С О Д Е Р Ж А Н И Е

ТЕМА 6. Моделирование сложных систем на основе аппарата сетей Петри.

6.1. Сеть Петри - как математическая структура и ориентиро-ванный граф.
6.2. Понятие состояния сетей Петри. Маркировка сетей Петри.
6.3. События, запуск переходов и выполнение сетей Петри.
6.4. Выполнение сетей Петри на основе решения матричных уравнений.
6.5. Свойства сетей Петри.
6.6. Дерево достижимости сетей Петри. Алгоритм построения дерева.

КОНСПЕКТ ЛЕКЦИЙ
ПО МОДЕЛИРОВАНИЮ НА ЭВМ

ТЕМА: МОДЕЛИРОВАНИЕ - ВСЕОБЩИЙ МЕТОД ПОЗНАНИЯ ОКРУЖАЮЩЕГО МИРА

Моделирование - метод исследования /познания/ окружающего мира, в котором некоторому изучаемому явлению поставлено в соответствие модель в виде также объекта, явления, процесса, которое может заменить натуру в процессе исследований.

Модель отражает некоторые свойства натуры, но всегда отличается от нее.

Фокус информативности есть модель.

В зависимости от целей моделирования могут быть выбраны разные модели для одного и того же объекта.

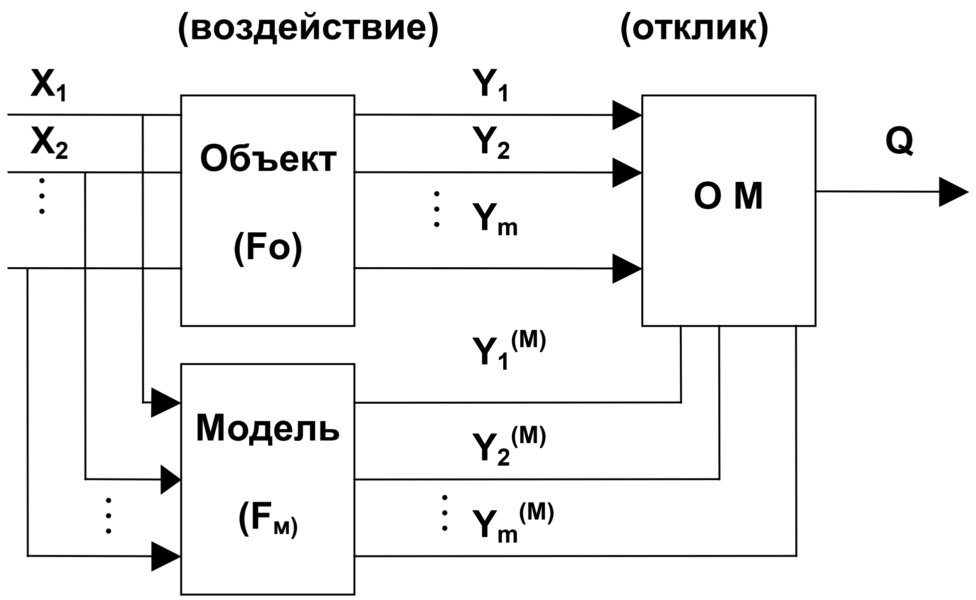
Моделирование используется для двух основных целей:

1. расширение наших знаний об окружающем мире;
2. разработка эффективных производственных процессов /проектирование механизмов, приборов, технологий, процессов/.

В моделировании очень важна роль модельного эксперимента. Для определения достоверности и выявления лучших экспериментов нужно много времени, причем очевидным является планируемость экспериментов /сокращается время исследования/. Существует два общих подхода к моделированию:

1. Классический /теоретико-аналитический/. Предполагается, что исследуемый объект, явление или процесс имеет строгое математическое описание, например оператор **Fo**. В процессе проектирования мы конструируем оператор **Fм**, он должен приближаться к реально существующему $\mathbf{F_0}$.

Теоретико-аналитический подход позволяет сконструировать несколько операторов **F′м, F″м,..., $\mathbf{F^{(n)}}$м**, каждый из которых может быть использован на том или ином этапе исследования. Степень приближения **Fм** к $\mathbf{F_0}$ определяется некоторым критериям соответствия, которые в целом называются функциями потерь /**Q**/. Одним из них является сумма среднеквадратичных отклонений. Для оценки этого показателя используются результаты многих экспериментов.



$$Q(t) = \sum_{i,t} \left| Y_i - Y_i^M \right|^2$$

Лучшей будет модель, где меньше потеря критериальных отклонений.

2. Современный /экспериментально-статистический/. Выделяет эксперимент как ведущее средство моделирования. Может применяться к особо сложным системам, которые могут и не иметь строгого математического описания $\mathbf{F_0}$.

Основным методом подхода является имитационное моделирование. В нем предполагается, что любой исследуемый объект может быть разделен на некоторое число компонентов, некоторые из них могут иметь математическое описание, а функционирование других компонентов может быть неизвестно. Т.о. осуществляется переход от общего к системе, а сама система неразрывно связана с внешней средой.

В экспериментально-статистическом подходе весь процесс исследования подразделяется на 2 этапа:

1/ Сбор статистики по воздействию на систему из внешней среды и реакции системы на эти воздействия, характеризующиеся значимыми отношениями: $\mathbf{R_1^*...R_5^*}$. Определяется также значимые свойства этих отношений $\mathbf{P_1^*...P_5^*}$.

Практически при имитационном моделировании взаимодействие системы со средой характеризуется конкретными моделями информационных потоков. Эти модели составляются в конце первого этапа в результате обработки статистической информации.

На втором этапе моделирования изменяется структура системы или параметры отдельных компонентов с целью определения лучшей организации. При моделировании генерируются входные потоки, отражающие внешнюю среду и исследуем выходные потоки, а также некоторые интегральные характеристики оценки качества работы системы.

Например:. среднее время обслуживания в системе, время реакции системы на конкретное задание, коэффициент простоя прибора в системе, вероятность появления отказа в обследовании заявки, коэффициент риска получения отказа и т.д.

Имитационная модель является планировщиком работы системы, особенно если изменяется, например, число приборов, или есть необходимость их изменения в связи с изменением числа заданий.

Преимущество экспериментально-статистического подхода является возможности решения достаточно сложных задач в проектировании и организации производственных процессов, когда сама система берет на себя функции сбора и обработки информации, а также планирования. В процессе моделирования сама модель может уточняться.

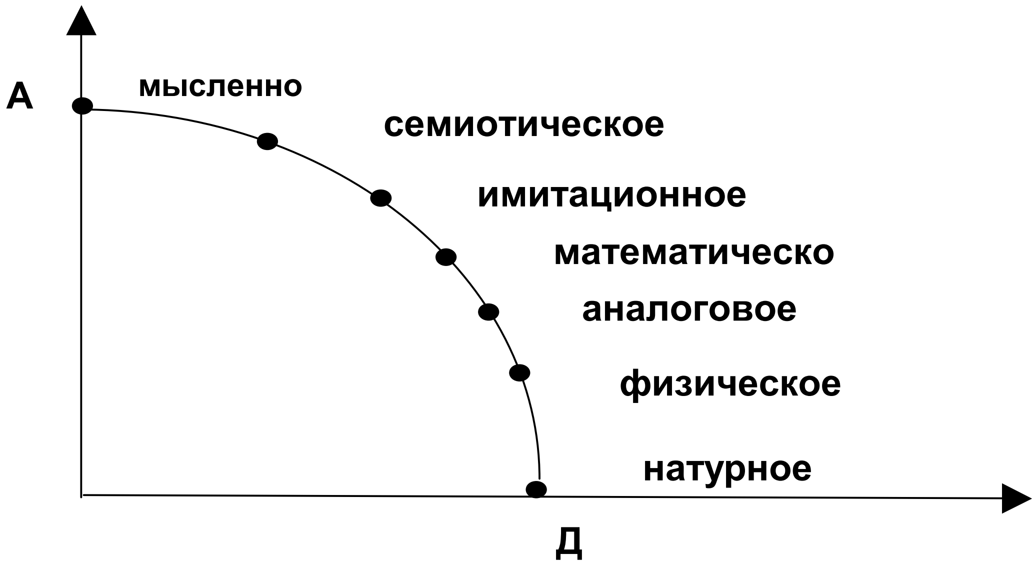
ТЕМА: ОБЩАЯ ХАРАКТЕРИСТИКА МЕТОДОВ И СРЕДСТВ МОДЕЛИРОВАНИЯ

2.1. Соотношение абстрактного и количественного при моделировании

Общей теории моделирования нет, поскольку в различной области знаний пользуются своими моделями. Тем не менее методы в научных исследованиях делятся на 7 основных:

1. мысленное моделирование
2. натурное
3. физическое
4. аналоговое
5. математическое
6. имитационное
7. семиотическое } современные, развивающиеся

Перечисленные методы позволяют исследовать изучаемые объекты, явления и процессы, с разной степенью абстракции А /приближения к мысленным представлениям/ и с разным уровнем достоверности Д, как достоверности измерений на модели. Достоверность измерений определяется отношением погрешности измерений в натуре к погрешностям измерений в модели. Если в модели можно измерить некоторые параметры в любой момент времени как в натуре, то такая модель достоверна. Приборы должны быть одного и того же класса точности. Перечисленные методы могут быть представлены в пространственной абстракции и достоверности.



2.2. Мысленное и натурное моделирование - основа научной и производственной деятельности человека.

Мысленное моделирование - широкая область, т.к. в мысленных моделях представлены наши знания.

Мысленные или познавательные модели /знания/ формируются как суммы мыслей образов всего человечества об изучаемом явлении, объекте и процессе.

К мысленным моделям относятся:

1. чувственно-наглядные
2. символьно-знаковые
3. математически-мысленные.

I - на основе интуитивных представлений, того что мы не можем видеть или принимаем за эталоны, образцы мысленных образов /модель атома Резерфорда, музыкальные произведения, и т.д./.

Эталон это уже модель /поведения и т.д./.

2 - с помощью условных знаков и обозначений позволяют представить структуру и организацию изучаемых явлений, объектов и процессов.

3 - с помощью условных знаков и символов отражаются строгие законы взаимодействия, имеющие место в оригинале. для их представления используются строгие математические теории.

До тех пор пока программа остается на бумаге-носителе она остается мысленной математической моделью. И только когда она "работает" на машине и дает решение модели она становится вещественной агрегатной моделью.

Наибольшей достоверностью обладает натурное моделирование, когда модельный эксперимент проводится непосредственно на изучаемом объекте, явлении или процессе.

Натурные модели:

1. натурный или производственный эксперимент
2. обобщенный производственный опыт
3. среднестатистические данные о явлениях в натуре.

В натурном эксперименте объект подвергается специальному испытанию. Он может быть исключен из производственного цикла.

Обобщенный производственный опыт позволяет создать некоторую эталонную модель производственного процесса, технологической установки, которые будут использоваться для сравнения при проектировании. Аналогичным образом в качестве эталонов модели используются среднестатистические данные о явлениях природы.

Физическое и аналоговое моделирование - сфера приложения теории подобия.

Физическое моделирование предусматривает использование модели одной физической природы с оригиналом. При этом имеет место обычно введение масштабов для линейных размеров, временной переменной и т.д.. Физическая модель отличается размерами от оригиналов. Размеры модели упрощают исследования, повышают наглядность и возможности вариации параметров модели.

Физические модели:

1. пространственные физические модели или компоновки
2. временные электро-динамические модели
3. пространственно-временные физические модели.

I-е в строительстве, учебном деле при конструировании РЭА.

2-е используется для исследования переходных процессов в электрических цепях. В качестве модели используются расчетные столы, на которых коммутируются электрические цепи, имитируется распределение параметров электрических цепей и позволяющее решать задачи управления энергосистемой.

3-е для исследования особо сложных систем, там где нет полного описания процесса /модели радиолокационных станций/.

Недостатки физического моделирования: высокие затраты на модели, сложность изменения параметров модели.

Преимущество: высокая достоверность полученных результатов исследований, поскольку воспроизводятся все процессы, которые имеют место в физическом исходном объекте.

Более широкие возможности у аналогового моделирования, когда исследование проводится на объектах я явлениях другой физической природы, чем природа в оригинале. В качестве модели используется электрические цепи, электрические процессы, для которых хорошо изучены и развиты методы электрических измерений. В качестве оригиналов выступает тепловая система, механические процессы, магнитные поля и т.д.

К аналоговым моделям относятся **R,L,C** цепи, моделирующие установки, электрические сетки и другие сеточные модели.

К аналоговым моделям относятся **R,L,C** цепи, моделирующие установки, электрические сетки и другие сеточные модели. R,L,C цепи используют для моделирования механических, физических систем с сосредоточенными параметрами. Аналогом таких систем является электрическая цепь. Устанавливается строгое соответствие между переменными и параметрами натуры и модели. Соответствие между переменными и параметрами натуры и модели, устанавливается на основе специальных соотношения, которые называются **критериями подобия**. Критерии могут быть сформированы на основе анализа уравнений натуры и модели, но могут и на основе анализа размерностей этих величин. На моделях 1-го типа исследуются разные процессы или строятся механические системы с высоким качеством переходных процессов.

Моделирующие установки используются для исследования потенциальных магнитных полей. В качестве модели аналога в моделирующих установках используется стационарное поле тока в проводящей среде.

Электрические сетки также исследуют потенциальные физические поля, но уже на основе решения разности уравнений. В узлах сетки производится измерение потенциала, которые дают представление о распределении физического поля. Всякие физические поля описываются дифференциальными уравнениями. /Лапласа, Пуассона, Фурье/. Разрешаются цифровые сетки как спецпроцессоры для гибридных схем.

Преимущество аналогового моделирования: возможности вариаций параметрами модели. Точность измерения невысокая, и вариации структурной модели требуют построения новых моделирующих установок. Более современный - метод математического моделирования на ЭВМ.

Математическое моделирование - сфера широкого применения математических методов и ЭВМ

Используется для исследования объектов, явлений и процессов. Применяется при накоплении определенных знаний об объекте. Если же явление не изучено, то применить математическое моделирование невозможно, тогда физическое или аналоговое. Математическое моделирование использует те или иные описания, решение которых производится на аналоговых или цифровых ВМ. К математическому моделированию относятся:

1. структурные математические моделирования /аналоговое и цифровое/

Аналогового типа решается на АЭВМ, причем структура самой модели /вычислительных средств/ повторяет структуру решаемой зависимости. В таких моделях используется структурный метод программирования, когда отдельные операции модели объединяются между собой в некоторую структурную схему решения уравнения. Кроме АВМ относятся также цифровые интегральные машины /ЦИМ/ - для решения дифференциальных уравнений и используют структурные методы программирования.

2. цифровые аналитические /детерминированные и стохастические /

Такие модели решаются на ЦВМ обычно с использованием численных методов решения уравнений и такие методы должны иметь строгое исходное математическое описание.

Детерминированные модели описываются широким набором математических функций и уравнений, решение которых может выполняться с возрастающей точностью.

Аналитические стохастические модели используют известные описания для распределения вероятности случайных переменных, и используя эти формулы в процессе исследования генерируются последовательности значений этих случайных величин. Например так исследуются входные и выходные потоки ОМО. Получаемые значения пара- метров или характеристик оценивают работу системы в статистическом смысле, т.е. определяется среднее значение.

3. математические имитационные модели

Широко используют метод статистических испытаний для решения тех или иных зависимостей, структур - этот метод **МОНТЕ-КАРЛО:** направленно провести эксперимент и существенно сократить время на решение задач моделирования. Решения , полученные на имитационных моделях имеют невысокую точность, однако имеются строгие оценки достоверности полученных результатов и при необходимости, увеличив объем выборок мы можем провести дополнительные испытания и получить более точные значения. При имитационном моделировании законы распределения вероятностей для тех или иных случайных величин могут быть неизвестными, но они уточняются в процессе моделирования, Результатом имитационного моделирования будет являться цифровая аналитическая стохастическая модель объекта. Преимущество имитационного моделирования: средства моделирования позволяют производить сбор и обработку данных и осуществлять прогнозы развития процессов в той или иной внешней среде. Экспериментально-статистический подход реализуется именно с помощью имитационных моделей.

Имитационное и семиотическое моделирование, современные методы исследования сложных систем

Имитационное моделирование реализуется на ЦЭВМ с помощью разработанных моделирующих систем и соответствующих языков имитационного моделирования. В имитационном моделировании существуют некоторые формальные описания или некоторые модели форм систем, которые представляют выбранное базовое описание сложных систем. К имитационным моделям относят:

1. формальную дискретную систему

2. формальную непрерывную систему

3. формальную непрерывно-дискретную систему.

Выбор зависит от целей, действительной сложности самой модели поставленных задач.

Формальной дискретной системой называется система, состояние которой изменяется лишь в некоторые дискретные моменты времени наступления событий, причем каждое событие может привести не только к изменению состояния, но и к изменению параметров системы, алгоритмов функционирования или даже структуры системы. В периоды между событиями состояние не изменяется. Переходы из состояния в состояние происходят мгновенно и нет таких моментов времени, когда состояние ни определено. Типичный пример - комбинационная схема ЭВМ.

Непрерывная формальная система - это система, функционирование которой с некоторого начального момента t_0 описывается системами дифференциальных уравнений, а непрерывное взаимодействие компонентов такой системы представляется в виде реализаций определенных функциональных зависимостей, характеристик состояния, в т.ч. непрерывного времени. Мгновенное состояние непрерывной системы в некоторый момент времени определяется множеством мгновенных значений характеристик состояний, или переменных, которые определены в левых частях системы дифференциальных уравнений.

Состояние непрерывной системы задается вектором состояния в примерной области характеристик состояния. Для непрерывных систем выделяется область допустимых состояний, когда ни одна из характеристик не выходит за пределы допустимых диапазонов.

Чисто непрерывных или чисто дискретных систем не существует, т.к. даже непрерывная система связана с событиями "пуск" и "останов" как минимум.

Более общим представлением является непрерывно-дискретная система, в которой отдельные события могут повлиять на параметры решаемых систем дифференциальных уравнений, могут порождать новые непрерывные процессы или завершить развитие их.

Непрерывно-дискретные системы - это непрерывные системы со вложенной дискретной моделью управления.

Для описания моделей сложных систем используются языки имитационного моделирования, которые ориентированы на некоторые программные готовые средства систем моделирования. Эти языки и системы имитационного моделирования могут быть ориентированы на исследование только дискретных, только непрерывных или непрерывно- дискретных систем.

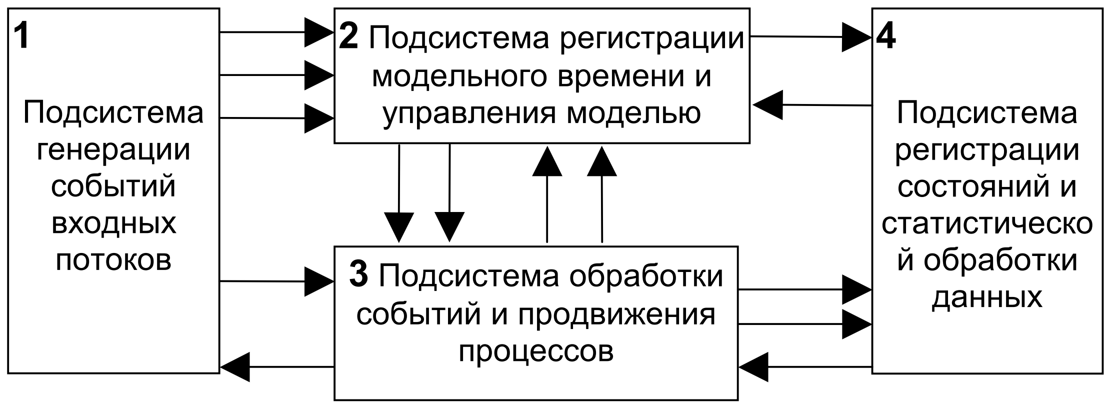
Для моделирования дискретных систем - **GPSS** /многоцелевая система моделирования, работающая в ОС ЕС. Язык **СИМУЛА** , **АЛСИМ** используется для описания дискретных систем АСУ производства.

Для исследования непрерывных систем языки: ДИНАМО, MIM,

которые исследуют динамические процессы в сложных динамических установках.

Для моделирования процессов в непрерывно-дискретных системах используют **НЕДИС, ИМФОР**.

Общая структура систем имитационного моделирования



Для исследования особо сложных систем в целях управления этими системами в настоящее время развивается новый метод семиотичного или знакового моделирования, которое позволяет воспроизвести или симитировать процесс принятия решения человеком. Особо сложные системы - это системы, которые не могут быть описаны аппаратом современной математики, даже в статистическом смысла. В этом случае модели представляются в текстовой форме, например, они описывают ситуацию в производстве, на рынке и т.д.. Они могут описывать в текстовой форме и принятые решения. В принципе эти модели должны описывать сущность некоторых процессов, чтобы промоделировать событие их и принять решение. Семиотические модели используются в системах ситуационного управления, когда некоторый оператор принимает решение в зависимости от ситуации /ЦУП для космических полетов/.

Типы семиотических моделей:

1. лингвистические /языковые/
2. информационно-знаковые /банки данных для принятия решений/
3. семантичные сети /логико-лингвистические модели для принятия решений/.

Они соответствуют трем уровням представления знаний об объекте:

1. декларативный,
2. процедуральный,
3. семантический.

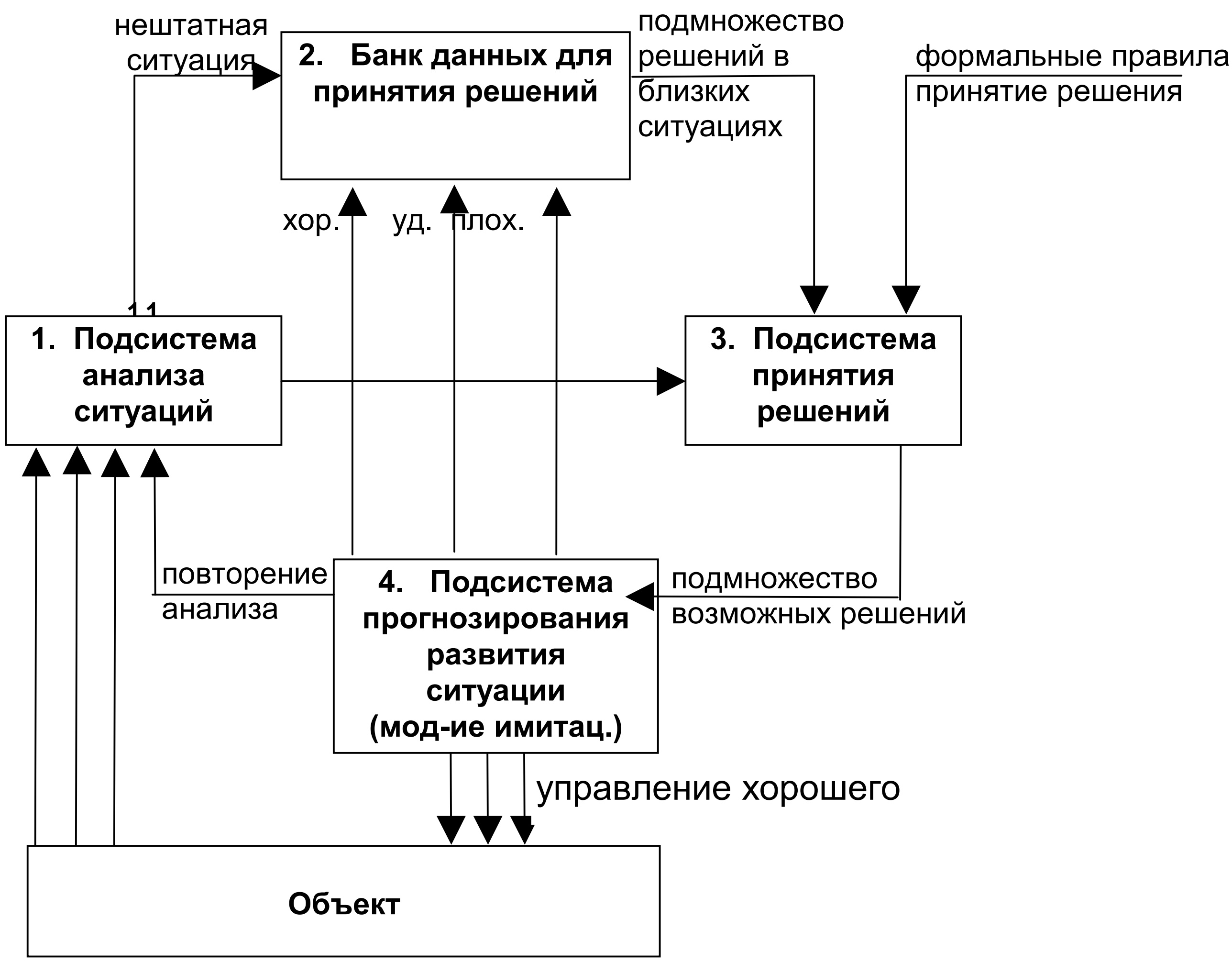
На 1-м уровне знания об объекте представляются в форме лингвистической модели, т.е. некоторые описания ситуаций, которые позволяют с помощью простых логических правил принять решение об управлении объектом. Основа таких моделей конструкция **IF...THEN** /продукция/.

На 2-м уровне используется информационно-знаковые модели в виде специальных организационных баз данных для принятия решении, в которых записывается множество ранее имевших место ситуаций, принятых решений и их последствия, причем база данных имеет средства для выбора "ближайших" ситуаций по отношению к любой текущей ситуации.

На 3-м уровне с помощью специальных сетей в виде мультиграфов описывается содержание (семантика) процессов в системе, которые позволяют на основе формальных представлений /правил/ принять правильное решение. Узлами таких графов являются понятия, а ребрами- отношения между понятиями, т.е. осуществляется переход от текстовых представлений к некоторым графическим описаниям, связывающих понятие. Варианты таких сетей - фреймы. **Фрейм** - минимум описания необходимого для выделения конкретного объекта.

Семиотические модели используются на разных этапах ситуационного управления. Общая структура такой системы управления представляется в виде: /см.рис./

На 1-м этапе ситуационного управления лингвистические модели используются для описания и анализа ситуаций. Логическая обработка этих описаний имеет целью выявление нештатных ситуаций.



На 2-м этапе для текущей ситуации с помощью банка данных для принятия решений формируется подмножество информационно-знаковых моделей поведения, которые ранее имели место в сходных ситуациях.

На основе формальных правил, отраженных в семантичных сетях, принимаются возможные решения, которые сравниваются с рядом решений, полученных из банка данных и вырабатывается ограниченное число лучших решений.

На 4-м этапе производится прогнозирование развития процесса в соответствии с каждым из подходящих решений. Оцениваются последствия решений и вырабатываются их интегральные характеристики. Решение, получившее наивысшую оценку реализуется для управления объектом.

Иерархия моделей информационно-вычислительных систем. /ЭВМ и ИВС./

Для проектирования этих систем появляется необходимость исследования процессов на самых различных уровнях /системном, логическом, физическом/. На каждом таком уровне ЭВМ выступает как комплекс технических средств, блоков, узлов, элементов, причем понятие элемент называется условным, поскольку на каждом более высоком уровне описания в качестве элемента выступает все более сложное устройство, т.е. элементом системы на данном уровне описания выступает некоторый неделимый компонент, функционирование которого не входит в рамки описания работы системы на данном уровне.

