

**Тема 1.** Модели, методы и оценка результатов компьютерное моделирования систем обработки информации и управления

**Лекция 1.** Компьютерное моделирование информационных систем

1. Компьютерное моделирование информационных систем
2. Особенности использования алгоритмических языков.
3. Подходы к разработке языков моделирования.
4. Классификации языков моделирования

## **Введение.**

Успех или неудача проведения имитационных экспериментов с моделями сложных систем существенным образом зависит от инструментальных средств, используемых для моделирования, т. е. набора аппаратно-программных средств, представляемых пользователю-разработчику или пользователю-исследователю машинной модели. В настоящее время существует большое количество языков имитационного моделирования — специальных языков программирования имитационных моделей на ЭВМ — и перед разработчиком машинной модели возникает проблема выбора языка, наиболее эффективного для целей моделирования конкретной системы.

### **1. Компьютерное моделирование информационных систем**

Использование современных ЭВМ, вычислительных комплексов и сетей является мощным средством реализации имитационных моделей и исследования с их помощью характеристик процесса функционирования систем  $S$ . Эффективность исследования системы  $S$  на программно-реализуемой модели  $M_m$  прежде всего зависит от правильности схемы моделирующего алгоритма, совершенства программы и только косвенным образом зависит от технических характеристик ЭВМ, применяемой для моделирования. Большое значение при реализации модели на ЭВМ имеет вопрос правильного выбора языка моделирования.

*Алгоритмические языки* при моделировании систем служат вспомогательным аппаратом разработки, машинной реализации и анализа характеристик моделей. Каждый язык моделирования должен отражать определенную структуру понятий для описания широкого класса явлений.

Основными моментами, характеризующими качество языков моделирования, являются: удобство описания процесса функционирования системы  $S$ ; удобство ввода исходных данных моделирования и варьирования структуры, алгоритмов и параметров модели, реализуемость статистического моделирования, эффективность анализа и вывода результатов моделирования,

простота отладки и контроля работы моделирующей программы, доступность восприятия и использования языка.

*Язык программирования* представляет собой набор символов, распознаваемых ЭВМ и обозначающих операции, которые можно реализовать на ЭВМ. На низшем уровне находится основной язык машины, программа на котором пишется в кодах, непосредственно соответствующих элементарным машинным действиям (сложение, запоминание, пересылка по заданному адресу и т. д.). Следующий уровень занимает автокод (язык *АССЕМБЛЕРА*) вычислительной машины. Программа на автокоде составляется из мнемонических символов, преобразуемых в машинные коды специальной программой — ассемблером.

*Компилятором* называется программа, принимающая инструкции, написанные на алгоритмическом языке высокого уровня, и преобразующая их в программы на основном языке машины или на автокоде, которые в последнем случае транслируются еще раз с помощью ассемблера.

*Интерпретатором* называется программа, которая, принимая инструкции входного языка, сразу выполняет соответствующие операции в отличие от компилятора, преобразующего эти инструкции в запоминающиеся цепочки команд. Трансляция происходит в течение всего времени работы программы, написанной на языке интерпретатора. В отличие от этого компиляция и ассемблирование представляют собой однократные акты перевода текста с входного языка на объектный язык машины, после чего полученные программы выполняются без повторных обращений к транслятору.

Программа, составленная в машинных кодах или на языке *АССЕМБЛЕРА*, всегда отражает специфику конкретной ЭВМ. Инструкции такой программы соответствуют определенным машинным операциям и, следовательно, имеют смысл только в той ЭВМ, для которой они предназначены, поэтому такие языки называются *машинно-ориентированными языками*.

Большинство языков интерпретаторов и компиляторов можно классифицировать как *процедурно-ориентированные языки*. Эти языки

качественно отличаются от *машинно-ориентированных языков*, описывающих элементарные действия ЭВМ и не обладающих проблемной ориентацией.

Все *процедурно-ориентированные языки* предназначены для определенного класса задач, включают в себя инструкции, удобные для формулировки способов решения типичных задач этого класса. Соответствующие алгоритмы программируются в обозначениях, не связанных ни с какой ЭВМ.

Язык моделирования представляет собой *процедурно-ориентированный язык*, обладающий специфическими чертами. Основные языки моделирования разрабатывались в качестве программного обеспечения имитационного подхода к изучению процесса функционирования определенного класса систем.

## **2. Особенности использования алгоритмических языков.**

Рассмотрим преимущества и недостатки использования для моделирования процесса функционирования систем *языков имитационного моделирования* (ЯИМ) и *языков общего назначения* (ЯОН), т. е. универсальных и процедурно-ориентированных алгоритмических языков.

Целесообразность использования *языков имитационного моделирования* (ЯИМ) вытекает из двух основных причин: 1) удобство программирования модели системы, играющее существенную роль при машинной реализации моделирующих алгоритмов; 2) концептуальная направленность языка на класс систем, необходимая на этапе построения модели системы и выборе общего направления исследований в планируемом машинном эксперименте. Практика моделирования систем показывает, что именно использование ЯИМ во многом определило успех имитации как метода экспериментального исследования сложных реальных объектов.

Языки моделирования позволяют описывать моделируемые системы в терминах, разработанных на базе основных понятий имитации. До того как эти понятия были четко определены и формализованы в ЯИМ, не существовало единых способов описания имитационных задач, а без них не было связи между различными разработками в области постановки имитационных экспериментов.

Высокоуровневые языки моделирования являются удобным средством общения заказчика и разработчика машинной модели  $M_m$ .

Несмотря на перечисленные преимущества ЯИМ, в настоящее время выдвигаются основательные аргументы как технического, так и эксплуатационного характера против полного отказа при моделировании от универсальных и процедурно-ориентированных языков.

