ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
ВЯТСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ФАКУЛЬТЕТ АВТОМАТИКИ И ВЫЧИСЛИТЕЛЬНОЙ ТЕХНИКИ
КАФЕДРА ЭЛЕКТРОННЫХ ВЫЧИСЛИТЕЛЬНЫХ МАШИН

В.С. Ростовцев

ТЕОРИЯ ПРИНЯТИЯ РЕШЕНИЙ

Учебное пособие

Специальность
230101.65
230100.62

Киров 2009

УДК 519.816(07)
Р 785

Ростовцев В.С. Теория принятия решений: Учебное пособие./В.С Ростовцев – Киров: Изд-во ВятГУ, 2009.- 136с.

Оглавление

1 Введение в теорию принятия решений

1.1 Основные понятия в области принятия решений

Всю сознательную жизнь человек принимает решения. Поэтому люди стали задумываться над тем, как они принимают решения и что необходимо сделать, чтобы принимать решения более эффективно.

История сохранила для нас много примеров замечательных решений, принятых известными полководцами, государственными деятелями, менеджерами, бизнесменами. Какие же приёмы и способы они использовали в своей деятельности? Эти умения и навыки в различные времена назывались по-разному, но самое удачное название для них - просвещённый здравый смысл. Он характеризуется тем, что люди старались учесть свой прежний опыт, хорошо уяснить проблему, получить всю необходимую информацию, тщательно продумывать все альтернативы и их последствия, принять во внимание разнообразные факторы, влияющие на результат выбора.

Немаловажную роль в принятии решений играла интуиция. Таким образом, принятие решений во все времена оставалось в большей мере искусством, чем наукой.

В наших решениях удивительным образом сочетаются интуиция и логическое мышление, эмоциональные оценки событий и холодный рациональный расчёт, стремление к безопасности и склонность к риску.

Понятие *«решение»* имеет два смысловых значения. В широком смысле под решением понимается *процесс* выбора одного или нескольких вариантов действий из множества возможных. В этом смысле термин решение означает процесс принятия решений. В узком смысле решение есть результат конкретного выбора варианта действий. В любом случае справедлива следующая формула: решение – это выбор альтернативы.

Понятие ***принятие решений*** затрагивает особый процесс человеческой деятельности, направленный на выбор наилучшего варианта действий [1,2].

Слова «принятие решений» в настоящее время используются очень широко. Говорят, что наилучший вариант решения может быть получен путем математических расчетов, и есть случаи, когда это возможно.

Для подавляющего большинства человеческих решений нельзя точно рассчитать и оценить последствия. Можно лишь предполагать, что определенный вариант решения приведет к наилучшему результату. Однако такое предположение может оказаться ошибочным, потому что никто не может заглянуть в будущее и знать все наверняка. Поэтому человеческие решения являются исключительно важным для практики и интересным для науки объектом исследования.

В процессе принятия решений люди могут играть разные роли. Будем называть человека, фактически осуществляющего выбор наилучшего варианта действий, ***лицом, принимающим решения (ЛПР)***.

Если решения принимаются малой группой, члены которой формально имеют равные права (жюри, комиссия), то человек является членом группы, принимающей решения. Главное в деятельности такой группы - достижение согласия при выработке совместных решений.

В процессе принятия решений человек может выступать в качестве эксперта, т.е. профессионала в той или иной области, к которому обращаются за оценками и рекомендациями все люди, включенные в этот процесс.

Кроме того, в принятии решений неявно участвует окружение ЛПР, сотрудники той организации, от имени которой ЛПР принимает решения. Обычно эта группа людей имеет общие взгляды, общие ценностные установки. Именно этой группе ЛПР в первую очередь объясняет логичность, разумность, обоснованность своего решения. В связи с этим, хотя ЛПР принимает индивидуальные решения, оно учитывает политику и предпочтения данной группы лиц.

Варианты действий принято называть ***альтернативами***. Альтернативы - неотъемлемая часть проблемы принятия решений: если не из чего выбирать, то нет и выбора. Следовательно, для постановки задачи принятия решений необходимо иметь хотя бы две альтернативы.

Альтернативы бывают независимыми и зависимыми. Независимыми являются те альтернативы, любые действия с которыми (удаление из рассмотрения, выделение в качестве единственно лучшей) не влияют на качество других альтернатив. При зависимых альтернативах оценки одних из них оказывают влияние на качество других. Имеются различные типы зависимости альтернатив. Наиболее простым и очевидным является непосредственная групповая зависимость: если решено рассматривать хотя бы одну альтернативу из группы, то надо рассматривать и всю группу. Так, при планировании развития города решение о сохранении исторического центра влечет за собой рассмотрение всех вариантов его реализации.

В современной науке о принятии решений считается, что варианты решений характеризуются различными показателями их привлекательности для ЛПР. Эти показатели называют признаками, факторами, атрибутами или ***критериями*** [2].

Критериями оценки альтернатив являются показатели их привлекательности (или непривлекательности) для участников процесса выбора.

Использование критериев для оценки альтернатив требует определения градаций качества: лучших, худших и промежуточных оценок. Иначе говоря, существуют шкалы оценок по критериям.

1. 2 История возникновения теории принятия решений

Научные основы теории принятия решения были заложены в период второй мировой войны. Его родоначальниками считаются Дж. фон

Нейман и О.Моргенштерн, которые в 1944 г. опубликовали книгу по теории игр. Позднее зарубежные специалисты: Р.Акофф, Ф.Эмери, Ст.Оптнер, Р.Льюс, Х.Райфа, Ст.Бир, Дж.Форрестер, Дж.Диксон, Л.Ластед, Р.Мюллер и др., а также отечественные специалисты: П.В.Авдулов, Ф.Ф. Аунапу, .Г.Афанасьев, А.Г.Венделин, Д.М.Гвишиани, В.М.Глушков, Л.В.Кантарович, О.И.Ларичев, И.М.Сыроежин и др. внесли существенный вклад в развитие и обогащение этой теории.

Под решением понимается набор воздействий (действий со стороны лица, принимающего решения (ЛПР)) на объект (систему, комплекс и т.д.) управления, позволяющий привести данный объект в желаемое состояние или достичь поставленной перед ним цели.

В теории различают бинарное решение, стандартное решение, многоальтернативное решение, непрерывное решение, инновационное решение и т.д. все эти виды решений дифференцируются в зависимости от количества альтернатив.

Принятие решений возможно на основании знаний об объекте управления, о процессах объективное в нем протекающих и могущих произойти с течением времени (иначе говоря, требуется наличие адекватной модели объекта), и при наличии множества показателей (критериев), характеризующих эффективность (качество, оптимальность и т.д.) принятого решения (иначе говоря требуется также наличие модели принятия и оценки принятого решения).

Теория принятия решения (ПР) является сложной междисциплинарной наукой в ее развитие внесли большой вклад экономисты, математики, психологи, социологи и т.д. В результате всех этих исследований можно выделить два основных направления развития ПР:

- теория принятия рациональных решений;

- психологическая теория принятия решений [26].

Первое направление теории ПР отвечает на вопросы: как принимать решения рационально, какие альтернативы оптимальны. Это направление, в свою очередь, развивается несколькими путями.

Во-первых, широкое использование математических методов и моделей. К основным, наиболее часто применимым можно отнести: линейные модели, транспортная задача, линейное программирование, динамическое программирование, теория игр, теория массового обслуживания, оптимальное программирование и т.д. Выбор метода принятия решения во многом, определяется характером и спецификой самого решения. В связи с этим чрезвычайно важным является классификация решений. Под моделью принятия решений понимается формальное представление процесса принятия решений.

Анализ литературы [1,2,3] позволяет использовать следующую классификацию решений (таблица 1.1).

Таблица 1.1- Классификация видов решений

Классификационный признак	Вид решения		
1.Степень структуризации исследуемой проблемы	Хорошо структури-рованное	Плохо структурированное	Не структурированное
2.По количеству этапов реализации решения	Статические (с одним этапом)	Динамические (много этапов)	
3. По уровню информированности о состоянии	В условиях определенности	В условиях риска	В условиях неопределенности

проблемы			
4. По количеству лиц, участвующих в процессе принятия решений	Один участник	Много участников	
5. По содержанию	Стратегические	Тактические	

На современном этапе развития важнейшей задачей является совершенствование методики и методологии решения в условиях определенности, риска и неопределенности. Условия принятия решений:

1. Выбор решения в условиях определенности предполагает, что результат каждого действия известен.

2. Выбор решения в условиях риска означает, что каждое действие приводит к одному из множества возможных частных исходов. При этом каждый исход имеет известную вероятность появления. Считается, что лицу принимающему решение (ЛПР) эти вероятности известны.

3. Выбор решения в условиях неопределенности происходит в том случае, когда то или иное действие имеют своим следствием множество возможных частных исходов, но вероятности этих исходов неизвестны.

Для решения задач в условиях определенности наиболее успешно могут применяться математические модели. Для решения задач в условиях риска лучше использовать методы теории вероятностей и математической статистики. В условиях неопределенности использовать методы математического моделирования крайне затруднительно. Более целесообразным является применение теории игр и "байесовский подход".

Следующим направлением теории ПР является кибернетический подход. Широко применяется логико-математическая формализация и моделирование. Ст. Бир придает большое значение исследованию операций, Д. Форрестер считает, что математическое моделирование полезно, но оно должно быть дополнено суждениями, основанными на интуиции.

Под моделью будем понимать некоторое упрощение реальности. Моделирование помогает лучше понять и описать реальность. В соответствии с целью применения модели можно разделить на следующие виды:

- описательная модель. Она имеет своей целью отождествить структуру системы. Такие модели могут быть построены на трех уровнях детализации: макромодель, микроаналитическая модель, модель микроповедения. К моделям подобным относятся модель процесса Маркова, теория очередей;

- модель-прогноз, позволяет предвидеть развитие действующей системы, в зависимости от различных гипотез относительно переменных величин, включенных в эту модель;

- нормативная или предписательная модель, содержит рекомендации к действию.

Модель — это представление, как правило, в математических терминах наиболее характерных черт изучаемого объекта или системы. Одним из самых распространенных инструментов для математического моделирования и исследования информационных процессов и систем являются сети Петри.