Для исследования ЭВМ информационных систем используются 7 уровней описания, но могут быть выделены и некоторые подуровни:

1. системный,
2. программно-алгоритмический или операционно-регистровый,
3 . функционально-автоматный ,
4. логический,
5. схемотехнический,
6. топологический,
7. компонентный или физический.

Эти уровни описания ЭВМ и соответствующие неделимые компоненты удобно представить в виде диаграммы (см. рис.1).

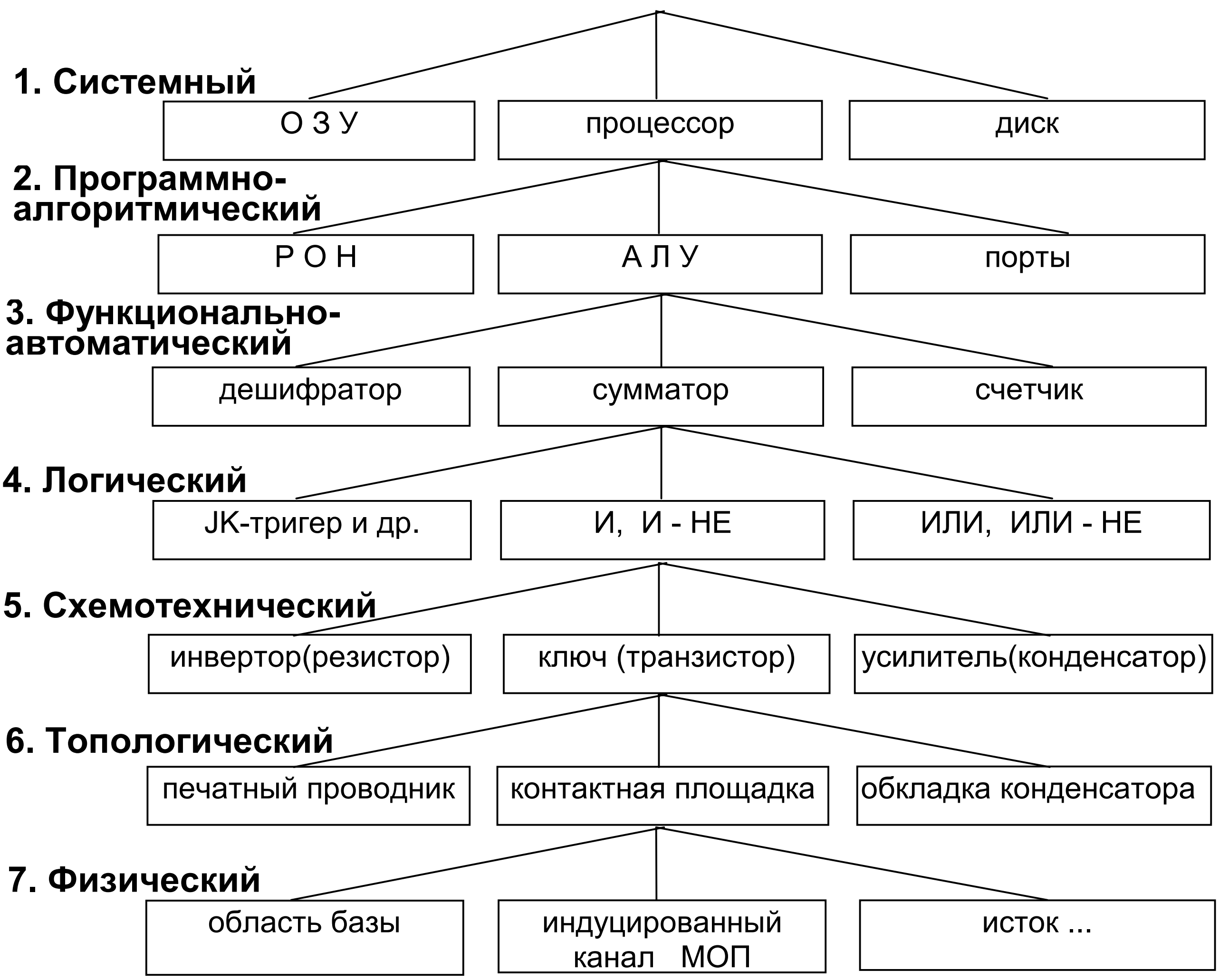
Модели 1-го уровня - системы массового обслуживания и сети с очередями, а также обобщенные динамические системы /ОДС/.

На 2-м уровне используются моделирующие программы /эмуляторы/ работы различных устройств, которые позволяют разработать и использовать программные средства для моделирования и выполнения на ЭВМ /одних на других/.

3-й использует представление устройств, как комплектацию автоматов: операционных и управляющих.

На 4-м используются уже элементарные автоматы или элементы логической схемы. Используются модели КС: статичные и динамичные. В (статике фиксируется только временная диаграмма и кодовые последовательности выходов схем. В динамике - переходные процессы.

На 5-м используют фазовые модели электронных компонентов /транзистора, ключа,.../, в которых в качестве непрерывных фаз переменных выступают токи и напряжения. Электронные системы рассматриваются как дискретные схемы. Цель - исследование АФХ или проектирование заданных характеристик.



6-й связан с переходом к интегральной технологии. Для каждой из интегральных схем необходимо сформировать конфигурацию печатных проводников по слоям. На этом уровне используются различные модели трассировки, позволяющие создать топологию слоя индивидуально на каждом кристалле.

На компонентном физическом уровне исследуются процессы в теле полупроводника. Это передача электрических зарядов в легированных областях и на их границах.

ТЕМА 3: СОСТОЯНИЕ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ ИССЛЕДУЕМЫХ ОБЪЕКТОВ И СИСТЕМ

3.1. Абстрактные математические модели. Теоретико-множественная модель технической системы.

Абстрактная математическая модель - математическая структура/система/, состоящая из множества **M** абстрактных математических объектов/чисел, векторов/ и множества **R** , задающего отношение между двумя и более объектами, т.е. модель задается парой или кортежем <**M,R**>.

$$\mathbf{M^{od}} = \langle \mathbf{M,R} \rangle,$$

$$\mathbf{M} = \{\mathbf{a, b, c,\dots, z}\}, \quad \mathbf{R} = \{\mathbf{r_1, r_2,\dots, r_к}\}.$$

отношения могут быть в виде операций, неравенств, уравнений, функций, объединенных два или более объектов.

$\mathbf{r_i \langle a,b \rangle}$ - бинарное,

$\mathbf{r_i \langle a,b,c \rangle}$ - териарное.

Существует два способа задания математических моделей:

1. задается аксиоматическим определением;
2. задается конструктивным определением.

Вторым способом задания является конструктивное определение. Здесь отношение между математическими объектами в модели задается уже известными определениями, но в качестве объектов такой модели выступает более сложные структуры/или объекты/.

Таким образом некоторая старая математическая модель переносится на новый уровень абстракции. Те же алгебраические определения могут быть перенесены на матрицы, векторы и т.д.. В прикладных математических моделях появляются новые объекты для исследования. В прикладных математических теориях используются конструкторское определение для задания модели. Одно из наиболее часто используемых абстрактных математических моделей является теоретико-множественная модель технической системы. Это базовое формальное описание некоторого объекта, который может быть подразделен на устройства, узлы, компоненты. Такой объект называется **технической системой**.

Можно задать общее определение системы , как связного целого, образованного взаимоподчинением и согласованию делений. Составляющих ее частей и компонентов. На разных уровнях описания в системе всегда можно выделить некоторый неделимый компонент или базовый элемент. В теоретико-множественной модели вводятся 3 первичные категории:

а/множество базовых элементов системы $\mathbf{M} = \{\mathbf{m_1, \dots, m_p}\}$

б/множество отношений этих элементов $\mathbf{R} = \{\mathbf{r_1, \dots r_к}\}$

в/ множество свойств элементов системы, которые могут проявляться во всех возможных отношениях $\mathbf{P} = \{\mathbf{p_1, \dots, p_s}\}$

Теоретико-множественная модель - базовая модель проектирования любой системы. На основе первичных категорий формируются вторичные категории:

а/ множество всех возможных структур системы **Q**

б/ множество всех возможных функций системы **F**

в/ множество всех возможных оценок качества системы Э /С/

Э - эффективность.

$\mathbf{Q = M \times R}$ - декартово произведение.

Из множества всех возможных структур выделяется подмножество подходящих структур $\mathbf{Qn \subset Q}$. Они задаются неопределенном **M** и реализуют требуемые значимые отношения **Rs** - как необходимые связи с внешней средой.

Проектирование заключается в выборе и синтезе подходящих структур. Множество всех возможных систем:

$\mathbf{S = Q \times P}$, **P** - множество всех возможных свойств. Но при применении систем при таком подходе важно выделить что множество подходящих систем использует множество подходящих структур, на которые должны быть обязательно проявлены множество всех значимых свойств **Ps** системы /задается в ТЗ на разработку системы/. Поэтому поиск подходящей системы ведется в классе подходящих структур, для которых $\mathbf{Ps \subset P}$.

При построении математической модели системы мы должны обеспечить переход к заданию всех вторичных категорий, задать функции системы, используя описания, а также задать формально оценки качества системы. Вводится понятие состояния системы, которое задается на множестве элементов, а точнее на множестве переменных, которые задают состояние каждого из элементов.

Мерность описания функционирования системы зависит от числа характеристик, которые учитываются в данном описании. Само же состояние на некоторый момент $\mathbf{t_к}$ определяется вектором $\overline{\mathbf{Z_K}}$, который записывается n-мерной области характеристик состояния. Описание между переменными характеристиками состояния зависят от конкретной структуры из множества подходящих структур, и определяется физическими законами. Чтобы задать состояние системы, необходимо привести все описания в некоторую фиксированную форму. Для дискретных систем общее описание функционирования должно быть представлено в виде:

$$\overline{Z_к} = f\left(\overline{Z_{к-1}},\ \overline{X_к},\ \overline{V_к}\right)$$

$\overline{X_к}$, $\overline{V_к}$ - вектора информационных входов и управляющих входных сигналов.

$\overline{Z_{к-1}}$ - состояние вектора на момент времени $t_{к-1}$.

H - закон функционирования системы /функционирования перехода/ - в виде уравнений, неравенств и т.д.. Простым примером - описание функционирования автомата. Несколько сложнее представляется состояние для непрерывных систем. Вектор состояния для непрерывных систем $\overline{Z}(t)$ представляется множеством мгновенных состояний этого вектора. Этот вектор задается на множестве характеристик состояния, каждая из которых представляет собой переменную, которая характеризует состояние элемента или группы элементов системы. Для расчета системы важно показать, что состояние системы находится в области допустимых состояний, причем эта **n**-мерная область допустимых состояний определяется минимальными и максимальными возможностями знания всех характеристик /рабочий ток срабатывания реле/.

Для непрерывных систем функционирование задается с помощью описания динамики работы системы, т.е. соответствующих дифференциальных уравнений или производных по времени. Тогда состояние системы определяется:

$$\frac{d\,\overline{Z}(t)}{d\,t} = H'\left[\overline{Z}(t), \overline{X}(t), \overline{V}(t)\right]$$

H′ - функция переходов в виде правых частей дифференциального уравнения, связанные между собой характеристики состояний $\overline{Z}(t)$, информационные сигналы $\overline{X}(t)$ и управляющие сигналы $\overline{V}(t)$.

Чтобы отразить реакцию системы на возбуждение, в модели представляются не только векторы состояния, но и все выходные сигналы системы, т.е. реакции системы.

Вектор выходных сигналов дискретной системы на момент времени $t_к$, определяется с помощью функции выходов:

$$\overline{Y}_k = G\left(\overline{Z}_k, \overline{X}_k, \overline{V}_k\right)$$

Это также система алгебраических и логических уравнений.

Аналогично для непрерывной системы выходные сигналы определяются с помощью функции выходов:

$$\overline{Y}(t) = G\left[\overline{Z}(t), \overline{X}(t), \overline{V}(t)\right]$$

Таким образом выбрав функции переходов и выходов для системы мы однозначно задаем функционирование системы:

$$F = \langle H, G \rangle$$

Модель функционирования **F** позволяет нам проверить, действительно ли данная система обеспечивает заданные значимые свойства, т.е. есть ли среди множества структур $\hat{Q}_n$ те, для которых множество:

$$P_s \subset \hat{P}_n$$

$\hat{P}_n$ -подмножество свойств, которые проявляются в подходящих структурах.

Среди важных интегральных характеристик применения системы, которые включаются в P_s, выделение оценки качества:

$$Э = \frac{P}{C} \qquad \frac{\text{/производительность системы/ /задач/час/}}{\text{/стоимость/ /количество корпусов/}}$$

Есть еще информационная производительность систем. Она определяется количеством бит обработки информации в секунду:

$$I = \frac{\text{бит}}{\text{сек}}$$

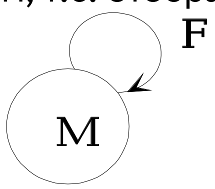
В некоторых случаях достаточно дать одну характеристику - стоимость. Оптимизация при проектировании заключается в том, чтобы из **Qn** выбрать те, которые дают максимум или минимум интегральных показателей.

ПОНЯТИЕ МОРФИЗМОВ В МОДЕЛИРОВАНИИ

При разработке моделей очень важно определить эквивалентные или подобные модели. Для построения таких моделей используются некоторые формальные преобразования, которые называются морфизмами .

Типы : - изоморфизм
- автоморфизм
- гомоморфизм

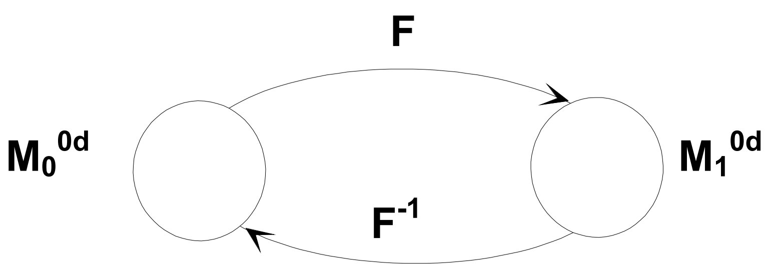
Эти преобразования используются для построения модели в соответствующих методах моделирования. Например, для натурного моделирования - автоморфизм, т.е. отображение самого на себя.



F - автоморфизм

F- некоторое планирование эксперимента, которое позволяет на те же самые элементы модели посмотреть с другой стороны, которая зависит от цели моделирования.

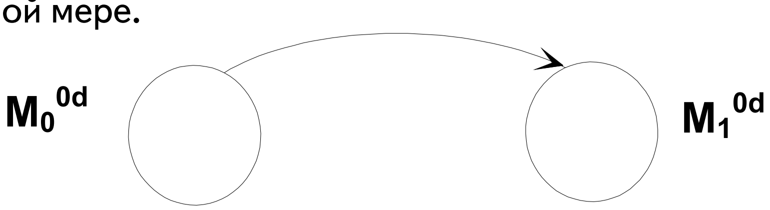
Для физического и аналогового моделирования свойственны изоморфизмы, когда две модели могут быть сопоставлены друг другу взаимооднозначно.



Это значит что $\exists$**F** и обратное $\mathbf{F^{-1}}$.

F называется в этом случае масштабным преобразованием.

Для математического моделирования свойственно однозначное отображение или гомоморфизм, которое позволяет сопоставить некоторую модель некоторой натуре, но обратного не $\exists$ или оно не может быть выполнено в полной мере.



При гомоморфизмах модель оказывается более простой, чем объект. Модель - это часть от целого или некоторое упрощение представления модели.

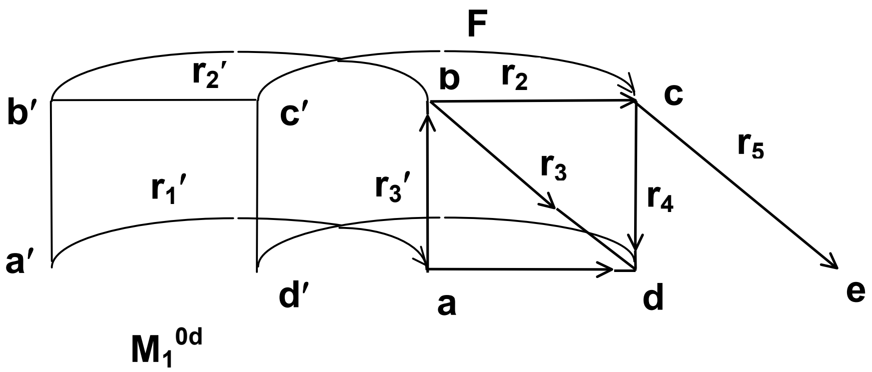
F - гомоморфизмы. Определяется прикладными теориями для каждого объекта в частности.

Существует математическая теория - идентификация, которая позволяет сформировать некоторую упрощенную математическую модель объекта.

Эти отображения задаются на множествах элементов моделей.

$$\mathbf{M_1^{0d}} = <\mathbf{M'}, \{\mathbf{r'_1}, \mathbf{r'_2}, \mathbf{r_3'}, \ldots, \mathbf{r'_к}\}>$$
$$\mathbf{M_0^{0d}} = <\mathbf{M}, \{\mathbf{r_1}, \mathbf{r_2}, \mathbf{r_3}, \ldots, \mathbf{r_к}\}>$$
$$\mathbf{M'} = \{\mathbf{a'}, \mathbf{b'}, \mathbf{c'}, \ldots, \mathbf{z'}\}$$
$$\mathbf{M} = \{\mathbf{a}, \mathbf{b}, \mathbf{c}, \ldots, \mathbf{z}\}$$

Для гомоморфизмов $\exists$ отображение, которое можно представить диаграммой /элементы отображаются узлами структур, а ребра - отношения элементов между собой/.



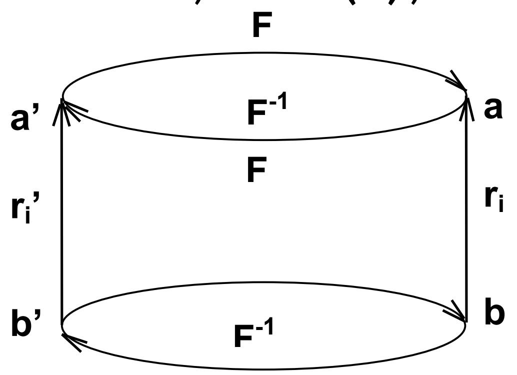
При гомоморфизме обратного отображения нет, поскольку оригинал более полный, чем модель.

2 модели $\mathbf{M_1^{0d}}$ и $\mathbf{M_0^{0d}}$ называются **изоморфными**, если $\exists$ некоторое отображение **F** элементов $\mathbf{M'} \rightarrow \mathbf{M}$, которое допускает обратное отображение $\mathbf{F^{-1}}$ множества элементов $\mathbf{M} \rightarrow \mathbf{M'}$, при котором $\forall$ отношение $\mathbf{r_i}$ задано на элементах множества $\mathbf{M}$ $/\mathbf{r_i} < \mathbf{a}, \mathbf{b} > /$ будет равносильно, т.е. одинаково И. или

Л., соответствующему отношению r_i, заданному на множестве **M'** , т.е. для сходственных отношений.

$r_i < a, b > \Leftrightarrow r'_i < a', b' >$

a* - образы элементов **a′ ;** **a = F (a′) , a′ = F (a)**



При этих отображениях операции между математическими объектами должны остаться равносильными /r_i и r'_i - одно и то же/.

Если дуальные объекты, то операции не соответствуют.

ЭТАПЫ ФОРМАЛИЗАЦИИ ОПИСАНИЙ, ИССЛЕДУЕМЫХ ОБЪЕКТОВ И СИСТЕМ

Построение строгой математической модели представляет сложный многоэтапный процесс. При этом используется ряд форм преобразований и некоторые содержательные преобразования описаний объекта.

Данные описания объекта есть модель. Конечным итогом разработки должна быть экспериментальная расчетная модель объекта, на которой можно производить эксперимент:

- вариации исходных данных/детерминированные воздействия/;
-вариации параметров объектов при воздействии исходных данных ;
- вариации структуры объекта при входных параметров и исходных данных понимается вариация алгоритмов функционирования.

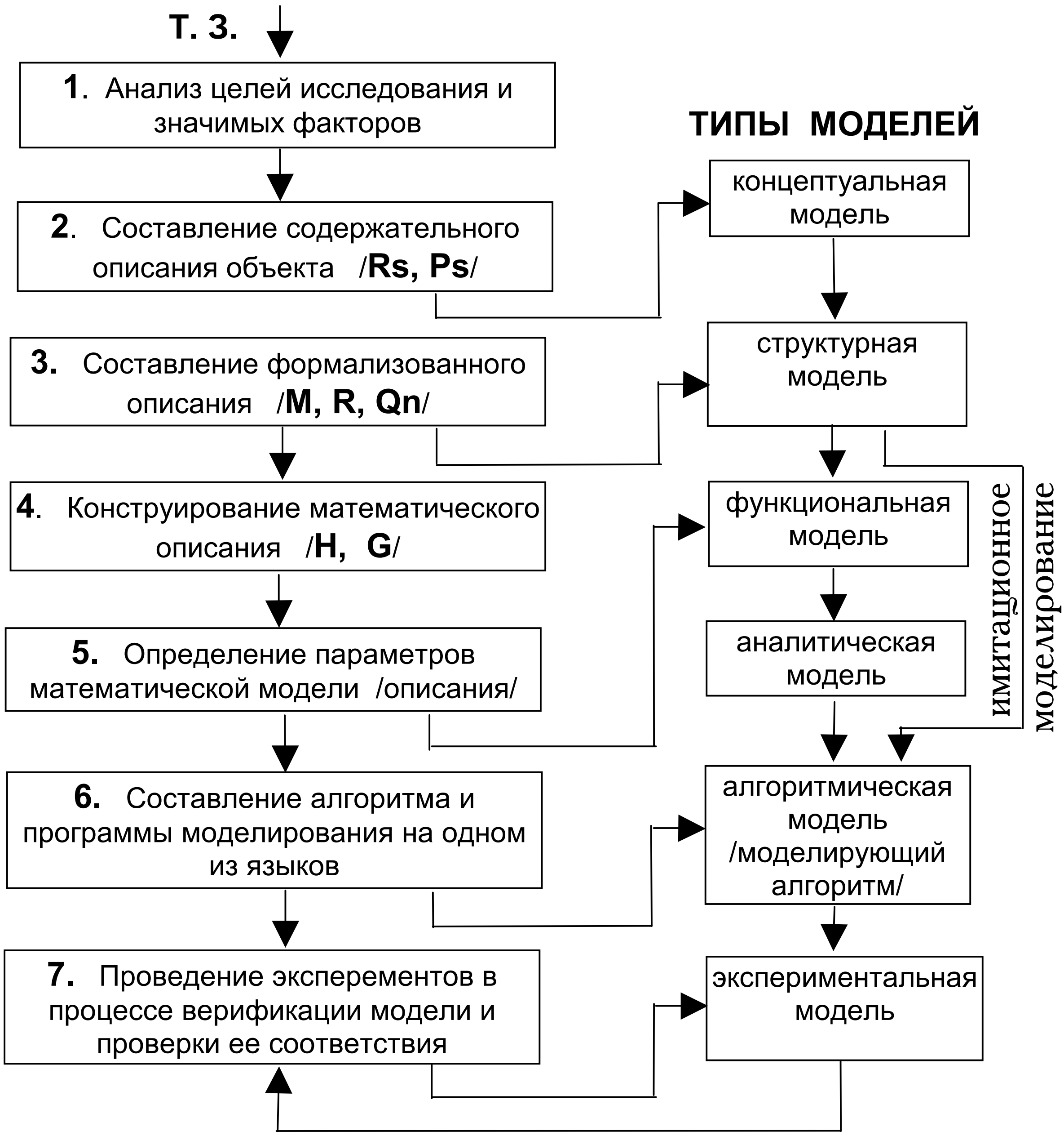
Такие конечные экспериментальные модели работают на ЭВМ.

Все этапы формирования описания выполняются в том случае, если объект новый, неисследованный. Если же объекты уже исследованы, то могут использоваться готовые разработки на тех или иных этапах.

Для разработки любой модели формируется техническое задание, в котором определяются цели построения модели и приводится перечень факторов, влияние которых на объект интерпретируют в первую очередь. При разных целях могут быть, созданы разные модели объекта. В процессе разработки можно выделить 7 этапов, представленных по следующей структуре.

1. Определяется возможность построения модели и анализируется существование модели. Если подходящих моделей нет, то разрабатывается новая модель объекта.

2. Специалисты-эксперты по объекту составляют содержательное описание его поведения, особенности его функционирования, множество входов и выходов, которые составляют значимые отношения объекта с внешней средой, т.е. формируется $\mathbf{P_s}$. $\mathbf{P_s}$ определяется н на основе натурного эксперимента или производственного опыта и часто определяются как желаемые. Такое определение объекта составляется специалистами-экспертами называется концептуальной моделью. Оно представляет собой текстовое описание. Для некоторых объектов только на этом этапе заканчивается составление модели, т.е. в базах знаний в основном будут накапливаться **n** модели объектов.



3. Выделяется множество базовых элементов модели **M** /проектируемого изделия/, на котором задаются всевозможные отношения, т.е. **R** , но очевидно выбор отношения элементов производится целенаправленно с использованием, например содержательного описания. При этом составляются так называемые формализованные описания, или его подходящих структур, представленных в виде структурных схем, функциональных схем, графических схем алгоритмов, кинематических схем и т.п., Функциональное описание, представленное с помощью графических средств и отражающие ее подходящие структуры называется структурной моделью описания. Структурная модель - это более строгая чем кинематическая модель, которая в дальнейшем применяет математический аппарат для описания процессов объекта.

4. Конструирование математической модели. Выделяя множество характеристик состояния и формируя операции между характеристиками состояния. Конечной целью является представление функций переходов и функций выходов, которые с математическим техническим заданием однозначно определяют переходы объекта из одного состояния в другое. С помощью **H, G** составляется функциональная модель. Для задания функции системы используются известные физические законы, положения прикладной теории для данного объекта и для мало- изученного объекта могут использоваться общие положения <u>теории</u> <u>идентификации</u> объекта. Эта теория включает: структурную и параметрическую идентификацию. В 1-ю часть включается множество методов, позволяющих сформировать математическое описание модели, т.е. определить класс уравнений, функций, которые могут использоваться для описания функционирования объекта. В конечном счете функционирование объекта представляется некоторым математическим описанием, которое задает **H** и **G** , а следовательно вектор состояний. Такое формализованное- математическое описание называется <u>функциональной моделью</u>. В функциональной модели пока еще не заданы числовые значения параметров /коэффициентов, функций, нелинейностей и т.д./, которые могли бы обеспечить соответствие процессов в модели и натуре.

5. Выбор и расчет параметров модели с целью соответствия процессов в модели и объекте. Параметры модели обычно определяются путем обработки ряда натурных экспериментов, в которых определены реакции системы на некоторые возмущения. В качестве входных воздействий на объект используется известный испытанный сигнал/гармонический, единичный скачок, единичный импульс, случайный сигнал с заданным распределением/. В конечном итоге формируется параметрическая или аналитическая модель объекта, которая может быть решена различными средствами и получены данные как результата тех или иных экспериментов. Расчетная аналитическая модель необходима для анализа поведения объекта и прогнозирования ситуации на будущее время. Использование модели аффективно для цели управления, но и не менее важно и проектирование /САПР/.

