

Применение моделирования для обучения в области компьютерных наук

Р.П. Романски
Технический Университет, София, Болгария
rrom@tu-sofia.bg

АННОТАЦИЯ

Существенная роль при обучении в области компьютерных наук выполняет применение современных методов изучения архитектурной организации и анализа системной производительности компьютерных систем (КС). В этом смысле, применение моделирования в процессе изучения базовых структур различных КС и организации компьютерных процессов позволяет разработать подходящее математическое описание исследуемого объекта и создать программное обеспечение для выполнения компьютерных экспериментов для оценивания основных характеристик системы. Процесс моделирования связан с организацией модельного эксперимента на основе проектирования компьютерной модели как последовательности трех этапов – концептуальная модель, математическая модель, программная модель. В настоящей работе рассматривается применение методов моделирования КС в процессе обучения. Основная цель заключается в определении обобщенной процедуры компьютерного моделирования как последовательность взаимосвязанных этапов и представлении основных этапов методологии модельного исследования.

Ключевые слова

компьютерное моделирование, компьютерные системы, обучение в области компьютерных наук, методология моделирования.

Введение

Для развития компьютерной техники и совершенствования архитектурной организации компьютерных систем (КС) необходимо непрерывное обучение и самосовершенствование компьютерных специалистов и студентов. При проведении этого обучения надо комбинировать формы традиционного обучения с возможностями самостоятельной подготовки, дистанционного обучения, практической разработки проектов и реализации экспериментов исследования [Butler, 2001; Ibbett, 2002; Lilja, 2001]. Существенная роль при обучении в области компьютерных наук выполняет применение современных методов изучения архитектурной организации и анализа системной производительности КС [Lilja, 2001]. В этом смысле, применение методов моделирования в процессе изучения базовых структур различных КС и организации компьютерных процессов позволяет разработать подходящее математическое описание исследуемого объекта и создать программное обеспечение для выполнения компьютерных экспериментов [Романски, 2001, Arons, 2000]. Анализ экспериментальных результатов моделирования [Брююль, 2002] позволяет оценить основные характеристики системы и производительность изучаемых КС.

Применение моделирования в процессе изучения КС позволяет исследовать особенности архитектуры и организацию вычисления и управления. Это можно осуществить на основе модельного эксперимента, организация которого предполагает проектирование компьютерной модели как последовательности трех компонентов (концептуальная модель, математическая модель, программная модель) и реализации этой модели в подходящей операционной среде. В настоящей работе рассматривается возможность применения разных методов исследования КС в процессе их изучения и в частности применение принципов моделирования для

исследования протекающих процессов, а также анализ системной производительности КС. Основная цель состоит в определении обобщенной процедуры компьютерного моделирования как последовательность взаимосвязанных этапов и представлении основных стадий методологии модельного исследования. Для этого в следующей части представлены общая формализация компьютерной обработки информации и особенности компьютерных вычислений в качестве объекта изучения. Применение принципов моделирования в процессе изучения КС связано с методологической организацией обучения в традиционном, дистанционном, либо распределенном смысле [Crosbie, 2000; Lilja, 2001; Romansky, 2002; Ye, 2002].

Компьютерные системы как объект изучения и методы исследования

Одной из основных задач специализированных курсов обучения в области компьютерных систем и исследования производительности является обучение будущих и настоящих компьютерных проектантов, разработчиков компьютерного оборудования и потребителей КС в правильном использовании технологических возможностей моделирования и измерения характеристик систем [Lilja, 2001]. Эти возможности применяют как в процессе оценивания эффективности новых компьютерных проектов, так и для проведения сравнительного анализа существующих систем. В процессе обучения ставится задача выяснения последовательности этапов исследования и возможности обработки экспериментальных результатов для получения адекватных оценок индексов производительности. Эту задачу можно уточнить в зависимости от конкретной области компьютерного обучения и особенностей принципов рассматриваемой компьютерной обработки информации.