Серьезные недостатки ЯИМ проявляются в том, что в отличие от широко применяемых ЯОН, трансляторы которых включены в поставляемое изготовителем математическое обеспечение всех современных ЭВМ, языки моделирования, за небольшим исключением, разрабатывались отдельными организациями для своих достаточно узко специализированных потребностей. Соответствующие трансляторы плохо описаны и приспособлены для эксплуатации при решении задач моделирования систем, поэтому, несмотря на достоинства ЯИМ, приходится отказываться от их практического применения в ряде конкретных случаев.

При создании системы моделирования на базе любого языка необходимо решить вопрос о синхронизации процессов в модели, так как в каждый момент времени, протекающего в системе (системного времени), может потребоваться обработка нескольких событий, т. е. требуется псевдопараллельная организация имитируемых процессов в машинной модели  $M_m$ . Это является основной задачей монитора моделирования, который выполняет следующие функции: управление процессами (согласование системного и машинного времени) и управление ресурсами (выбор и распределение в модели ограниченных средств моделирующей системы).

### **3. Подходы к разработке языков моделирования.**

К настоящему времени сложились два различных подхода к разработке языков моделирования: непрерывный и дискретный — отражающие основные особенности исследуемых методом моделирования систем. Поэтому ЯИМ делятся на две самостоятельные группы, которые соответствуют двум видам

имитации, развивавшимся независимо друг от друга: для имитации непрерывных и дискретных процессов.

Для моделирования непрерывных процессов могут быть использованы не только АВМ, но и ЭВМ, последние при соответствующем программировании имитируют различные непрерывные процессы. При этом ЭВМ обладают большей надежностью в эксплуатации и позволяют получить высокую точность результатов, что привело к разработке языков моделирования, отображающих модель в виде блоков таких типов, которые играют роль стандартных блоков АВМ (усилителей, интеграторов, генераторов функций и т. п.).

Заданная схема моделирующего алгоритма преобразуется в систему совместно рассматриваемых дифференциальных уравнений. Моделирование в этом случае сводится, по сути дела, к отысканию численных решений этих уравнений при использовании некоторого стандартного пошагового метода.

Универсальная ЭВМ — устройство дискретного типа, а поэтому должна обеспечивать дискретную аппроксимацию процесса функционирования исследуемой системы  $S$ . Непрерывные изменения в процессе функционирования реальной системы отображаются в дискретной модели  $M_0$ , реализуемой на ЭВМ, некоторой последовательностью дискретных событий, и такие модели называются *моделями дискретных событий*. Отдельные события, отражаемые в дискретной модели, могут определяться с большой степенью приближения к действительности, что обеспечивает адекватность таких дискретных моделей реальным процессам, протекающим в системах  $S$ .

*Архитектуру ЯИМ*, т. е. концепцию взаимосвязей элементов языка как сложной системы, и технологию перехода от системы  $S$  к ее машинной модели  $M_m$  можно представить следующим образом:

- 1) объекты моделирования (системы  $S$ ) описываются (отображаются в языке) с помощью некоторых атрибутов языка;
- 2) атрибуты взаимодействуют с процессами, адекватными реально протекающим явлениям в моделируемой системе  $S$ ;

3) процессы требуют конкретных условий, определяющих логическую основу и последовательность взаимодействия этих процессов во времени;

4) условия влияют на события, имеющие место внутри объекта моделирования (системы  $S$ ) и при взаимодействии с внешней средой  $E$ ;

5) события изменяют состояния модели системы  $M$  в пространстве и во времени.

В большинстве случаев с помощью машинных моделей исследуются характеристики и поведение системы  $S$  на определенном отрезке времени, поэтому одной из наиболее важных задач при создании модели системы и выборе языка программирования модели является реализация двух функций:

1) корректировка временной координаты состояния системы («продвижение» времени, организация «часов»);

2) обеспечение согласованности различных блоков и событий в системе (синхронизация во времени, координация с другими блоками).

#### 4. Классификации языков моделирования

Для машинного моделирования системы  $S$  пригодны три способа проведения вычислений, в основе которых лежит применение цифровой, аналоговой и гибридной вычислительной техники рис. 1.

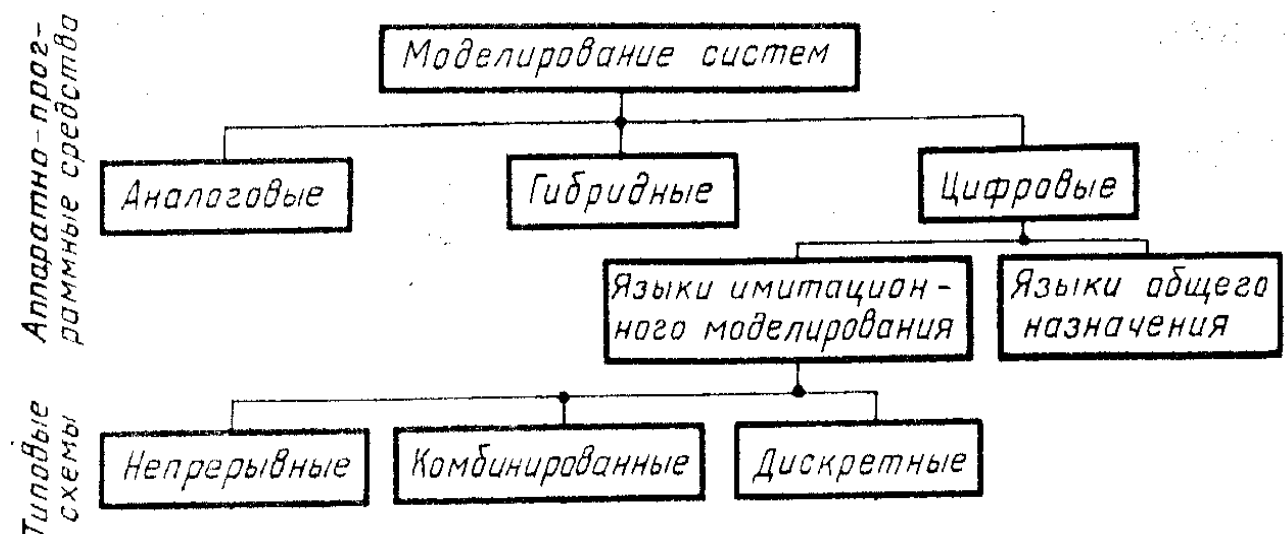


Рис. 1.