Одним из хороших приемов является аналитическое решение, но его можно выполнить только для простейших случаев. Основным методом решения аналитических моделей является моделирование на ЭВМ.

6. Выбирается численный метод для решения исходных описаний и составляется общий алгоритм решения модели. Такой алгоритм называется моделирующим алгоритмом. В ряде случаев его можно построить , минуя 4 и 5-й этапы. Метод имитационного моделирования позволяет прямо на основе структурной модели составить моделирующий алгоритм и переходить в дальнейшем к эксперименту. Результаты этих экспериментов позволяют уточнить аналитическую модель.

7. Формируется экспериментальная машинная модель в виде отлаженной программы /на языке высокого уровня или в машинных кодах/, которая может быть запущена и выполнена при любых заданных исходных данных, вариации которых определяются планом эксперимента. Экспериментальная модель - та, на которой можно изменив исходные данные можно получить новое решение модели. Дня аналоговых моделей вместо алгоритма модели используется некоторая структура операционных блоков, которые непосредственно составляются на основе аналитической модели.

На 6-м этапе для АВМ используется структурный метод программирования, а экспериментальная модель уже выглядит в виде аппаратуры. На 7-м этапе экспериментальная модель проверяется на достоверность и адекватность/соответствие/ объекту. Процесс про- верки функционирования модели с целью определения ее достоверности, т.е. соответствия исходным предположениям называется процессом модели верификации. В этом процессе надо убедиться, что не существует таких исходных данных, при которых модель вышла бы за рамки существующих предположений о поведении объекта. Существуют критериальные оценки адекватности, а есть и параметрические оценки соответствия, в которых соответствие устанавливается путем сравнения параметров объекта и модели.

НЕКОТОРЫЕ ПРИМЕРЫ РАЗРАБОТКИ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ ПРОСТЕЙШИХ ОБЪЕКТОВ И СИСТЕМ

Первый пример простейшей дискретной системы - модель АПБ

а/ Техническое задание на разработку модели.

Необходимо разработать математическую модель, отражающую состояние автомата ПБ, который бы работал на транспортном средстве и выдавал билеты в соответствии с суммой оплаты, причем оплата может производиться монетами любого достоинства /1, 2, 3, 5/. В дальнейшем предусмотреть реализацию модели в различных исполнениях.

б/ Концептуальная модель. АПБ срабатывает в некоторый дискретный момент времени $\mathbf{t_1}$, $\mathbf{t_2}$,..., $\mathbf{t_{к-1}}$, $\mathbf{t_к}$, наступление которого заранее задать нельзя.

Эти моменты связаны с поступлением входных сигналов и изменением состояния автомата.

Вектор информации входных сигналов включает:

$$\overline{\mathbf{X}} = \{\mathbf{1,2,3,5}\}$$

В таком автомате целесообразно использовать управляющий сигнал, который связан с попыткой получения билета.

$$\overline{\mathbf{V}} = \{\mathbf{0,1}\}$$ в любой из моментов времени $\mathbf{t_к}$.

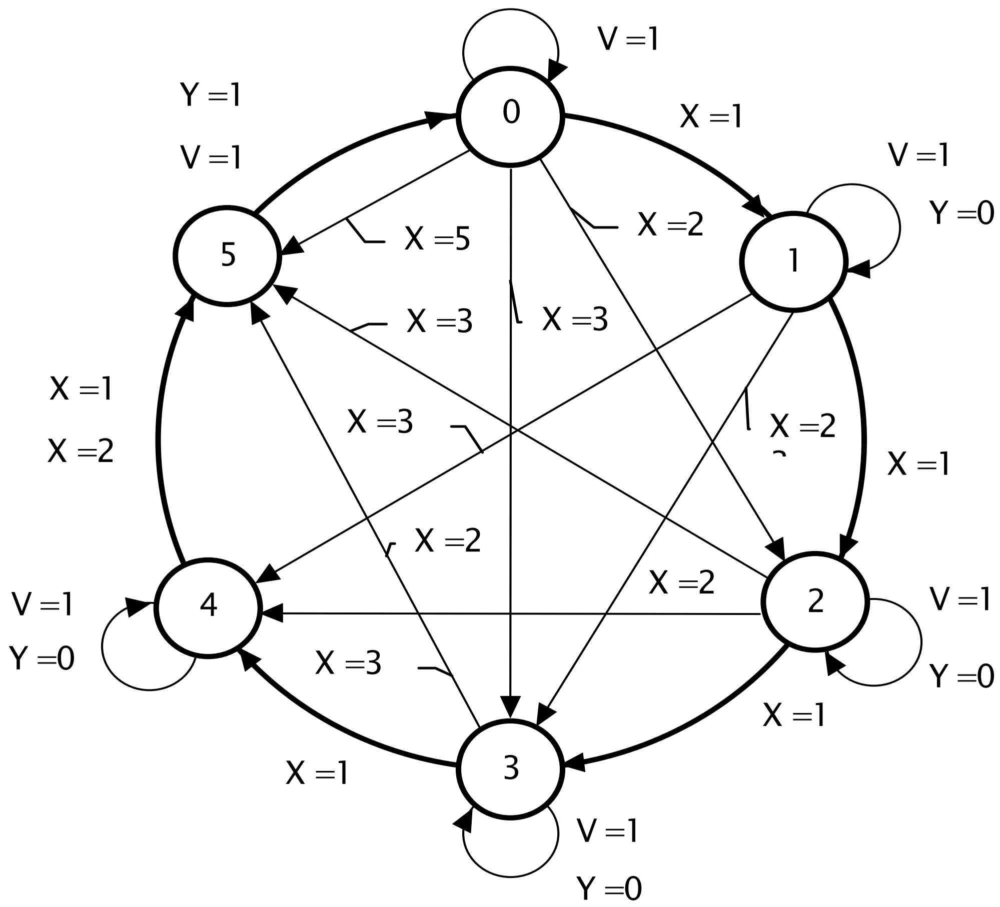
Вектор состояния автомата **Z** связан с промежуточной суммой оплаты за билет:

$$\overline{\mathbf{Z}} = \{\mathbf{0,1,2,3,4,5}\}$$

Выдача будет производиться если $\overline{\mathbf{V}} = \mathbf{1}$, а $\overline{\mathbf{Z}} = \mathbf{5}$. При этом формируется выходной сигнал автомата:

$$\overline{\mathbf{Y}} = \{\mathbf{0,1}\}.$$

в/ <u>Структурная модель</u>. Обязательно связана с графовым представлением. Для данного АПБ его функционирование достаточно однозначно можно представить с помощью графа переходов между состояниями автомата. Отразим на этой структуре все возможные переходы. Данная структурная модель строго и однозначно отражает функционирование АПБ и независимо от типа реализации, эта структурная модель обязательно будет верной.



г/ <u>Функциональная модель</u>. **H** , **G**.

Функционирование может быть определено логическими выражениями, которые должны определить состояние автомата /функции переходов и состояний/.

$$
\underset{(\text{Н})}{Z(t_к)=}\begin{cases} 1)\ \text{если} & Z(t_{к-1})+X(t_к)<5,\ \text{то}\ Z(t_к)=Z(t_{к-1})+X(t_к); \\ 2)\ \text{если} & Z(t_{к-1})+X(t_к)\geq 5\ \text{и}\ V(t_к)=0,\ \text{то}\ Z(t_к)=5; \\ 3)\ \text{если} & Z(t_{к-1})+X(t_к)\geq 5\ \text{и}\ V(t_к)=1,\ \text{то}\ Z(t_к)=0. \end{cases}
$$

Эти модели будут представлены в виде **IF THEN**. В данном описании никак не определены функции выходов, т.к. это функции переходов

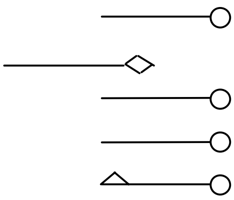
$$
Y(t_к)=\begin{cases} 1)\ \text{если} & Z(t_к)\geq 5\ \text{и}\ V(t_к)=1,\ \text{то}\ Y(t_к)=1; \\ 2)\ \text{иначе} & Y(t_к)=0 \end{cases}
$$

Аналитическую или параметрическую модель целесообразно составить при некотором конкретном задании входных сигналов. Входной сигнал может быть определен как случайный.

ПРИМЕР ПРОСТЕЙШЕЙ НЕПРЕРЫВНОЙ СИСТЕМЫ. МОДЕЛЬ ЭМР

ТЗ: Разработать расчетную модель малогабаритного электромагнитного реле /ЭМР/ для использования в качестве исполненного элемента в электрических схемах путем включения его в плечо триггера или любой логической схемы.

$I_{р\,ср} < 5 \div 10$ mA
$U_п = (3{,}5 \div 5)$ В
$t_{ср} = (20 \div 50)$ мс



Реле должно переключать токи исполнительного механизма, **Iкп = 4-5 А** и иметь один слаботочный разомкнутый контакт.

Концептуальная модель: МГР включает электромагнитную систему из обмотки, сердечника, якоря. В обмотке постоянное напряжение питания, в результате чего якорь занимает определенное состояние и при этом срабатывают контакты. Для возврата якоря и перемещения контактов в малогабаритном реле имеется механическая система: рычагов и якоря, обладающего некоторой массой, возвратной пружины и демпфер - элемент, сухого трения, который выполняет полезную роль в ряде случаев, поскольку гасит колебания в системе.

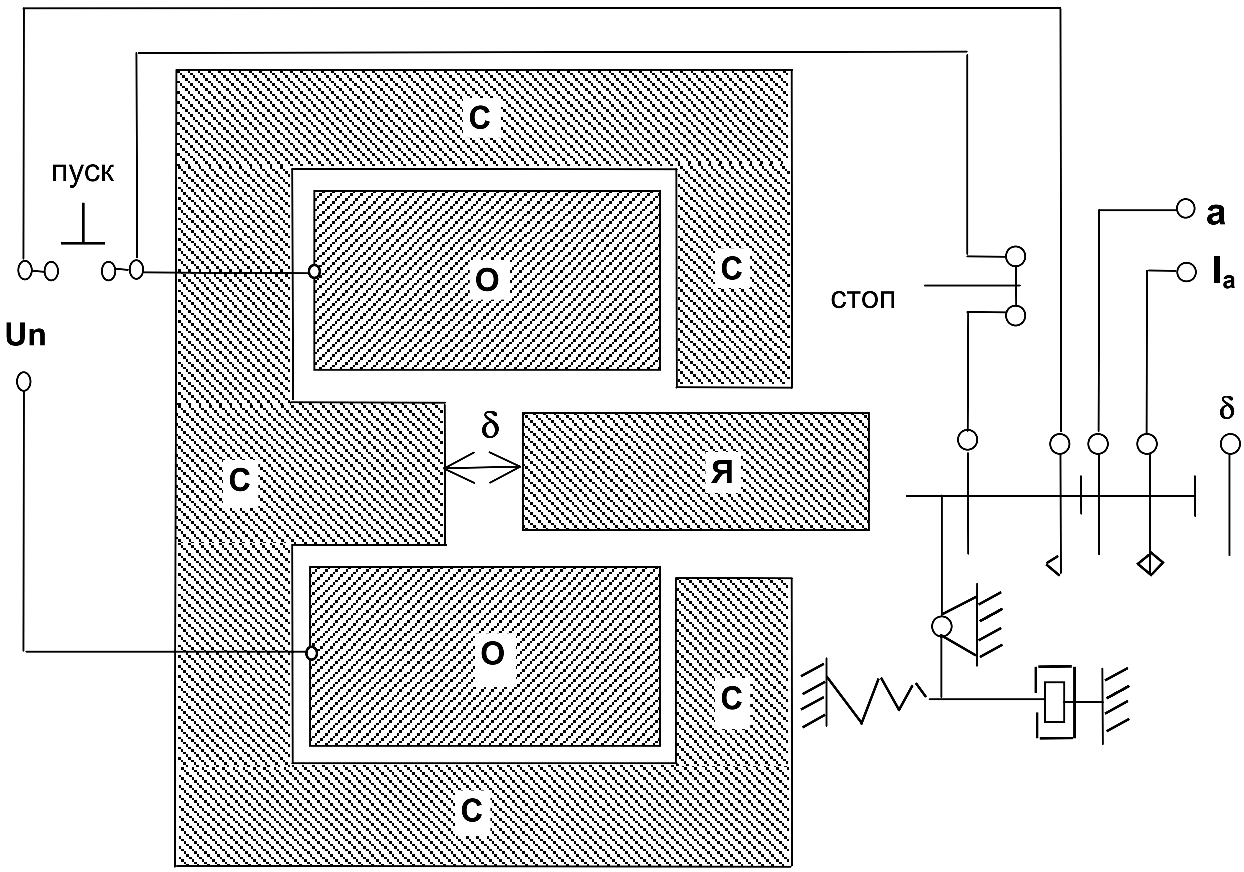
В модели исследуется демпфер обязательно, хотя в реальном реле может и не быть. Это позволяет идеализировать работу всех остальных элементов, выделив характеристику трения. Механическая система связана с контактными группами. Для исполнения эффекта замыкания ЭМР должно быть обеспечено зазором в магнитопроводе не равным нулю.

Перемещение якоря связано с координатой $\mathbf{X}/\mathbf{X}_{min}$ **и** $\mathbf{X}_{max}$**/.** В обмотке имеется исходный рабочий ток, который изменяется с течением времени /$\mathbf{i}_p$ / и должен обеспечить срабатывание.

<u>Структурная модель</u>: Проектируемое реле может быть представлено следующим образом /структурной схемой/ которая сочетает в себе структуру электромагнитной системы в разрезе и кинематическую схему механизма. /см.сл.стр./.

На структурной схеме мы определяем детали в символьной форме, а также сопротивление обмотки, масса якоря, жесткость пружины, коэффициент сухого трения, напряжение питания.

<u>Функциональная модель</u>: При построении такой модели мы ориентируемся на физические законы, которые описывают поведение данной системы.



Динамика механической системы с демпфером и якорем описывается уравнением Д'аламбера:

$$m_я \frac{d^2x}{dt^2} + K_т \frac{dx}{dt} + C_п X = F_{эм}(I, X)$$

момент инерции якоря   линейное трение   закон Гука

Для определения механического состояния системы необходимо иметь две характеристики состояния X и $\frac{dx}{dt} = V$. Воздействующая сила на якорь в электромагнитной системе зависит от тока и положения якоря. Эта зависимость определена экспериментально и получено, что

$$F_{ЭМ}(I, X) = K_{ЭМ} \frac{I^2 x}{x^2}$$

оно справедливо для всех электромагнитных систем в которых есть воздушный зазор.

Динамика ЭМС описывается переходными процессами в индуктивной цепи

$L_0(I, X)\frac{di}{dt} + R_0 i = Un$ - нелинейно, т.к. L зависит от I и x.

Для дальнейшего исследования целесообразно ввести некоторые ограничения, которые приведут к упрощению модели - линеаризация модели.

1. Зазор не равен нулю. Если и магнитная система в рабочем режиме находится в условиях насыщения, то в этом случае можно считать, что индуктивная обмотка не зависит от тока.

L_0 можно представить как функцию $L_0(x)$.

2. Если перемещение якоря Δx окажется значительно меньше чем размеры самого якоря, а следовательно $\Delta x = \varnothing$ тогда $L_0 = const$.

$$L^*_0 \frac{di}{dt} + R_0 i = Uл$$

Нам необходимо выделить функцию перехода системы. Это требует выделения характеристик состояний.

В качестве таковых:

$$\bar{z}(t) = \{x(t), v(t), i(t)\}$$

Для непрерывной системы функцию переходов мы можем записать

$$\frac{d\,\overline{Z}(t)}{dt} = H'\left[\overline{Z}(t), \overline{X}(t), \overline{V}(t)\right]$$ - система диф. уравнений

$$\begin{cases} \dfrac{dx}{dt} = V \\ \dfrac{dV}{dt} = -\dfrac{K_T}{m_я} V - \dfrac{C_n}{m_я} X + \dfrac{K_{эм}}{m_я} \cdot \dfrac{I^2}{x^2} \end{cases}$$ - тогда

$$\frac{di}{dt} = -\frac{R_0}{L_0} i + \frac{1}{L_0} U_n$$

<u>**G** - Функция выходов</u>.

Если бы ЭМР было только исполняющим механизмом, то выходным сигналом реле было бы перемещение, но реле является и переключающим элементом, то в этом случае вектор выходных сигналов $\overline{V}(t)$ может быть представлен кодирующим способом, с помощью кодирования состояния контактов, одним из выходных сигналов является состояние контактов

$$\underset{вкл}{y_1(t)} = \begin{Bmatrix} 1) & a \to 1 \\ 2) & б \to 0 \end{Bmatrix}$$

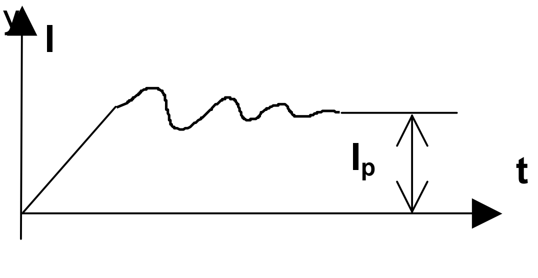
$$\underset{п.р.к}{y_2(t)} = \begin{Bmatrix} 1 & \text{замкнут} \\ 2 & \text{разомкнут} \end{Bmatrix}$$

А условие переходов будет следующее:

если $X_{р.я.} \geq X_{п.с.}$, то $y_1 = 1$ и $y_2 = 1$

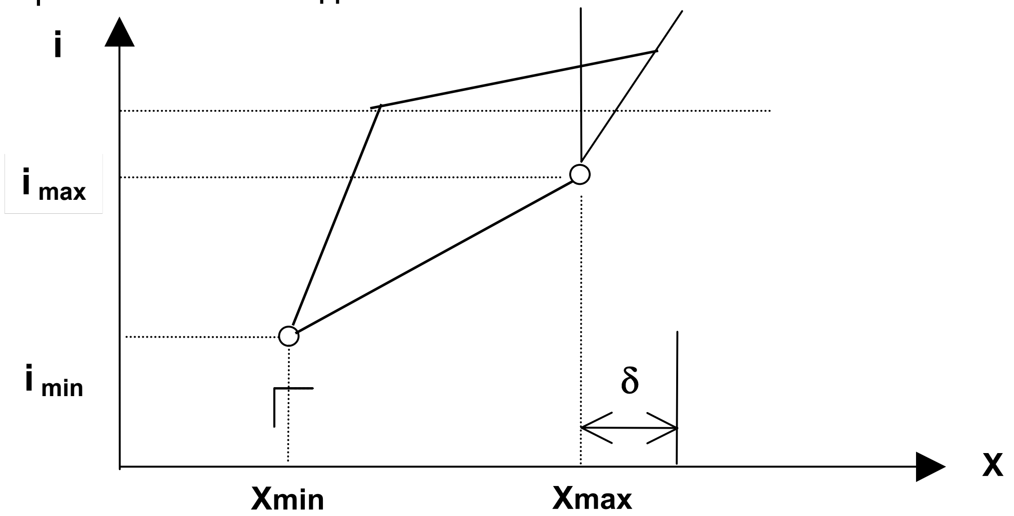
При таком представлении функции **G** очевидно, что ЭМР является простейшей непрерывной дискретной системой.

<u>Аналитическая модель</u>: На этом этапе необходимо рассчитать параметры модели, которые бы обеспечили соответствие. Это соответствие будет определено по исполнении технических требований ТЗ. Определяются $m_я$, $K_т$, $C_п$, $K_{эм}$, L_0, R_0, $U_п$. После выбора параметров на расчетной модели мы должны доказать, что время переключения переходного процесса меньше или равно заданному.



Необходимо выделить множество значимых отношений из значимых свойств объекта и каким образом это обеспечивается.

В конечном итоге проектирования или исследований должно быть обеспечить функционирование реле в области допустимых состояний /зависит надежность и качество/. Это строят на экспериментальной модели.



Для ЭМР область допустимых состояний удобно задавать на плоскости /**I**, **x**/.

МОДЕЛИ ДЕТЕРМИНИРОВАННЫХ И СТОХАСТИЧЕСКИХ ОБЪЕКТОВ

Рассматриваемый объект ЭМР является детерминированным, т. е. его состояние и выходные сигналы могут быть рассчитаны с точностью средств вычислений и метода численного решения уравнений.

Но не для всех объектов возможно такое однозначное вычисление состояний и выходных сигналов. Это стохастические объекты, в которых имеют место различные источники случайных воздействий, и т.о. его состояние можно лишь определить если известно распределение вероятности состояний. Поведение человека непредсказуемо, он стохастический объект.

Детерминированный объект - состояние которого $\bar{z}(t)$ для любого момента времени **t** и выходной сигнал $\bar{Y}(t)$ могут быть однозначно определены и рассчитаны с точностью средств измерения /вычислений/. Детерминированный объект описывается моделями в виде дифференциальных уравнений, алгоритмических уравнений, которые могут быть решены для любого **t** .

Стохастический объект - состояние которого $\bar{z}(t)$ и выходные сигналы $\bar{Y}(t)$ для любого момента времени **t** могут быть заданы лишь распределениями вероятностей состояний на множестве возможных состояние и вероятностями появления тех или иных сигналов на множестве всех возможных сигналов. Таким образом может быть задано наиболее вероятное состояние математического ожидания и соответствующие характеристики.

Стохастические объекты разбиваются на 4 класса по сложности описаний:

<u>1 класс</u> - стохастические объекты со случайными входами

$$p(\bar{x}), p(\bar{v}).$$

<u>2 класс</u> - стохастические объекты со случайными выходами

$$p(\bar{y}).$$

<u>3 класс</u> - стохастические объекты со случайными переходами

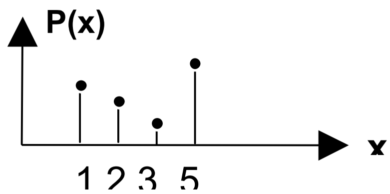
$$(p(\bar{z})).$$

<u>4 класс</u> - стохастические объекты со случайными начальными состояниями $p(\bar{z}_0)$.

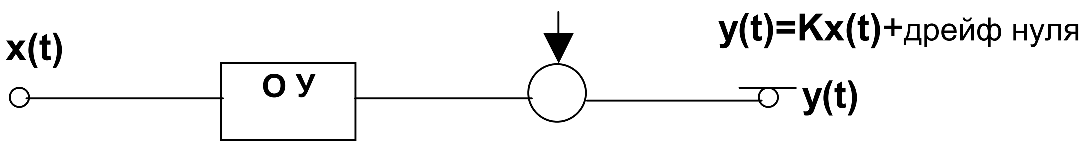
АПБ относятся к первому классу, ОУ - со случайными выходами, система массового обслуживания - 3 класс, 4 класс - объект, в который нельзя влезть.

<u>Примеры стохастических объектов и соответственных моделей</u>

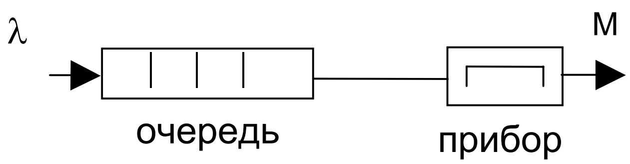
1. 1 класс - АПБ



2. 2 класс - ОУ.



3. 3 класс - СМО.



λ, M - интенсивности входных и выходных информационных потоков.

Это истинно стохастическая система со случайными вероятными входными, выходными сигналами, переходными состояниями, начальными состояниями.

4. Измеритель частиц является объектом со случайным состоянием.

5. Сигналы от объектов, измеряемы в радиолокации и гидроакустике как сигналы с неизвестной начальной фазой.

<u>ТРЕТИЙ ПРИМЕР РАЗРАБОТКИ МОДЕЛИ БЮРО МТА</u>

Техническое задание

Разработать модель истинно стохастического объекта - СМО для оценки качества организации бюро МТА, находящейся внешней среде, т.е. с конкретным адресом.

В модели необходимо учесть такие интегральные показатели качества:

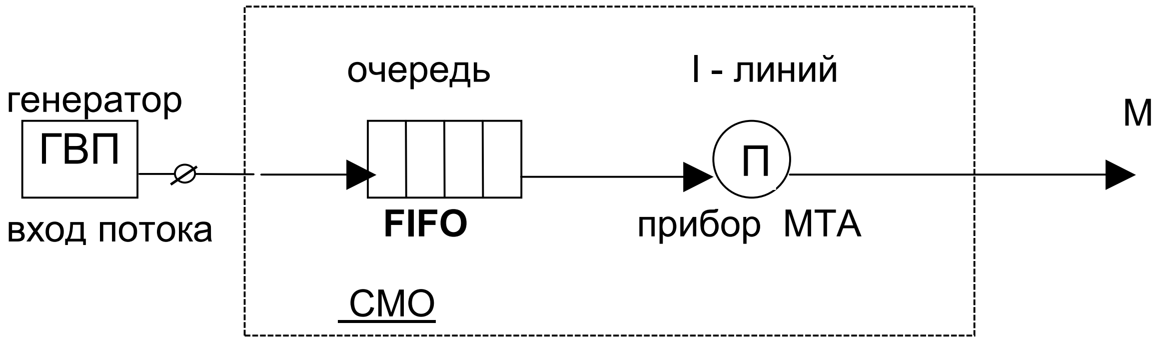
1. Коэффициент простоя приборов.
2. Среднее время пребывания абонента в очереди.
3. Максимальная или средняя длина очереди.

Концептуальная модель.

Бюро состоит из нескольких линий и, **l**- аппаратов, представляющих прибор, и общей очереди абонента на установление связи , БЮРО САМО не позволяет организовать больше очередь, т.е. имеется риск потери абонентов /заявок/. Необходимо определить **l** оптимальное с минимальными коэффициентами простоя аппарата, линий и минимальными потерями заявок. Предполагается, что в течении суток на длительных интервалах времени на входе бюро имеется стационарный поток заявок. Все данные по ТЗ входят в концептуальную модель.

Структурная модель.

С помощью стандартных изображений для СМО бюро МТА можно представить:



Система состоящая из очереди заявок и прибора с обследуемой заявкой / **l** - заявкой/ называется СМО. Входят сюда все определения по концептуальной модели.

Функциональная модель /имитационная/.

Для данного типа объектов создаются имитационные модели, которые позволяют воспроизводить процессы в СМО. Это процессы взаимодействия заявок с ресурсами и между собой.

Имитационная модель строится в 2 этапа.

1 - натурный эксперимент
2 - имитационный эксперимент } обработка данных

Для бюро МТА на 1-м этапе проводится измерительный эксперимент, в котором на некотором интервале **T** /несколько часов/, на котором мы считаем информационные потоки стационарными, производится измерение момента времени входа и выхода абонента.

Каждый входящий - входная заявка, каждый выходящий - обслуживающая заявка. Результаты измерений сводятся -в таблицу эксперимента.