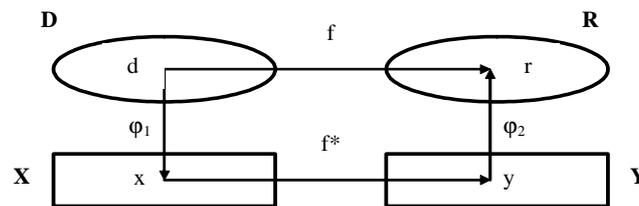


Рис. 1. Информационное поддержание компьютерной обработки.

В общем, компьютерная обработка связана с реализацией определенных функций для преобразования входных данных в виде окончательных решений. Это определяет два уровня функционального преобразования информации (рис. 1):

- математическое преобразование информации – реальная обработка данных в виде математических объектов и представляется обобщенной функцией $f: D \rightarrow R$, которая изображает элементы множества данных D в элементах множества результатов R ;
- компьютерная реализация обработки – представляет конкретную реализацию $f^*: X \rightarrow Y$ математической функции f в зависимости от компьютерного и программного оборудования на базе подходящего физического представления реальных информационных объектов.

В результате можно записать обобщенную функциональную модель компьютерной обработки $r = f(d) \equiv \varphi_2\{f^*[\varphi_1(d)]\}$, где функции φ_1 и φ_2 являются вспомогательными для кодирования и декодирования информации.

Рассматривая КС как объект изучения, надо иметь в виду, что компьютерная обработка состоит из процессов, каждый из которых можно представить в виде структуры $I = \langle t, A, T \rangle$, где: t – начальный момент возникновения процесса; A – дефинирующие атрибуты; T – трасса процесса. Последний компонент формального описания определяет временную последовательность событий e_j для обращения данного процесса к элементам системного ресурса $S = \{S_1, S_2, \dots, S_n\}$.

Последовательность временных этапов и нагрузка системного ресурса позволяют определить профиль процесса вычисления (рис. 2).

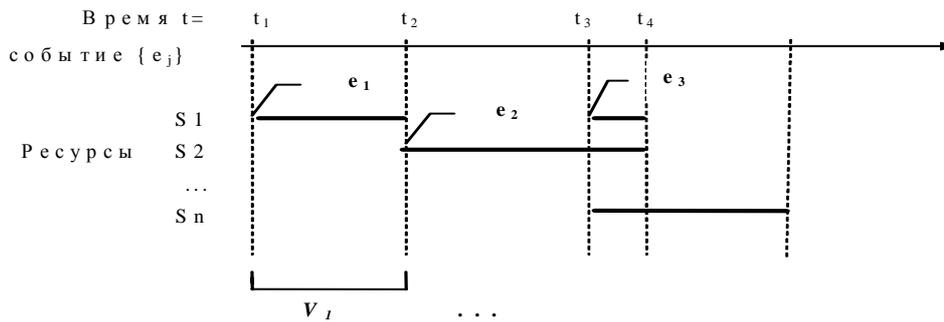


Рис. 2. Примерный профиль компьютерного процесса.

Поддерживание разных процессов при организации компьютерной обработки формирует системную нагрузку компьютерной среды. Для каждого момента ($t=1,2,\dots$) ее можно представить вектором $V(t)=V_t=\langle v_1, v_2, \dots, v_n \rangle$, элементы которого выражают свободное ($v_j=0$) или занятое ($v_j=1$) устройство $S_j \in S$ ($j=1,2,\dots,n$).

При изучении КС необходимо определить набор базовых системных параметров, которые отражают сущность компьютерной обработки, а также разработать методику исследования поведения системного ресурса и протекающих процессов. В качестве основных системных параметров (индексы производительности) можно исследовать, например, рабочую нагрузку каждого элемента системного ресурса, общую системную нагрузку КС, время ответа при решении комплекса задач в мультипрограммном режиме, степень устойчивости (стойкости) оборудования, стоимость компьютерной обработки, эффективность планирования параллельных или псевдопараллельных процессов и т.д.

