой графики для представления результатов анализа;
- анализа «если что»;

- искусственного интеллекта на уровне экспертных систем.

Системы оперативной аналитической обработки OLAP (OnLine Analysis Processing) для принятия решений используют следующие средства:

- мощную многопроцессорную вычислительную технику в виде специальных OLAP-серверов;
- специальные методы многомерного анализа;
- специальные хранилища данных Data Warehouse.

Реализация процесса принятия решений заключается в построении информационных приложений. Выделим в информационном приложении типовые функциональные компоненты, достаточные для формирования любого приложения на основе БД.

PS (Presentation Services) - средства представления. Обеспечиваются устройствами, принимающими ввод от пользователя и отображающими то, что сообщает ему компонент логики представления PL, плюс соответствующая программная поддержка. Может быть текстовым терминалом или X-терминалом, а также персональным компьютером или рабочей станцией в режиме программной эмуляции терминала или X-терминала.

PL (Presentation Logic) - логика представления. Управляет взаимодействием между пользователем и ЭВМ. Обрабатывает действия пользователя по выбору альтернативы меню, по нажатию кнопки или выбору элемента из списка.

BL (Business or Application Logic) - прикладная логика. Набор правил для принятия решений, вычислений и операций, которые должно выполнить приложение.

DL (Data Logic) - логика управления данными. Операции с базой данных (SQL-операторы SELECT, UPDATE и INSERT), которые нужно выполнить для реализации прикладной логики управления данными.

DS (Data Services) - операции с базой данных. Действия СУБД, вызываемые для выполнения логики управления данными, такие как манипулирование данными, определения данных, фиксация или откат транзакций и т.п. СУБД обычно компилирует SQL-приложения.

FS (File Services) - файловые операции. Дисковые операции чтения и записи данных для СУБД и других компонент. Обычно являются функциями ОС.

Среди средств разработки информационных приложений можно выделить следующие основные группы:

- традиционные системы программирования;
- инструменты для создания файл-серверных приложений;
- средства разработки приложений «клиент-сервер»;
- средства автоматизации делопроизводства и документооборота;
- средства разработки Интернет/Инtranет-приложений;
- средства автоматизации проектирования приложений.