Цель представления системы в виде сети Петри и последующего анализа этой сети состоит в получении важной информации о структуре и динамическом поведении моделируемой системы. Эта информация может использоваться для оценки моделируемой системы и выработки предложений по её усовершенствованию. Впервые сети Петри предложил немецкий математик Карл Адам Петри.

1.3 Шкалы оценок

В принятии решений принято различать шкалы непрерывных и дискретных оценок, шкалы количественных и качественных оценок. Так, для критерия «стоимость» может быть использована непрерывная количественная шкала оценок (в денежных единицах). Для критерия «наличие дачи» может быть качественная двоичная шкала: есть либо нет. Кроме категорий «качественные - количественные», «непрерывные — дискретные» в принятии решений различают следующие типы шкал.

1. ***Шкала порядка*** — оценки упорядочены по возрастанию или убыванию качества. Примером может служить шкала экологической чистоты района около места жительства:
   - очень чистый район;
   - вполне удовлетворительный по чистоте;
   - экологическое загрязнение велико.
2. ***Шкала равных интервалов*** — интервальная шкала. Для этой шкалы имеются равные расстояния по изменению качества между оценками. Например, шкала дополнительной прибыли для предпринимателя может быть следующей: 1 млн., 2 млн., 3 млн. и т.д. Для интервальной шкалы

характерно, что начало отсчёта выбирается произвольно, так же, как и шаг (расстояние между оценками) шкалы.

3. ***Шкала пропорциональных оценок*** - идеальная шкала. Примером является шкала оценок по критерию стоимости, отсчет в которой начинается с установленного значения (например, с нулевой стоимости).

В принятии решений чаще всего используются порядковые шкалы и шкалы пропорциональных оценок.

Процесс принятия решений. Не следует думать, что принятие решений есть одномоментный акт. Очень часто это достаточно длинный и мучительный процесс. Г. Саймон [1] выделяет в нем три этапа: поиск информации, поиск и нахождение альтернатив и выбор лучшей альтернативы.

На первом этапе собирается вся доступная на момент принятия решения информация: фактические данные, мнение экспертов. Там, где это возможно, строятся математические модели; проводятся социологические опросы; определяются взгляды на проблему со стороны активных групп, влияющих на ее решение.

Второй этап связан с определением того, что можно, а что нельзя делать в имеющейся ситуации, т. е. с определением вариантов решений (альтернатив). И уже третий этап включает в себя сравнение альтернатив и выбор наилучшего варианта (или вариантов) решения.

1.4 Популярные методики принятия решений

Высокий интерес к процессам принятия решений в личной и деловой жизни людей побудил многих исследователей к разработке собственных рекомендаций о том, как следует принимать решения и как лучше организовать этот процесс.

Известный американский социолог М. Рубинштейн предложил десять правил принятия управленческих решений [3]:

1. Прежде чем вникать в детали, постарайтесь получить представление о проблеме в целом.

2. Не принимайте решения, пока не рассмотрите все возможные варианты.

3. Сомневайтесь - даже самые общепринятые истины должны вызывать недоверие, и не нужно бояться отводить их.

4. Старайтесь взглянуть на стоящую перед вами проблему с самых разных точек зрения, даже если шансы на успех кажутся минимальными.

5. Ищите модель или аналогию, которая поможет вам лучше понять сущность решаемой проблемы. Так, диаграммы и схемы позволят охватить сложную и обширную проблему буквально одним взглядом.

6. Задавайте как можно больше вопросов. Правильно заданный вопрос порой может радикально изменить содержание ответа.

7. Не удовлетворяйтесь первым решением, которое вам придёт в голову. Постарайтесь найти его слабые места. Попробуйте найти другие решения данной проблемы и сравните их с первым.

8. Перед принятием окончательного решения поговорите с кем-нибудь о своих проблемах.

9. Не пренебрегайте своими чувствами. Ведущая роль логического мышления анализе проблем не подлежит сомнению, но нельзя преуменьшать значение чувств и интуиции.

10. Помните, что каждый человек смотрит на жизнь и возникающие повседневно проблемы со своей, особенной точки зрения.

Американский психолог Спенсер Джонсон предложил оригинальную методику принятия личных и деловых решений, названную им система принятия верных решений «да» или «нет».

С. Джонсон указывает, что большинство ошибок в процессе принятия решений возникает из-за того, что люди используют неправильную систему мышления. Он предлагает прежде, чем принять решение, ответить на шесть вопросов – три вопроса практические и три вопроса личностные. Отвечая на каждый вопрос «да» или «нет», мы сможем лучше разобраться в ситуации и в самих себе, изменить ранее принятое решение и повысить его качество.

Практические вопросы:

1. *Удовлетворяю ли я подлинную потребность.* Любое решение направлено на достижение определённой цели. Поэтому когда выбор цели ограничивается только тем, что нам нужно, то принятие решения сильно облегчается.

2. *Узнал ли я варианты выбора.* Чтобы принять оптимальное решение, необходимо располагать полной информацией об альтернативах.

3. *Всё ли я продумал до конца.* Требуется тщательно продумать последствия каждой альтернативы.

Личностные вопросы:

1. *Честен ли я сам с собой.* Перед принятием решения необходимо сделать реальное описание текущей или будущей ситуации, не закрывая глаза на неприятные и неожиданные последствия и факты, противоречащие нашим интересам, взглядам и опыту.

2. *Доверяю ли я своей интуиции.* Принимая решение необходимо не только прислушиваться, но и доверять своей интуиции. Интуиция – это подсознательное знаний, основанное на личном опыте.

3. *Достоин ли я лучшего.* Рассматривая возможные вариантывыбора, мы часто ограничиваем круг своих возможностей только лишь потому, что не верим, что достойны лучшего. Чтобы добиться успеха, необходимо освободиться от этого опасного заблуждения, которое ставит искусственные преграды на пути оптимальному решению.

Отвечая последовательно на все перечисленные вопросы, ЛПР получает возможность улучшить предварительно принятое решение. Если на все вопросы вы ответили «да», то велика вероятность того, что ваше решение – это удачный выбор.

Если хотя бы на один вопрос вы ответили «нет», то, скорее всего, ваше решение окажется неудачным и его следует пересмотреть.

Литература

1. Ларичев О.И. Теория и методы принятия решений, а также хроника событий в волшебных странах: Учеб./О.И. Ларичев- М.: Логос, 2000.-296 с.- (Учеб. XXI века).
2. Грешилов А.А. Как принять наилучшее решение в реальных условиях/ А.А. Грешилов. - М.: Радио и связь, 1991. – 320 с.
3. Кулагин О.А. Принятие решений в организациях. 2006

2 Применение множества Эджворта – Парето

Варианты выбора альтернатив представлены в таблице 1.1. При оценке альтернатив использованы два критерия: стоимость и привлекательность.

Введем следующее определение. Назовем альтернативу **А** доминирующей по отношению к альтернативе **В**, если по всем критериям оценки альтернативы **А** не хуже, чем альтернативы **В**, а хотя бы по одному критерию оценка **А** лучше. При этом альтернатива **В** называется доминируемой.

Таблица 2.1 - Оценки альтернативных вариантов туров

Альтернатива	Критерий	
	Стоимость	Привлекательность, новые впечатления
1.Океанские острова	Небольшая	Малая
2.Скандинавия	Высокая	Большая
3.Египет	Небольшая	Большая

Предположим, что по какой-то причине поездка в Египет оказалась невозможной (например, из-за участившихся нападений на туристов). Туры в Скандинавию и на Океанские острова не находятся в отношении доминирования. По одному из критериев лучше альтернатива 2, по другому — альтернатива 1.

Введем следующее определение: альтернативы относятся к множеству Эджворта-Парето (Э-П), если каждая из них превосходит любую другую по какому-то из критериев [1].

Множество Эджворта-Парето названо так по именам ученых, впервые обративших внимание на альтернативы, не уступающие друг другу по критериальным оценкам, т. е. на альтернативы, не находящиеся в отношении доминирования. Альтернативы, принадлежащие множеству Э-П, принято называть несравнимыми. Их действительно невозможно сравнить непосредственно на основе критериальных оценок. Но если решение должно быть принято (например, супруги должны из многих туров выбрать один), то сравнение альтернатив, принадлежащих множеству Э-П, возможно на основе дополнительной информации.

Нетрудно убедиться, что множество Э—П включает в себя наиболее «контрастные» альтернативы, сложные для сравнения. Если стоит задача выбора одной лучшей альтернативы, то она обязательно принадлежит множеству Э-П. Поэтому во многих методах принятия решений очень важен этап выделения множества Э-П из всего множества заданных альтернатив.

Один из возможных способов решения этой задачи состоит в попарном сравнении альтернатив и исключении доминируемых. Задача выделения множества Э—П обычно рассматривается как предварительная. За ней следует наиболее существенный этап принятия решений.

Рассмотрим второй пример. Имеются три альтернативных варианта создания информационных систем (ИС) с различными характеристиками стоимости и надежности.

Альтернатива	Стоимость	Надежность
ИС1	В	С
ИС2	В	В
ИС3	С	С

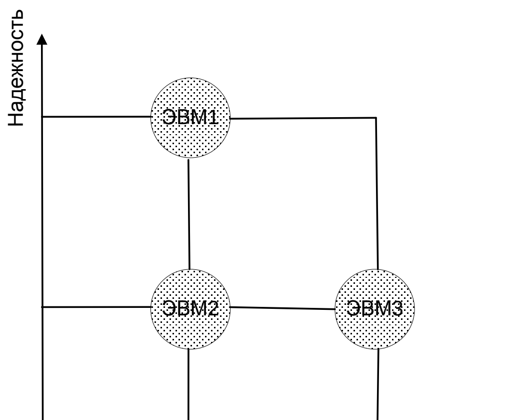
К множеству Эджворта-Парето относятся те альтернативы, которые не доминируют по какому-либо критерию, т.е. они являются несравнимыми.

ИС3 доминирует над ИС1 по критерию стоимости. ИС2 и ИС3 не доминируют друг над другом и относятся к множеству Эджворта – Парето. Поиск оптимального варианта из множества Эджворта – Парето невозможен без дополнительной информации. Например, при ограничении бюджета на стоимость разработки ИС можно выбрать один вариант.