Типичный курс обучения в области анализа и исследования производительности КС должен обсуждать основные теоретические и практические проблемы в следующих направлениях:

- возможности исследования производительности компьютерного оборудование и эффективности компьютерных процессов;
- применение эффективных методов исследования (измерение, моделирование);
- технологические особенности измерения параметров системы (benchmark, monitoring);
- технологические особенности и организация моделирования (аналитическое, симуляционное и др.);
- методы анализа экспериментальных результатов.

Все это связано с применением данного метода исследования и выбором подходящего инструментария. В этом смысле на рис. 3 представлена примерная классификация методов исследования КС и процессов. Можно определить три основные группы:

1. Программные смеси – представляют математические зависимости для оценки производительности процессора на базе коэффициентов применения отдельных операционных классов. Позволяют оценить нагрузку процессора статистическим анализом после выполнения типовых программ.
2. Методы подсчета – позволяют получить достоверную информацию о протекании компьютерных процессов на основе непосредственной регистрации определенных значений доступных параметров КС [Fabre, 2002; Pandey, 2003]. Для этого необходимо использовать или разработать подходящее средство подсчета (монитор) и организовать выполнение эксперимента по подсчету. Надо отметить, что современные операционные системы имеют собственные системные мониторы, которые можно использовать на программном или микропрограммном уровне.

3. Методы моделирования – применяются в том случае, когда отсутствует реальный объект эксперимента. Исследование структуры или протекающих процессов в КС осуществляется на базе компьютерной модели. Она отражает самые важные аспекты поведения структурных и системных параметров в зависимости от поставленной цели. Для разработки модели надо выбрать самый подходящий метод моделирования, позволяющий получить максимальную адекватность и достоверность [Crosbie, 2000; Sargent, 2003; Ye, 2002].



Рис. 3. Классификация методов исследования КС и процессов.

Традиционный процесс обучения предполагает проведение основного курса лекций совместно с набором аудиторных упражнений и/или лабораторным практикумом. В области компьютерных наук при изучении организации КС и принципов управления компьютерными процессами (на низком и на высоком уровне), а также при анализе системной производительности, часто возникает необходимость в разработке компьютерных моделей во время выполнения лабораторных задач в классе или при самостоятельной реализации проектов. Для удачного выполнения этих практических работ и для получения нужных практических умений необходимо определить последовательность этапов и представить технологические особенности разработки моделей. Это позволит обучаемым приобрести необходимые знания о разработке адекватных и достоверных компьютерных моделей исследования, оценки и сравнительного анализа системной производительности разных компьютерных архитектур. В результате этого далее предложена обобщенная процедура проведения моделирования, а также методологическая схема модельного исследования КС и процессов.

Процедура компьютерного моделирования при исследовании КС и процессов

Основная задача компьютерного моделирования при исследовании КС и процессов заключается в получении информации об индексах производительности. Планирование модельного эксперимента в процессе обучения осуществляется на основе следующих этапов:

1. сбор эмпирических данных для конкретных значений базовых системных параметров;
2. структурирование и обработка эмпирической информации и разработка функциональной схемы модели;

3. определение априорной информации и дефиниционные области рабочих параметров для разработки подходящей математической модели объекта-оригинала;
4. реализация модельных экспериментов, накопление модельной информации и ее последующий анализ.

Обобщенная формализованная процедура модельного исследования для организации модельного эксперимента показана на рис. 4.