Для моделирования систем используются как универсальные и процедурно-ориентированные ЯОН, так и специализированные ЯИМ. При этом ЯОН предоставляют программисту-разработчику модели  $M_m$  больше возможностей в смысле гибкости разработки, отладки и использования модели. Но гибкость приобретается ценой больших усилий, затрачиваемых на программирование модели, так как организация выполнения операций, отсчет системного времени и контроль хода вычислений существенно усложняются.

Имеющиеся ЯИМ можно разбить на три основные группы, соответствующие трем типам математических схем: непрерывные, дискретные и комбинированные. Языки каждой группы предназначены для соответствующего представления системы  $S$  при создании ее машинной модели  $M_m$ .

В основе рассматриваемой классификации в некоторых ЯИМ лежит принцип формирования *системного времени*. Так как «системные часы» предназначены не только для продвижения системного времени в модели  $M_m$ , но также для синхронизации различных событий и операций в модели системы  $S$ , то при отнесении того или иного конкретного языка моделирования к определенному типу нельзя не считаться с типом механизма «системных часов».

Непрерывное представление системы  $S$  сводится к составлению уравнений, с помощью которых устанавливается связь между зависимыми и независимыми переменными модели. Примером такого непрерывного подхода является использование дифференциальных уравнений. Причем в дальнейшем дифференциальные уравнения могут быть применены для непосредственного получения характеристик системы.

Представление системы  $S$  в виде типовой схемы, в которой участвуют как непрерывные, так и дискретные величины, называется *комбинированным*. Состояние модели системы  $M(S)$  описывается набором переменных, некоторые из которых меняются во времени непрерывно. Законы изменения непрерывных компонент заложены в структуру, объединяющую дифференциальные



уравнения и условия относительно переменных. Предполагается, что в системе могут наступать события двух типов:

- 1) события, зависящие от состояния системы;
- 2) события, зависящие от времени.

События первого типа наступают в результате выполнения условий, относящихся к законам изменения непрерывных переменных.

Для событий второго типа процесс моделирования состоит в продвижении системного времени от момента наступления события до следующего аналогичного момента.

В рамках *дискретного подхода* можно выделить несколько принципиально различных групп ЯИМ.

Первая группа ЯИМ подразумевает наличие списка событий, отличающих моменты начала выполнения операций. Продвижение времени осуществляется по событиям, в моменты наступления которых производятся необходимые операции, включая операции пополнения списка событий.

При использовании ЯИМ второй группы после пересчета системного времени, в отличие от схемы языка событий, просмотр действий с целью проверки выполнения условий начала или окончания какого-либо действия производится непрерывно. Просмотр действий определяет очередность появления событий.

Третья группа ЯИМ описывает системы, поведение которых определяется процессами. В данном случае под процессом понимается последовательность событий, связь между которыми устанавливается с помощью набора специальных отношений. Динамика заложена в независимо управляемых программах, которые в совокупности составляют программу процесса.

## **Лекция 2.** Методы планирования компьютерных экспериментов

1. Машинный эксперимент.
2. Стратегическое планирование экспериментов
3. Тактическое планирование экспериментов

## **Введение.**

Имитационное моделирование является по своей сути машинным экспериментом с моделью исследуемой или проектируемой системы. Основная цель экспериментальных исследований с помощью имитационных моделей состоит в наиболее глубоком изучении поведения моделируемой системы. Для этого необходимо планировать и проектировать не только саму модель, но и процесс ее использования, т. е. проведение с ней экспериментов на ЭВМ. Весь комплекс вопросов планирования экспериментов с имитационными моделями для их успешного решения рационально разбить на стратегическое и тактическое планирование.

### **1. Машинный эксперимент.**

Машинный эксперимент с моделью системы  $S$  при ее исследовании и проектировании проводится с целью получения информации о характеристиках процесса функционирования рассматриваемого объекта. Эта информация может быть получена как для анализа характеристик, так и для их оптимизации при заданных ограничениях, т. е. для синтеза структуры, алгоритмов и параметров системы  $S$ . В зависимости от поставленных целей моделирования системы  $S$  на ЭВМ имеются различные подходы к организации имитационного эксперимента с машинной моделью  $M_m$ .

Основная задача планирования машинных экспериментов — получение необходимой информации об исследуемой системе  $S$  при ограничениях на ресурсы (затраты машинного времени, памяти и т. п.). К числу частных задач, решаемых при планировании машинных экспериментов, относятся задачи уменьшения затрат машинного времени на моделирование, увеличения точности и достоверности результатов моделирования, проверки адекватности модели и т. д.

Эффективность машинных экспериментов с моделями  $M_m$  существенно зависит от выбора плана эксперимента, так как именно план определяет объем

и порядок проведения вычислений на ЭВМ, приемы накопления и статистической обработки результатов моделирования системы  $S$ .

Поэтому основная задача планирования машинных экспериментов с моделью  $M_m$  формулируется следующим образом: необходимо получить информацию об объекте моделирования, заданном в виде моделирующего алгоритма (программы), при минимальных или ограниченных затратах машинных ресурсов на реализацию процесса моделирования.

Таким образом, при машинном моделировании рационально планировать и проектировать не только саму модель  $M_m$  системы  $S$ , но и процесс ее использования, т. е. проведение с ней экспериментов с использованием инструментальной ЭВМ.

К настоящему времени в различных областях знаний сложилась теория планирования экспериментов, в которой разработаны достаточно мощные математические методы, позволяющие повысить эффективность таких экспериментов.