Упражнения

1. Какие альтернативы (ЭВМ), из приведенных ниже, относятся к множеству Эджворта-Парето и почему?

Альтернатива	Стоимость	Надежность
ЭВМ1	Средняя	Высокая
ЭВМ2	Средняя	Средняя
ЭВМ3	Низкая	Средняя



Ответ: если требуется выбрать наилучшую альтернативу, то необходимо искать доминирующие альтернативы. ЭВМ3 доминирует над ЭВМ2 и ЭВМ1 доминирует над ЭВМ2. К множестве Э-П относятся ЭВМ1 и ЭВМ3. Это несравнимые альтернативы. Для окончательного выбора необходима дополнительная информация.

2. Какие альтернативы (ЭВМ), из приведенных ниже, относятся к множеству Эджворта-Парето и почему?

Альтернатива	Стоимость	Надежность
ЭВМ1	Средняя	Высокая
ЭВМ2	Низкая	Высокая
ЭВМ3	Низкая	Средняя

3. Какие альтернативы (ЭВМ), из приведенных ниже, относятся к множеству Эджворта-Парето и почему?

Альтернатива	Стоимость	Надежность
ЭВМ1	Средняя	Высокая
ЭВМ2	Низкая	Высокая
ЭВМ3	Низкая	Средняя

Литература

4. Ларичев О.И. Теория и методы принятия решений, а также хроника событий в волшебных странах: Учеб./О.И. Ларичев- М.: Логос, 2000.- 296 с.- (Учеб. XXI века).

3 Теория полезности в выборе альтернатив

3.1 Аксиомы рационального поведения

В [1,2] вводится пять аксиом и доказывается существование функции полезности. Дадим содержательное представление этих аксиом. Обозначим через **x, y, z** различные исходы (результаты) процесса выбора, а через **p**, **q** вероятности тех или иных исходов. Введем определение лотереи. **Лотереей** называется игра с двумя исходами: исходом **x**, получаемым с вероятностью **p**, и исходом **y**, получаемым с вероятностью **1-p** (рисунок 3.1).

Примером лотереи является подбрасывание монеты. При этом, как известно, с вероятностью **p**=0,5 выпадает орел или решка. Пусть **x**=10 у.е. и **y**=10 у.е. (т. е. мы получаем 10 у.е. при выпадении орла и платим столько же при выпадении решки).

Ожидаемая (или средняя) цена лотереи определяется по формуле

$$p \times x+(1-p) \times y.$$

Приведем аксиомы рационального выбора.

Аксиома 1. Исходы **x, y, z** принадлежат множеству **A** исходов.

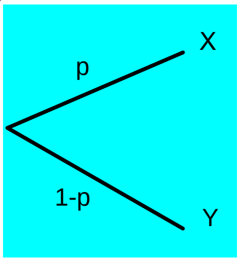


Рисунок 3.1

Аксиома 2. Пусть **P** означает строгое предпочтение (похожее на отношение > в математике); **R** – нестрогое предпочтение (похожее на отношение ≥); **I** - безразличие (похожее на отношение =). Ясно, что **R** включает **P** и **I**. Аксиома 2 требует выполнения двух условий.

Представление лотереи условий:

- связанности: либо **xRy**, либо **yRx**, либо то и другое вместе;
- транзитивности: из **xRy** и **yRz** следует **xRz**.

Аксиома 3. Две представленные на рисунке 3.2 лотереи находятся в отношении безразличия. Справедливость этой аксиомы очевидна. Она записывается в стандартном виде как ((**x, p, y**)**q, y**) **I** (**x, pq, y**).

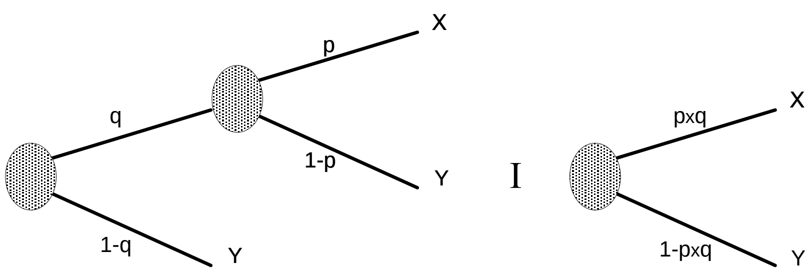


Рисунок 3.2

Аксиома 4. Если **xIy**, то (**x, p, z**) **I** (**y, p, z**).

Аксиома 5. Если **xPy**, то **xP(x, p, y)Py**.

Аксиома 6. Если **xPyPz**, то существует вероятность **p**, такая что **y**(**x, p, z**).

Все приведенные выше аксиомы достаточно просты для понимания и кажутся очевидными.

В предположении, что аксиомы выполняются, была доказана теорема [1]. Если аксиомы 1 - 6 удовлетворяются, то существует численная функция **U**, определенная на **A** (множество исходов) и такая, что:

- **xRy** тогда и только тогда, когда **U**(**x**) > **U**(**y**);
- **U**(**x, p, y**) = **p**×**U**(**x**)+(l-**p**) ×**U**(**y**).

Функция **U**(**x**) измеряется на шкале интервалов. Функция **U**(**x**) - единственная с точностью до линейного преобразования (например, если **U**(**x**)> **U**(**y**), то и **a**×**U**(**x**) > **a**×**U**(**y**), где **a** — целое положительное число).

3.2 Задачи с вазами

Теория полезности экспериментально исследовалась в так называемых задачах с вазами (или урнами). Ваза — это непрозрачный сосуд, в котором находится определенное (известное лишь организатору эксперимента) количество шаров различного цвета. Задачи с вазами типичны для группы наиболее простых задач принятия решений — задач статистического типа. Для решения этих задач надо знать элементарные начала теории вероятностей [1]. Человек делает выбор в этих задачах, основываясь на расчетах. Варианты действий выражены в наиболее простом виде.

Типовая задача для испытуемого может быть представлена следующим образом. Перед испытуемым ставится ваза, которая может быть вазой 1-го или 2-го типа. Дается следующая информация: сколько имеется у экспериментатора ваз 1-го и 2-го типов; сколько черных и красных шаров в вазах 1-го и 2-го типов, какие выигрыши ожидают испытуемого, если он угадает, какого типа ваза; какие проигрыши ожидают его, если он ошибется. После получения такой информации испытуемый должен сделать выбор: назвать, к какому типу принадлежит поставленная перед ним ваза.

Пусть, например, экспериментатор случайно выбирает вазу для испытуемого из множества, содержащего 700 ваз 1-го типа и 300 ваз 2-го типа. Пусть в вазе 1-го типа содержится 6 красных шаров и 4 черных. В вазе 2-го типа содержится 3 красных и 7 черных шаров. Если перед испытуемым находится ваза 1-го типа, и он угадает это, то получит выигрыш 350 условных единиц (у. е.), если не угадает, его проигрыш составит 50 у. е. Если перед ним ваза 2-го типа и он это угадает, то получит выигрыш 500 у. е., если не угадает, его проигрыш составит 100 у. е. Испытуемый может предпринять одно из следующих действий:

d1 — сказать, что ваза 1-го типа;

d2 — сказать, что ваза 2-го типа.

Условия задачи можно представить в таблице 3.1.

Таблица 3.1 -Представление задачи с вазами

Тип вазы	Вероятность выбора вазы данного типа	Выигрыш при действии	
		d1	d2
1	0,7	350	-100
2	0,3	-50	500

Что же делать человеку? Теория полезности отвечает: оценить среднюю (ожидаемую) полезность каждого из действий и выбрать действие с максимальной ожидаемой полезностью. В соответствии с этой рекомендацией можно определить среднее значение выигрыша для каждого из действий:

U(**d1**)= 0,7 * 350 - 0,3 * 50=230 у.е.;

U(**d2**)= 0,3 * 500 - 0,7 * 100=80 у.е.

Следовательно, разумный человек выберет действие **d1**, а не действие **d2**.

Из этого примера следует общий рецепт действий для рационального человека: определить исходы, помножить их на соответствующие вероятности, получить ожидаемую полезность и выбрать действие с наибольшей полезностью.

Задачи с вазами помогут нам познакомиться с построением деревьев решений и принятием решений с их помощью.

3.3 Деревья решений

Приведенная таблица 3.1 может быть представлена в виде дерева решений (рисунке 3.3). На этом дереве квадратик означает место, где решение принимает человек, а светлый кружок — место, где все решает случай. На ветвях дерева написаны уже знакомые нам значения вероятностей, а справа у конечных ветвей — значения исходов (результаты).

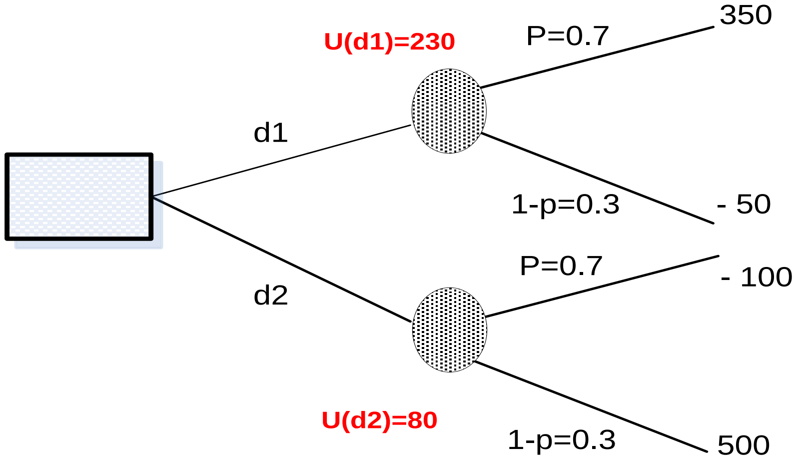


Рисунок 3.3

Для чего нужно дерево решений? Мы можем использовать его для представления своих возможных действий и для нахождения последовательности правильных решений, ведущих к максимальной ожидаемой полезности. Чтобы показать это, усложним задачу. Предоставим человеку, выбирающему между действиями **d1** и **d2**, дополнительные возможности. Пусть он может до своего ответа вытащить за определенную плату один шар из вазы, причем после вытаскивания шар кладется обратно в вазу. Плата за вытаскивание одного шара 60 у.е.