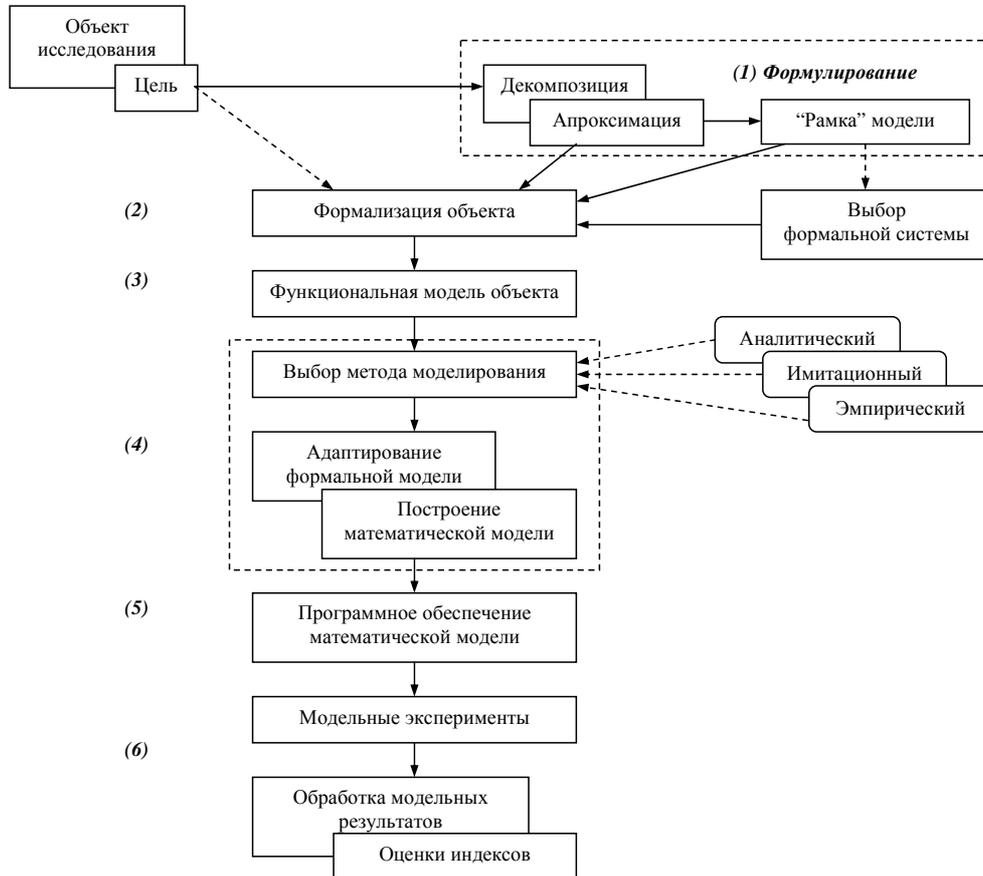


Рис. 4. Процедура модельного исследования.

Первоначальная цель определяется необходимостью исследования реального объекта (система или процесс). Основные этапы процедуры следующие:

1. Определение базовой концепции построения модели декомпозированием объекта на подсистемы и введение допустимой степени идеализации для некоторых аспектов поведения системных процессов.
2. Математическая формализация структуры и взаимосвязи в исследованном объекте на базе подходящей формальной системы.
3. Математическое описание функционирования реальной системы и разработка подходящей функциональной модели в зависимости от цели моделирования.
4. Реализация математической модели с использованием самого подходящего метода моделирования.
5. Описание созданной математической модели средствами подходящей программной среды (специализированной или универсальной).
6. Выполнение экспериментов на базе созданной модели и последующая обработка и интерпретация модельной информации для оценки параметров объекта исследования.

Основные методы компьютерного моделирования следующие:

1. *Аналитические методы* – используют математические средства для описания компонентов реальной системы и протекающих процессов. На базе выбранного математического подхода математическая модель

обычно строится как система уравнений, позволяющая легко программировать, но для реализации необходимы высокая точность формулировок и принятых рабочих гипотез, а также значительная верификация.

2. *Симуляционные (имитационные) методы* – поведение реального объекта подражается программному имитатору, который при своей работе использует реальную рабочую нагрузку (эмуляция), либо программную модель рабочей нагрузки (симуляция). Такие модели позволяют исследовать сложные системы и получение достоверных результатов, но выполняются во времени и это определяет основной недостаток метода – значительное потребление машинного времени.
3. *Эмпирические методы* – это количественные приемы для регистрации, накопления и анализа информации функционирования реального объекта, на базе которых можно построить статистическую модель для его исследования. Обычно применяются линейные или нелинейные уравнения для представления взаимосвязи выбранных параметров (например, из множества первичных факторов) и для вычисления статистических характеристик.