Таблица измерений /натурного/ экспеmремента. I=1 /один автомат/

№ опыта або-нента	Состав-ное время прибы-тия мин.	Интер-вал време-ни между абонен-тами мин.	Время исполь-зования прибора мин.	Время занятия прибо-ра мин.	Время осво-божде-ния прибо-ра мин.	Время нахож дения в сис-теме мин.	Время пребы-вания в очере-ди мин.	Время прос-тоя прибо-бора мин.	Длина очере-ди на мо-мент при-бытия шт.
1	2	3	4	5	6	7	8	9	10
1	00.00	---	3	00.00	00.03	3	0	0	0
2	00.01	1	5	00.03	00.08	7	2	0	1
3	00.04	3	1	00.08	00.09	5	4	0	1
4	00.06	2	7	00.09	00.16	10	3	0	2
5	00.07	1	4	00.16	00.20	13	9	0	3
6	00.12	5	3	00.20	00.23	11	8	0	2
7	00.14	2	3	00.23	00.26	12	9	0	3
8	00.18	4	6	00.26	00.32	14	8	0	3
9	00.21	3	2	00.32	00.34	13	11	0	3
10	00.25	4	2	00.34	00.36	11	9	0	3
11	00.30	5	1	00.36	00.37	7	6	0	3
12	00.34	4	2	00.37	00.39	5	3	0	

Объем данной таблицы велик /200-500 опытов/, что определяется необходимым объемом выборки для получения статистики значимых оценок. Во время измерения фиксируется только 3-я и 4-я колонки, которые отражают экспериментальные значения 2-х независимых случайных величин. Из этой таблицы на основе натурного эксперимента определяются оценки качества обслуживания, которые характеризуют существующую организацию бюро МТА.

<u>ОЦЕНКИ КАЧЕСТВА РАБОТЫ ПРИБОРА</u>

1. Среднее время нахождения в системе

$$T_S = \frac{\sum t_i}{N}; \quad (7)$$

2. Среднее время пребывания в очереди

$$T_Q = \frac{\sum t_i}{N}; \quad (8)$$

3. Коэффициент простоя

$$K = \frac{\sum t_i}{T} \cdot 100\%; \quad (9)$$

T - время наблюдения.

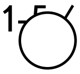
4. Средняя длина очереди

$$\nu_{ср.} = \frac{\sum \nu_i}{N}; \quad (10)$$

По этой же таблице можно провести модельный эксперимент при **I** = 2,3,... и проведем расчет, соответствия времен 5-10 будут меняться, I-1 остаются прежними.

Появятся новые оценки качества работы прибора.

Для более широких исследований можно создать имитационную модель, которая в простейшем случае будет представлять собой некоторый механический генератор случайных воздействий /СВ/, который будет производить 2 потока: поток входных заявок и поток времени обслуживания.

Для интервала между несколькими событиями /1-5/  и для времен обслуживания /1-7/. 

В предположении, что эти интервалы распределяются равномерно, мы используем этот механический генератор из 5 шариков и 7 кубиков. Они помещаются в ящик, из которого в данном опыте извлекается по одному шарику и кубику, потом их возвращаем.

При имитации очень важно, чтобы генератор воспроизводил то распределение, которое в действительности имеет место в потоке.

Для исследования на ЭВМ может использоваться стандартный генератор псевдослучайных чисел с заданными распределениями. А СМО описывается на языке моделирования.

АНАЛИТИЧЕСКИЕ МОДЕЛИ ИНФОРМАЦИОННЫХ ПОТОКОВ.

§1. Интерпретация потоков однородных событий.

Потоки в технических системах могут быть самыми различными. Например: - потоки прикладных программ

- потоки телеграмм и писем
- потоки телеграмм и телефонных сообщений

- потоки пассажиров и грузов.

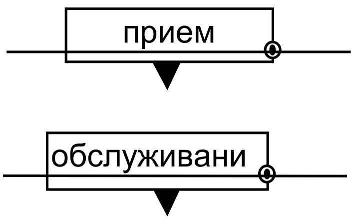
Потоки могут быть входные, и потоки результатов обработки.

Все эти потоки , как входные так и выходные, могут быть представлены единообразно, если ввести понятие события в потоке.

<u>Под событием</u> в информационном потоке понимается появление одной или группы заявок входных или обслуженных за интервал времени $\Delta t \to 0$, который принимается как шаг регистрации системы времени.

Любое сообщение принимается за некоторый интервал времени Δt, но событие приема сообщения фиксируется в конце этого интервала как событие приема данной заявки.

Сообщение об обслуживании может поступать в некоторое Δt и тоже фиксируется.



Шаг регистрации имеет особое значение для системы. Он должен быть, намного меньше чем интервал времени между двумя заявками.

Таким образом информационные потоки представляются как потоки событий, для которых фиксируется в точку момент наступления t_i события. События равноправны или однородны.

Если все заявки, с приемом которых связано событие в потоке, являются равноправными по характеру обслуживания, то такие события называются <u>однородными</u>.

Для задания информационного потока одних событий можно перечислить все моменты времени наступления события.

$$\frac{t_0, t_1, t_2, \ldots, t_k, \ldots}{T}.$$

Эти измерения произведены на некотором интервале **T**. В измерениях присутствует параметр, который не зависит от событий- в потоке; t_o - особая точка, начало интервала наблюдения. Остальные t_i t_i связаны с **i** событием. Запись времен из списка событий иногда называют <u>трассированной записью</u>. Эту запись можно использовать при моделировании, как одну из реализации потока. Но хранить их много нецелесообразно.

Моменты времени /события/ в потоке можно описать соответствующими детерминированными уравнениями.

Однако в реальных условиях подавляющее число потоков является стохастичным, когда события наступят в случайные моменты времени. Для таких потоков удобнее исследовать интервалы времени между событиями в потоке.

Вводим вместо **t** интервалы

$\tau_1, \tau_2, \ldots, \tau_k, \ldots$

$\tau_1 = t_1 - t_0$

$\tau_2 = t_2 - t_1$ и т.д.

Эти интервалы являются конкретными реализациями некоторых случайных величин. В общем случае для нестационарного потока он может быть задан совместной функцией распределения интервалов времени между событиями в потоке.

$\mathbf{F}(\tau_1, \tau_2, \ldots, \tau_k)$ - интегральный закон или функция совместного распределения к случайным величинам.

$$\mathbf{F}(\tau_1, \tau_2, \ldots, \tau_k) = \mathbf{P}\,(\xi_1 < \tau_1, \xi_2 < \tau_2, \ldots, \xi_k < \tau_k),$$

где ξ_k - экспериментальные значения СВ, которые могут быть получены в опытах; τ- линейно изменяемая переменная, по которой строится распределение.

Поток однородных событий называется рекуррентным или с ограниченным последствием, если интервалы времени между событиями в потоке-независимые случайные величины /СВ/. В этом случае функция плотности совместного распределения вероятности, т.е. дифференциальный закон представляется в виде

$$\mathbf{f}(\tau_1, \tau_2, \ldots, \tau_k) = \mathbf{f}(\tau_1) \cdot \mathbf{f}(\tau_2) \cdot \ldots \cdot \mathbf{f}(\tau_k).$$

Среди рекуррентных потоков для исследования в основном используются стационарные рекуррентные потоки.

Поток однородных событий называется стационарным, если распределение вероятности интервалов времени между любыми событиями в потоке одинаковы при **k≥2**, т.е. для стационарного рекуррентного потока можно записать

$$\mathbf{f}(\tau_2) = \mathbf{f}(\tau_3) = \ldots = \mathbf{f}(\tau_k) = \mathbf{f}(\tau)\cdot$$

$\tau_2, \tau_3, \ldots$ - есть некоторые реализации величины τ .

Исключение составляет начальный интервал: $\mathbf{f}(\tau_1)$, поскольку он задан $\mathbf{t_0}, \mathbf{t_1}$.

Стационарный поток однородных событий обладает свойством отсутствия последствия, если имеет место равенство:

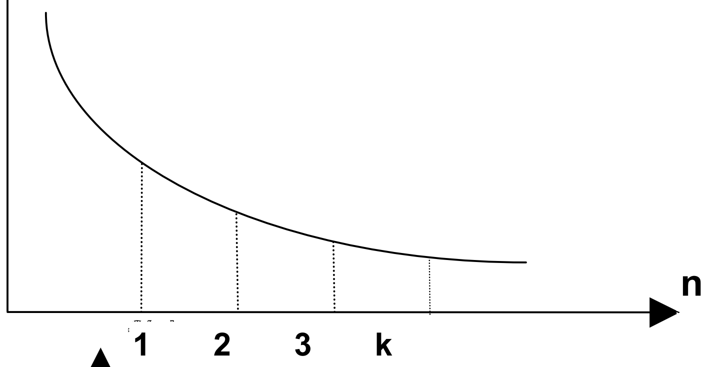
$$\mathbf{f}(\tau_1) = \mathbf{f}(\tau),$$

т.е. на начальном интервале имеет место то же самое распределение.

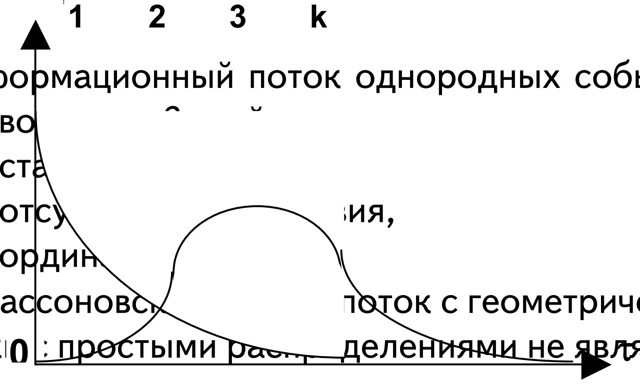
Существуют ординарные и неординарные потоки.

Ординарным информационным потоком называется такой поток, события в котором связаны только с появлением одной заявки, а вероятность появления за $\Delta\mathbf{t}$ стремится к $\varnothing$.

В неординарных потоках возможно появление 2, 3 и более заявок. Для задания таких потоков необходимо определить распределение числа заявок в одном событии потока **Q(n)**.



Информационный поток однородных событий называется простейшим, если удовлетво...



1/ ста...

2/ отсу... ия,

3/ ордин...

* пуассоновс... поток с геометрическим распределением. Другие потоки с простыми распределениями не являются простейшими.

Основные характеристики информационных потоков.

Любой информационный поток достаточно полно представляется своими распределениями вероятностей для интервала времени между событиями в потоке.

Поток может быть задан интервалом распределением **F(τ)**, который задается как вероятность того, что некоторое измеряемое значение **ξ < τ.**

$$F(\tau) = P(\xi < \tau)$$

Большей выразительностью обладает дифференциальный закон распределения или функция плотности распределения вероятности

$$\mathbf{f}(\tau) = \frac{\mathbf{dF}(\tau)}{\mathbf{d}\tau};$$

Она существует при больших значениях лишь в том интервале, когда функция меняет свою крутизну

Фактически распределени **F(τ)** и **f(τ)** являются моментами потока. Однако для качественных сравнений используются некоторые характеристики потока.

К основным характеристикам потока относятся:

1. математическое ожидание интервала **t** между событиями в потоке **M(τ)**

Оно вычисляется как I центральный момент при непрерывном распределении.

$$\mathbf{M}(\tau) = \int_0^\infty \tau\, \mathbf{dF}(\tau) = \int_0^\infty \tau \cdot \mathbf{f}(\tau)\mathbf{d}\tau.$$

Математическое ожидание можно оценить с помощью данной отдельной выборки для потока, т.е. массива измерений, выполненных в натуре.

N - объем выборки.

Математическое ожидание оценивается некоторым:

$$\mathbf{m}_\tau = \frac{\mathbf{1}}{\mathbf{N}}\sum_{\mathbf{i}=\mathbf{1}}^{\mathbf{N}} \tau_{\mathbf{i}};$$ /единица времени/.

2. Интенсивность потока λ. Для стационарных потоков она константа.

$$\lambda = \frac{\mathbf{1}}{\mathbf{M}(\tau)}$$ $\frac{\text{заявок/событий/}}{\text{в ед. времени}}$

3. Дисперсия распределения интервалов временя относительно математического ожидания.

$$\mathbf{D}(\tau) = \int_0^\infty \left[\mathbf{F} - \mathbf{M}(\tau)\right]^2 \mathbf{dF}(\tau) = \int_0^\infty \left[\tau - \mathbf{M}(\tau)\right]^2 \mathbf{f}(\tau)\mathbf{d}\tau.$$

Дисперсию можно, оценить по данным отдельной выборки с помощью среднеквадратичного отклонения

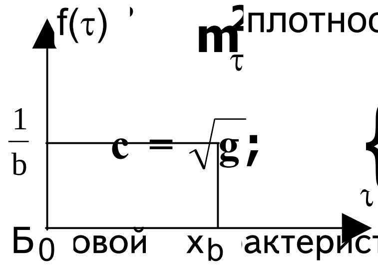
$$\delta_\tau^{\mathbf{2}} = \frac{\mathbf{1}}{\mathbf{N - 1}}\sum_{\mathbf{i}=\mathbf{1}}^{\mathbf{N}} \left(\tau_{\mathbf{i}} - \mathbf{m}_\tau\right)^{\mathbf{2}}$$ /единица времени в квадрате/

4. Коэффициент вариаций потока **(g)**

$$\mathbf{g} = \frac{\mathbf{D}(\tau)}{\mathbf{M}^2(\tau)};$$

По данным отдельной выборки:

$$g = \frac{\delta_\tau^2}{m_\tau^2}; \quad \text{или} \quad c = \frac{\delta_\tau}{m_\tau};$$



плотность распределении

интервалы $\tau[0,b]$

$c = \sqrt{g};$ $\left\{ \begin{array}{l} g \geq 1 \\ g < 1 \end{array} \right.$ пре/но близко к ней/

$x = b$

Б овой актеристикой для нестационарного потока является интенсивность $\lambda(t)$.

Для каждого момента времени:

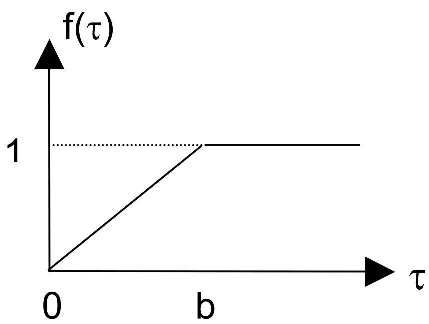
$$\lambda(t) = \lim_{\Delta t \to 0} \frac{L(t,\ t + \Delta t)}{\Delta t};$$

L - число заявок потока.

$\lambda \ll 1$ при правильном или неправильном выборе Δt. В этом λ рассматривается как вероятность появления хотя бы одной заявки за шаг Δt или в единицу времени.

ПРИМЕРЫ МОДЕЛЕЙ НФОРМАЦИОННЫХ ПОТОКОВ.

1. Стационарный поток с равным распределением.



$$\int_0^b f(\lambda)\,d\tau = 1$$

$$F(\tau) = \int_0^b \frac{1}{b}\,d\tau = 1$$

Рассмотрим основные характеристики данного потока.

1. Математическое ожидание:

$$M(\tau) = \int_0^b \tau f(\tau) d\tau = \int_0^b \tau \frac{1}{b} d\tau = \frac{b}{2}$$

2. Интенсивность потока:

$$\lambda = \frac{1}{M(\tau)} = \frac{b}{2};$$

3. Дисперсия:

$$D(\tau) = \int_0^b [\tau - M(\tau)]^2 f(\tau) \cdot d\tau = \frac{1}{b}\frac{\tau^3}{3}\Big|_0^b - \frac{\tau^2}{2}\Big|_0^b + \frac{b}{4}\tau\Big|_0^b = \frac{b^2}{3} - \frac{b^2}{2} + \frac{b^2}{4} = \frac{b^2}{12}$$

4. Коэффициент вариаций потока:

$$g = \frac{D(\tau)}{M^2(\tau)} = \frac{b^2 \cdot 4}{12 \cdot b^2} = \frac{1}{3};$$

Если по данным измерений мы получим **g**$\approx \frac{1}{3}$, мы предполагаем, что распределение в таком потоке равномерно.

Измеренные характеристики не дают нам возможности сделать заключение о свойстве последствия, т.е. является ли поток с ограниченым последствием, или он простейший /отсутствие последствия/.

Чтобы сделать такое заключение необходимо исследовать распределение времен до 1-го события в данном потоке $f_1(\tau_1)$.

Для анализа этого распределения удобно воспользоваться формулой Пальма:

$$f_1(\tau_1) = \lambda\left[1 - \int_0^{\tau_1} f(\tau) d\tau\right] = \frac{2}{b} - \frac{2}{b^2}\tau_1;$$

$f(\tau)$ - функция плотности распределения для любого события.

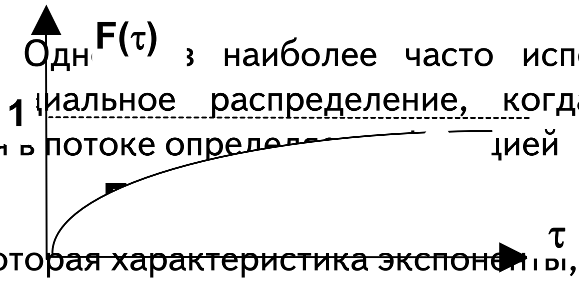
Функция плотности распределения отлична от обычного распределения.

$$M(\tau_1) = \int_0^b \tau_1 \cdot f_1(\tau_1) d\tau_1;$$

$$M(\tau_1) = \int_0^b \tau_1\left[\frac{2}{b} - \frac{2}{b^2}\cdot\tau_1\right] d\tau_1 = b - \frac{2}{3}b = \frac{b}{3}$$

Математическое ожидание до 1-го события в потоке оказывается меньше, чем менаду любыми другими. Таким образом поток с равном, распределением не является простейшим и обладает свойством ограниченного последствия.

<u>2. Стационарный поток о экспоненциальным распределением (Пуассона)</u>



Одн[illegible]з наиболее часто используемых моделей потока является экспоне[illegible]иальное распределение, когда вероятность наступления одного события в потоке определя[illegible]цией

α - некоторая характеристика экспоненты,

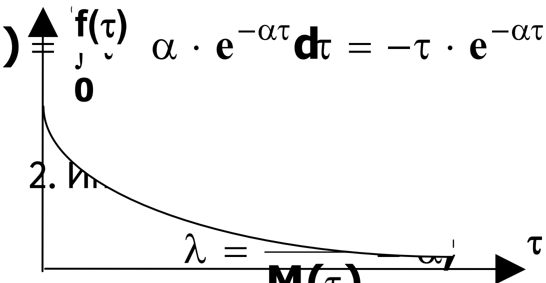
$e^{-\alpha\tau}$ - вероятность наступления ни одного события в потоке. Таким образом распределяется вероятность безотказной работы в системе.

Рассмотрим функцию плотности распределения:

$$f(\tau) = \frac{dF(\tau)}{d\tau} = -(-\alpha)e^{-\alpha\tau} = \alpha \cdot e^{-\alpha\tau};$$

1. Математическое ожидание:

$$M(\tau) = \int_0^{\infty} \tau \cdot \alpha \cdot e^{-\alpha\tau} d\tau = -\tau \cdot e^{-\alpha\tau}\Big|_0^{\infty} - \int_0^{\infty}(-e^{-\alpha\tau})d\tau = -\frac{1}{2} \cdot e^{-\alpha\tau}\Big|_0^{\infty} = \frac{1}{\alpha}$$



2. И[illegible]

$$\lambda = \frac{1}{M(\tau)} = \alpha;$$

3. Дисперсия:

$$D(\tau) = \int_0^{\infty}\left[\tau - \frac{1}{\alpha}\right]^2 \cdot \alpha \cdot e^{-\alpha\tau}\, d\tau = \frac{1}{\alpha^2};$$

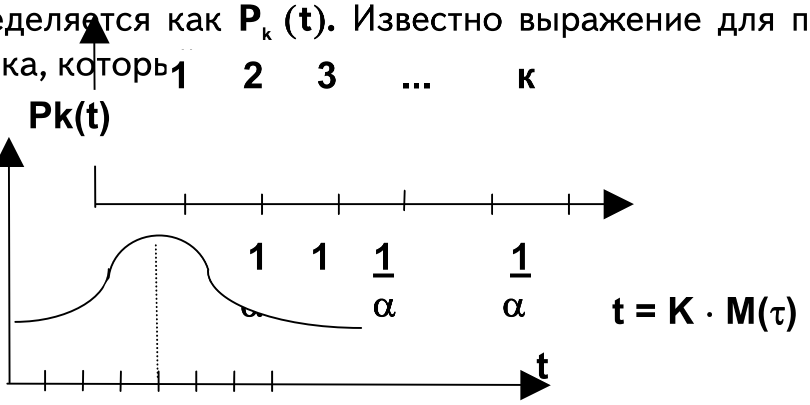
4. Коэффициент вариаций потока:

$$g = \frac{D(\tau)}{M^2(\tau)} = 1$$ /это свидетельствует об экспоненциальном распределении/

$$f_1(\tau_1) = \alpha \cdot \left[1 - \int_0^{\tau_1} \alpha \cdot e^{-\alpha\tau}\, d\tau\right] = \alpha\left[1 + e^{-\alpha\tau}\Big|_0^{\tau_1}\right] = \alpha \cdot e^{-\alpha\tau_1};$$

$\tau_1 \to \tau$ следовательно экспоненциальный поток является простейшими и обладает свойством отсутствия последствия.

Важной оценкой для любого информационного потока является вероятноость появления **K**-событий потока за некоторый заданный интервал времени **t**, причем **t** > или < математического ожидания. Такая вероятность определяется как $P_k(t)$. Известно выражение для простейшего пуассоновского потока, которы



$$\frac{1}{\alpha} = M(\tau)K \cdot \frac{1}{\alpha}$$

Это вероятность появления именно **K** событий.

3. Специальные обобщения для заданий распределения в потоках.

Очень удобна для анализа модель Пуассоновского потока, однако в практических системах часто коэффициент вариаций **g≠1** , и тогда использование простейшего потока в качестве модели приводит к грубым или неточным результатам.

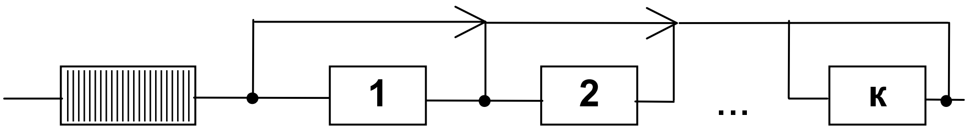
Используя преимущество экспоненциальной формы распределения для различных случа[...] пе[...]е модели/распределения/, которые базиру[...] нц[...] представлении, но в таких потоках учитыва[...]жные ситуации, чем в простом потоке. Такие ситуации получили название фазы. А общее число ситуаций/фаз/ определяет порядок распределения. В одном из таких специальных обобщений для заданий редких потоков вводится несколько фаз /1.2.3,..., к /, представляемых во времени как **1**/α.

$$\boxed{P_k(t) = \frac{\alpha \cdot t^{k-1}}{(k-1)!}\, e^{-\alpha t},}\quad \alpha = \lambda$$

На 1-м интервале ожидания появление одного события в потоке. А на 2-м интервале ситуация изменяется. 1-го события еще не было, но надо ожидать и 2-е.

В сложных потоках может быть ситуация ожидания **к** - события, даже если ни одного не было.

Интерпритация такой системы - воспроизводится работа многофункциональной системы, в которой приборы обследования распределены последовательно.



Использование фаз удобно тем, что при исследовании для появившихся событий в потоке регистрируется не время его, а фаза появления события.

3.1. Стационарный поток со специальным распределением Эрланга

В специальном составном распределении Эрланга для информационных потоков введено рассматриваемое ранее понятие фазы, причем дм каждой модели потока заранее вводится порядок потока. Используя общее экспоненциальное распределение в данном потоке учитываются вероятности появления одного события эа некоторый интервал **К** .

Интегральный закон распределения в модели:

$$\mathbf{F}(\tau) = \mathbf{1} - \sum_{\mathbf{i}=\mathbf{0}}^{\mathbf{k}-\mathbf{1}} \frac{(\alpha \cdot \tau)^{\mathbf{i}}}{\mathbf{i!}} \cdot \mathbf{e}^{-\alpha \cdot \tau}$$, где - фаза **i=0, k-1**

$$\mathbf{F}(\tau) = \mathbf{1} - \mathbf{e}^{-\alpha\tau} - \frac{\alpha\tau}{\mathbf{1}}\mathbf{e}^{-\alpha\tau} - \cdots - \frac{(\alpha\tau)^{\mathbf{k}-\mathbf{1}}}{(\mathbf{k}-\mathbf{1})!}\mathbf{e}^{-\alpha\tau};$$

Рассмотрим функцию плотности данного распределения,

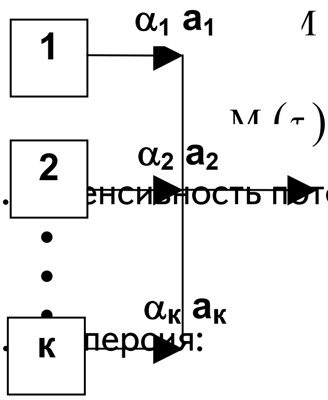
$$\mathbf{f}(\tau) = \frac{\mathbf{d}\,\mathbf{F}(\tau)}{\mathbf{d}\tau},$$

$$\mathbf{f(\tau)} = \alpha \cdot \mathbf{e}^{-\alpha\cdot\tau} - \alpha\cdot \mathbf{t}^{\alpha\cdot\tau} + \alpha^{2}\cdot\tau\cdot \mathbf{e}^{-\alpha\cdot\tau} - \frac{2\alpha^{2}\cdot\tau}{2}\cdot \mathbf{e}^{-\alpha\cdot\tau} + \cdots + \alpha\frac{(\alpha\cdot\tau)^{k-1}}{(k-1)!}\mathbf{e}^{-\alpha\cdot\tau}$$

$$f(\tau) = \alpha\,\frac{(\alpha\,\tau)^{k-1}}{(k-1)!}\,e^{-\alpha\tau}$$ для потока Эрланга **к** - го порядка .

Рассмотренные функции однозначно задают поток .

1. Математическое ожидание:



$M(\tau) =$ $a_1, \ldots, a_k$ - вероятности

$M(\tau) = (1/$ $\alpha_2\ a_2\ \lambda\ \alpha_1 \neq \alpha_2 \neq \ldots \neq \alpha_K$

2. Интенсивность потока:

/в общем случае/.

За время Δt считаем:

$a_1 + a_2, \ldots + a_K = 1$

3. Дисперсия:

$$D(\tau) = \frac{K}{\alpha^{2}}$$

4. Коэффициент вариаций потока:

$$g = \frac{D(\tau)}{M^{2}(\tau)} = \frac{K\cdot\alpha^{2}}{\alpha^{2}\cdot K^{2}} = \frac{1}{K}; \quad K > 1 \Rightarrow g < 1.$$

Можно показать, что Эрлангов поток не является простейшим, т.е. является стационарным рекурентным потоком с ограниченным последствием. Такая модель в представлении вычислительных систем используется для представления работы внешних устройств ввода-вывода.

4. Стационарный поток с гиперэкспоненциальным распределением

В некоторых случаях "частых" потоков коэффициент вариаций оказывается больше 1. В этом случае можно использовать другую модель потока, также использующую понятие фазы.