Дерево решений с двумя его основными ветвями представлено на рисунке 3.4. Вот теперь вопрос о том, какое решение следует принимать, стал сложнее: необходимо решить, стоит ли вынимать шар и какой ответ дать после вытаскивания красного или черного шара. При принятии этих решений нам окажет существенную помощь известный в теории вероятностей (и в теории статистических решений) способ подсчета изменения вероятностей событий после получения дополнительной информации.

Вернемся к описанию задачи. Вероятность вытащить красный шар из вазы 1-го типа $\mathbf{p_К(B1)}$= 0,6, а из вазы 2-го типа $\mathbf{p_К}$ (**B2**)=0,3. Зная все условные вероятности (зависящие от условия), а также вероятности **P1** и **P2** выбора ваз 1-го и 2-го типа (таблица 3.1), можно поставить следующие вопросы.

Первый вопрос: каковы вероятности вытащить красный и черный шары? Для ответа на этот вопрос произведем простые вычисления. Вероятность вытащить красный шар $\mathbf{p_К(B1)}$= 0,7 × 0,6 = 0,42, если ваза окажется 1-го типа, $\mathbf{p_К(B2)}$=0,3 ×0,3 = 0,09, если ваза окажется 2-го типа. Следовательно, вероятность вытащить красный шар в общем случае $\mathbf{P_К}$= 0,51. Аналогичным образом можно посчитать, что вероятность вытащить черный шар **Рч** =0,49.

Второй вопрос более сложный. Пусть вытащенный шар оказался красным (черным). Какое действие следует выбрать: **d1** или **d2**? Для ответа на этот вопрос нужно знать вероятности принадлежности ваз к 1-му и 2-му типам после получения дополнительной информации. Эти вероятности позволяет определить знаменитая формула Байеса.

Например, мы вытащили красный шар. Какова после этого вероятность того, что перед нами стоит ваза 1-го типа?

Приведем все обозначения вероятностей:

$\mathbf{Рк(B_1)}$ - вероятность вытащить красный шар из вазы 1-го типа;

$\mathbf{Рч(B_1)}$- вероятность вытащить черный шар из вазы 1-го типа;

$\mathbf{Рк(B_2)}$ - вероятность вытащить красный шар из вазы 2-го типа;

$\mathbf{Рч(B_2)}$ - вероятность вытащить черный шар из вазы 2-го типа;

$\mathbf{P(B_1)}$ - вероятность того, что ваза 1-го типа;

$\mathbf{p(B_2)}$ - вероятность того, что ваза 2-го типа;

P(B1/к) - вероятность того, что ваза окажется 1-го типа после вытаскивания красного шара;

P(B1/ч) - вероятность того, что ваза окажется 1-го типа после вытаскивания черного шара;

P(B2/к) - вероятность того, что ваза окажется 2-го типа после вытаскивания красного шара;

P(B2/ч) - вероятность того, что ваза окажется 2-го типа после вытаскивания черного шара.

Формула Байеса позволяет оценить **p(Bi/к)** и **p(Bi/ч)**, где **i**=1,2, используя все прочие вероятности. Например:

P(B1/к)= [P(B1/к) × P(B1)] / [P(B1/к) × P(B_1)+ Pк(B_2) × P(B2)]

Для нашей задачи: P(B1/к)=0,82; P(B1/ч) =0,57; P(B2/к)=0,18; P(B2/ч)= 0,43.

Пример решения задачи. Провести «сворачивание дерева решений», приведенного на рисунке 3.4. ЛПР должен выбрать одно из решений: **d1** – выбрать вазу 1-го типа; **d2** - выбрать вазу 2-го типа. Вытаскивать шар или нет?

Ваз 1-го типа - 60, а ваз 2-го типа - 40. В вазе 1-го типа 2 красных и 3 черных шара. В вазе 2-го типа 3 красных и 2 черных шара. Чтобы определить, какая ваза перед ЛПР, ему предлагается вытащить из вазы шар и снова положить его обратно за плату 100 у.е.

Решение

Вероятность вытащить красный шар из вазы 1-го типа **Рк** (**В1**)=0,4.

Вероятность вытащить красный шар из вазы 2-го типа **Рк** (**В2**)=0,6.

Вероятность вытащить красный шар из вазы 1-го типа с учетом количества ваз **Рк** (**В1**)=2/5 ×0,6=0,24.

Вероятность вытащить черный шар из вазы 1-го типа с учетом количества ваз **Рч** (**В1**)=3/5 ×0,6=0,36.

Вероятность вытащить красный шар из вазы 2-го типа **Рк** (**В2**)=3/5 × 0,4=0,24.

Вероятность вытащить черный шар из вазы 2-го типа **Рч** (**В2**)=2/5 × 0,4=0,16.

Вероятность вытащить красный шар из ваз **Рк** =0,24+0,24=0,48.

Вероятность вытащить черный шар из ваз **Рч** =0,52.

Вероятность того, что это ваза 1-го типа **Р(В1)**=60/100=0,6.

Вероятность того, что это ваза 2-го типа **Р(В2)**=40/100=0,4.

Вероятность того, что это ваза 1-го типа после вытаскивания красного шара **Р(В1**/к).

Вероятность того, что это ваза 1-го типа после вытаскивания черного шара **Р(В1**/ч).

Вероятность того, что это ваза 2-го типа после вытаскивания красного шара **Р(В2**/к).

Вероятность того, что это ваза 2-го типа после вытаскивания черного шара **Р(В2**/ч).

По формуле Байеса находятся вероятности **P(B1/к)**, **P(B1/ч)**, **(B2/к)**, **P(B2/к)**.

$P(B1/к)= [Pк (B1) \times P(B1)] / [Pк (B1) \times P(B1)+ Pк (B2) \times P(B2)]$

$P(B1/к)=0,4\times0,6 / [0,4\times0,6 +0,6 \times0,4]=0,24/(0,24+0,24)=0,5$ $P(B1/ч)=0,5$

$P(B2/к)= [Pк (B2) \times P(B2)] / [Pк (B2) \times P(B2)+ Pк (B1) \times P(B1)]$

$P(B2/к)=[0,6\times0,4] / [0,6\times0,4 +0,4 \times0,6]=0,24/(0,24+0,24)=0,5$ $P(B2/ч)=0,5$

$P(B1/ч)= Pч (B1) \times P(B1)/[Pч (B1) \times P(B1)+ Pч (B2) \times P(B2)]$

$P(B1/ч)=0,6\times0,6/[0,6\times0,6 +0,4 \times0,4]=0,36/(0,36+0,16)=0,69$

$P(B1/ч)=0,69$ $P(B2/ч)=0,31$

По дереву решений идти от конечных веток к его корню. Там, где есть случайность (кружок), найти среднее значение функции полезности.

Там, где квадратик (этап принятия решения), выбрать ветку с наибольшей ожидаемой полезностью, а другую отсечь двумя черточками.

Так как полезность верхней ветки выше (100 больше 34,6), то шар вытаскивать нецелесообразно и ЛПР должен выбрать вазу 2-го типа.

Теперь мы имеем всю информацию, необходимую для принятия решений. На рисунке 3.4 показаны две основные ветви дерева решений, причем верхняя ветка повторяет дерево решений на рисунке 3.3. Квадратик 1 слева соответствует первому решению — вытаскивать шар или нет. В случае отказа от вытаскивания шара используется верхняя основная ветвь. Решению вытаскивать шар соответствует нижняя ветвь, начинающаяся со случайного события (кружок). В квадратиках 2, 3, 4 принимаются решения о выборе одной из двух стратегий: **d1** или **d2**. Далее все решает элемент случайности (кружки).

Над кружками и прямоугольниками рассчитываются значения полезности.