Несмотря на то что цели экспериментального моделирования на ЭВМ и проведения натуральных экспериментов совпадают, между видами экспериментов существуют различия, поэтому для планирования эксперимента наиболее важное значение имеет следующее:

- 1) простота повторения условий эксперимента на ЭВМ с моделью  $M_m$  системы  $S$ ;
- 2) возможность управления экспериментом с моделью  $M_m$ , включая его прерывание и возобновление;
- 3) легкость варьирования условий проведения эксперимента (воздействий внешней среды  $E$ );
- 4) наличие корреляции между последовательностью точек в процессе моделирования;
- 5) трудности, связанные с определением интервала моделирования  $(0, T)$ .

Преимуществом машинных экспериментов перед натурным является возможность полного воспроизведения условий эксперимента с моделью

исследуемой системы  $S$ . Сравнить две альтернативы возможно при одинаковых условиях, что достигается, например, выбором одной и той же последовательности случайных чисел для каждой из альтернатив. Существенным достоинством перед натурными является простота прерывания и возобновления машинных экспериментов, что позволяет применять последовательные и эвристические приемы планирования, которые могут оказаться нереализуемыми в экспериментах с реальными объектами. При работе с машинной моделью  $M_m$  всегда возможно прерывание эксперимента на время, необходимое для анализа результатов и принятия решений об его дальнейшем ходе (например, о необходимости изменения значений параметров модели  $M_m$ ).

Недостатком машинных экспериментов является то, что часто возникают трудности, связанные с наличием корреляции в выходных последовательностях, т. е. результаты одних наблюдений зависят от результатов одного или нескольких предыдущих, и поэтому в них содержится меньше информации, чем в независимых наблюдениях. Так как в большинстве существующих методов планирования экспериментов предполагается независимость наблюдений, то многие из этих методов нельзя непосредственно применять для машинных экспериментов при наличии корреляции.

Рассмотрим основные понятия теории планирования экспериментов.

В связи с тем что математические методы планирования экспериментов основаны на кибернетическом представлении процесса проведения эксперимента, наиболее подходящей моделью последнего является абстрактная схема, называемая «черным ящиком». При таком кибернетическом подходе различают входные и выходные переменные:  $x_1, x_2, \dots, x_k, y_1, y_2, \dots, y_i$ .

В зависимости от того, какую роль играет каждая переменная в проводимом эксперименте, она может являться либо фактором, либо реакцией.

Если цель эксперимента — изучение влияния переменной  $x$  на переменную  $y$ , то  $x$  — фактор, а  $y$  — реакция. В экспериментах с машинными моделями  $M_m$

системы  $S$  фактор является независимая или управляемой (входной) переменной, а реакция — зависимой (выходной) переменной.

Каждый фактор  $x_i$ ,  $i=1, k$  может принимать в эксперименте одно из нескольких значений, называемых *уровнями*. Фиксированный набор уровней факторов определяет одно из возможных состояний рассматриваемой системы. Одновременно этот набор представляет собой условия проведения одного из возможных экспериментов.

Каждому фиксированному набору уровней факторов соответствует определенная точка в многомерном пространстве, называемом *факторным пространством*. Эксперименты не могут быть реализованы во всех точках факторного пространства, а лишь в принадлежащих допустимой области.

Существует вполне определенная связь между уровнями факторов и реакцией (откликом) системы.

Функцию, связывающую реакцию с факторами, называют *функцией реакции*, а геометрический образ, соответствующий функции реакции,— *поверхностью реакции*.

При планировании экспериментов необходимо определить основные свойства факторов. Факторы при проведении экспериментов могут быть управляемыми и неуправляемыми, наблюдаемыми и ненаблюдаемыми, изучаемыми и неизучаемыми, количественными и качественными, фиксированными и случайными.

*Фактор* называется *управляемым*, если его уровни целенаправленно выбираются исследователем в процессе эксперимента.

*Фактор* называется *наблюдаемым*, если его значения наблюдаются и регистрируются.

Но неуправляемый фактор также можно наблюдать. *Наблюдаемые неуправляемые факторы* получили название *сопутствующих*.

*Фактор* относится к *изучаемым*, если он включен в модель  $M_m$  для изучения свойств системы  $S$ , а не для вспомогательных целей, например для увеличения точности эксперимента.

*Фактор* будет *количественным*, если его значения — числовые величины, влияющие на реакцию, а в противном случае *фактор* называется *качественным*.

*Фактор* называется *фиксированным*, если в эксперименте исследуются все интересующие экспериментатора значения фактора, а если экспериментатор исследует только некоторую случайную выборку из совокупности интересующих значений факторов, то *фактор* называется *случайным*.

Каждый фактор может принимать в испытании одно или несколько значений, называемых *уровнями*, причем фактор будет управляемым, если его уровни целенаправленно выбираются экспериментатором. Для полного определения фактора необходимо указать последовательность операций, с помощью которых устанавливаются его конкретные уровни. Такое определение фактора называется *операциональным* и обеспечивает однозначность понимания фактора.

Основными требованиями, предъявляемыми к факторам, являются требование управляемости фактора и требование непосредственного воздействия на объект.

Под управляемостью фактора понимается возможность установки и поддержания выбранного нужного уровня фактора постоянным в течение всего испытания или изменяющимся в соответствии с заданной программой.

Требование непосредственного воздействия на объект имеет большое значение в связи с тем, что трудно управлять фактором, если он является функцией других факторов.

При планировании эксперимента обычно одновременно изменяются несколько факторов. Основные требования к факторам — совместимость и независимость.

Совместимость факторов означает, что все их комбинации осуществимы, а независимость соответствует возможности установления фактора на любом уровне независимо от уровней других.

## 2. Стратегическое планирование экспериментов

Применяя системный подход к проблеме планирования машинных экспериментов с моделями систем, можно выделить две составляющие планирования: стратегическое и тактическое планирование.

*Стратегическое планирование* ставит своей целью решение задачи получения необходимой информации о системе  $S$  с помощью модели  $M_m$  реализованной на ЭВМ, с учетом ограничений на ресурсы. По своей сути стратегическое планирование аналогично внешнему проектированию при создании системы  $S$ , только здесь в качестве объекта выступает процесс моделирования системы.