Такой поток частых событий может быть интерпретирован следующей схемой. Некоторое число паралельных генераторов формирует каждый свою заявку

В таком потоке фаза связана с номером генератора.

Функция распределения вероятности.

$$F(\tau) = 1 - \sum_{i=1}^{K} a_i \cdot e^{-\alpha_i \tau};$$

$$\sum_{i=1}^{K} a_i = 1$$ - свойство коэффициентов.

Функция плотности распределения:

$$f^{(K)}(\tau) = \sum_{i=1}^{K} a_i \alpha_i \cdot e^{-\alpha_i \tau};$$

Для анализа такое распределение оказываетоя сложным,т.к. выражения для математических ожиданий, суммарная общая интенсивность потока сложны и зависят от многих параметров.

Всякое .экспоненциальное распределение базируется на одном и том же параметре α . Для упрощения модели такого /гиперэкспонен- пиального/ потока вводится понятие непрерывной фазы /φ/, причем сам поток рассматривается как поток второго порядка. При этом φ рассматривается как вероятность появления заявки с выхода 1-го генератора. Кроме того для данного гиперэкспоненциального потока вводится некоторая эквивалентная интенсивность, которая, как и для всех экспоненциальных потоков, задается

$\alpha_1 = 2\varphi\alpha$; $\alpha_2 = 2(1-\varphi)\alpha$.

$\lambda < \alpha$; $a_1 = \varphi$, $a_2 = 1-\varphi$.

Вводятся два параметра α и φ , которые рассчитываются по трем. В таком представлении для потока используется следующее распределение:

$$F^{(2)}(\tau) = 1 - \varphi \cdot e^{-2\varphi\alpha\tau} - (1-\varphi) \cdot e^{-2(1-\varphi)\alpha\tau};$$

$$f^{(2)}(\tau) = 2\varphi^2\alpha e^{-2\varphi\alpha\tau} + 2(1-\varphi)^2\alpha e^{-2(1-\varphi)\alpha\tau}.$$

В данном потоке второго порядка **g** может быть любым.

Для данного распределения мы можем посчитать обычным аналитическим способом основные характсристики.

1. Математическое ожидание:

$$M(\tau) = \frac{1}{\alpha}.$$

2. Интенсивность потока:

$$\lambda = \alpha .$$

3. Дисперсий:

$$\mathbf{D}(\tau) = \frac{\mathbf{1}}{\alpha^{\mathbf{2}}}\left(\mathbf{1} + \frac{(\mathbf{1} - \mathbf{2}\varphi)^{\mathbf{2}}}{\mathbf{2}\varphi(\mathbf{1} - \varphi)}\right).$$

4. Коэффициент вариаций потока:

$$\mathbf{g} = \mathbf{1} + \frac{(\mathbf{1} - \mathbf{2}\varphi)^{\mathbf{2}}}{\mathbf{2}\varphi(\mathbf{1} - \varphi)};$$

φ вычисляется для каждого конкретного **g.** Можно показать, что φ может быть любым.

Если $\varphi = \mathbf{0}$ $\quad \mathbf{g} \to \infty$

$\varphi = \frac{\mathbf{1}}{\mathbf{2}}$ $\quad \mathbf{g} = \mathbf{1}$

5. Стационарный поток с кусочно-степенным распределением

В некоторых случаях практические кривые распределения для стационарных потоков трудно поддаются описанию рассматриваемыми моделями. Тогда можно использовать общую форму кусочно-степенного распределения, когда производится непосредственная апроксимация практических кривых полиномами 5 -ой степени по участкам.

Для ряда сложных информационных потоков не удается найти непрерывное распределение. Но можно использовать практические кривые; гистограмму $\mathbf{f}(\tau)$ и кумулятивное интеДанная Модель с введением φ

В стационарных потокахдостаточно близко воспроизводит ием практические кривые распределенiработу центрального процессора. ıми-рующих функций. Причем кривая разбивается на **n** участков, на из которых аппроксимирующая функция - полином **S -** степени.

В общем случае интегральный закон распределения можно за- писать в виде:

$$\mathbf{F}(\tau) = \sum_{\mathbf{i}=\mathbf{0}}^{\mathbf{n}} \sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{S}} \mathbf{bi, k}(\tau - \mathbf{ai})^{\mathbf{k}} \cdot \eta(\tau - \mathbf{ai});$$

ai; $\mathbf{b}_{i,k}$ - коэффициенты полинома.

$\eta(\tau - \mathbf{a})$ - единочная функция Хависайда.

Плотность распределения в этом случае:

$$\mathbf{a}(\tau) = \sum_{\mathbf{i}=\mathbf{0}}^{\mathbf{n}} \mathbf{b_{i,o}}\delta(\tau - \mathbf{a_i}) + \sum_{\mathbf{i}=\mathbf{0}}^{\mathbf{n}} \sum_{\mathbf{k}=\mathbf{1}}^{\mathbf{S}} \mathbf{b_{i,k}} \cdot \mathrm{K}(\tau - \mathbf{a_i})^{\mathbf{k}-\mathbf{1}} \cdot \eta(\tau - \mathbf{a}),$$

где

$$\delta(\tau - a_i) = \frac{d\eta(\tau - a_i)}{d\tau};$$

$\sum\sum b_{i,k}(\tau - a_i)^k \cdot \delta(\tau - a_i) \neq 0$ только при $\tau = a_i$ следовательно первое слагаемое.

Частным случаем является вырожденный поток:

$$\Rightarrow M(\tau) = \int_0^\infty \tau \cdot \delta(\tau - a)^{\neq 0 \text{ при } \tau = \alpha \text{ и}} d\tau = \eta(\tau - a)\big|\tau = a$$

1. Математическое ожидание: $M(\tau) = a$.
2. Интенсивность потока: $\lambda = \frac{1}{a}$.
3. Дисперсия: $D(\tau) = 0$.
4. Коэффициент вариаций потока: $g = 0$.

...венных синхронизирующих импульсов с периодом a и частотой

$F(\tau) = \eta(\tau - a)$

$f(\tau) = \delta(\tau - a)$

Такие модели потока в общем случае используется в крайних случаях, когда не подходит ни одна другая модель,. Но вырожденные потоки применяются довольно часто.

6. Стационарный поток с нормальным распределением /Гаусса/

В редких случаях /$g << 1$/ вместо распределения Эрланга при $к>>1$ применявт нормальное /Гауссовское/ распределение.

Оно составляется с учетом как бы известных характеристик математического ожидания $M(\tau)$ и дисперсии $D(\tau) = \sigma^2$:

$$F(\tau) = \frac{1}{\delta\sqrt{2\pi}} \int_0^\infty e^{-\frac{[\tau - M(\tau)]^2}{2\delta^2}} d\tau \Rightarrow f(\tau) = \frac{1}{\delta\sqrt{2\pi}} e^{-\frac{[\tau - M(\tau)]^2}{2\delta^2}}$$

/применяется редко/.

МАРКОВСКИЕ МОДЕЛИ СМО

КЛАССИФИКАЦИЯ СМО

Для функционирования СМО необходимо в первую очередь задать модели входных и выходных потоков. Если эти потоки известны, то мы имеем досгаточную полную информацию для анализа поведения системы.

<u>Классификация Кандалла</u> /формульное описание СМО/. Формула Кандалла для СМО имеет пять полей:

$$\underset{1}{\mathbf{G}}^{\mathbf{e}} / \underset{2}{\mathbf{G}}^{\mathbf{n}} / \underset{3}{\mathbf{n}} / \underset{4}{\mathrm{N}} / \underset{5}{\mathrm{M}}$$

1 - входной информационный поток ;

2 - выходной информационный поток ;

3 - число приборов в системе ;

4 - число заявок в очереди системы ;

5 - число заявок, которые могут иметь место во входном потоке;

Если $\mathrm{N}, \mathrm{M} = \infty$, то эти поля можно опустить.

G - некоторое распределение для рекурентного потока однородных событий с возможным появлением групповых заявок. .

e - максимальное число заявок в группе. Если **e = 1**, то оно опускается.

Для известного распределения в потоках вместо **G** записывают конкретные символы:

M **-** для экспоненциального/Пуассоновского/ потока;

$\mathbf{E}^{(k)}$- для распределения Эрланга k-го порядка;

H **-** для гиперэкспоненциального распределения 2-го порядка;

$\mathbf{Q}^{(S)}$- для кусочно-степенного распределения в потоке

/**s** -максимальная степень полинома/;

D - для вырожденного потока.

n- число линий обслуживаемых в приборе.

Простейшей моделью СМО является <u>М/М/1.</u>

Реально вичислительнне системы представляются: $\mathbf{E}^{(k)}$/H/n.

Если число мест в очереди ограничено /имеется вероятность потери заявки/ необходимо выдавать поле **N** :

M/M/1/N

Во многих известных теориях СМО применяется и другая классификация СМО.

В 1-ю очередь СМО разделяются на 3 класса:

I. СМО с отказами.

II. СМО с ожиданием.

III. СМО с ограниченным ожиданием.

Эти классы определяются по критическому времени ожидания $\mathrm{T}_{\mathbf{ож}}^{\mathbf{кр}}$ обслуживаемой заявки в системе. Считается, что по истечению этого времени ваявка выходит из системы /теряется/.

Для 1 - $\mathrm{T}_{\mathbf{ож}}^{\mathbf{кр}} = \mathbf{0}$ /в этой системе нет специальных средств для организации ожидания в очереди для обслуживания/.

Пример: Системы измерения для некоторых частиц, процессов, которые нельзя остановить.

Системы контроля на технологических линиях, производящие выборочный контроль.

Для таких систем важными характеристиками являются:

1. вероятность получения отказа в системе;
2. среднее число заявок, получающих отказы в единицу времени.

Для системы контроля:

1. вероятность пропуска дефектной детали;
2. критерий риска не виявления партии дефектных деталей.

СМО II класса имеют возможность сохранять заявки в очереди неограниченное время $T_{\mathbf{ож}}^{\mathbf{кр}} = \infty$ /идеализированная СМО/.

Для этих СМО в качестве интегральных характеристик организации работы используются;

1. среднее время пребывания заявки в системе;
2. средняя длина очереди в СМО;
3. коэффициент простоя прибора;
4. максимальная длина очереди и т.д.

III класс СМО - реальные технические системы более сложные по описанию, где $T_{\mathbf{ож}}^{\mathbf{кр}} = \mathbf{N}$ /конечно/, т.е. имеется некоторая вероятность потери заявки. Эти СМО описываются всеми характеристиками /I и II/.

СМО делятся на:

1. однолинейные
2. многолинейные

По числу фаз обслуживания в приборе:

1. однофазные;
2. многофазные.

Дисциплина обслуживания заявок в СМО

На характер процессов изменения состояний в СМО значительно влияет принцип организации очереди, который определяет, каким образом выбирается заявки из очереди.

Все принципы упра } ıечисло линий и/ можно разбить на группы.
Первая группа: ор } ıв приборе где выбор заявок из очереди осуществляется в порядке поступления.

1. /**FIFO**/ /либо **FCFS**/

Преимущества: 1. равноправность заявок;
2. простота организации очереди.

Недостатки : 1. короткие задания должны долго ждать в очереди;

2. возможные срочные задания в системе будут обослуживаться долго.

2. **LIFO** - имеет особое значение при обработке прерываний,

Преимущества: возможность обслуживания прерываний во-всех системах.

Вторая группа дисциплин обслуживания ваявок.

Использует дополнительную информацию о времени выполнения задания /обработки заявки/. Каждая заявка поступившая в систему должна нести в себе информацию о необходимом времени для ее обследования. В этом случае возможно 2 типа очереди:

1. **SτF**
2. **LτF**

1-я позволяет быстро обследовать короткие задания. Сокращается средняя длина очереди в системе, а соответственно минимизируется средство для организации очереди.

Недостатки: длинные /"тяжелые"/ могут находится в системе очень большое время или даже потеряться в системе /отказ/.

2-я позволяет реализовать в первую очередь наиболее сложные задания, используя начальный этап эксплуатации линии СМО с повышенной надежностью.

Недостаток: очередь быстро растет за счет накопления коротких заданий.

Третья группа обследует заявки по вычесляемому оставшемуся времени пребывания в системе.

В такой СМО имеются средства для фиксации времени пребывания заявки и вычисления оставшегося времени до некоторого критерия $T_{ож}^{кр}$, которые известны для всей системы, или для типов заявок.

В этой группе имеется 2 дисциплины:

1. **SRTF**
2. **LRTF**

1-я обеспечивает примерное усреднение времен пребывания заявок в системе с достаточно большим средним временем, но меньшим $T^{кр}$.

Преимущество : количество заявок, получивших отказы сокращается, и соответственно уменьшается вероятность получения отказа.

Недостаток: короткие задания находятся в системе достаточно долго, и время реакция в системе оказывается значительным.

2-я /дисциплина сабботажа /организованного невыполнения//. Преимущество: наибольшее число заявок получивших отказы в системе.

Четвертая группа обеспечивает выбор заявок из очереди в случайном порядке.

Носит общее название **RAND**, но при выборе заявок задается то или иное распределение. Одним из вариантов является равномер- вое распределение. Выбор того или иного распределения вероятности позволяет необходимым образом уменьшить или увеличить среднее время пребывания заявки в системе, уменьшить дисперсию/разброс/ пребывания заявок в системе. Одним из вариантов такой дисциплины является процесс хеширования для последних файлов, когда периодически в последнем файле записи меняется местами в случайном порядке. Для организации очереди чаще всего используется последовательный набор ячеек памяти, в котором заявки упорядочиваются по необходимому критерию, т.е. используются процесс сортировки для заданного массива.

Пятая группа реализует обследование заявок с прерыванием

Т.е. ваявка обследуемая в данном моменте и находящаяся в приборе может быть снята с обследования и прибор будет предоставлен другой ваявке.

1. **RR** /циклическое повторение/.

2. **EP** /внешний приоритет/.

1-я предусматривает выделение некоторого интервала времени $\mathbf{h_R}$ для обследования каждой заявки. Если ва это время заявка может быть обследована, то выводит из системы, если нет то ставится в очередь **FIFO**. Заявка не несет в себе информацию о времени обследования.

В системе рассматривается идея как-бы "паралельной" обработки заданий. Поскольку каждое из заданий достаточно быстро получает доступ к прибору и начинает выполнятся.

Недостаток : длинные задания обследываются долго, чем в обычной **FIFO** системе. Эта дисциплина используется для обработки диалогового режима в многотерминальной системе. Эта дисциплина называется разделением времени процессора.

2-я - прерывание по приоритетам. Заявка, поступающая в систему, несет в себе информацию о приоритете, т.е. некотором количественном показателе, который сравнивается в системе и в случае, если обслуживаемая заявка имеет более низкий приоритет, то она прерывается, а появившаяся заявка принимается в прибор на обследование. Возможные заявки о меньшим приоритетом находятся в очереди и выбираются в соответствии с приоритетом.

Если заявки имеют одинаковый приоритет, то организуется локальная очередь типа **FIFO**.

Кодовый показатель приоритета меньшей величины соответствует высшему приоритету. Прерванные заявки могут:

1. теряться в системе;
2. становиться в очередь и обследуется сначала;
3. становится в очередь и дообследуется.

PEP - обследование по приоритетам, но без прерываний.

Возможны и другие дисциплины, в частности различные комбинации уже рассматриваемых.

Выбор дисциплины обследования является центральным вопросом проектирования, который существенно влияет на качество организации системы. Если нет дополнительных требований и особых целей применяется **FIFO**.

$\mathbf{h_R}$ - фиксированы для некоторой конфигурации системы.

При усложнении системы, когда для обследования используется множество приборов, мультипроцессорные системы, возникает случай конфликтов за появляющейся заявкой.

В этом случае необходима организация ресурсов для исключения конфликтов. Организация распределяется по ресурсам. Для управления ресурсами используется 3 группы дисциплин, линий прибора.

1. В порядке освобождения линий или приборов

Включает дисциплину **FIFO** непрерывно ведущую, когда каждая освободившаяся оканчивается последней в области ресурсов, а первая из освободившихся получает заявку на обслуживание. Все линии и приборы эксплуатируются равномерно. Линии оказываются равноправны. Простота организации.

LIFO. Последний из обвободившихся приборов сразу же принимает заявку на обследование. Эта дисциплина обеспечивает неравномерность в эксплуатации приборов, когда первые из загруженных будут эксплуатироваться интенсивно. Зато остальные линии являются мало загруженными, чем обеспечивается надежность системы /за счет замены выбывших линий/.

2. Предоставление линий /или приборов/ в случайном порядке

RAND - когда задается на множестве линий распределение их вероятностей для обследования или выборки. Позволяет варьировать загрузку и эксплуатацию приборов.

3. Предоставление линий в соответствии с приоритетом, но без прерываний

PEP. Освободившаяся линия с большим приоритетом захватывает заявку. Преимущество: отдельные линии с большими приоритетами эксплуатируются интенсивно.

МАРКОВСКИЕ СЛУЧАЙНЫЕ ПРОЦЕССЫ. ОСНОВНЫЕ ПОНЯТИЯ И ОПРЕДЕЛЕНИЯ.

В процессе обработки входных заявок СМО переходит из одного состояния в другое, причем число состояния и физический смысл может определятся исследователем, в зависимости от поставленной задачи проектирования.

СМО - система с дискретными состояниями. Для многоприборной системы понятие состояния может быть связано с состоянием ресурсов. В некоторых можно выделить ПРВВ, ПРД, ЦПР. Тогда можно ввести состояния:

свободен 0

занят 1

В таком случае в системе возможно 8 состояний

Z_0 = все свободны **0 0 0**

$\vdots$ $\vdots$

Z_7 **1 1 1**

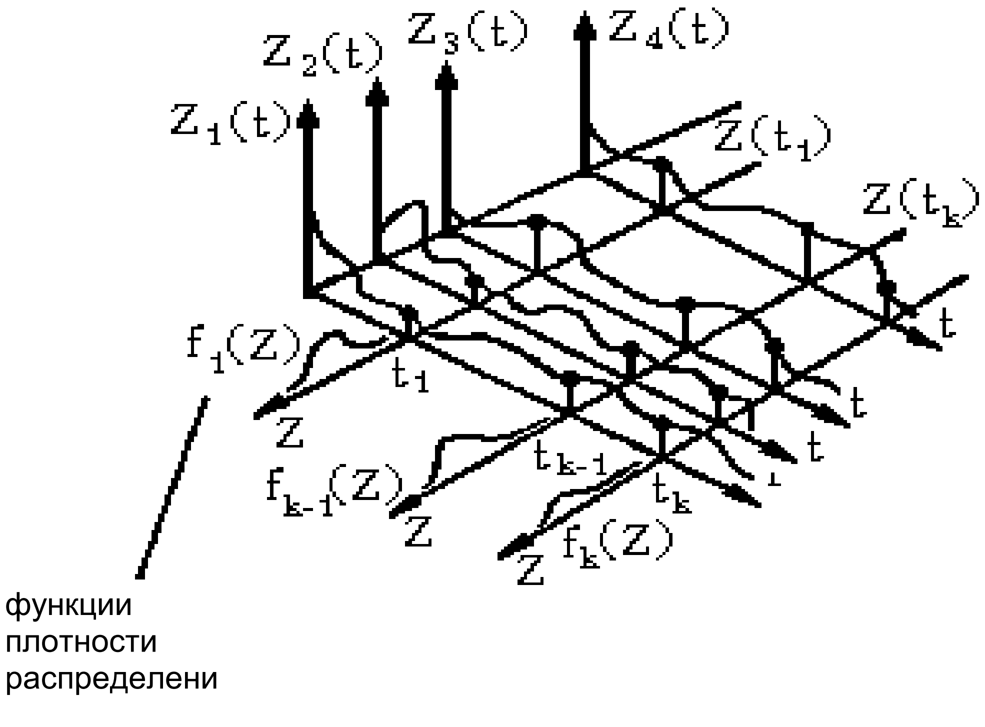
Но переход из состояния в состояние в такой СМО выполняется в случайном порядке, т.е. развивается некоторый стохастический процесс.

В классическом пледставлении общее состояние СМО связывается с общим числом заявок, находящихся в системе /в приборе и в очереди/.

Для описания процессов, происходящих в СМО часто используются так называемые Марковские случайные процессы с ограниченным последствием.

ОПРЕДЕЛЕНИЕ 1. Стохастическим или случайным процессом называется такая функция **Z(t)**, которая в каждый момент времени $\mathbf{t_k}$, принимает значение некоторой случайной величины $\mathbf{Z(t_k)}$, которая называется также сечением случайного процесса. В общем случае каждое сечение случайного процесса может иметь свое распределение вероятности.

Исследование случайного процесса обычно выполняется на множестве его реализаций: $\mathbf{Z_1(t)}$, $\mathbf{Z_2(t)}$, ..., $\mathbf{Z_n(t)}$, которые должны быть получены одновременно, синхронно при исследовании **n** совершенно одинаковых идентичных объектов. Множество этих реализаций называется ансамблем, и характеристики каждого сечения могут быть получены путем обработки соответствующих данных в данном сечении по множеству реализаций.



Если функции плотности распределения разные, то это нестационарный случайный процесс.

Если "волна распределения" будет строгой /стоячей/, то имеет место стационарный случайный процесс.

ОПРЕДЕЛЕНИЕ 2. Если в каждом сечении случайного процесса $\mathbf{Z(t_к)}$ случайная величина принимает лишь конечное или счетное множество значений $\{\mathbf{Z_1, Z_2, \ldots, Z_j}\}$, то такой дискретный процесс со случайными состояниями $\mathbf{t_1, t_2}$ называется цепью.

Переход из состояния в состояние обязательно мгновенный.

ОПРЕДЕЛЕНИЕ 3. Если все моменты времени, в которых возможны переходы из состояния в состояние, образуют конечное или счетное множество моментов $\{\mathbf{t_1, t_1, \ldots, t_к}\}$, то такой случайный процесс с дискретным временем называется последовательностью.

ОПРЕДЕЛЕНИЕ 4. Если переходы из состояния в состояние осуществляются в любой момент времени, т.е. нельзя выделить какого-либо даже очень малого шага $\Delta\mathbf{t}$ для задания всех возможных моментов переходов, то такой случайный процесс является процессом с непрерывным временем.

Процессы в СМО обязательно являются цепью, но их развитие может быть представлено как в непрерывном, так и в дискретном времени. Для аналитического исследования часто используется непрерывное время, в то же время для моделирования с помощью ЭВМ используем дискретное время, т.е. представляем процесс в виде цепи с дискретным временем.

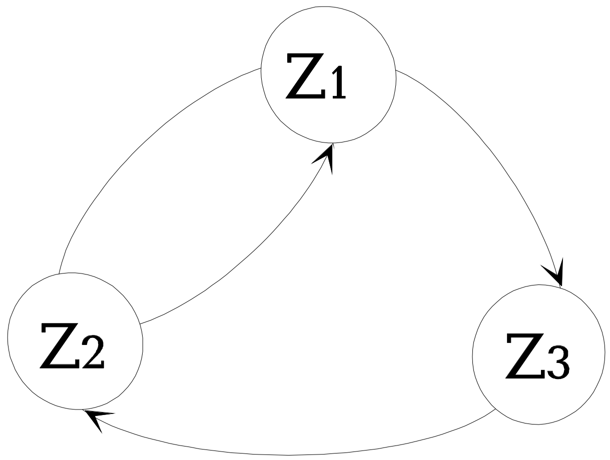
ОПРЕЛЕНИЕ 5. Марковским случайным процессом называется цепь с дискретными состояниями $\{\mathbf{Z_1, Z_2, \ldots, Z_1, \ldots, Z_{j, \ldots}}\}$, в котором распределение вероятностей состояний в момент времени $\mathbf{t_k}$, т.е. $\mathbf{f_k(z)}$, зависит только от того, в каком состоянии находятся процесс в момент времени $\mathbf{t_{k-1}}$, и не зависит от того в каком состоянии находилась система в моменты $\mathbf{t_1, t_2, \ldots, t_{k-1}}$, причем:

$$\left| \mathbf{t_1 < t_2 < t_3 < t_{k-1} < t_k} \right|.$$

Функция распределения $\mathbf{f_k(z)}$ /является/ зависит от того, в каком состоянии находилась система в момент $\mathbf{t_{k-1}}$.

$$\frac{\mathbf{f_k(z)}}{\mathbf{z(t_k)}} = \mathbf{F}\left(\mathbf{z(t)}_{\mathbf{t = t_{k-1}}}\right).$$

Для задания марковской цепи необходимо в принципе задать множество распределений $\mathbf{f_k(z)}$ для некоторых $\mathbf{Z_{k-1}}$. А фактически это отражается в некотором графе переходов из одного состояния в другое, для которого отображаются переходы для пары $\mathbf{t_{k-1} \rightarrow t_k}$.



Кроме графа переходов для каждого сечения может быть задана и таблица переходов, в которой фиксируются вероятности переходов $\mathbf{P_{ij}}$ из $\mathbf{Z_i} \rightarrow \mathbf{Z_j}$

В общем случае и граф переходов **G(k)** марковской цепи и таблица переходов **T(к)** отражают то распределение, которое имело место в **к** -ом сечении. И хотя множество распределений одинаково для любых пар моментов, но история процесса отражается уже многими графами и многими таблицами.

СВОЙСТВО ОДНОРОДНЫХ МАРКОВСКИХ ЦЕПЕЙ

Простейшим видом марковской цепи является однородная марковская цепь, в которой вероятности перехода из состояния в состояние $\mathbf{P_{ij}}$ не зависят от времени, а следовательно марковская цепь для любых сечений случайного процесса представляется одним и тем же **G**, одной и той же таблицей **T.**

Некоторые определения, которые позволят сформулировать свойства однородных марковских цепей /для рассчета и анализа в процессах /.

ОПРЕДЕЛЕНИЕ 1 Если в одной марковской цепи можно из любого состояния $\mathbf{Z_i}$ перейти в любое другое состояние за как угодно большой интервал времени, то такая марковская цепь называется *неприводимой*. В графе такой марковской цепи будут отсутствовать "висячие" вершины /в которые входят стрелки, но не выходят/.

ОПРЕДЕЛЕНИЕ 2 Состояние однородной марковской цепи $\mathbf{Z_j}$ называется возвратным, если вероятность вернуться в него по истечении некоторого большого времени равна 1.

ОПРЕДЕЛЕНИЕ 3 Средним временем возвращения $\bar{\mathbf{t}}_{\mathbf{j}}$ в некоторое возвратное состояние $\mathbf{Z_j}$ называется *среднее время между многими очередными возвращениями в это состояние*. Если имеет место цикличность, то тогда речь идет о стахастическом процессе, а можно использовать более простые средства анализа.

ОПРЕДЕЛЕНИЕ 4 Если все состояния однородной марковской цепи апериодичные, то такая марковская цепь называется *апериодической*, и в ней не развиваются различные циклические процессы.

ОПРЕДЕЛЕНИЕ 5 Если времена возвращения t_{rj} не одинаковы /не повторяются/, то такое возвратное состояние Z_j называется *опериодичным.*

СВОЙСТВО 1:

Если марковская цепь однородна, неприводима и апериодична, то для каждого состояния цепи существуют некоторые предельные вероятности состояний P_j, которые называются равновесными, и которые не зависят от времени и от начальных условий, т.е.

$$P_j = \lim_{k \to \infty} P_j(K)$$

Вероятностный переходный процесс как-бы закончился и наступил равновесный, не зависящий от $P_j(0)$.