Основной задачей компьютерного моделирования является создание адекватной модели, при помощи которой достаточно точно можно представить структуру исследуемой системы и протекающих процессов [Sargent, 2003]. Разработка компьютерной модели включает три последовательных уровня – концептуальная модель (идейная концепция структурирования модели), математическая модель (изображение концептуальной модели средством математической формальной системы) и программная модель (программная реализация математической модели с подходящей языковой средой). На каждом уровне компьютерного моделирования необходимо проверять адекватность модели, чтобы обеспечить достоверность конечной модели и точность результатов модельных экспериментов. Специфика отдельных этапов процедуры моделирования определяет применяемые подходы и средства оценки адекватности. Эти особенности нашли место в разработанной методологии компьютерного моделирования, которая представлена ниже.

Методология модельного исследования

В процессе компьютерного моделирования, независимо от применяемого метода, можно определить обобщенную методологическую схему модельного исследования (рис. 5). Предложенная формализованная методологическая последовательность предусматривает несколько основных фаз, представленных ниже. В основном, она представляет итерационную процедуру для получения необходимой достоверности разрабатываемой компьютерной модели на базе формулировки начальной модельной гипотезы и ее последовательной модификации. Такой подход является удачным при исследовании сложных систем, а также и при отсутствии достаточной априорной информации для исследуемого объекта.

Стадия “Формулирование”

На первом этапе разработки модели необходимо точно и ясно определить объект моделирования, условия и гипотезы исследования, а также критерии оценки модельной эффективности. Это позволит разработать концептуальную модель и дефинировать ее абстрактными терминами и понятиями. Обычно абстрактное описание определяет начальные принципы модельного построения (основные аппроксимации, дефиниционные области переменных, критерии эффективности и типы ожидаемых результатов). На этой стадии можно определить следующие подэтапы:

1. *Дефинирование и анализ поставленной задачи.* Включает ясно определенную сущность задачи исследования и планирование необходимых мероприятий. На базе анализа проблемы определяется объем предполагаемых действий и необходимость декомпозиции задачи.

2. *Уточнение типа начальной информации.* Эта информация позволяет получить корректные выходные результаты моделирования и поэтому надо обеспечить необходимый уровень достоверности оценок.
3. *Введение допущений и гипотез.* Это необходимо при отсутствии достаточной информации для реализации модели. Допущения заменяют отсутствующие данные, либо отсутствие их полноты. Гипотезы относятся к типу возможных результатов или к среде реализации исследуемых процессов. В процессе моделирования эти гипотезы и допущения могут быть приняты, отброшены, либо модифицированы.
4. *Определение основного содержания модели.* На базе применяемого метода моделирования отчитывается особенность реального объекта, поставленная задача и средства ее решения. Результаты этого подэтапа включают формулировку базовой концепции модели, формализованное описание реальных процессов и выбор подходящей аппроксимацией.
5. *Определение модельных параметров и выбор критериев эффективности.* На этом подэтапе определяются первичные и вторичные факторы, входные воздействия и ожидаемые на выходе реакции модели, что является особенно важным для достижения необходимой точности математического описания. Уточнение критериев эффективности связано с дефинированием функциональных зависимостей оценки реакции системы при изменении модельных параметров.
6. *Абстрактное описание модели.* Фаза общего формулирования концептуальной модели заканчивает построение абстрактной модели в подходящей среде абстрактных терминов – например, в виде структурной схемы, как потоковой диаграммы (Data Flow Diagram), в виде графической схемы (State Transition Network) и т.д. Это абстрактное представление позволит легко построить математическую модель.