При стратегическом планировании машинных экспериментов с моделями систем возникает целый ряд проблем, взаимно связанных как с особенностями функционирования моделируемого объекта (системы  $S$ ), так и с особенностями машинной реализации модели  $M_m$  и обработки результатов эксперимента. В первую очередь к таким относятся проблемы построения плана машинного эксперимента; наличия большого количества факторов; многокомпонентной функции реакции; стохастической сходимости результатов машинного эксперимента; ограниченности машинных ресурсов на проведение эксперимента.

При построении плана эксперимента необходимо помнить, что целями проведения машинных экспериментов с моделью  $M_m$  системы  $S$  являются либо получение зависимости реакции от факторов для выявления особенностей изучаемого процесса функционирования системы  $S$ , либо нахождение такой комбинации значений факторов, которая обеспечивает экстремальное значение реакции.

Проблема стратегического планирования машинных экспериментов — наличие большого количества факторов. Это одна из основных проблем реализации имитационных моделей на ЭВМ, так как известно, что в факторном анализе количество комбинаций факторов равно произведению числа значений



всех факторов эксперимента. Если факторы  $x_i$   $i=1, k$ , являются количественными, а реакция  $y$  связана с факторами некоторой функцией, то в качестве метода обработки результатов эксперимента может быть выбран регрессионный анализ. Когда при моделировании требуется полный факторный анализ, то проблема большого количества факторов может не иметь решения. Достоинством полных факторных планов является то, что они дают возможность отобразить всю поверхность реакции системы, если количество факторов невелико. Эффективность этого метода существенно зависит от природы поверхности реакции.

Следующей проблемой стратегического планирования машинных экспериментов является многокомпонентная функция реакции. В имитационном эксперименте с вариантами модели системы  $S$  на этапе ее проектирования часто возникает задача, связанная с необходимостью изучения большого числа переменных реакции. Эту трудность в ряде случаев можно обойти, рассматривая имитационный эксперимент с моделью по определению многих реакций как несколько имитационных экспериментов, в каждом из которых исследуется (наблюдается) только одна реакция.

Существенное место при планировании экспериментов с имитационными моделями, реализуемыми методом статистического моделирования на ЭВМ, занимает проблема стохастической сходимости результатов машинного эксперимента. Эта проблема возникает вследствие того, что целью проведения конкретного машинного эксперимента при исследовании и проектировании системы  $S$  является получение на ЭВМ количественных характеристик процесса функционирования системы  $S$  с помощью машинной модели  $M_m$ . В качестве таких характеристик наиболее часто выступают средние некоторых распределений, для оценки которых применяют выборочные средние, найденные путем многократных прогонов модели на ЭВМ, причем чем больше выборка, тем больше вероятность того, что выборочные средние приближаются к средним распределений. Сходимость выборочных средних с ростом объема выборки называется *стохастической сходимостью*.

Применяя системный подход к проблеме стратегического планирования машинных экспериментов, можно выделить следующие этапы:

- 1) построение структурной модели;
- 2) построение функциональной модели.

При этом структурная модель выбирается исходя из того, что должно быть сделано, а функциональная — из того, что может быть сделано.

Структурная модель плана эксперимента характеризуется числом факторов и числом уровней для каждого фактора. Число элементов эксперимента где  $k$  — число факторов эксперимента;  $q$  — число уровней  $i$ -го фактора,  $i=1, \dots, k$ . При этом под элементом понимается структурный блок эксперимента, определяемый как простейший эксперимент в случае одного фактора и одного уровня.

Функциональная модель плана эксперимента определяет количество элементов структурной модели  $N_\phi$ , т. е. необходимое число различных информационных точек. При этом функциональная модель может быть полной и неполной.

Функциональная модель называется полной, если в оценке реакции участвуют все элементы, т. е.  $N_\phi = N_c$ , и неполной, если число реакций меньше числа элементов, т. е.  $N_\phi < N_c$ . Основная цель построения функциональной модели — нахождение компромисса между необходимыми действиями при машинном эксперименте (исходя из структурной модели) и ограниченными ресурсами на решение задачи методом моделирования.

Таким образом, использование при стратегическом планировании машинных экспериментов с  $M_m$  структурных и функциональных моделей плана позволяет решить вопрос о практической реализуемости модели на ЭВМ исходя из допустимых затрат ресурсов на моделирование системы  $S$ .

### **3. Тактическое планирование экспериментов**

*Тактическое планирование* представляет собой определение способа проведения каждой серии испытаний машинной модели  $M_m$ , предусмотренных планом эксперимента. Для тактического планирования также имеется аналогия с внутренним проектированием системы  $S$ , но опять в качестве объекта рассматривается процесс работы с моделью  $M_m$ .

Тактическое планирование эксперимента с машинной моделью  $M_m$  системы  $S$  связано с вопросами эффективного использования выделенных для эксперимента машинных ресурсов и определением конкретных способов проведения испытаний модели  $M_m$ , намеченных планом эксперимента, построенным при стратегическом планировании.

Тактическое планирование машинного эксперимента связано прежде всего с решением следующих проблем:

- 1) определения начальных условий и их влияния на достижение установившегося результата при моделировании;
- 2) обеспечения точности и достоверности результатов моделирования;
- 3) уменьшения дисперсии оценок характеристик процесса функционирования моделируемых систем;
- 4) выбора правил автоматической остановки имитационного эксперимента с моделями систем.

Первая проблема при проведении машинного эксперимента возникает вследствие искусственного характера процесса функционирования модели  $M_m$ , которая в отличие от реальной системы  $S$  работает эпизодически, т. е. только когда экспериментатор запускает машинную модель и проводит наблюдения.