96=0.6*200-0.4*60

100=0.4*400-0.6*100

70=0.5*200-0.5*60

150=0.5*400-0.5*100

55=0.31*400-0.69*100

119.4=0.69*200-0.31*60

134.6=0.48*150+0.52*119.4

34.6=134.6-100

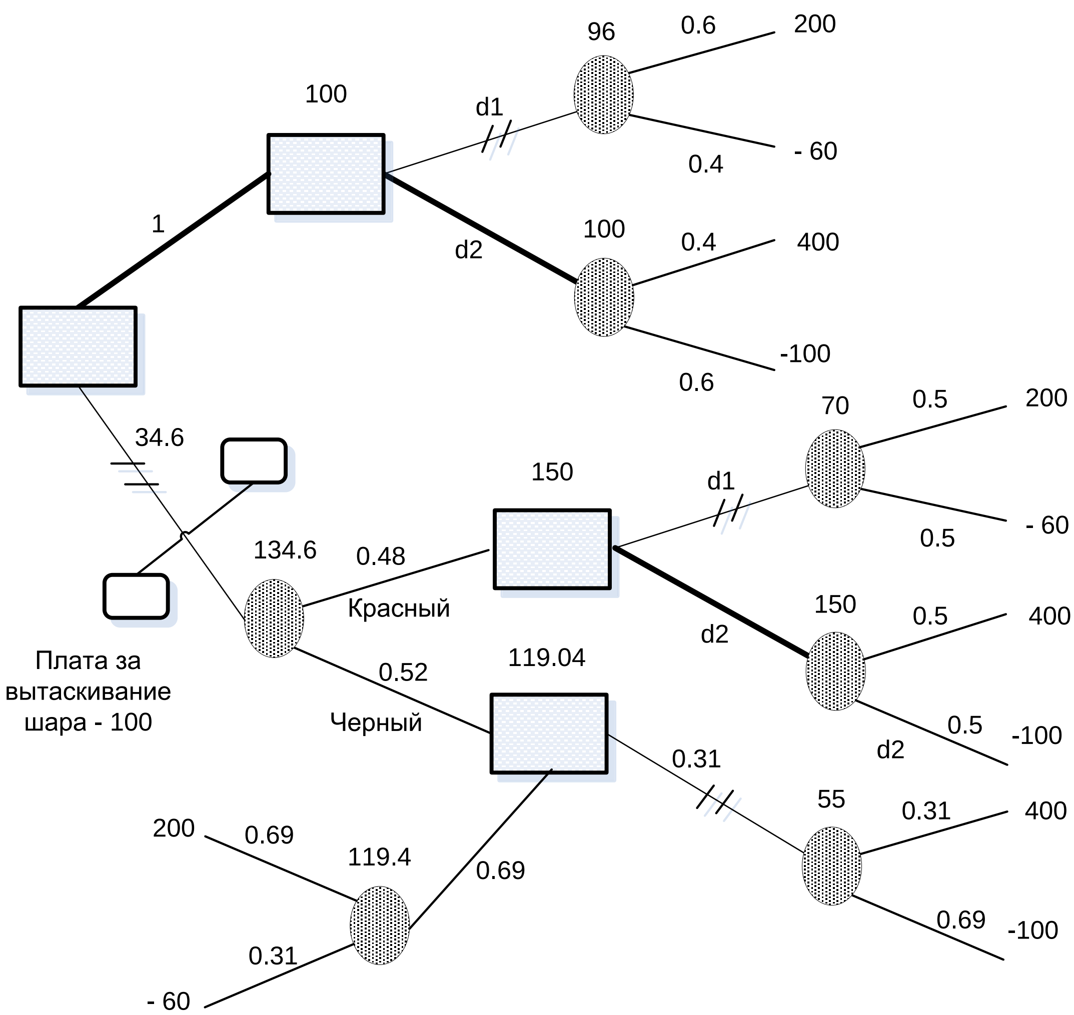


Рисунок 3.4

Есть три простых правила выбора оптимальной (по критерию максимума ожидаемой полезности) последовательности решений на основе дерева решений:

- идти от конечных ветвей дерева к его корню;
- там, где есть случайность (кружок), находится среднее значение;
- там, где есть этап принятия решений (квадратик), выбирается ветвь с наибольшей ожидаемой полезностью, а другая отсекается двумя черточками.

На этом рисунке над кружками указаны средние значения полезности, двумя черточками отсечены ветви с меньшим значением ожидаемой полезности. Наилучший вариант действий: шар не вытаскивать и выбирать действие **d1**. Этот вариант соответствует самому верхнему пути дерева решений на рисунке 3.4. Такая процедура нахождения оптимального пути на деревьях решений получила название «сворачивания» дерева решений.

Деревья решений при заданных числовых значениях вероятностей и исходов позволяют осуществить выбор той стратегии (последовательности действий), при которой достигается наибольший выигрыш, т.е. достигается максимум функции полезности **ЛПР**.

3.4 Метод построения дерева решений

Задача. Перед ЛПР находится 700 ваз 1-го и 300 ваз 2-го типа. Затем ему предъявляется одна из ваз и он должен угадать, к какому типу относится ваза: **d1** – ваза 1-го типа; **d2** – ваза 2-го типа. Если он угадает вазу 1-го типа, получает 350 у.е., а если не угадает – штраф 50 у.е. Аналогично для второй вазы (таблица 3.2).

Таблица 3.2 -Представление задачи с вазами

Тип вазы	Количество	Вероятность выбора	Выигрыш, у.е.	
			d1	d2
1	700	0,7	350	-100
2	300	0,3	-50	500

В соответствии с теорией полезности оценивается ожидаемая полезность по каждому из вариантов:

U(d1) = $0.7 \times 350 + 0.3 \times (-50) = 230$

U(d2) = $0.3 \times 500 + 0.7 \times (-100) = 80$

Следовательно, перед испытуемым ваза 1-го типа, так как **U**(**d1**) > **U**(**d2**).

Таблицу можно представить в виде дерева решений:

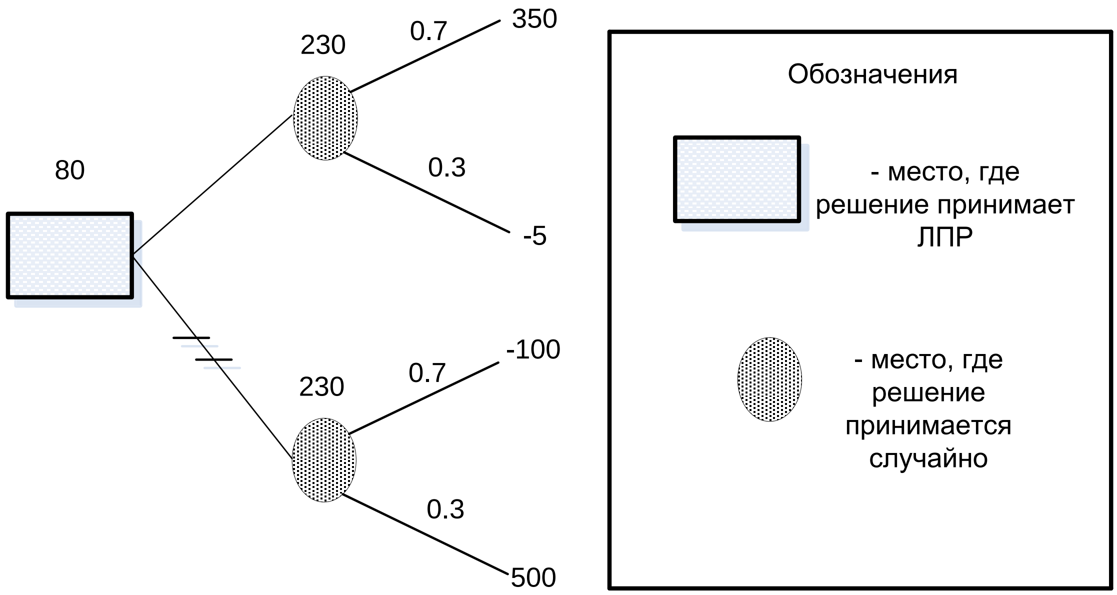


Рисунок 3.5

Процесс определения оптимальной ветки называется *«сворачиванием дерева решений»*. Отбрасываемая ветка перечеркивается двумя линиями.

Упражнения

Задача 1. В соответствии с теорией полезности оценить ожидаемую полезность действия **d1** или **d2** для задачи с вазами. Построить дерево решений. Формулировка задачи с вазами. Перед Вами 2 вазы (ваза типа **А** и ваза типа **Б**). Ваз типа **А** –55 штук, а ваз типа **Б** – 45. Если Вы угадали вазу типа **А**, то получаете 300 у.е., а если не угадали –платите штраф 170 у.е. Если Вы угадали вазу типа **Б**, то получаете 200 у.е., а если не угадали – платите штраф 120 у.е.

Ответ: в соответствии с теорией полезности определить среднее значение выигрыша каждого из действий.

U(**А**)=0,55 × 300+ 0,45 × (-170)= 165-76,5=88,5 у.е.

U(**Б**)= 0,45 × 200+0,55 ×(-120)=90- 66=24 у.е.

Следовательно, **ЛПР** должен выбрать вазу типа **А**.

Задача 2. Существует два проекта реализации информационных систем (ИС), приведённые в таблице 3.3

Таблица 3.3 -Представление задачи с вазами

Проект 1		Проект 2	
Прибыль	Вероятность	Прибыль	Вероятность
2000	0,2	1200	0,1
3000	0,25	2000	0,2
1500	0,15	3500	0,25
4000	0,3	4500	0,3
1000	0,1	1500	0,15

Величина планируемого дохода в каждом случае не определена и приведена в виде распределения вероятностей.

Ответ: Для выбора оптимального варианта подсчитывается математическое ожидание дохода от каждого проекта.

М (проект 1) = 2000 × 0,2 + 3000 ×0,25 + 1500 × 0,15 + 4000 × 0,3 + 1000 ×0,1 = 400 + 750 + 225 + 1200 + 100 = 2675

М (проект 2) = 1200 × 0,1 + 2000 × 0,2 + 3500 ×0,25 + 4500 ×0,3 + 1500 ×0,15 = 120 + 400 + 875 + 1350 + 225 = 2970

Таким образом, предпочтительнее выбрать проект 2, так как ***М (проект 2)*** > ***М (проект 1)***.

Задача 3. В соответствии с теорией полезности оценить ожидаемую полезность действия **d1** или **d2** для задачи с вазами. Построить дерево решений. Перед Вами 2 вазы (ваза типа **А** и ваза типа **Б**). Ваз типа **А** – 60 штук, а ваз типа **Б** – 40. Если Вы угадали вазу типа **А**, то получаете 200 у.е., а если не угадали –платите штраф 70 у.е.   Если Вы угадали вазу типа **Б**, то получаете 100 у.е., а если не угадали – платите штраф 20 у.е.

Задача 4. В соответствии с теорией полезности оценить ожидаемую полезность действия **d1** или **d2** для задачи с вазами. Построить дерево решений. Перед Вами 2 вазы (ваза типа **А** и ваза типа **Б**). Ваз типа **А** –200 штук, а ваз типа **Б** – 400. Если Вы угадали вазу типа **А**, то получаете 500 у.е., а если не угадали –платите штраф 270 у.е. Если Вы угадали вазу типа **Б**, то получаете 600 у.е., а если не угадали – платите штраф 320 у.е.

Задача 5. В соответствии с теорией полезности оценить ожидаемую полезность действия **d1** или **d2** для задачи с вазами. Построить дерево решений. Перед Вами 2 вазы (ваза типа **А** и ваза типа **Б**). Ваз типа **А** – 60 штук, а ваз типа **Б** –40. Если Вы угадали вазу типа **А**, то получаете 200 у.е., а если не угадали –платите штраф 70 у.е. Если Вы угадали вазу типа **Б**, то получаете 100 у.е., а если не угадали – платите штраф 80 у.е.

Задача 6. Как нужно изменить поощрения и наказания по угадыванию типа вазы, чтобы **ЛПР** выбрал нижнюю ветку в предыдущей задаче на рисунке 3.4.

Литература

5. Ларичев О.И. Теория и методы принятия решений, а также хроника событий в волшебных странах: Учеб./О.И. Ларичев- М.: Логос, 2000.- 296 с.- (Учеб. XXI века).
1. Фишберн П. Теория полезности для принятия решений/ П. Фишберн- М.: Наука, 1978.