СВОЙСТВО 2:

Если все состояния однородной марковской цепи возвратны, а средние времена t_{rj} конечны, то для такой марковской цепи имеет место стационарное распределение вероятностных состояний:

$$\begin{cases} 1.\ \sum\limits_{j=0}^{\infty(n)} Pj = 1; \\ 2.\ Pj = \sum\limits_{j=0}^{n(\infty)} Pij\, Pi \end{cases}$$

- сумма произведений другого состояния на вероятность перехода из i-го состояния в j-тое

Равновесие вероятности состояний марковской цепи и вероятности переходов позволяют вычислить важные характеристики однородного марковского процесса:

$$\begin{cases} t_{rj} \\ t_j \end{cases}$$

- время возвращения в j-тое состояние

- среднее время пребывания в каждом j-ом состоянии

Но эти времена задаются с учетом шага испытаний Δt.

Текущее состояние процесса на марковском графе отображается точкой. На каждом шаге Δt мы точку снимаем и разрываем переход. Среднее время возврата в каждое состояние:

$$\overline{t_{rj}} = \frac{1}{P_j}\,\Delta t$$

Δt выбирается как минимальный шаг, при котором возможен физический переход.

Для определения среднего времени $\overline{t_j}$ воспользуемся некоторыми фрагментами цепи, которые подвергались испытаниям на каждом шаге.

Среднее время будет видимо определяться вероятностью переходов в то же самое состояние P_{jj}, которая должна согласовываться с вероятностью выхода / l- P_{jj} /.

Среднее время пребывания в j-том состоянии может определить как геометрическое расределение вероятности возвращения в то же состояние на k шагах испытаний. В этом случае вероятность того, что в Z_j будем находить k шагов будет:

$$P\,(t_j = k\,\Delta t) = P_{jj}^{(k-1)}\,(1-P_{jj})$$

Это геометрическое дискретное распределение характеризует результаты испытаний с /k-l/ неуспехом и одним успехом, и наоборот.

Для среднего времени при геометрическом распределении:

$$\overline{t_j} = \frac{P_{jj}}{1-P_{jj}}\,\Delta t,$$

$\frac{P_{jj}}{1-P_{jj}}$ - задает среднее время для k.

Геометрическое распределение является простейшим распределением без последствия. Именно оно характеризует процесс обследования заявок в СМО, поскольку Z_j мы можем рассмотреть как состояние обследования j-той заявки. Тогда среднее время ее обследования распределяется геометрически.

Если выходной поток СМО имеет более сложное распределение чем экспоненциальное или геометрическое, а для описания процессов СМО используется однородная марковская цепь, то в этом случае имеет место некоторое логическое несоответствие, и такая модель называется *полумарковской*. Такая модель называется с вложенным марковским процессом.

Если же выходной поток простейший, то можно назвать эту модель марковской и использовать однородные цепи для анализа процесса.

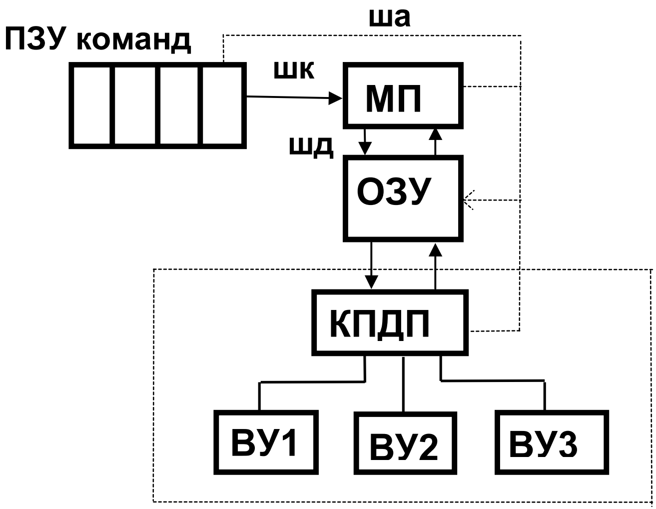
ПРИМЕР РЕШЕНИЯ МАРКОВСКОЙ МОДЕЛИ МИКРО-ЭВМ КАК СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ С ОЖИДАНИЕМ

Цель исследования на модели - определение возможного свободного времени оперативной памяти /ОП/ микро-ЭВМ, работающей в системе управления в режиме реального времени для осуществления необходимого долгого ввода-вывода данных в память по каналу ПД без остановки вычислительного процесса.

Для оценки такого возможного времени микро-ЭВМ рассматривается как СМО с неограниченным ожиданием, причем каждая команда - некоторая заявка, которая характеризует ПЗУ. В этом случае ПЗУ - как некоторая очередь команд.

Структура такой СМО изображена на рисунке 1.

Если команды следуют последовательно друг за другом, то считаем, что это очередь **FIFO** заявок, если имеют место переходы, то это дисциплина **RAND**. Известно, что при выполнении некоторого класса команд нет обращений к ОП. Например, это команды с непосредственной адрессацией /К580ИК80/, прямой регистровой и с неявкой. Во время цикла выполнения такой команды, ОЗУ - свободно. Его можно занять для ввода-вывода.



Чередование выполнения команд с различным типом адрессаций есть стахостический процесс, который в часном случае может обладать марковским свойством.

Выделим следующие дискретные состояния системы в соответствии с целью исследования.

$$\begin{cases} Z_1 & \text{- выполнение команды с прямой адрессацией} \\ Z_2 & \text{- выполнение команды с косвенной адрессацией} \\ Z_3 & \text{- выполнение команд с непосредственной адрессацией с прямой регистровой и неявной} \\ Z_4 & \text{- с другой через указатель стека} \end{cases}$$

Очень важно для системы выделить некоторый шаг испытаний Δt. Очевидно, что для такой системы Δt должно быть меньше или равно минимальному времени возможного перехода системы из одного состояния в другое.

$\Delta t = 2$мкс /как время выполнения самой короткой команды в машине/.

Для перехода к марковской модели процессов необходимо взять за основу следующие постулаты.

1-Й ПОСТУЛАТ: Тип адрессации каждой очередной команды не зависит от типа адрессации предыдущей

2-Й ПОСТУЛАТ: Время выполнения очередной команды распределяется то ли равномерно, то ли экспоненциально.

Вероятности переходов от одного типа адрессации к другому в микро-ЭВМ не зависят от времени.

Это дает нам право использовать однородную марковскую цепь для исследования процессов микро-ЭВМ. В результате исследований на модели были получены вероятности переходов $\mathbf{P}_{ij}$ и составлен марковский граф процесса.

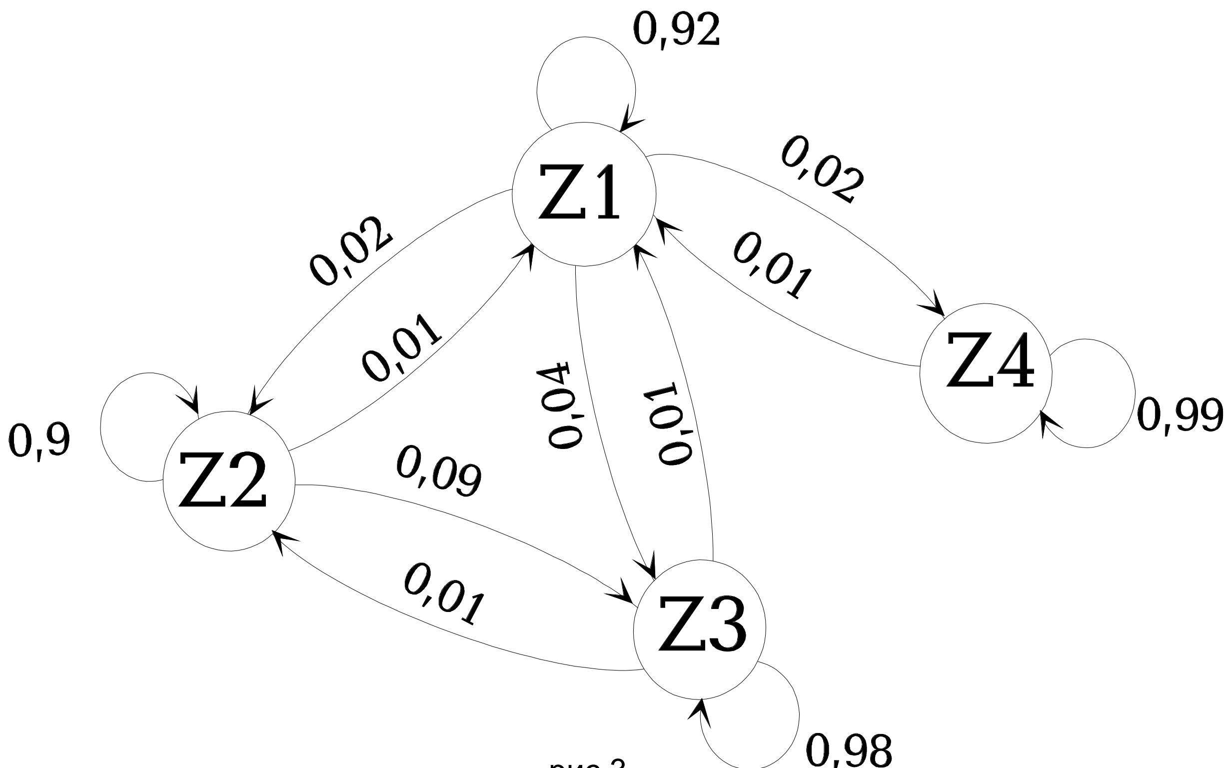


рис.3

По марковскому графу в соответствии со 2-м свойством однородных марковских цепей можно составить уравнения для равновесных вероятностей состояний и таким образом получить распределение вероятности.

Но кроме того необходимо оценить характеристики:

$\overline{\mathbf{t}}_{\mathbf{j}}$ и $\overline{\mathbf{t}}_{\mathbf{j}}$. И тогда $\overline{\mathbf{t}}_{\mathbf{з}}$ - это то время, которое мы сможем выделить для ввода-вывода даных в некотором эквивалентном цикле.

<u>1.</u> Составим алгебраические уравнения относительно вероятностей состояний; для чего удобно пользоваться таблицей переходов из состояния в состояние.

$Z_i \diagdown Z_j$	1	2	3	4	!
1	0.92	0.02	0.04	0.02	=1
2	0.01	0.9	0.09	-	=1
3	0.01	0.01	0.98	-	=1
4	0.01	-	-	0.99	=1

$$Pj = \sum_{j=0}^{n} Pij\ Pi$$

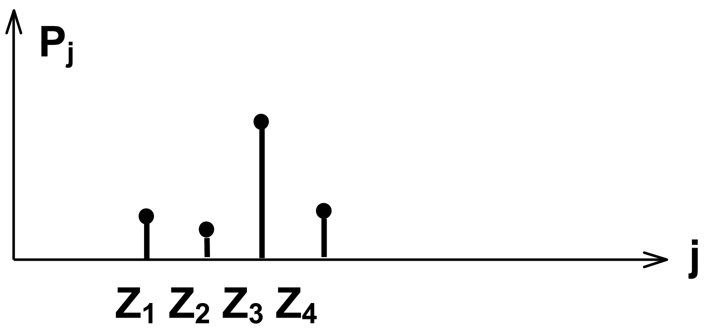
$P_2=0.02P_1+0.9P_2+0.01P_3;$
$P_3=0.04P_1+0.09P_2+0.98P_3;$
$P_4=0.02P_1+ \qquad +0.99P_4;$

решаем только четыре уравнения

$1=P_1+P_2+P_3+P_4;$
$10P_2=2P_1+P_3;$
$2P_3=4P_1+9P_2;$
$P_4=2P_1;$
$1=P_1+P_2+P_3+P_4;$
$10P_2=2P_1+2P_1+\frac{9}{2}P_2;$
$1=P_1+P_2+2P_1+\frac{9}{2}P_2+2P_1;$
$P_4=\frac{2}{9};$ $\qquad P_1=\frac{1}{9};$
$P_3=\frac{2}{9}+\frac{9}{2}\cdot\frac{8}{99}=\frac{58}{99};$

$\frac{11}{2}P_2=4P_1;$
$P_2=\frac{8}{11}P_1;$
$P_2=\frac{8}{99};$

подстановка в уравнение должна сходиться



ВЫВОД: Вероятность выполнения команд с интересующим нас типом адрессации велика.

2. В процессе рассчета нас интересует время, которое мы можем определить, оценивая средние времена пребываний в каждом из состояний.

$$\overline{tj} = \frac{Pjj}{1-Pjj}\Delta t \qquad \overline{t3} = \frac{0.98}{1-0.98}\cdot 2 = 98 \text{ мкс.}$$

$$\overline{trj} = \frac{1}{Pj}\Delta t \qquad \overline{t4} = \frac{0.99}{1-0.99}\cdot 2 = 198 \text{ мкс.}$$

$$\overline{t_2} = 18 \text{ мкс.} \qquad \overline{t_1} = 23 \text{ мкс.}$$

$$T_\Sigma = \sum_{j=1}^{n} \overline{t_j}; \qquad T_\Sigma = 337 \text{ мкс.}$$

$$\xi = \frac{\overline{t_3}}{T_\Sigma} = \frac{98}{337}$$ - доля свободного времени $\approx 0.3;$

т. е. 30% времени работы мы можем использовать для ввода-выода данных.

3. Определяем средние времена возвращений:

$$\overline{t_{r1}} = \frac{9}{1} \cdot 2 = 18 \text{ мкс;} \qquad \overline{t_{r2}} = \frac{99}{8} \cdot 2 = 24.75 \text{ мкс;}$$

$$\overline{t_{r3}} = \frac{99}{58} \cdot 2 = 3.3 \text{ мкс;} \qquad \overline{t_{r4}} = 9 \text{ мкс.}$$

Мы предполагаем, что вероятности переходов заданы /P_{ij} /, но в реальных исследованиях весь расчет практически выполняется в обратном порядке. Дело в том, что в моделирующей системе мы накапливаем множество времен пребывания системы в некотром j-том состоянии /t_j /, организуя счетчики и работая с таблицей состояний. Из этой таблицы можно посчитать t_{rj}. Обрабатывая t_j и t_{rj} получаем средние значения $\overline{t_{rj}}$ и $\overline{t_j}$. Зная шаг испытаний Δt по этим характеристикам можно найти равновесные вероятности состояний:

$$P_j = \frac{1}{\overline{t_{rj}}} \Delta t;$$

а также:

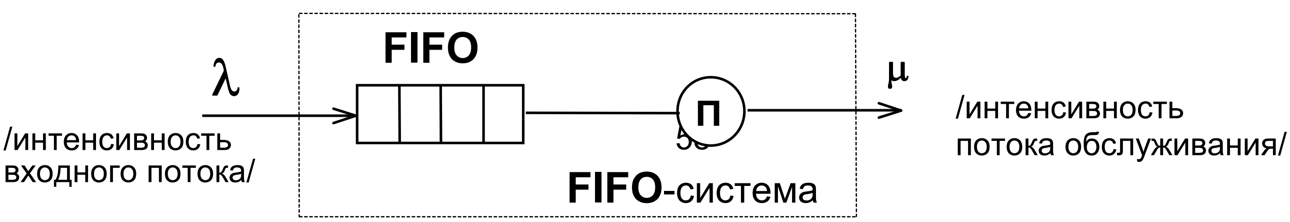
$$P_{jj} = \frac{\overline{t_j}}{\overline{t_j} + \Delta t};$$

а вероятности переходов из одного состояния в другое определяются путем решения алгебраических уравнений, но в несколько ином ключе, находим p_{ij}, которое задаем затем на марковском графе.

РЕШЕНИЕ НЕПРЕРЫВАЕМОЙ МОДЕЛИ СМО С ОДНИМ ЦЕНТРОМ ОБСЛУЖИВАНИЯ

Среди непрерываемых дисциплин обслуживаемых заявок одной из ведущих является дисциплина **FIFO**.

Рассмотрим работу простой системы с одним центром обслуживания /прибором, линией/, который использует организацию очереди **FIFO**. При этом отнесем ее к классу M/M/I.



$\mathbf{T}$**обсл**. $= \dfrac{\mathbf{1}}{\mu}$ /среднее время обслуживания/

Для системы очень важным показателем является коэффициент нагрузки системы ρ, который определяется:

$$\rho = \frac{\lambda}{\mu}$$; меньше 1.

Только в этом случае система может находится в некотором равновесном состоянии и среднее число заявок будет колебаться относительно некоторого фиксированого среднего значения.

В качестве состояния $\mathbf{Z}_j$ принимается такое, когда в системе находится $\mathbf{j}$ заявок, причем / $\mathbf{j}$ -1/ заявка в очереди и одна в очереди на обслуживание.

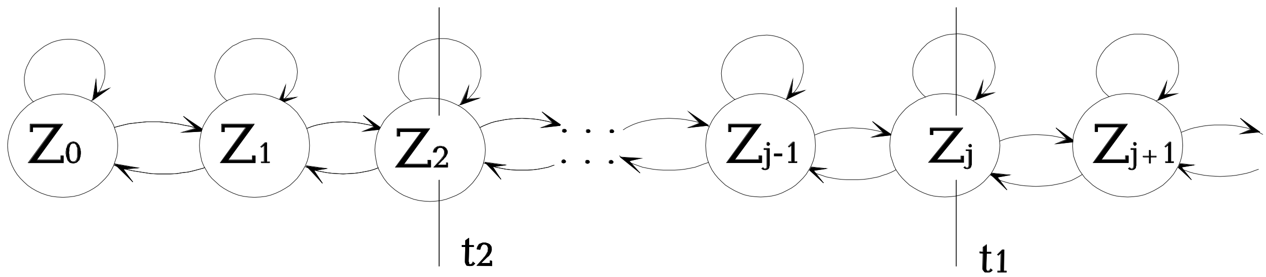
Очевидно, что за некоторый шаг регистрации $\Delta\mathbf{t} = \mathbf{h}$ в данной системе может происходить лишь 3 события:

1. С интенсивностью λ на входе системы появляется входная заявка. Тогда с вероятностью $\lambda\mathbf{h}$ система перейдет из $\mathbf{Z}_j \to \mathbf{Z}_{j+1}$.
2. Но может оказаться, что интенсивностью μ выйдет обслуженная заявка. Тогда с вероятностью $\mu\mathbf{h}$ система перейдет из $\mathbf{Z}_j \to \mathbf{Z}_{j-1}$.
3. За время $\mathbf{h}$ не появляется ни входная, ни выходная заявки. Тогда с вероятностью $\mathbf{1} - \lambda\mathbf{h} - \mu\mathbf{h}$ система перейдет из $\mathbf{Z}_j \to \mathbf{Z}_j$.

Таким образом мы можем построить марковский граф процесса для такой системы. Однако, вводится ограничение, что шаг $\Delta\mathbf{t} = \mathbf{h}$ должен быть настолько малым, чтобы вероятность появления двух событий за этот шаг одного потока была равна 0.

Очевидно, что для модели M/M/I такие условия удовлетворяются при одинаковом шаге регистрации.

Представим марковский граф для **FIFO**-системы:



Текущее состояние отражается точкой на марковском графе. Причем при более интенсивном входном потоке она перемещается вправо. Такой граф называется процессом "рождения и гибели", отражая при этом свойства появления новых объектовв системе или гибели старых. Можно показать, что данная цепь действительно обладает марковским свойством.

В момент t_1 система находится в j-ом состоянии. При этом существую лишь вероятности переходов в соседнее состояние. Если система в другой момент t_2 находится в любом другом, то вероятности переходов из соседнего в то-же самое не меняются /кроме Z_0/.

Таким образом, мы можем считать, что вероятности переходов не зависят от времени.

Рассмотрим какими уравнениями описывается динамический стохастический процесс в данной системе.

Запишем уравнение для вероятности переходов в каждом из состояний на момент времени $t+h$, если в момент t система находилась в другом состоянии.

Пусть в некоторый $t+h$ момент система оказалась в состоянии Z_j, тогда:

$$\begin{cases} P_j(t+h) = (1-\lambda h-\mu h)P_j(t) + \lambda h P_{j-1}(t) + \mu h P_{j+1}(t) \\ j = 1,\ 2,\ 3\ldots \\ P_0(t+h) = (1-\lambda h)P_0(t) + \mu h P_1(t) \end{cases}$$

Воспользуемся конечными разностями:

$$\begin{cases} \dfrac{P_j(t+h) - P_j(t)}{h} = (\lambda + \mu)P_j(t) + \lambda P_{j-1}(t) + \mu P_{j+1}(t) \\ \dfrac{P_0(t+h) - P_0(t)}{h} = \lambda P_0(t) + \mu P_1(t) \end{cases}$$

$$\lim_{h\to 0} \frac{P_j(t+h) - P_j(t)}{h} = \frac{dP_j(t)}{dt}$$

Таким образом, мы перешли к описанию динамического стахостического процесса уравнениями Колмогорова.

$$\begin{cases} \dfrac{dP_j(t)}{dt} = -(\lambda + \mu)P_j(t) + \lambda P_{j-1}(t) + \mu P_{j+1}(t),\ j = 1,\ 2,\ 3\ldots \\ \dfrac{dP_0(t)}{dt} = -\lambda P_0(t) + \mu P_1(t) \end{cases}$$

Для однородных марковских цепей мы можем говорить о завершении переходного процесса и о существовании предельных равновесных вероятностей состояний.

СВОЙСТВО 1

В этом случае соответствующие производные равны нулю. Левые части равны нулю, и мы переходим к системе алгебраических уравнений, которые описывают процесс в однородной марковской цепи. Эти уравнения еще называются уравнениями баланса токов.

$$\begin{cases} P_{j+1} = \dfrac{\lambda+\mu}{\mu} P_j - \dfrac{\lambda}{\mu} P_{j-1}, \quad j = 1, 2, 3\ldots \\ P_1 = \dfrac{\lambda}{\mu} P_0 \end{cases}$$

Эти вероятности уже не зависят от времени. Рассмотрим решение данной системы алгебраических уравнений, введя базовый параметр для СМО:

$$\rho = \frac{\lambda}{\mu}$$

$$\{P_1 = \rho P_0; \qquad P_{j+1} = (1+\rho)P_j - \rho P_{j-1}$$

Для марковского графа $P_0 = 1 - \rho$ /очевидно/.

$$P_1 = \rho(1-\rho); \quad P_2 = (1+\rho)P_1 - \rho P_0 = (1+\rho)\rho P_0 - \rho P_0 = \rho^2 P_0 = \rho^2(1-\rho);$$

$$P_3 = \rho^3(1-\rho) \text{ и т.д.}; \quad P_{j+1} = \rho^{j+1}(1-\rho), \quad j = 1, 2, 3\ldots$$

В такой системе, описываемой однородной марковской цепью имеет место геометрическое распределение вероятностей состояний относительно параметра ρ. ρ можно рассматривать как некоторую вероятность перехода в соседнее состояние /в правую сторону по марковскому графу/.

Известна характеристика для геометрического распределения, т.е. некоторое среднее состояние системы.

$$\bar{j} = \frac{\rho}{1-\rho}$$

/ среднее число - оно означает среднее число заявок в системе/. Такая характеристика обозначается N_s.

Для **FIFO:** $$N_s = \bar{j} = \frac{\rho}{1-\rho}$$

Таким образом, мы получили одну из центральных характеристик СМО для простейшего варианта **FIFO**-системы.

АНАЛИЗ ХАРАКТЕРИСТИК FIFO-СИСТЕМЫ

Любая СМО характеризуется несколькими основными параметрами. Среди них особое место занимают четыре:

1. N_S;

2. D_N - дисперсия числа заявок в системе;

3. T_S - среднее время пребывания заявки в системе;

4. $Tw(\hat{ts})$ - время реакции системы на любое $(\hat{ts})$ задания.

Для **FIFO**: $$Ns = \frac{\rho}{1-\rho} = \frac{\lambda}{\mu(1-\rho)} = \frac{\lambda}{\mu - \lambda}$$

Из этих соотношений мы видим, что $\rho \neq 1$, т.к. система переключается и не будет никакого равновесия стахастического процесса.

2. D_N определяется как дисперсия для геометрического распределения:

$$D_N = \frac{\rho}{(1-\rho)^2};$$

3. T_S определяется использованием некоторой базовой формулы, справедливой для всех СМО - <u>это формула Литла (Закон Ома для СМО):</u>

$$Ns = \lambda\, Ts; \quad Ts = \frac{1}{\lambda} Ns$$

Формула Литла справедлива для всех распределений в потоках и определения всех типов СМО.

Для **FIFO**: $$Ts = \frac{1}{\lambda}\frac{\rho}{1-\rho} = \frac{1}{\mu(1-\rho)} = \frac{1}{\mu - \lambda};$$

$$\textbf{4. } Tw(\hat{ts}) = \hat{ts} + TQ$$

T_Q- среднее время пребывания в очереди /зависит от дисциплины обследования очереди/.

Для **FIFO**: $$TQ = \frac{1}{\mu} Ns = \frac{81\rho}{\mu(1-\rho)} = \frac{\rho}{\mu(1-\rho)};$$

Имеет место важный переход:

Для всех систем справедливы соотношения, которые можно использовать в качестве контроля:

$$Ts = Tw\left(\hat{ts} = \frac{1}{\mu}\right).$$

Среднее время пребывания в системе - это реакция системы на среднее задание. Для **FIFO**:

$$\mathbf{Tw}\left(\hat{\mathbf{t}}_{\mathbf{s}}=\frac{\mathbf{1}}{\mu}\right)=\frac{\mathbf{1}}{\mu}+\frac{\rho}{\mu(\mathbf{1}-\rho)}=\frac{\mathbf{1}-\rho+\rho}{\mu(\mathbf{1}-\rho)}=\frac{\mathbf{1}}{\mu(\mathbf{1}-\rho)}.$$

Рассматриваемые характеристики **FIFO**-системы справедливы лишь для простейшего потока обслуживаний /экспоненциального/ с коэффициентом вариаций равным 1.

Если распределение выходного потока любое, то необходимо учитывать коэффициент вариаций выходного потока $\mathbf{g} \neq \mathbf{1}$.

Для более общего случая расчета **FIFO**-системы используются известные соотношения полученные на основе формулы Хингена-Полачика: формула $\mathbf{T_s}$ оценивает при любом $\mathbf{g}$ выходного потока.

1. $$\boxed{\mathbf{Ts}=\frac{\mathbf{1}}{\mu}+\frac{\rho(\mathbf{1}+\mathbf{g})}{\mathbf{2}\mu(\mathbf{1}-\rho)};}$$ 1-я характеристика для **FIFO**.

Можно показать, что предыдущие соотношения для $\mathbf{T_s}$ есть лишь частный случай этой общей формулы при $\mathbf{g}=\mathbf{1}$.