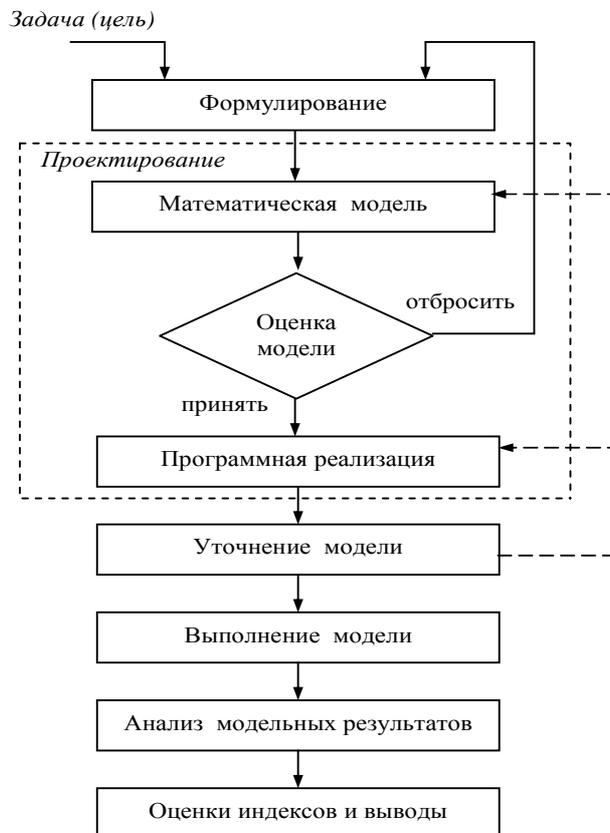


Рис. 5. Методологическая схема модельного исследования.

Стадия “Проектирование”

Проектирование компьютерной модели связано с разработкой математической модели и ее программным описанием.

Математическая модель является представлением структуры исследуемого объекта и протекающих процессов в подходящем математическом виде $Y = \Phi(X, S, A, T)$, где: X – множество внешних воздействий; S – множество системных параметров; A – отражает функциональное поведение (алгоритмы функционирования); T – время работы. Таким образом поведение (реакция) объекта Y моделирует набор функциональных воздействий Φ , представляющих аналитические зависимости (детерминированные или вероятностные). В этом смысле, математическая модель является описанием абстрактной модели средствами выбранной математической системы, оценивая принятые гипотезы и аппроксимации, начальные условия и дефинированные параметры исследования. При разработке математической модели возможно применить известные математические формулы, зависимости или математические законы (например, вероятностные распределения), а также комбинировать и дополнить их. Самые распространенные для цели моделирования теоретические математические системы предоставляют возможность представить математическую модель и в графическом виде – сети Петри [Music, 2003], цепи Маркова [Benveniste, 2003], системы массового обслуживания [Роскес, 2001] и др. На базе определенных на предыдущей стадии критериев, созданную математическую модель необходимо оценить с целью достижения необходимой степени достоверности и адекватности, и после этого можно утвердить или отбросить ее.

Программная модель представляет собой реализацию математического описания программным языком – для этого выбираются подходящие технические и технологические средства. В процессе программной реализации на базе математической модели разрабатывается логическая структурно-функциональная схема модели. Для построения этой схемы можно использовать традиционные блок-схемы, либо графические средства, которые представляются специализированной средой моделирования – как например в GPSS (General Purpose Simulation System) [Stahl, 2001]. Программная реализация модели является задачей разработки программного обеспечения и в этом смысле подчиняется принципам технологии программирования.

Стадия “Уточнение”

Действия этой стадии предназначены для полной валидации проектированной модели и утверждения ее адекватности. Существенное значение для их эффективности имеет оценка текущей адекватности на предыдущих стадиях. В этом смысле процесс уточнения модели надо рассматривать как совокупность распределенных действий на всех предыдущих стадиях компьютерного моделирования. В общем плане, стадию уточнения можно представить как итеративную процедуру (рис. 6), позволяющую последовательную модификацию начальной версии разрабатываемой модели.