4 Сетевое планирование в задачах принятия решений

На практике приходится сталкиваться с задачами планирования разных работ, процесс выполнения которых нельзя отразить в формульных зависимостях. Например, строительство крупного объекта, разработка сложного радиоэлектронного комплекса и т.п.

Характерная особенность таких проектов – взаимная обусловленность проводимых работ, выраженная в требовании соблюдать определенный порядок выполнения работ. Следовательно, подготовка решений при разработке проектов должна проводиться с учетом многих факторов, отражающих ограниченность сырьевых, денежных, энергетических ресурсов и ограниченного времени, отводимого на предполагаемые работы. Одним из методов, широко применяемых в исследовании проблем, является метод сетевого планирования [1].

Метод сетевого планирования позволяет решать как прямые, так и обратные задачи исследования операций. Первые связаны с оценкой последствий выбора определенного решения, вторые - с поиском наилучших решений.

4.1 Способы представления комплекса работ

Исходными данными в сетевом планировании являются перечни проводимых работ, установленные для каждой из них сроки, а также описания существующих между ними зависимостей, выраженных отношениями предшествования. Эти данные можно представить следующими способами:

в виде таблицы с последовательностью работ;

в виде Гантт-карты (рисунок 4.1);

в виде сводного сетевого графика (ССГ).

В соответствии с теорией расписаний, расписанием можно назвать документ, содержащий сведения о количестве и номенклатуре выполняемых работ, включая их этапы; о моментах начала и окончания каждой работы, о затратах времени и выделяемой трудоемкости. Расписания можно задавать разными способами, среди которых наиболее наглядным является геометрический на базе графика Гантта (Гантт-карты).

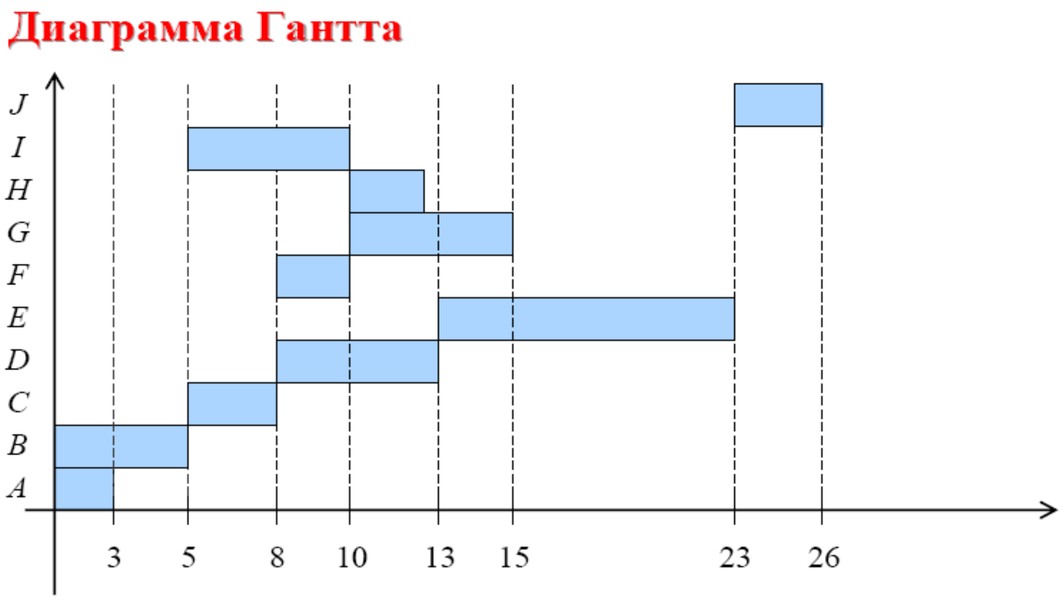


Рисунок 4.1-Диаграмма Гантта

4.2 Сводный сетевой график комплекса работ

Правила построения ССГ. Работы изображаются сплошными линиями, а связи между событиями изображаются пунктирными линиями (рисунок 4.2).

Фрагмент сетевого графика работ

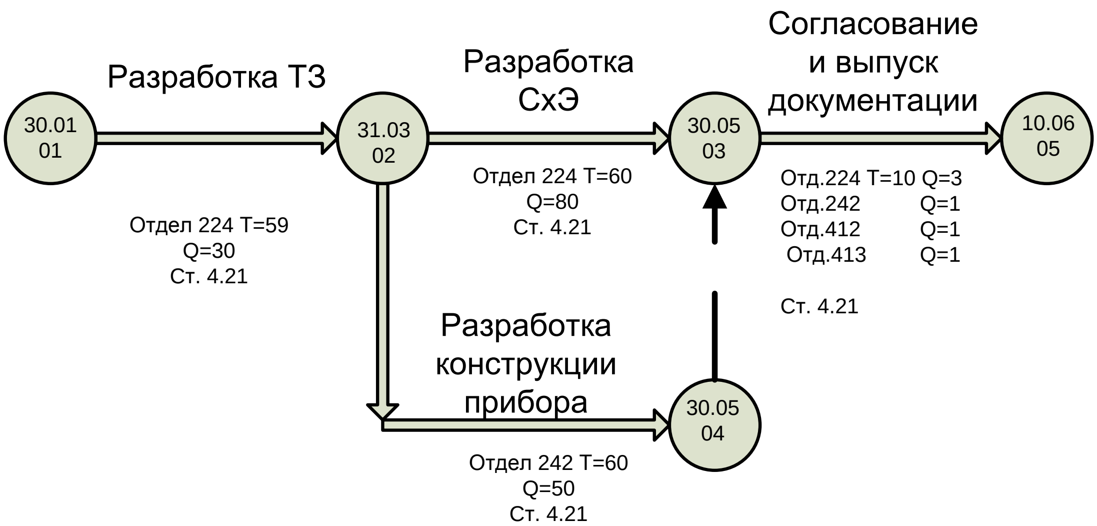


Рисунок 4.2- Фрагмент сводного сетевого графика

События изображаются кружками. Под линией, соединяющей два события, обозначаются продолжительность, трудоемкость выполнения работы и исполнители. Пример изображения фрагмента ССГ разработки электронного модуля приведен на рисунке 4.2.

При выпуске конструкторской документации задействован отдел нормоконтроля 413 и отдел размножения и учета КД (отдел 412). Номера событий записываются в нижней части кружка, а ее срок начала или окончания – в верхней части кружка. ССГ представляет собой ориентированный граф, дугами которого являются работы. Путем в графе называют последовательность дуг, выбранных таким образом, что конец каждой предыдущей дуги совпадает с началом последующей.

Последовательными называются работы, лежащие на некотором пути графа. Параллельные работы не могут составить путь и изображаются несколькими дугами, соединяющие вершины, например: события 03 и 04.

Связи между событиями, изображенные пунктирными линиями, еще называют фиктивными работами.

В любом ССГ имеется начальное событие (01) и конечное событие (05).

Каждая работа представляет собой процесс, связанный с затратами труда, материальных ресурсов и времени. Трудоемкость обозначается буквой Q, а продолжительность работы буквой T. Стадия разработки определяются в соответствии со стандартом.

Путь, имеющий в ССГ наибольшую продолжительность, называется критическим, а работы находящиеся на нем – критическими.

Особенность критических работ определяется необходимостью их четкого выполнения. Срыв таких работ приводит к срыву плана.

Каждый ССГ имеет свой номер, например: подсистема 96.

Стыковка разных графиков производится следующим образом.

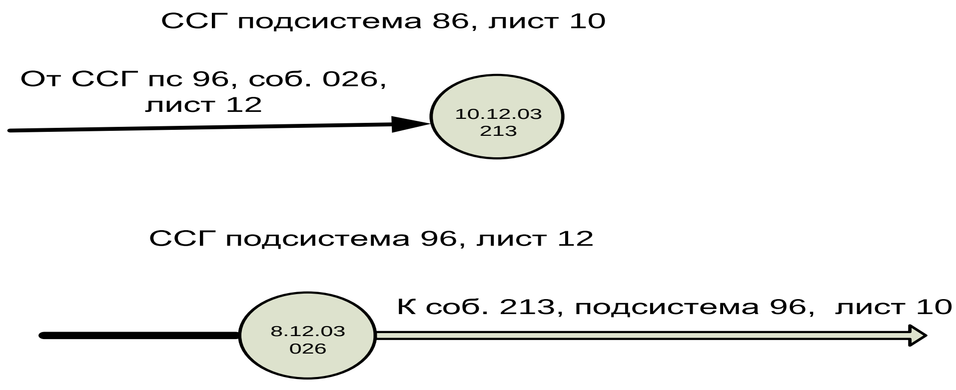


Рисунок 4.3- Стыковка графиков

Задача 1. В соответствии с директивным сетевым графиком (ДСГ) разработать три ССГ и согласовать их по срокам. Найти критический путь на каждом ССГ. Для этого создаются три группы студентов.

Группа 1 разрабатывает ССГ (подсистема 11) по проектированию системного блока с АЦП

Группа 2 разрабатывает ССГ (подсистема 12) по выбору типа монитора, клавиатуры и мышки

Группа 3 разрабатывает ССГ (подсистема 21) по проектированию программы информационно-измерительного комплекса (ИИК) на базе персонального компьютера.

Перечень работ по подсистемам 11 и 12 приведен в таблице 4.1

Таблица 4.1

Наименование работы	Продолжительность (дней)	Трудоемкость (ИТР/дней)
Разработка ТЗ	45	30
Согласование ТЗ	10	5
Выпуск ТЗ	3	2
Разработка СхЭ блоков, прибора	60	50
Разработка КД блоков, прибора	60	50
Разработка тестов блоков и прибора	60	50
Изготовление макетного образца	40	30
Настройка макетного образца	30	60
Отладка тестов блоков и прибора	40	40
Отладка интерфейса «системный блок-монитор, клавиатура, мышь»	20	40
Испытания прибора с использованием АЦП и имитацией внешней среды	30	60
Доработка СхЭ, КД по результатам испытаний	15	20

4.3 Правила построения сетевых моделей

Между парой событий можно изобразить только одну работу. При необходимости двух параллельных работ вводится промежуточное событие и фиктивная работа с нулевой продолжительностью. Из сети исключают тупиковые события, из которых не начинается ни одна работа, кроме завершающего события. Устраняют пересечения. Нумерацию событий проводят слева направо и сверху вниз.