Решение второй проблемы тактического планирования машинного эксперимента связано с оценкой точности и достоверности результатов моделирования (при конкретном методе реализации модели, например, методе статистического моделирования на ЭВМ) при заданном числе реализаций (объеме выборки) или с необходимостью оценки необходимого числа реализаций при заданных точности и достоверности результатов моделирования системы  $S$ .

Проблемой выбора количества реализаций при обеспечении необходимой точности и достоверности результатов машинного эксперимента тесно связана и третья проблема, а именно проблема уменьшения дисперсии. В настоящее время существуют методы, позволяющие при заданном числе реализаций увеличить точность оценок, полученных на машинной модели  $M_m$ , и, наоборот, при заданной точности оценок сократить необходимое число реализаций при статистическом моделировании. Эти методы используют априорную информацию о структуре и поведении моделируемой системы  $S$  и называются методами уменьшения дисперсии. При подходе к уменьшению дисперсии задача состоит в специальном построении моделирующего алгоритма системы  $S$ , позволяющего получить положительную корреляцию, например, за счет управления генерацией случайных величин. Вопрос об эффективности использования метода уменьшения дисперсии может быть решен только с учетом необходимости дополнительных затрат машинных ресурсов (времени и памяти) на реализацию подхода, т. е. теоретическое уменьшение затрат машинного времени на моделирование вариантов системы (при той же точности результатов) должно быть проверено на сложность машинной реализации модели.

Последней из проблем, возникающих при тактическом планировании имитационных экспериментов является выбор правил автоматической остановки имитационного эксперимента. Простейший способ решения проблемы — задание требуемого количества реализаций  $N$  (или длины интервала моделирования  $T$ ). Однако такой детерминированный подход неэффективен, так как в его основе лежат достаточно грубые предположения о распределении выходных переменных, которые на этапе тактического планирования являются неизвестными.

Другой способ — задание доверительных интервалов для выходных переменных и остановка прогона машинной модели  $M_m$  при достижении заданного доверительного интервала, что позволяет теоретически приблизить время прогона к оптимальному. При практической реализации введение в

модель  $M_m$  правил остановки и операций вычисления доверительных интервалов увеличивает машинное время, необходимое для получения одной выборочной точки при статистическом моделировании.

Правила автоматической остановки могут быть включены в машинную модель такими способами:

1) путем двухэтапного проведения прогона, когда сначала делается пробный прогон из  $N^*$  реализаций, позволяющий оценить необходимое количество реализаций  $N$  (причем если  $N^* \geq N$ , то прогон можно закончить, в противном случае необходимо набрать еще  $N - N^*$  реализаций);

2) путем использования последовательного анализа для определения минимально необходимого количества реализаций  $N$ , которое рассматривается при этом как случайная величина, зависящая от результатов  $N - 1$  предыдущих реализаций (наблюдений, испытаний) машинного эксперимента.

Таким образом, чем сложнее машинная модель  $M_m$ , тем важнее этап тактического планирования машинного эксперимента, выполняемый непосредственно перед моделированием на ЭВМ системы  $S$ . Процесс планирования машинных экспериментов с моделью  $M_m$  итерационен, т. е. при уточнении некоторых свойств моделируемой системы  $S$  этапы стратегического и тактического планирования экспериментов могут чередоваться.

### **Лекция 3. «Обработка результатов компьютерного моделирования»**

1. Особенности статистической обработки результатов на ЭВМ.
2. Корреляционный анализ результатов моделирования.
3. Регрессионный анализ результатов моделирования.
4. Дисперсионный анализ результатов моделирования.

## **Введение.**

Успех имитационного эксперимента с моделью системы существенным образом зависит от правильного решения вопросов обработки и последующего анализа и интерпретации результатов моделирования. Особенно важно решить проблему текущей обработки экспериментальной информации при использовании модели для целей автоматизации проектирования систем.

### **1. Особенности статистической обработки результатов ЭВМ**

При выборе методов обработки существенную роль играют три особенности машинного эксперимента с моделью системы  $S$ .

1. Возможность получать при моделировании системы  $S$  на ЭВМ большие выборки позволяет количественно оценить характеристики процесса функционирования системы, но превращает в серьезную проблему хранение промежуточных результатов моделирования. Эту проблему можно решить, используя рекуррентные алгоритмы обработки, когда оценки вычисляются по ходу моделирования.

2. Сложность исследуемой системы  $S$  при ее моделировании на ЭВМ часто приводит к тому, что априорное суждение о характеристиках процесса функционирования системы, например о типе ожидаемого распределения выходных переменных, является невозможным. Поэтому при моделировании систем широко используются непараметрические оценки и оценки моментов распределения.

3. Блочность конструкции машинной модели  $M_m$  и отдельное исследование блоков связаны с программной имитацией входных переменных для одной частичной модели по оценкам выходных переменных, полученных на другой частичной модели. Если ЭВМ, используемая для моделирования, не позволяет воспользоваться переменными, записанными на внешние носители, то следует представить эти переменные в форме, удобной для построения алгоритма их имитации.

При исследовании сложных систем и большом числе реализаций  $N$  в результате моделирования на ЭВМ получается значительный объем информации о состояниях процесса функционирования системы. Поэтому необходимо так организовать процесс вычислений фиксации и обработку результатов моделирования, чтобы оценки для искомым характеристик формировались постепенно по ходу моделирования, т. е. без специального запоминания всей информации о состояниях процесса функционирования системы  $S$ .

Если при моделировании процесса функционирования конкретной системы  $S$  учитываются случайные факторы, то и среди результатов моделирования присутствуют случайные величины. В качестве оценок для искомым характеристик рассчитывают средние значения, дисперсии, корреляционные моменты и т. д.

При обработке результатов моделирования можно подойти к оценке вероятностей возможных значений случайной величины, т. е. закона распределения. Область возможных значений случайной величины  $\eta$  разбивается на  $n$  интервалов. Затем накапливается количество попаданий случайной величины в эти интервалы  $m_k$ ,  $k=1, n$ . Оценкой для вероятности попадания случайной величины в интервал с номером  $k$  служит величина  $m_k/N$ . Таким образом, при этом достаточно фиксировать  $n$  значений  $m_k$  при обработке результатов моделирования на ЭВМ.