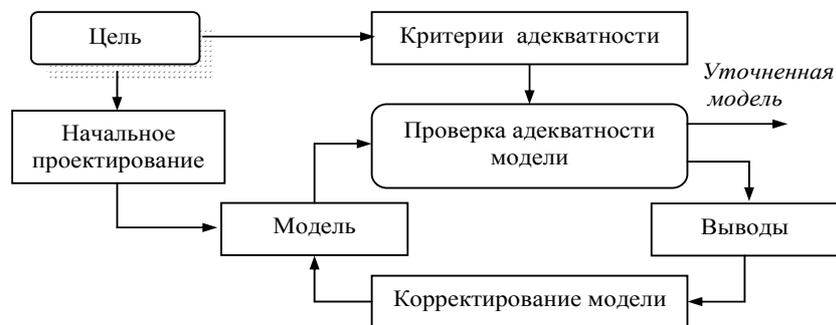


Рис. 6. Итеративная процедура для уточнения модели.

Основной целью проверки модельной достоверности является определение уровня точности соответствия при представлении процессов реального объекта и механизма регистрации модельных результатов. В общем плане, компьютерная

модель представляет совокупность отдельных компонентов и в этом смысле особенно важно правильно планировать проверки адекватности.

Стадия “Выполнение”

Это этап реализации созданной модели (решение численным методом либо выполнение во времени). Самая главная цель – получение максимальной информации для минимальных затрат машинного времени. Предусмотрены два подэтапа:

1. *Планирование модельного эксперимента* – определение значения управляемых факторов и правила регистрации наблюдаемых факторов при выполнении модели. Выбор конкретного плана эксперимента зависит от поставленной цели исследования при оптимизации времени выполнения. Для получения эффективного плана обычно применяются статистические методы (полный план, однофакторный план, рандомизированный план и т.д.), позволяющие удалить совместное влияние наблюдаемых факторов и оценить допустимую экспериментальную ошибку.
2. *Реализация эксперимента* – подготовка входных данных, компьютерная реализация экспериментального плана и сохранение экспериментальных результатов. Реализацию эксперимента можно выполнить следующим образом: контрольное моделирование (для проверки работоспособности и чувствительности модели и оценки модельного времени); рабочее моделирование (действительная реализация разработанного плана эксперимента).

Стадия “Анализ и интерпретация модельных результатов”

При реализации плана модельного эксперимента накапливается информация (результаты моделирования), которую необходимо анализировать для получения оценки и выводов о поведении исследуемого объекта. Это определяет два аспекта – выбор методов для анализа экспериментальной информации и применение подходящих способов интерпретации полученных оценок. Последнее особенно важно для формирования корректных выводов исследования. В смысле первого аспекта обычно применяют статистические методы – дискриптивные анализы (подсчет граничных значений параметров, математического ожидания, дисперсии и средне-квадратической ошибки; определение расщепления для выбранного фактора; вычисление гистограммы и др.); корреляционный анализ (определение уровня факторной взаимосвязи); регрессионный анализ (исследование причинной взаимосвязи в группе факторов); дисперсионный анализ (для установливания относительного влияния определенных факторов на базе экспериментальных результатов).

Результаты анализа модельных данных можно представить в числовом или табличном виде, при помощи графических зависимостей, диаграмм, гистограмм и пр. Чтобы выбрать подходящие графические средства существенное значение имеет использованный метод анализа, а также субъективные умения экспериментатора для оформления результатов эксперимента.

Заключение

Основная цель организации каждого модельного эксперимента – это реализация эффективного моделирования. Она связывается с машинным временем – значительный объем обработки в модели повышает стоимость моделирования и понижает эффективность. Для эффективности исследования существенное значение имеет быстрая валидизация модели и достижение сходимости. Для каждой реальной системы часто приходится создавать множество разных моделей, отличающихся способом декомпозиции и уровнем детализации, методом моделирования, средствами программной реализации и т.д. В процессе выбора оптимального варианта только оценка точности и адекватности является недостаточной. Из множества сходимых моделей надо выбрать самый эффективный вариант, который тратит минимальное время на реализацию.