2. $\mathbf{Ns}=\lambda\mathbf{Ts}$ - формула Литла =

$$=\frac{\lambda}{\mu}+\frac{\lambda\rho(\mathbf{1}+\mathbf{g})}{\mathbf{2}\mu(\mathbf{1}-\rho)}=\boxed{\rho+\frac{\rho^{\mathbf{2}}(\mathbf{1}+\mathbf{g})}{\mathbf{2}(\mathbf{1}-\rho)}}.$$

3. $\mathbf{Tw}\left(\hat{\mathbf{t}}_{\mathbf{s}}\right)=\mathbf{t_s}+\mathbf{T_Q}$

$$\boxed{\mathbf{T_Q}=\frac{\rho(\mathbf{1}+\mathbf{g})}{\mathbf{2}\mu(\mathbf{1}-\rho)}}$$

FIFO-система - это система пакетной обработки заданий.

Особенность: передав задание на обработку, оно выполняется в неопределенное время.

РЕШЕНИЕ ПРЕРЫВАЕМОЙ МАРКОВСКОЙ МОДЕЛИ СМО С ОДНИМ ЦЕНТРОМ ОБСЛУЖИВАНИЯ И RR-ДИСЦИПЛИННОЙ ОЧЕРЕДЬЮ УПРАВЛЕНИЯ /RR-СИСТЕМЫ/

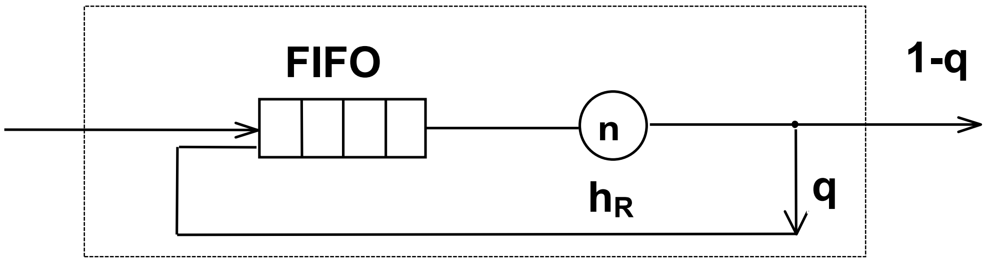
В некоторых системах заявка, находящаяся в приборе, может быть снята с обслуживания, а прибор будет передан другой заявке. Это прерываемые дисциплины обслуживания /**RR**-дисциплины/.

Заявка, снятая с обслуживания снова попадает в очередь. А прибор предоставляется каждой заявке периодически на фиксированное время.

h_R - шаг времени предоставления прибора.

Естественно, что в такой системе обслуживания заявка может с некоторой вероятностью **q** снова вернуться в очередь системы, не получив полного обслуживания и с вероятностью **1-q** может быть обследована.

Система с такой организацией очереди называется **RR**-системой, ее структуру можно представить в следующем виде:



Если любое задание $t_s \approx k \cdot h_R$, то очевидно, что эта заявка **k-1** раз вернется в систему и на **k**-ом шаге она выйдет из системы. Анализ процессов в такой системе достаточно сложный и он основывается на исследовании 4-й характеристики СМО, т.е. времени реакции системы на задание:

$$T_W\left(\hat{t}_s\right)$$

$$1.\ T_S = T_W\left(t_s = \frac{1}{\mu}\right)$$

поэтапно мы выйдем на все характеристики

$$2.\ N_S = \lambda \cdot T_S$$

$$3.\ D_N$$

оценить можем дисперсию

Рассмотрим некоторые предпосылки для анализа базовой характеристики **RR**-системы

$$T_W(t_s)$$

при некотором $t_s = k \cdot h_R, \quad k \neq 1$. Тогда оно **k**-раз будет принято прибором и **k-1**-раз вернется в очередь **FIFO**, т.е. можно выделить **k**-циклов нахождения и обследования этой заявки в системе.

Длительность цикла $t_{i,s}$ - общее время пребывания заявки в системе

$$T_W\left(\hat{t}_s\right) = \sum_{i=1}^{k} t_{i,s}$$

в исходном положении в системе находится **j**-заявок.

Z_j - состояние.

1-я заявка находится в приборе, а **(j-1)** - в очереди.

В момент прибытия t_s - задания **j-1**-заявка находится впереди в очереди.

Тогда время 1-го цикла обслуживания будет выглядеть:

$$t_{i,s} = \alpha \cdot h_R + (j-1)h_R + h_R$$

$$0 \langle \alpha \langle 1$$

1-е слагаемое - время завершения обследования заявки находящейся в приборе.

В ряде случаев этим слагаемым можно пренебречь.

$(j-1)h_R$ - обслуживание заявок, находящихся впереди тестового задания.

h_R-шаг времени прибора для t_s-задания, первый раз предоставленного прибору.

На втором цикле обслуживания впереди t_s-задания появляются заявки, формируемые двумя источниками.

<u>1-й источник: входной поток заявок</u>

λ - интенсивность этого потока.

Общее число заявок, которые появляются представленные 2-м обследованием t_s-задания, определяет длительность первого цикла

$$\lambda \cdot t_{1,s}$$

Каждая из этих заявок получит время h_R.

<u>2-й источник: часть заявок</u>, которые находятся в 1-м цикле впереди T_S-задания; с вероятностью **q** будут возвращены в очередь и не получат полного обследования.

Количество таких заявок определяется:

/вероятность/ × /общее число заявок/

$$q \cdot l_1; \quad l_1 = j_1.$$

Но для l_2 мы уже не знаем сколько их будет.

Каждая из этих заявок получит h_R для обслуживания.

Таким образом, длительность 2-го цикла обследования будет равна

$$t_{2,s} = q \cdot l_1 \cdot h_R + \lambda \cdot t_{1,s} . h_R + h_R$$

Все остальные циклы синхронны, т.е. нет $\alpha \cdot h_R$.

$q \cdot l_1 \cdot h_R$ - время обслуживания возвращенных заданий;

$\lambda \cdot t_{1,s} \cdot h_R$ - время обслуживания новых заявок, появившихся извне. В течение 1-го цикла за t_s-заданием и оказывается впереди t_s-задания.

h_R - обслуживание t_s-задания.

Аналогичным образом можно записать времена и для всех остальных обслуживаний t_s-заданий.

$$t_{3,s} = q \cdot l_2 \cdot h_R + \lambda \cdot t_{2,s} \cdot h_R + h_R;$$

$$t_{4,s} = q \cdot l_3 \cdot h_R + \lambda \cdot t_{3,s} \cdot h_R + h_R;$$

...

$$t_{k,s} = q \cdot (l_{k-1} \cdot h_R) + \lambda \cdot (t_{k-1,s} \cdot h_R) + h_R.$$

Прежде чем суммировать времена циклов, введем некоторые преобразования:

$$l_1 \cdot h_R = t_{1,s} \quad \text{при } \alpha = 0, \ l_1 = j$$

$$l_2 \cdot h_R = t_{2,s}$$

...

$$l_{k-1} \cdot h_R = t_{k-1,s}$$

Произведем замену и вычислим $t_{k,s}$:

$$t_{k,s} = (q + \lambda \cdot h_R) t_{k-1,s} + h \cdot l$$

Введем новый эквивалентный параметр $q_R = q + \lambda \cdot h_R$, который называется <u>вероятностью появления заявок в **RR**-системе</u>. Второе слагаемое можно рассматривать как вероятность появления нового задания в системе за h_R.

Получим упрощенные уравнения системы; для $t_{1,s}$ и $t_{2,s}$ запись уравнения остается прежней, а для остальных циклов сделаем некоторые преобразования.

$$t_{3,s} = q_R \cdot t_{2,s} + h_R;$$

$$t_{4,s} = q_R \cdot t_{3,s} + h_R;$$

$$t_{k,s} = q_R \cdot t_{k-1,s} + h_R.$$

Выразим все времена через $t_{2,s}$:

$$t_{3,s} = q_R \cdot t_{2,s} + \frac{(1 - q_R)}{(1 - q_R)} \cdot h_R;$$

$$t_{4,s} = q_R \cdot \left(q_R \cdot t_{2,s} + h_R\right) + h_R = q_R^2 \cdot t_{2,s} + (1 + q_R) \cdot h_R =$$

$$= q_R^2 \cdot t_{2,s} + \frac{(1 - q_R^2)}{(1 - q_R)} \cdot h_R;$$

$$t_{5,s} = q_R^3 \cdot t_{2,s} + (1 + q_R + q_R^2) \cdot h_R = q_R^3 \cdot t_{2,s} + \frac{1 - q_R^3}{1 - q_R} \cdot h_R;$$

$$t_{k,s} = q_R^{k-2} \cdot t_{2,s} + \frac{1-q_R^{k-2}}{1-q_R} \cdot h_R$$

h_R выбираем столь малое, чтобы $q_R \le 1$.

Суммируем эти времена:

$$T_W\left(\hat{t_s}\right) = \sum_{i=1}^{k} t_{i,s};$$

$$T_W\left(\hat{t_s}\right) = t_{1,s} + t_{2,s} + q_R \cdot t_{2,s} + q_R^2 \cdot t_{2,s} + \ldots + q_R^{k-2} \cdot t_{2,s} +$$

$$+\frac{1-q_R}{1-q_R} \cdot h_R + \frac{1-q_R^2}{1-q_R} \cdot h_R + \ldots + \frac{1-q_R^{k-2}}{1-q_R} \cdot h_R = t_{1,s} + \frac{1-q_R^{k-1}}{1-q_R} \cdot$$

$$\cdot t_{2,s} + \frac{h_R}{1-q_R} \cdot \underbrace{\left(1-1+1-q_R+1-q_R^2+\ldots-q_R^{k-2}\right)}_{\frac{1-q_R^{k-1}}{1-q_R}} =$$

$$= t_{1,s} + \frac{1-q_R^{k-1}}{1-q_R} \cdot t_{2,s} + \frac{(k-1)h_R}{1-q_R} - \frac{1-q_R^{k-1}}{1-q_R} \cdot h_R =$$

$$= t_{1,s} + \frac{(k-1)h_R}{1-q_R} + \frac{1-q_R^{k-1}}{1-q_R} \cdot \left(t_{2,s} - \frac{h_R}{1-q_R}\right)$$

Проведем преобразования в скобках:

$$t_{2,s} - \frac{h_R}{1-q_R} = \lambda \cdot t_{1,s} \cdot h_R + q \cdot l_1 \cdot h_R + h_R - \frac{h_R}{1-q_R} =$$

$$= h_R \cdot \left[\lambda \cdot t_{1,s} + q \cdot l_1 + \left(1 - \frac{1}{1-q_R}\right)\right] = h_R \cdot \left[\lambda \cdot t_{1,s} + q \cdot l_1 - \frac{q_R}{1-q_R}\right] =$$

$$= h_R \cdot \left[\lambda \cdot t_{1,s} + q \cdot l_1 - \frac{q}{1-q_R} - \frac{\lambda \cdot h_R}{1-q_R}\right] = -h_R \cdot \frac{q}{1-q_R} +$$

$$+ h_R \cdot \left(\lambda \cdot t_{1,s} + q \cdot l_1 - \frac{\lambda \cdot h_R}{1-q_R}\right)$$

Эти преобразования направлены на получение формы удобной для некоторого предельного перехода к системе работающей с $h_R \to 1$.

$$T_W\left(\hat{t}_s\right) = t_{1,s} + \frac{(k-1)h_R}{1-q_R} - \frac{1-q_R^{k-1}}{(1-q_R)^2} \cdot h_R +$$

$$+\frac{1-q_R^{k-1}}{1-q_R} \cdot h_R \cdot \left(\lambda \cdot t_{1,s} + q \cdot l_1 - \frac{\lambda \cdot h_R}{1-q_R}\right)$$

Анализ для **RR**-системы этого выражения достаточно сложен. Поэтому чаще пользуются не моделью **RR**-системы, а ее идеальным представлением при $h_R \to 1$.

АНАЛИЗ ПРОЦЕССОВ В RR-СИСТЕМЕ И ЕЕ ИДЕАЛИЗИРОВАННОЙ МОДЕЛИ – PS-СИСТЕМЫ

Поскольку h_R достаточно мало, то мы можем перейти к идеализированной модели **RR**-системы, так называемой **PS**-системы /разделение времени процессора/.

В такой системе предполагается, что $h_R \to 1$, это означает, что все задания прерываются, а это значит, что $q \to 1$ и в идеальном случае каждая заявка получает свою долю обслуживания в системе. Получается, что все заявки обследуются как-бы паралельно. Но поток одинаковый по скорости обслуживания обратно пропорционален числу заявок.

Преимуществом **RR**-системы, как и **PS**-системы, является то, что короткие задания в этой системе обслуживаются быстро.

Рассмотрим переход в оценках характеристик **RR**-системы при ее идеализации, т.е. $h_R \to 1$.

1. $h_R \to 1$, $q \to 1$, $(1-q) \to 1$

Однако это не значит, что из системы не будут выходить обслуженные заявки. На выходе системы будет иметь место некоторый поток заявок с интенсивностью μ.

$$\rho = \frac{\lambda}{\mu} = const,$$

для **RR**-системы ρ определяется:

$$\rho = \frac{\lambda \cdot h_R}{1-q} = const;$$

$\lambda \cdot h_R$ – вероятность появления заявки входного потока;

$1 - q$ – вероятность появления заявки в выходном потоке.

$$\mu = \frac{1-q}{h_R} = const.$$

2. $h_R \to 1,\ \ k \to \infty,\ \ k \cdot h_R = \hat{t}_s = \text{const}.$

3. $q \to 1$, то $q_R \to 1$.

$$\underbrace{\frac{1-q_R^{k-1}}{1-q_R}}_{0} = \underbrace{1+q_R+q_R^2+\ldots+q_R^{k-2}}_{\underbrace{k-1}_{\infty}}; \qquad \frac{1-q_R^{k-1}}{1-q_R} \cdot h_R = \hat{t}_s.$$

Проанализируем слагаемые, которые входят в общее выражение **Tw:**

$T_W = (\hat{t}_S)$ для **RR** - системы

$t_{1,2} \to \infty$

$$1-q_R = 1-q-\lambda \cdot h_R = (1-q) - \underbrace{\frac{\lambda \cdot h_R (1-q)}{(1-q)}}_{\frac{\lambda \cdot h_R}{(1-q)}=\rho} = (1-q)(1-\rho);$$

$$1-q_R \to 0.$$

В выражении $T_W\ (\hat{t}_s)\ \ t_{1,s} \to 0.$ Преобразуем 2-е и 3-е слагаемое:

$$\frac{(k-1)h_R}{1-q_R} - \frac{h_R \cdot q}{1-q_R} h_R \cdot \underbrace{\frac{1-q_R^{k-1}}{1-q_R}}_{\frac{1-q_R^{k-1}}{1-q_R}=k-1} = \frac{(k-1)h_R}{1-q_R} \cdot (1-q) =$$

$$= \frac{(k-1)h_R}{(1-q)(1-\rho)}(1-q) = \frac{\overbrace{(k-1)h_R}^{(k-1)h_R=\hat{t}_s}}{(1-\rho)} = \frac{\hat{t}_s}{(1-\rho)};$$

Преобразуем последнее слагаемое:

$$\frac{\lambda \cdot h_R}{1-q_R} = \frac{\overbrace{\lambda \cdot h_R}^{\frac{\lambda \cdot h_R}{(1-q)}=\rho}}{(1-q)(1-\rho)} = \frac{\rho}{(1-\rho)}.$$

Тогда:

$$T_W(\hat{t}_s) = \frac{\hat{t}_s}{1-\rho} + \hat{t}_s\left(j - \frac{\rho}{1-\rho}\right),$$ где

j- начальное число заявок.

Можно предположить, что $j = N_s$. Учитывая, что среднее число заявок $N_S = \frac{\rho}{1-\rho}$, то выражение в скобках $\rightarrow 0$.

Тогда, для среднего начального состояния

$$T_W(\hat{t}_s) = \frac{\hat{t}_s}{1-\rho}$$ - классическая оценка для **PS**-системы.

Анализируя это выражение: для коротких заданий время реакции системы может быть как угодно мало. Можно перейти к другой характеристике **PS**-системы:

<u>2.</u> $$T_S = T_W\left(t_s = \frac{1}{\mu}\right) = \frac{1}{\mu \cdot (1-\rho)}.$$

Оценка T_s совпала с соответствующей характеристикой для **FIFO**-системы.

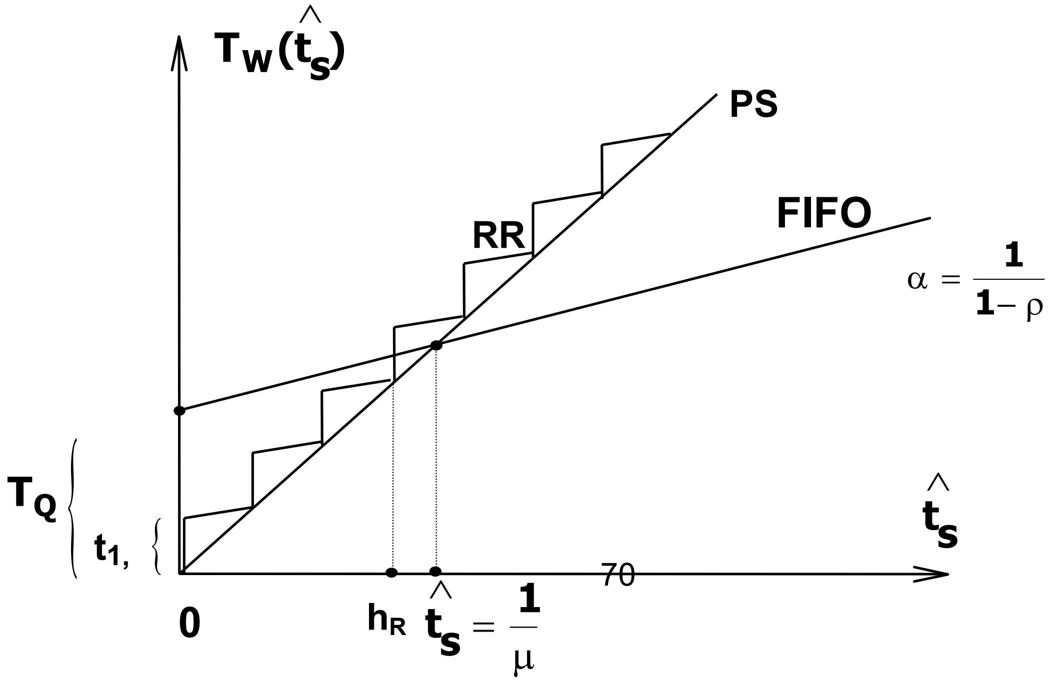
<u>3.</u> $$N_S = \lambda \cdot T_S; \quad \frac{\lambda}{\mu} = \rho;$$

$$N_S = \lambda \cdot \frac{1}{\mu(1-\rho)} = \frac{\rho}{1-\rho}.$$

Среднее число заявок определяется так, как мы определили в формуле и является такой ка и **FIFO**.

В целом характеристики **PS**-системы лучше для коротких заданий, но хуже для длинных и наоборот для **FIFO**.

Характеристики времени реакции системы для **FIFO**, **RR** и **PS** -систем удобно сравнить.



Точка пересечения - характеристики систем совпадают.

Ступеньки не одинаковы. Если **$j=0$,** то ступеньки пойдут по прямой **PS**.

Система с дисциплиной **SJF** очень близка по характеру обслуживания к **PS** - системе, но требуется дополнительный анализ.

МОДЕЛИРОВАНИЕ СЛОЖНЫХ СИСТЕМ НА ОСНОВЕ АППАРАТА СЕТЕЙ ПЭТРИ

Сеть Пэтри - как математическая структура и ориентированный граф

Для представления и исследования паралельных процессов с отражением механизмов взаимодействия все шире используется язык сетей Пэтри.

Это математический аппарат описания процессов, которые позволяют воспроизводить поведение заявок в системе, использование ресурсов, взаимодействие заявок между собой и заявок с ресурсами.

Удобство сетей Пэтри - этот язык может быть использован на самых различных уровных описания /при моделировании ЭВМ, системный уровень/.

В отличии от марковских цепей в сетях Пэтри отражается специфика реализации взаимодействия и отражается состояние различных ресурсов цепи.

Имеются простые переходы от модели на основе стохастической модели сетей Пэтри к марковской цепи.

ОПРЕДЕЛЕНИЕ 1. Сети Пэтри - математическая структура, представляемая четверкой

$$СП=<P, T, I, Q>,$$

I - множество позиций сети $P=\{P_1...P_n\}$;

$T=\{t_1,t_2,...t_m\}$- множество переходов в сети, на которых задаются следующие функции:

функции $I(t_j)$представлены $I=\{I(t_1), I(t_2),..., I(t_n)$ входов-переходов;

$Q(t_j)$- функции выходов-переходов.

$$Q(t_j)= \{Q(t_1),...,Q(t_n)\}$$

Эти функции представляются как некоторые векторы, которые отражают конфигурацию системы и определены на всех позициях сети.

Вектор $\bar{I}(t_j)$ отображает все позиции сети на входы реакций

$\mathbf{P \rightarrow t_j}$,

$\mathbf{\overline{Q}(t_j)}$ - как отображение всех позиций сети на выходы $\mathbf{t_j}$ переходов.

Сеть Пэтри может быть задана четырьмя множителями, как математическая структура.

ПРИМЕР:

$\mathbf{P=\{P_1, P_2, P_3\}}$	$\mathbf{T\{t_1, t_2, t_3, t_4\}}$
$\mathbf{I(t_1)=\{P_1\}}$	$\mathbf{Q(t_1)=\{P_2\}}$
$\mathbf{I(t_2)=\{P_2\}}$	$\mathbf{Q(t_2)=\{P_3\}}$
$\mathbf{I(t_3)=\{P_1, P_2\}}$	$\mathbf{Q(t_3)=\{P_1, P_3\}}$
$\mathbf{I(t_4)=\{P_3\}}$	$\mathbf{Q(t_4)=\{P_2, P_2\}}$.

Наиболее удобно и наглядно представление сети Пэтри своим графом.

ОПРЕДЕЛЕНИЕ 2. Граф сети Пэтри: $\mathbf{/G/=\langle V, A\rangle}$ представляет собой двудольный мультиграф, задаваемый парой $\mathbf{\langle V, A\rangle}$, где

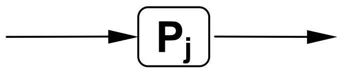
V - множество вершин графа, разделяемое на два непересекаемые подмножества: **P** и **T;**

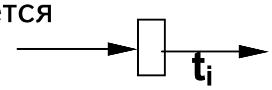
$$\mathbf{V=P\cup T;\ P\cap T=0.}$$

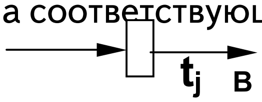
А каждый переход связан с позицией и наоборот.

A - множество дуг равное $=\{\alpha_1,\ldots,\ \alpha_s\}$, соединяющих переходы с позициями, либо позиции с переходами /**P** с **T**/.

G - мультиграф, т.к. одна и та же позиция может быть несколько раз связана с одним и тем же переходом.

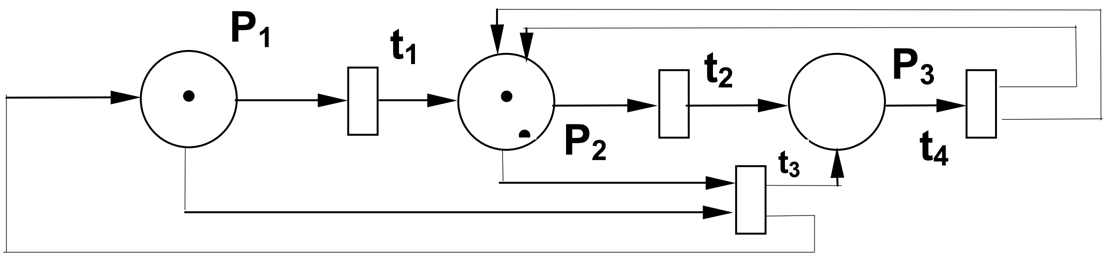
Соответствующая позиция обозначается 

,
а соответствующий переход - планкой 

- мгновенный,


- временной.

Рассмотрим граф общей сети Пэтри, которая задана как математическая структура /согласно вышеизложенного примера/.



Позиции, представленные функциями входов $\mathbf{I(t_i)}$, являются входными позициями по отношению к переходу $\mathbf{t_i}$, а позиции представленные функциями выходов $\mathbf{Q(t_i)}$, являются выходными позициями для $\mathbf{t_i}$.

Граф сети Пэтри обладает свойством дуальности; соответственно графом сети Пэтри мы можем представить дуальный граф, в котором все позиции заменены переходами и наоборот /P $\rightleftarrows$ T /.

Тогда **I, T** будут заданы на множестве переходов относительно позиций: $\mathbf{I(P_i), Q(P_i)}$.

Дуальные графы используются в связи с несколько иной интерпритацией состояния системы.

ПОНЯТИЕ СОСТОЯНИЯ СЕТЕЙ ПЭТРИ. МАРКИРОВКА СЕТЕЙ ПЭТРИ

Объекты сетей Пэтри чаще всего представляются точками /фишками/ в той или иной позиции сети, хотя могут использоваться специальные элементы.

Фишки: ресурс или процессор **V** , память **V**, задание и так далее.

Размещение фишек на позициях сети Пэтри отражает текущее мгновенное состояние системы.

Так представляются дискретные системы. Но в состоянии системы отражается не только состояние ресурсов, а и фазы продвижения заявок в системе.

Множество фаз продвижения одной заявки в системе называется процессом, а сама заявка и ее движение называется трансактом.

Соответствующее состояние сети Пэтри отражается маркировкой **M**. Это вектор, который задается на множестве **P** и отражает число фишек на позициях сети.

ОПРЕДЕЛЕНИЕ. Маркировкой сети Пэтри называется отображение множества позиций **P** сети Пэтри на множество неотрицательных целых чисел:

$$\mathbf{P \Rightarrow \{0,1,2,\ldots,N\}}.$$

Причем соответствующий элемент маркировки $\mathbf{m_i}$ отражает число фишек, находящихся в позиции $\mathbf{P_i}$ /на рассмотренной в лекции сети Пэтри задана маркировка $\mathbf{M^0 = /1,2,0/}$, /... ,

Для пояснения работы модели на основе сети Пэтри обязательно использование специальной таблицы назначения фишек /либо это определено словесно/.

Позиция сети	Назначение фишки
$\mathbf{P_1}$	Заявка, генерируемая извне /колличество фишек равно колличеству заявок/
$\mathbf{P_2}$	Заявка, обслуживаемая системой /колличество фишек равно колличеству заявок/

P_3	Свободный прибор в системе /колличество фишек равно колличеству заявок/

<u>СОБЫТИЯ, ЗАПУСКИ ПЕРЕХОДОВ И ВЫПОЛНЕНИЕ СЕТЕЙ ПЭТРИ</u>

Собственно развитие процессов в сетях Пэтри иммитируется перемещением фишек из одной позиции в другую. Каждая новая маркировка отражает некоторую мгновенную фазу развития процессов. Движение фишек обеспечивается за счет запусков и срабатывания переходов. Мгновенный переход t_j^m запускается и срабатывает на один и тот же момент времени, т.е. переход срабатывает мгновенно. В то же время другие переходы /$t_s^В$/ /временные/ для времени срабатывания и запуска разделены некоторым случайным временем τ с известным распределением. <u>В зависимости от организации моделирования в сетях Пэтри могут развиваться синхронные или асинхронные процессы.</u>

В 1-м случае все моменты запусков, переходов определяются извне, т.е. задаются внешней средой, т.е. должна существовать некоторая дополнительная схема синхронизации работы модели.