Для оценки среднего значения случайной величины  $\eta$  накапливается сумма возможных значений случайной величины  $y_k$ ,  $k=1, N$ , которые она принимает при различных реализациях. Тогда среднее значение

$$\bar{y} = (1/N) \sum_{k=1}^N y_k.$$

При этом ввиду несмещенности и состоятельности оценки

$$M[\bar{y}] = M[\eta] = \mu_\eta;$$
$$D[\bar{y}] = D[\eta]/N = \sigma_\eta^2/N.$$



В качестве оценки дисперсии случайной величины  $\eta$  при обработке результатов моделирования можно использовать

$$S_b^2 = \sum_{k=1}^N (y_k - \bar{y})^2 / N$$

При обработке результатов машинного эксперимента с моделью  $M_m$  наиболее часто возникают следующие задачи: определение эмпирического закона распределения случайной величины, проверка однородности распределений, сравнение средних значений и дисперсий переменных, полученных в результате моделирования, и т. д. Эти задачи с точки зрения математической статистики являются типовыми задачами по проверке статистических гипотез.

Задача определения эмпирического закона распределения случайной величины наиболее общая из перечисленных, но для правильного решения требует большого числа реализаций  $N$ . В этом случае по результатам машинного эксперимента находят значения выборочного закона распределения  $F_s(y)$  (или функции плотности  $f_s(y)$ ) и выдвигают нулевую гипотезу  $H_0$ , что полученное эмпирическое распределение согласуется с каким-либо теоретическим распределением. Проверяют эту гипотезу  $H_0$  с помощью статистических критериев согласия Колмогорова, Пирсона, Смирнова и т. д., причем необходимую в этом случае статистическую обработку результатов ведут по возможности в процессе моделирования системы  $S$  на ЭВМ.

Критерий согласия Колмогорова основан на выборе в качестве меры расхождения  $U$  величины  $D = \max[F_s(y) - F(y)]$ .

Из теоремы Колмогорова следует, что  $\delta = D\sqrt{N}$  при  $N \rightarrow \infty$  имеет функцию распределения

$$F(z) = P\{\delta < z\} = \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2 z^2}, z > 0.$$

Если вычисленное на основе экспериментальных данных значение  $5$  меньше, чем табличное значение при выбранном уровне значимости  $u$ , то

гипотезу  $Y_0$  принимают, в противном случае расхождение между  $F_0(y)$  и  $F(y)$  считается неслучайным гипотеза  $H_0$  отвергается.

Критерий Колмогорова для обработки результатов моделирования целесообразно применять в тех случаях, когда известны все параметры теоретической функции распределения. Недостаток использования этого критерия связан с необходимостью фиксации в памяти ЭВМ для определения  $D$  всех статистических частот с целью их упорядочения в порядке возрастания.

Критерий согласия Пирсона основан на определении в качестве меры расхождения  $U$  величины

$$\chi^2 = \sum_{i=1}^d (m_i - N p_i) / (N p_i),$$

где  $m_i$  — количество значений случайной величины  $\eta$ , попавших в  $i$ -й подынтервал;  $p_i$  — вероятность попадания случайной величины  $\eta$  в  $i$ -й подынтервал, вычисленная из теоретического распределения;  $d$  — количество подынтервалов, на которые разбивается интервал измерения в машинном эксперименте.

При  $N \rightarrow \infty$  закон распределения величины  $U$ , являющейся мерой расхождения, зависит только от числа подынтервалов и приближается к закону распределения  $\chi^2$  (хи-квадрат) с  $(d-r-1)$  степенями свободы, где  $r$  — число параметров теоретического закона распределения.

Из теоремы Пирсона следует, что, какова бы ни была функция распределения  $F(y)$  случайной величины  $\eta$ , при  $N \rightarrow \infty$  распределение величины  $\chi^2$  имеет вид

$$F_k(z) = P\{\chi^2 < z\} = 1 / [2^{k/2} \Gamma(k/2)] \int_0^z e^{-t/2} t^{(k/2-1)} dt, \quad z > 0$$

где  $\Gamma(k/2)$  — гамма-функция;  $z$  — значение случайной величины  $\chi^2$ ,  $k = d - r - 1$  — число степеней свободы. Функции распределения  $F_k(z)$  табулированы.

По вычисленному значению  $U = \chi^2$  и числу степеней свободы  $k$  с помощью таблиц находится вероятность  $P\{\chi^2 \geq \chi^2\}$ . Если эта вероятность превышает

некоторый уровень значимости  $\gamma$ , то считается, что гипотеза  $H_0$  о виде распределения не опровергается результатами машинного эксперимента.

Для принятия или опровержения гипотезы выбирают некоторую случайную величину  $U$ , характеризующую степень расхождения теоретического и эмпирического распределения, связанную с недостаточностью статистического материала и другими случайными причинами. Закон распределения этой случайной величины зависит от закона распределения случайной величины  $\eta$  и числа реализаций  $N$  при статистическом моделировании системы  $S$ . Если вероятность расхождения теоретического и эмпирического распределений  $P \{ U_T \geq U \}$  велика в понятиях применяемого критерия согласия, то проверяемая гипотеза о виде распределения  $H_0$  не опровергается. Выбор вида теоретического распределения  $F(y)$  (или  $f(y)$ ) проводится по графикам (гистограммам)  $F_0(y)$  (или  $f_0(y)$ ), выведенным на печать или на экран дисплея.

Хотя рассмотренные оценки искомых характеристик процесса функционирования системы  $S$ , полученные в результате машинного эксперимента с моделью  $M_m$ , являются простейшими, но охватывают большинство случаев, встречающихся в практике обработки результатов моделирования системы для целей ее исследования и проектирования.