Существенное значение для достижения достаточной эффективности модели имеет и применяемый язык программной реализации, а также полнота формальной системы абстрактного представления концептуальной модели, простота терминов описания, разработывание оптимального плана и др. Применение универсальных программных систем отличается отсутствием специфических языковых операторов и поэтому они являются подходящими прежде всего для аналитического моделирования. Для реализации симуляционных моделей удачно использовать специализированные языковые среды.

Литература

- [**Брююль 2002**] Брююль А. SPSS: искусство обработки информации. Анализ статистических данных. Санкт Петербург: ДиаСофт, 2002, – 608 с.
- [**Романски, 2001**] Романски Р. Математическое моделирование и исследование стохастических временных характеристик процессов компьютерной обработки данных // Информационные технологии. – Москва, Россия, 2001, No 2, - С. 51 - 55.
- [**Arons, 2000**] Arons H., van Asperen E. Computer assistance for model definition // Proceedings of the 32nd Winter Simulation Conference. – Florida, USA, December 2000. - P. 399-408.
- [**Benveniste, 2003**] Benveniste A., Fabre E., Haar St. Markov nets: probabilistic models for distributed and concurrent systems // IEEE Transactions on Automatic Control. November 2003, vol. 48, No 11. - P. 1936-1950.
- [**Butler, 2001**] Butler J.E., Brockman J. B. A Web-based learning tool that simulates a simple computer architecture // ACM SIGCSE Bulletin. June 2001, vol. 33, No. 2. - P. 47-50.
- [**Crosbie, 2000**] Crosbie R. E. A model curriculum in modeling and simulation: Do we need it? Can we do it? // Proceedings of the 32nd Winter Simulation Conference. December 2000. -P. 1666-1668.
- [**Fabre, 2002**] Fabre E., Pigourier V. Monitoring distributed systems with distributed algorithms // Proceedings of the 41st IEEE Conference on Decision and Control. – vol. 1. 10-13 December 2002 - P. 411-416.
- [**Ibbett, 2002**] Ibbett R.N. WWW Visualisation of Computer Architecture Simulations // Proceedings of the 7th Annual Conf. on Innovation and Technology in Computer Science Education. June 2002. - P. 247.
- [**Lilja, 2001**] Lilja D.J. Comparing Instructional Delivery Methods for Teaching Computer Systems Performance Analysis // IEEE Trans. on Education. February 2001, vol. 44, No 1, - P. 35-40.
- [**Music, 2003**] Music G., Zupancic B., Matko D. Petri net based modeling and supervisory control design in Matlab // Proceedings of the IEEE Conference EUROCON 2003 "Computers as a Tool". – vol. 1. 22-24 Sept. 2003. – Slovenia. - P. 362-366.
- [**Pandey, 2003**] Pandey S., Ramamritham K., Chakrabarti S. Monitoring the dynamic Web to respond to continuous queries // Proceedings of the 12th International Conference on World Wide Web. – Hungary, May 2003, - P. 659-668.
- [**Pockec, 2001**] Pockec P., Mardini W. Modeling with queues: an empirical study // Proceedings of the Canadian Conference on Electrical and Computer Engineering. – vol. 1. 13-16 May 2001. - P. 685-689.
- [**Romansky, 2002**] Romansky R. et all. An Organization of Informational Network InfoNet for Distributed e-Learning // Proceedings of the 3rd International Conference on Computer Systems and Technologies (e-Learning). 20-21 June 2002. Sofia, Bulgaria. - P. IV.4-1 – IV.4-6.
- [**Sargent, 2003**] Sargent R.G. Verification and validation of simulation models // Proceedings of the 2003 Winter Simulation Conference. – vol. 1. 7-10 December 2003. - P. 27-48.
- [**Stahl, 2001**] Stahl, I. GPSS: 40 years of development // Proceedings of the 33rd Winter Simulation Conference. December 2001. - P. 577-585.
- [**Ye, 2002**] Ye D, Xiaofei Xu, Yuliu Chen. Integrated modeling methodology for virtual enterprises // Proceedings of the 10th Conference on Computers, Communications, Control and Power Engineering. – vol. 3. October 2002. - P. 1603-1606.