Наиболее эффективно в сетях Пэтри представляются асинхронные процессы. Но при этом в качестве базовых средств описания должны использоваться стохастические сети. События запусков при асинхронном развитии процессов, определяются условиями, которые возникают в сетях. Если некоторые переходы оказываются разрешенными в той или иной маркировке, то они сразу же срабатывают.

Условием разрешения некоторого t_j перехода в маркировке M' является условие:

$$\overline{M}' \geq \bar{I}(t_j).$$

Дело в том, что на всех возможных маркировках сети Пэтри задается два важных отношения:

1. Отношение эквивалентности: $M' = M''$
2. Отношение покрываемости: $M' \geq M''$

$$\left.\begin{array}{l} M' = M'' \\ M' \geq M'' \end{array}\right\}$$

<u>ОПРЕДЕЛЕНИЕ 4.</u> Маркировка M' сети Пэтри называется эквивалентной маркировке M'', если для всех элементов m_j маркировок выполняется условие:

$$\underbrace{M'(P_i)}_{m_i} = \underbrace{M''(P_i)}_{m_i}.$$

Это означает, что число фишек во всех позициях сети одинаково для каждой из этих маркировок.

<u>ОПРЕДЕЛЕНИЕ 5.</u> Маркировка **M'** сети Пэтри называется покрывающей по отношению к **M''** той же сети, если для всех элементов соответствующих маркировок выполняется условие:

$$\mathbf{M'(P_i) \geq M''(P_i)}$$
$$\mathbf{M' \geq M''}$$

Это означает, что хотя бы в одной позиции сети маркировки **M'** имеется хотя бы на единицу большее число фишек.

Аналогичное свойство типа покрываемости может быть использовано и для разрешения переходов в некоторой маркировке **M'**. При этом задается отношение между маркировкой и функцией входов

$$\mathbf{M'^{\,m} \geq I(t_j)^{\,m}}.$$

Выделим важное условие для выполнения сетей:

<u>Переход может быть запущен, если он разрешен.</u>

<u>ОПРЕДЕЛЕНИЕ 6.</u> Переход $\mathbf{t_j}$ в некоторой маркировке **M'** и называется разрешенным, если количество фишек во всех входных позициях данного перехода превышает или равно числу дуг, соединяющих эти позиции с данным переходом.

Это правило интерпритируется известным отношением:

$$\mathbf{M'^{\,m} \geq I(t_j)^{\,m}}.$$

Слева отражено положение фишек.

Справа - конфигурация сети.

Таким образом, для исследования процессов запусков переходов рассмотрим простой пример двух конкурирующих процессов /или кооп-ов/, в которых каждая выходная обслуженная заявка является рекламой для привлечения заявокиз очереди другого кооп-ва.

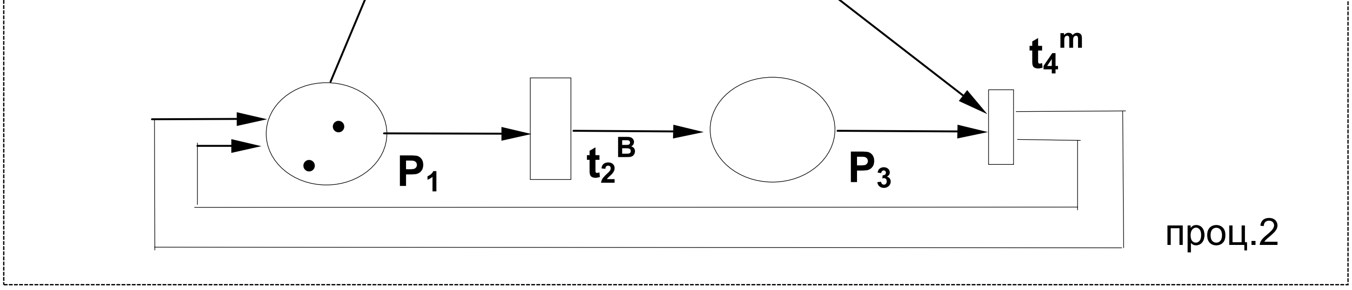


Таблица 1. Назначений фишек

$\mathbf{P_1}$ - заявка в очереди процесса 2;

$\mathbf{P_2}$ - заявка в очереди процесса 1;

$\mathbf{P_3}$ - обслуженная заявка процесса 2;

$\mathbf{P_4}$ - обслуженная заявка-реклама процесса 1.

Таблица 2. Событий запусков, срабатываний переходов

$\mathbf{t_1^m}$ - прием заявки в очередь первого процесса;

$\mathbf{t_2^B}$- прием заявки на обслуживание прибором второго процесса /запуск/;

Событие срабатывания - завершение обслуживания заявки прибором процесса 2;

$\mathbf{t_3^B}$- прием заявки на обслуживание прибором процесса 1;

Событие срабатывания - завершение обслуживания заявки прибором процесса 1;

$\mathbf{t_4^m}$- прием заявки в очередь процесса 2.

Данная модель представленна стохастической сетью Пэтри, в которой: $\mathbf{t_2^B}$ и $\mathbf{t_3^B}$ иммитируют работу приборов, которые обследуют заявки с некоторой интенсивностью $\boldsymbol{\mu_1}$, $\boldsymbol{\mu_2}$. В данной системе развиваются асинхронные процессы взаимодействия кооператоров, но взаимодействие это на основе конфликта /конкуренции/.

Пусть исходная маркировка $\mathbf{M^0}$ /две заявки в очереди второго кооператора и одна у первого/, т.е.

$$\mathbf{M^0(2,1,0,0)}$$

В данной маркировке разрешены сразу 2 перехода:

$$\mathbf{t_2^B} \text{ и } \mathbf{t_3^B}$$

Можем записать:

$$\mathbf{I(t_2^B)}=/0,1,0,0/.$$
$$\mathbf{I(t_4^M)}=/1,0,0,1/.$$
$$\mathbf{I(t_2^B)}=/0,1,0,0/.$$
$$\mathbf{I(t_1^M)}=/1,0,0,1/.$$

$\mathbf{M^0}$/2,1,0,0/, т.е. проверив условие, получим, что $\mathbf{t_2}$, $\mathbf{t_3}$ оба разрешены, оба запустятся.

Предположим, что первым срабатывает прибор $\mathbf{t_2^B}$ /процесс 2/. Появляется новая маркировка:

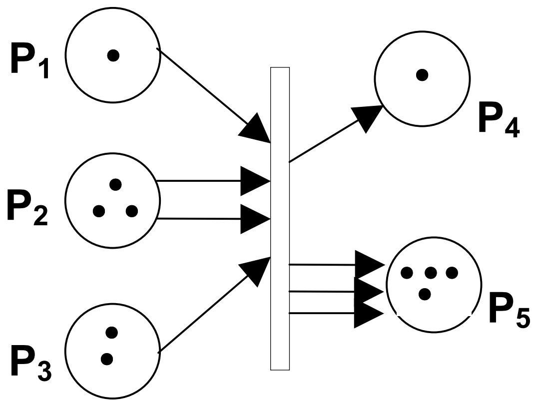
$$t_2^{в} \rightarrow M'(1,1,1,0).$$

В этой маркировке не является разрешенным $t_4^{м}$, а также $t_3^{в}$, но он до сих пор не обслужил свою заявку. $t_4^{м}$ срабатывает сразу, поскольку имеется заявка-реклама в позиции P_3, которая привлекает недообслуженную заявку из очереди процесса 1.

$$t_4^{м} \rightarrow M''(3,0,0,0).$$

В дальнейшем, срабатывает трижды $t_2^{в}$, $\rightarrow M'''$/0,0,3,0/,что завершает победу второго процесса.

…о реализовать правило срабатывания



M' /1,3,2,0,1/

I(tj)=/1,2,1,0,0/

M''/0,1,1,1,4/

/после срабатывания t_j перехода/

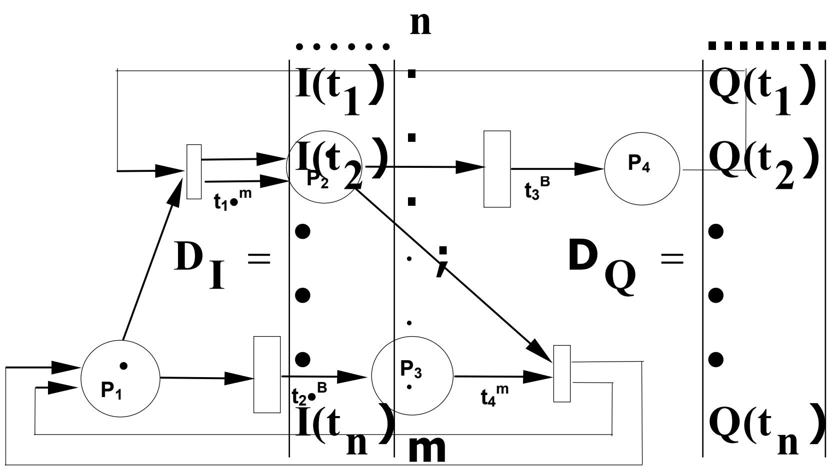
ВЫПОЛНЕНИЕ СЕТЕЙ ПЭТРИ НА ОСНОВЕ РЕШЕНИЯ МАТРИЧНЫХ УРАВНЕНИЙ

Движение фишек в сети может быть представлено в модели достаточно простыми математическими средствами.

собственно сеть Пэтри отражается в машине двумя матрицами:

D_I- матрица входов

D_Q- матрица функций выходов



М - мгновенный
В - временный
Тогда:

$$D_I = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{vmatrix} \quad D_Q = \begin{vmatrix} 0 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 2 & 0 & 0 & 0 \end{vmatrix}$$

<u>Пример:</u>

Используются специальные вектора запуска перехода:

$\bar{\delta}(t_j)$ - вектор запуска одного перехода t_j;

$\bar{\delta}(G)$ - вектор запуска последовательности переходов **G** .

Этот вектор задается колличеством запусков, но при этом теряется их последовательность.

n - число позиций
m - число переходов

<u>Пример:</u>

$\bar{\delta}$ **($t_2\bullet^B$)=(1000)**
$\bar{\delta}$ **(G)=(1201)**$\Rightarrow$
G=<$t_2\bullet^B$, $t_1\bullet^m$, t_4^m, $t_1\bullet^m$) либо
G=<$t_2\bullet^B$, t_4^m, $t_1\bullet^m$, $t_1\bullet^m$)

G

$M^0 \to M'$ - здесь порядок переходов роли не играет.

Для того, чтобы быть уверенными в разрешении переходов: $M'' \geq I(t_j)$ целесообразно использовать векторы запуска по одному переходу $\bar{\delta}(t_j)$.

Баланс движения фишек в сетях Пэтри определяется следующим матричным уравнением:

$$M'' = M' - \overbrace{\bar{\delta}(G) D_I}^{\text{количество фишек, снимаемых с позиций сети при запуске.}} + \underbrace{\bar{\delta}(G) D_Q}_{\text{количество фишек размещающихся в выходныхпозициях переходов в результате запуска.}}$$

M' - исходная маркировка /до запуска/.

M'' - конечная маркировка /после запуска/.

$$M'' = M' + \bar{\delta}(G)\underbrace{(D_Q - D_I)}_{D_R}$$

D_R - результатирующая матрица для сети Пэтри

D_R отражает конфигурацию сети Пэтри и не зависит от маркировок.

Условие разрешения переходов можно проверить только с помощью матрицы D_I.

$M'' = M' + \bar{\delta}(G) \bullet D_R$ - окончательно.

<u>Пример:</u> /для представленной сети Пэтри/:

$$D_R = D_Q - D_I = \begin{pmatrix} -1 & 0 & 1 & 0 \\ -1 & 2 & 0 & -1 \\ 0 & -1 & 0 & 1 \\ 2 & -1 & -1 & 0 \end{pmatrix}$$

Пусть $M = \overbrace{/2,1,0,0/}^{n}$

$G = \langle t_2 \bullet^{B};\ t_4^{m} \rangle \Rightarrow M'' = /3,0,0,0/.$

$\Updownarrow$

$$\bar{\delta}(G) = /\overbrace{1,0,0,1}^{m}/$$

$$M'' = /\ 2,1,0,0\ / \ + \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 1 & 0 \\ -1 & 2 & 0 & -1 \\ 0 & -1 & 0 & 1 \\ 2 & -1 & -1 & 0 \end{pmatrix} =$$

$$= /2,1,0,0/ + 1/-1,0,1,0/ + 1/\ 2,-1,-1,0/ = /3,0,0,0/$$

Используя матричные операции в модели на ЭВМ, легко выполнить сети Пэтри, но должна быть задана конфигурация сети.

При исследованиях иногда решают задачу найти последовательность запусков **G** , которая привела из состояния **M'** в состояние **M''** .

Таких последовательностей может быть несколько.

В соответствующую модель должна быть заложена процедура выполнения сетей Пэтри.

СВОЙСТВА СЕТЕЙ ПЭТРИ.

1. Безопасность и ограниченность
2. Абсолютная сохраняемость или сохраняемость по отношению к вектору взвешивания
3. Активность /пассивность/ фрагмента сети или перехода
4. Достижимость или покрываемость той или иной маркировкисети

ОПРЕДЕЛЕНИЕ 1. Позиция сети Пэтри называется безопасной, если элемент маркировки $m_i = M(p_i) \leq 1$ всегда!, т.е. $m_j = \{0,1\}$.

Сеть Пэтри называется безопасной, если все ее позиции безопасны /в соответствующей позиции сети не может появиться более одной фишки/.

ОПРЕДЕЛЕНИЕ 2. Позиция сети Пэтри P_i называется **N**- ограниченной, если $m_i \leq N$ любой маркировки $/N>1/$. Сеть Пэтри называется **N**- ограниченной, если все позиции сети **N**- ограничены.

ОПРЕДЕЛЕНИЕ 3. Сеть Пэтри с начальной маркировкой M^0 называется абсолютно-сохраняемой, если соответствующей **M'** сети справедливо:

$$\sum_{i=1}^{n} M^{0}(p_i) = \sum_{i=1}^{n} M'(p_i).$$

В нашем примере: Сеть Пэтри - триограничена и строгосохраняема.

ОПРЕДЕЛЕНИЕ 4. Сеть Пэтри называется сохраняемой по отношению к некоторому вектору взвешивания $\mathbf{W} = /\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\ /$, заданному на множестве позиций **P** сети, где $\mathbf{W}_i$=1,2,..., если соответственно **M'** сети справедливо:

$$\sum_{i=1}^{n} w_i M^0(p_i) = \sum_{i=1}^{n} w_i M'(p_i),$$

таким образом, $\mathbf{w}_i$ - весовой коэффициент фишки в **i**-й позиции.

В сетях Пэтри с $\mathbf{M^0}$ может оказаться такая ситуация, в которой переход $\mathbf{t}_j$ будет неразрешенным соответственно последовательности запусков.

В этом случае, переход $\mathbf{t}_j$ называется пассивным, а соответствующий фрагмент сети называется пассивным фрагментом.

С понятием пассивности связано понятие тупика сети Пэтри.

ОПРЕДЕЛЕНИЕ. Если в сети Пэтри с $\mathbf{M^0}$ существует некоторая последовательность запуска **G**, приводящая к **M'**, когда все переходы сети оказываются пассивными, то в такой сети Пэтри могут иметь место тупики, а называется тупиковым состоянием.

Для исследования сетей Пэтри необходимо определить, что некоторая **M'** достижима из $\mathbf{M^0}$.

ОПРЕДЕЛЕНИЕ. **M'** называется достижимой в сети Пэтри с начальной $\mathbf{M^0}$, если существует некоторая последовательность запусков **G**, превращающая сеть Пэтри в данную **M'**.

Говорят, что **M'** принадлежит дереву достижимости сети Пэтри с начальной $\mathbf{M^0}$:

$\mathbf{M'} \subset \mathbf{R(СП, M^0)}$.

При исследовании сетей Пэтри важно выделять специальное свойство: наличие конфликта в сети.

Этим свойством могут обладать 2 или более перехода сети.

ОПРЕДЕЛЕНИЕ. Если в некоторой маркировке **M'** сети Пэтри существует 2 или более разрешенных перехода,срабатывание одного из которых приводит к снятию условий разрешения для других переходов, то такие переходы называются конфликтующими, а сеть Пэтри - конфликтной.

Для разрешения конфликтов в сетях используют специальные средства.

Например, для стохастических сетей Пэтри используют ключи распределения $\mathbf{R}_i$.

Рассмотренная ранее сеть Пэтри является конфликтной /переходы $\mathbf{t}_1^m$ и $\mathbf{t}_2^в$ конфликтующие, $\mathbf{t}_3^в$ и $\mathbf{t}_4^m$ - тоже/. В этом примере конфликт разрешался приоритетным путем /мгновенные переходы имеют приоритет выше, чем временные/.

Но признак приоритета может устанавливаться извне, независимо мгновенный это переход или временной.

В случае установления приоритетов для разрешения приоритетов переходов необходимо дополнительное условие разрешения более приоритетных переходов.

Наиболее эффективным является использование ключей, причем $\sum \mathbf{p_i} = \mathbf{1}$. Они рассматриваются как вероятности срабатывания одного из конфликтующих переходов.

В соответствующей модели должны быть выбраны и исследованы механизмы разрешения конфликтов.

ДЕРЕВО ДОСТИЖИМОСТИ СЕТЕЙ ПЕТРИ, АЛГОРИТМ ПОСТРОЕНИЯ ДЕРЕВА

Для анализа свойств сетей Петри наиболее удобно использовать граф представления множества достижимости сетей Петри /ДДСП/. В этом дереве представлены все достижимые состояния сетей Петри.

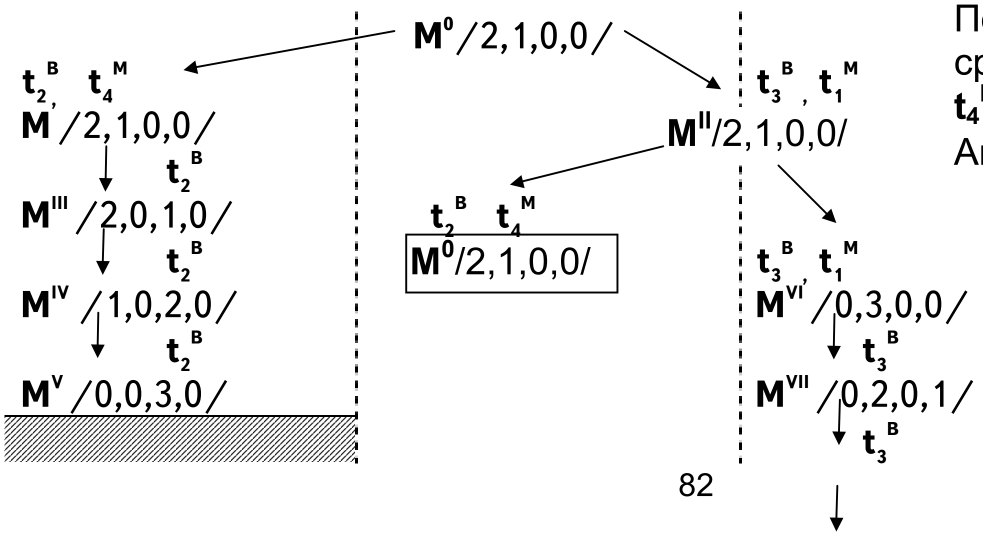
Для стохастических сетей Петри ДДСП несколько сокращается, т.к. в них отражаются только реальные состояния.

Мгновенное состояние сети Петри в таком дереве не представлено.

Общий путь построения ДДСП **R** /СП, $\mathbf{M^0}$/ заключается в определении всех разрешенных переходов в соответствующей маркировке с последующим анализом соответствующего очередного состояния /маркировки/, достигающихся при независимых автоматических последовательностей запусков всех разрешенных переходов предыдущей маркировки.

Для нашего примера ДДСП имеет вид:

Начальная маркировка - корневая точка ДДСП.



После запуска $\mathbf{t_2^В}$ сразу срабатывает $\mathbf{t_4^М}$.
Аналогично $\mathbf{t_3^В \rightarrow t_1^М}$

M^{VIII} /0,1,0,2/

$t_3^{в}$

M^{IX} /0,0,3,0/

дублирующая вершина дерева, после нее ДДСП строить необязательно.

В M^{V} нет разрешения переходов и она называется терминальной или конечной.

Аналогично M^{IX} завершает победу 1-го процесса.

Продолжение построения ДДСП с повторением маркировки M^{0} не имеет смысла, т.к. ДДСП будет иметь бесконечное множество одинаковых фрагментов.

ДДСП всегда представляется конечным графом. Ограничение ДДСП достигается за счет:

1. появления пассивных маркировок /терминальных вершин/ завершающих соответствующие ветви ДДСП;
2. появление маркировок тождественных некоторым ранее полученным маркировкам ДДСП /дублирующие вершины/ и дальше ДДСП не строится, т.к. ветви будут повторятся бесконечное число раз;
3. появление маркировок M^{Z}, которые находятся в отношениях покрываемости с некоторыми уже имеющимися маркировками ДДСП /$M^{Z} > M^{Y}$/. Это значит, что в одной или нескольких розициях P_i имеет место увеличение числа фишек. Такое увеличение можно считать неограниченным. В этом случае для позиции P_i вводится специальный символ ω, отражающий свойство накопления фишек в данной позиции. К ω фишки не прибавляются, следовательно появляются дублирующие вершины.

Постоение ДДСП для простых сетей Петри может быть выполнено вручную, а для более сложных - машинным способом.

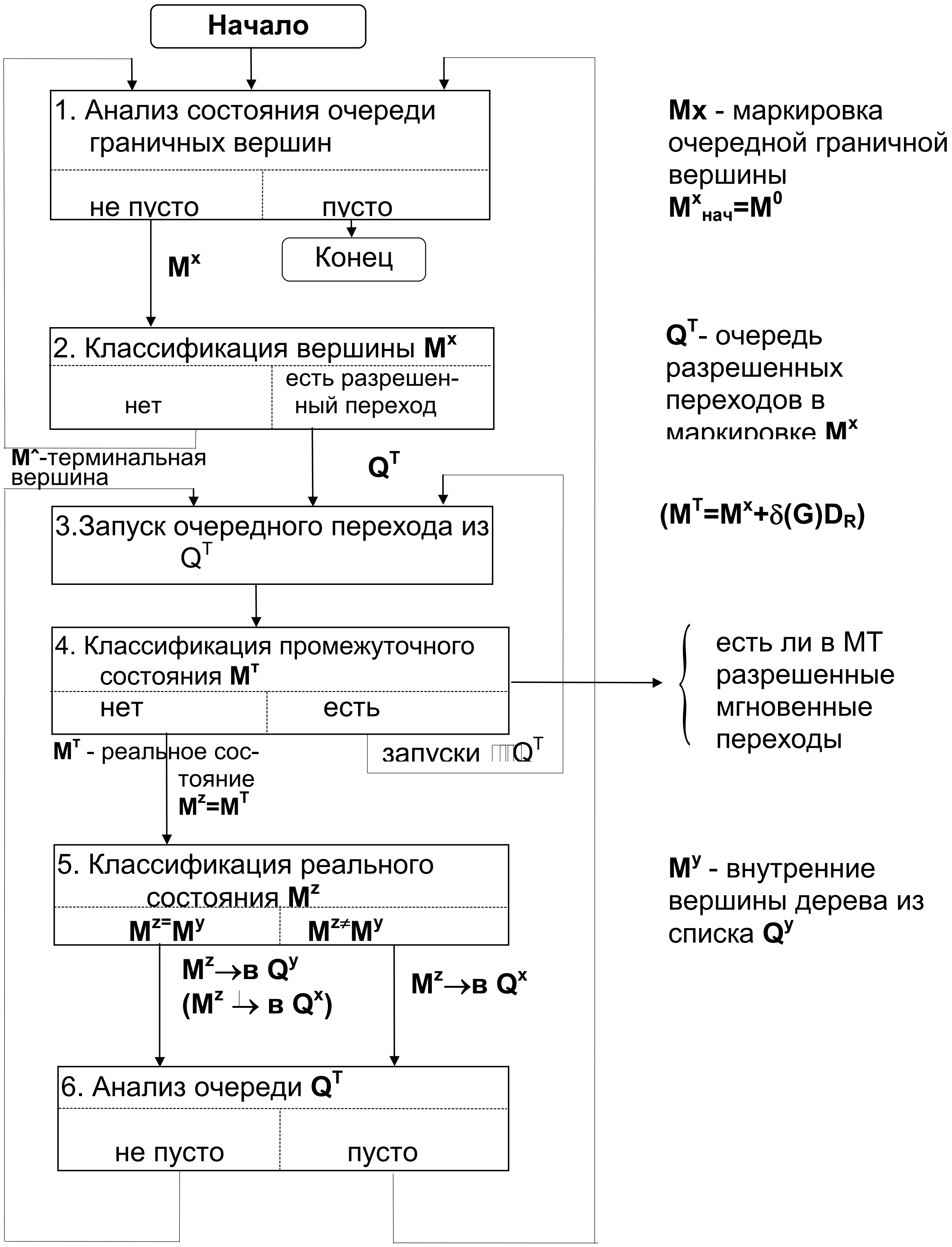
В имитационных моделях выделяют специальный режим для построения ДДСП.

Для работы модели формируется несколько списков /очередей/, например очередь граничных вершин Q^{x}.

Все остальные вершины дерева должны быть преобразованы алгоритмом из граничных вершин в следующие три типа вершин:

1. терминальные;
2. дублирующие;
3. внутренние вершины ДДСП;

Обобщенный алгоритм работы блоков имитационной модели в режиме построения ДДСП имеет вид:



Начало
1. Анализ состояния очереди граничных вершин
не пусто
пусто
Конец
Mx
2. Классификация вершины Mx
нет
есть разрешенный переход
M^-терминальная вершина
QT
3.Запуск очередного перехода из QT
4. Классификация промежуточного состояния MT
нет
есть
MT - реальное состояние
Mz=MT
запуски QT
5. Классификация реального состояния Mz
Mz=My
Mz≠My
Mz→в Qy
(Mz → в Qx)
Mz→в Qx
6. Анализ очереди QT
не пусто
пусто
Mx - маркировка очередной граничной вершины
Mxнач=M0
QT- очередь разрешенных переходов в маркировке Mx
(MT=Mx+δ(G)DR)
есть ли в MT разрешенные мгновенные переходы
My - внутренние вершины дерева из списка Qy

Таким образом в списке $\mathbf{Q}^y$ будут записаны все вершины ДДСП .

Рассмотрим более сложные функции блоков:

2 -проверяет условие разрешения переходов, если разрешеннные переходы есть, то относится $\mathbf{M}^x$ к внутренним вершинам.

5 -проверяется условие покрываемости $/\mathbf{M}^z > \mathbf{M}^y/$ и при этом вводится параметр ω.

Если $/\mathbf{M}^z > \mathbf{M}^y/$, то $\mathbf{M}^z \rightarrow$ в $\mathbf{Q}^x$.