Таблица 4.2- Задача перевода предприятия на самообслуживание

Работа	Содержание работы	Длительность, в днях
0,1	Составление сметы	15
1,2	Приобретение оборудования	16
1,3	Подбор кадров	6
2,4	Монтаж оборудования	6
3,5	Подготовка кадров	5
4,6	Оформление зала	8
5,6	Доставка товаров	6
5,8	Заказ и получение формы	14
5,7	Заказ и получение ценников	8
6,8	Выкладка товаров	2
7,8	Заполнение ценников	4
8,9	Генеральная репетиция	3

График работ, построенный в соответствии с таблицей 4.2, приведен на рисунке 4.4. Длительность выполнения работ указаны на дугах графа.

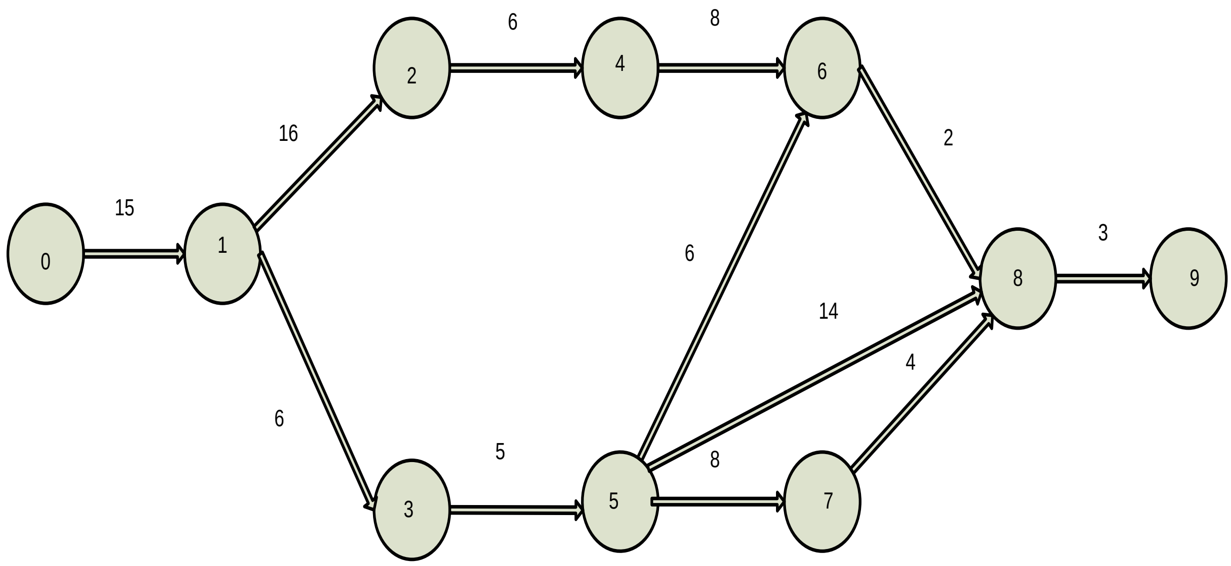


Рисунок 4.4

4.4 Анализ сетевых моделей

Анализ проводится с целью выявления резервов. Графический метод дает большую наглядность. Обнаруженные резервы позволяют гибко управлять комплексом работ [1].

4.4.1 Определение минимального времени выполнения всего комплекса работ (4 варианта путей):

L1=(0,1)(1,2) (2,4)(4,6)(6,8)(8,9)=50,

L2=(0,1)(1,3) (3,5)(5,6)(6,8)(8,9)=37,

L3=(0,1)(1,3) (3,5)(5,8)(8,9)=43,

L4=(0,1)(1,3) (3,5)(5,7)(7,8)(8,9)=41.

Поскольку многие работы, лежащие на этих путях выполняются параллельно, общий срок перевода предприятия на самообслуживание будет определяться путем максимальной продолжительности, называемым *критическим. Ткр=50 дней.*

Длительность L2 =37 минимальна, но не позволяет выполнить все работы.

5.4.2 Определение полных резервов времени:

R(L1)=Ткр-T1=0,

R(L2)=Ткр-T2=13 дней,

R(L3)=Ткр-T3=7 дней,

R(L4)=Ткр-T4=9 дней.

В пределах этих резервов времени с выполнением некоторых работ можно не спешить, так как общий срок выполнения комплекса работ не увеличивается. Если длительность работ, лежащих на критическом пути увеличилась, то общий срок возрастет.

4.4.3 Оптимизация сетевого графика.

Путем решения системы уравнений может быть найден оптимальный сетевой график, приведенный на рисунке 4.5. В этом варианте длительности всех четырех путей стали равными 43,3 дня.

Решение задачи оптимизации состоит в последовательном переносе средств с некритических работ на критические, переходе от одного пути к другому до тех пор, пока все работы не будут критическими и не будут иметь резервов, а длительности всех путей станут равными [1].

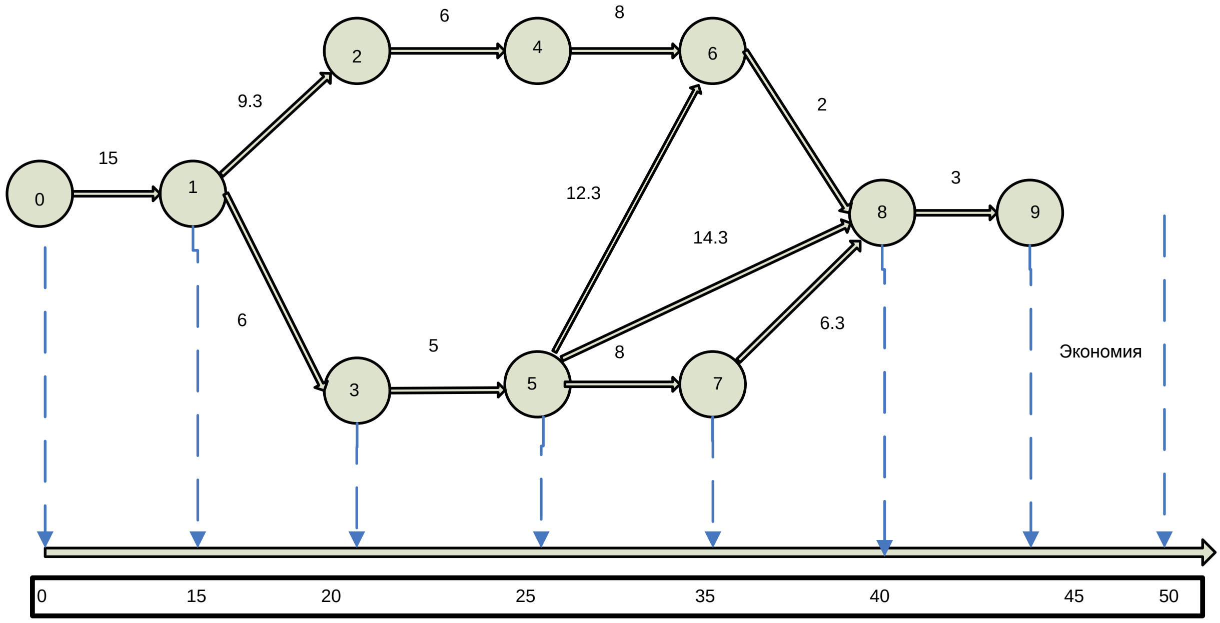


Рисунок 4.5

Упражнения

1. Построить сетевой график выполнения курсового проекта. Найти критический путь. Провести оптимизацию графика выполнения курсового проекта.

Литература

1.Фомин Г.П. Математические методы и модели в коммерческой деятельности: Учеб./Г.П. Фомин.- М.: Финансы и статистика, 2001.-544с.

5 Применение теории цепей Маркова

5.1 Основные понятия теории марковских цепей

Конечный стохастический процесс полностью описывается деревом логических возможностей и заданием весов всех его ветвей. Одним из типов таких процессов являются марковские цепи [1].

Конечный стохастический процесс с функциями исходов f1,f2,…,fn называется ***марковской цепью***, если исходное состояние f_0 фиксировано [5].

Марковская цепь характеризуется тем, что вероятности перехода P i j, задающие вероятность перехода системы из состояния Si в состояние Sj, определяется для всех упорядоченных пар состояний. Кроме того, должно быть задано исходное состояние, в котором, по предположению, находится наша система в начальный момент времени.

На основании этих данных можно построить для любого (конечного) числа шагов марковской цепи дерево логических возможностей и вероятностную меру на нем.

Пример. В регионе никогда не бывает двух ясных дней подряд. Если сегодня ясно, то завтра с одинаковой вероятностью пойдет дождь или снег. Если сегодня снег (или дождь), то с вероятностью ½ погода не изменится. Если все же она изменяется, то в половине случаев снег заменяется дождем или наоборот, и лишь в половине случаев на следующий день будет ясная погода. Допустим исходное состояние – в регионе ясная погода. Буквами Д, Я, С обозначаются условно дождь, ясный день, снег. Условия задачи удобно представить в виде квадратной матрицы следующего вида:

$$\begin{array}{c} \\ Д \\ Я \\ С \end{array}\begin{array}{c} \begin{array}{ccc} Д & Я & С \end{array} \\ \left| \begin{array}{ccc} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{array} \right| \end{array}$$

Числа в первой строке матрицы представляют собой вероятности различной погоды после дождя. Числа во второй строке – вероятности различной погоды после ясного дня, а в третьей строке – после снега. Воспользовавшись этой матрицей можно построить дерево логических возможностей для погоды, например на три последовательных дня, и определить вероятностную меру на нем.

Построенное дерево представлено на рисунке 5.1. Это дерево позволяет сделать прогноз о возможности дождя в каждый из трех последующих дней: P[f1=Д]=1/2, P[f2=Д]=3/8, P[f3=Д]=13/32.