## **2. Корреляционный анализ результатов моделирования.**

С помощью *корреляционного анализа* исследователь может установить, насколько тесна связь между двумя (или более) случайными величинами, наблюдаемыми и фиксируемыми при моделировании конкретной системы  $S$ . Корреляционный анализ результатов моделирования сводится к оценке разброса значений  $\eta$  относительно среднего значения  $\bar{y}$ , т. е. к оценке силы корреляционной связи. Существование этих связей и их тесноту можно для схемы корреляционного анализа  $y = M[\eta/\xi = x]$  выразить при наличии линейной связи между исследуемыми величинами и нормальности их совместного распределения с помощью коэффициента корреляции.

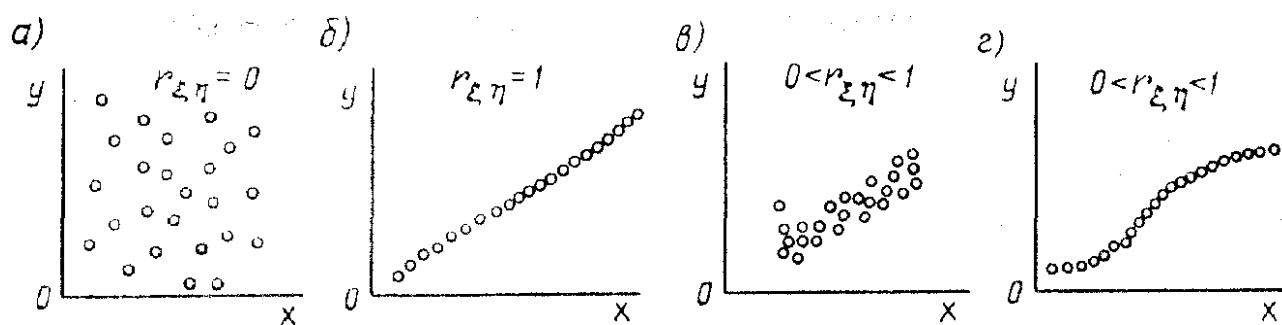


Рис.1. Различные случаи корреляции переменных

Для того чтобы оценить точность полученной при обработке результатов моделирования системы  $S$  оценки  $r_{\xi\eta}$ , целесообразно ввести в рассмотрение коэффициент

$$w = \ln [(1 + r_{\xi\eta}) / (1 - r_{\xi\eta})] / 2,$$

причем  $w$  приближенно подчиняется гауссовскому распределению со средним значением и дисперсией:

$$\mu_w = \ln [(1 + r_{\xi\eta}) / (1 - r_{\xi\eta})] / 2$$

$$\sigma_w^2 = 1 / (N - 3)$$

Из-за влияния числа реализаций при моделировании  $N$  на оценку коэффициента корреляции необходимо убедиться в том, что  $0 \leq r_{\xi\eta} \leq 1$  действительно отражает наличие статистически значимой корреляционной зависимости между исследуемыми переменными модели  $M_m$ . Это можно сделать проверкой гипотезы  $H_0: r_{\xi\eta} = 0$ . Если гипотеза  $H_0$  при анализе отвергается, то корреляционную зависимость признают статистически значимой. Очевидно, что выборочное распределение введенного в рассмотрение коэффициента  $w$  при  $r_{\xi\eta} = 0$  является гауссовским с нулевым средним  $\mu_w = 0$  и дисперсией  $\sigma_w^2 = (N - 3)^{-1}$ .

При анализе результатов моделирования системы  $S$  важно отметить то обстоятельство, что даже если удалось установить тесную зависимость между двумя переменными, то отсюда еще непосредственно не следует их причинно-следственная взаимообусловленность. Возможна ситуация, когда случайные  $\xi$  и  $\eta$  стохастически зависимы, хотя причинно они являются для системы  $S$  независимыми. При статистическом моделировании наличие такой

зависимости может иметь место, например, из-за коррелированности последовательностей псевдослучайных чисел, используемых для имитации событий, положенных в основу вычисления значений  $x$  и  $y$ .

Таким образом, корреляционный анализ устанавливает связь между исследуемыми случайными переменными машинной модели и оценивает тесноту этой связи. Однако в дополнение к этому желательно располагать моделью зависимости, полученной после обработки результатов моделирования.

### **3. Регрессионный анализ результатов моделирования.**

*Регрессионный анализ* дает возможность построить модель, наилучшим образом соответствующую набору данных, полученных в ходе машинного эксперимента с системой  $S$ . Под наилучшим соответствием понимается минимизированная функция ошибки, являющаяся разностью между прогнозируемой моделью и данными эксперимента. Такой функцией ошибки при регрессионном анализе служит сумма квадратов ошибок.

### **4. Дисперсионный анализ результатов моделирования.**

При обработке и анализе результатов моделирования часто возникает задача сравнения средних выборок. Если в результате такой проверки окажется, что математическое ожидание совокупностей случайных переменных  $\{y^{(1)}\}$ ,  $\{y^{(2)}\}$ , ...,  $\{y^{(n)}\}$  отличается незначительно, то статистический материал, полученный в результате моделирования, можно считать однородным (в случае равенства двух первых моментов). Это дает возможность объединить все совокупности в одну и позволяет существенно увеличить информацию о свойствах исследуемой модели  $M_m$ , а следовательно, и системы  $S$ . Попарное использование для этих целей критериев Смирнова и Стьюдента для проверки нулевой гипотезы затруднено в связи с наличием большого числа выборок при моделировании системы. Поэтому для этой цели используется *дисперсионный анализ*.

Дисперсионный анализ позволяет вместо проверки нулевой гипотезы о равенстве средних значений выборок проводить при обработке результатов моделирования проверку нулевой гипотезы о тождественности выборочной и генеральной дисперсий.

Возможны и другие подходы к анализу и интерпретации результатов моделирования, но при этом необходимо помнить, что их эффективность существенно зависит от вида и свойств конкретной моделируемой системы  $S$ .