Расчет производится следующим образом. На втором ярусе (второй день) имеется два исхода Д с вероятностными мерами 1/2×1/2=1/4 и ½*1/4=1/8. Их сумма дает 3/8. На третьем ярусе (третий день) имеется шесть исходов Д с вероятностными мерами 1/2×1/2×1/2=1/8, 1/2×1/4×1/2=1/16, 1/2×1/4×1/4=1/32, 1/2×1/4×1/2=1/16, 1/2×1/4×1/2=1/16, 1/2×1/2×1/4=1/16. Их сумма дает 13/32.

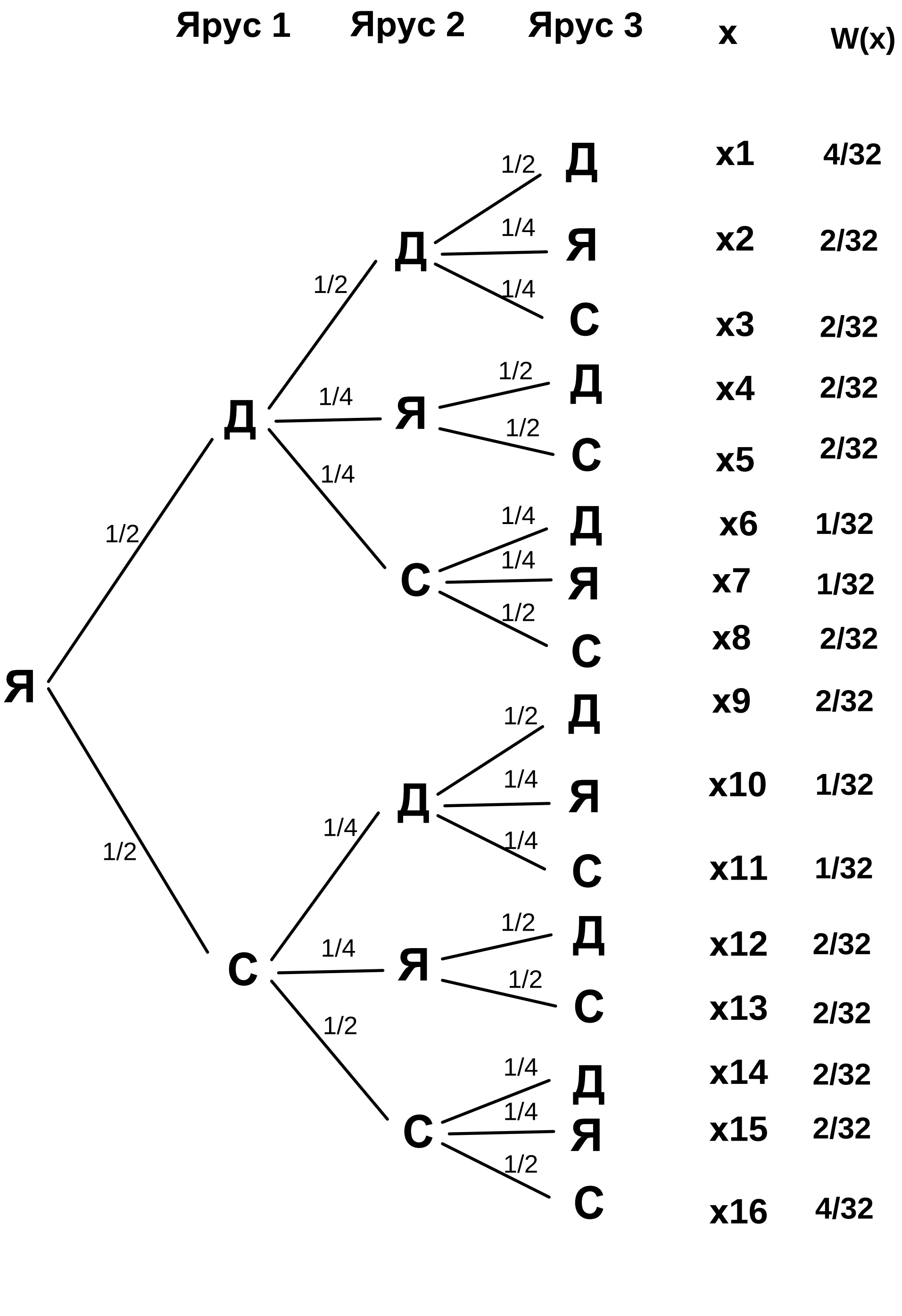


Рисунок 5.1-Дерево логических возможностей

5.2 Марковские цепи и диаграммы переходов

Исходом каждого эксперимента служит один исход из конечного множества возможных исходов a1,a2,…, an, причем в каждом эксперименте вероятность исхода aj либо вовсе не зависит от исходов предшествующих экспериментов, либо зависит от исхода эксперимента, непосредственно предшествующего данному исходу. Эта зависимость задается величиной Pij, представляющей вероятность исхода aj при условии, что предшествующий эксперимент имел исход ai. Исходы a1,a2,…, an называются состояниями, а числа Pij – вероятностями перехода. Вероятности перехода можно представить двумя различными способами: матрицей и диаграммой перехода. Для марковской цепи с состояниями a1,a2,…, an такая матрица перехода системы P имеет вид

	a1	a2	…	an
a1	a11	a12		an
a2	a21	a22		a2n
…				
an	an1	n2		ann

Пример матрицы вероятностей перехода, и диаграммы перехода приведен на рисунке 5.2. Стрелки, идущие от каждого состояния, указывают состояния, в которое это состояние может переходить в рассматриваемом процессе.

P =

	a1	a2	a3
a1	0	1	0
a2	0	½	½
a3	1/3	0	2/3

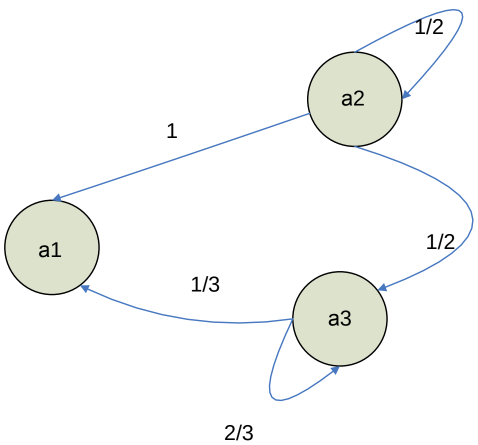


Рисунок 5.2 – Диаграмма перехода

В матрице P сумма элементов каждой строки равна единице. Это должно быть справедливо для любой матрицы вероятностей перехода.

Основной вопрос при изучении марковских цепей состоит в следующем. Пусть процесс начинается из состояния i. ***Какова вероятность***

того, что через n шагов он перейдет в состояние j? Обозначим эту вероятность $P_{ij}^{(n)}$. Здесь параметр n обозначает не степень числа P, а количество шагов. Более того, нас интересует эта вероятность для всех возможных начальных состояний i и всех возможных конечных состояний j.

Это удобно представить в виде матрицы. Для марковской цепи с тремя состояниями матрицу перехода можно представить в следующем виде:

$P^{(n)}$ =

$P_{11}^{(n)}$	$P_{12}^{(n)}$	$P_{13}^{(n)}$
$\mathbf{P_{21}^{(n)}}$	$P_{22}^{(n)}$	$P_{23}^{(n)}$
$P_{31}^{(n)}$	$P_{32}^{(n)}$	$P_{33}^{(n)}$

Пример. Для марковской цепи с вероятностями перехода, приведенными на рисунке 5.2, найти вероятности различных состояний через три шага для случая, когда процесс начинается из состояния a1.

Решение. Вначале строится дерево и производится поиск вероятностной меры (см. рисунок 5.3).

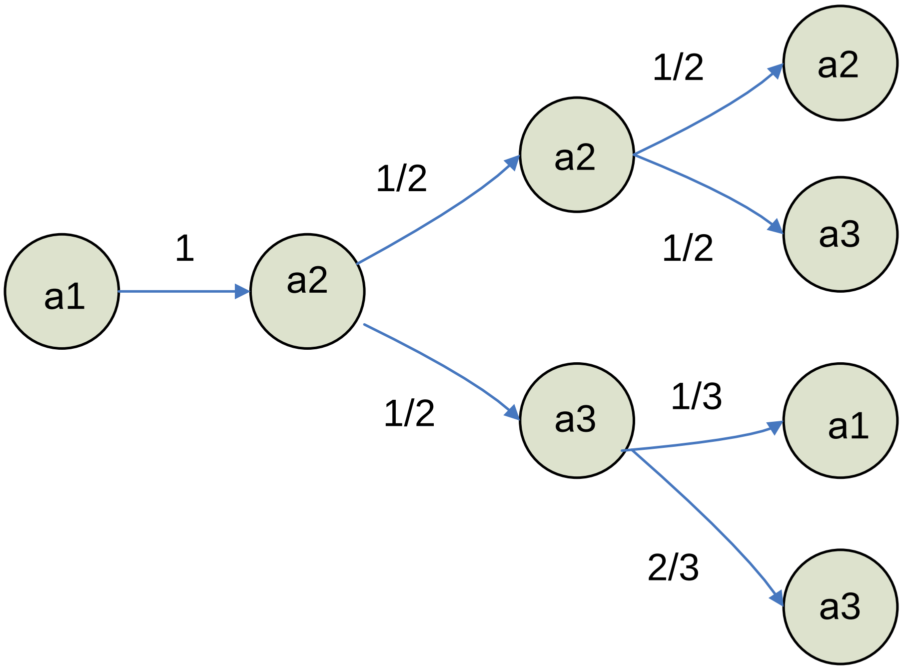


Рисунок 5.3

Например, вероятность $P_{13}^{(3)}$ есть сумма всех весов, приписанных введенной вероятностной мерой тем путям дерева, которые оканчиваются состоянием a_3 : $P_{13}^{(3)} = 1\times1/2\times1/2 + 1\times1/2\times2/3 = 7/12$; $P_{12}^{(3)} = 1\times1/2\times1/2 = 1/4$;

$P_{11}^{(3)} = 1\times1/2\times1/3 = 1/6$. Для того, чтобы найти вероятности $P_{21}^{(3)}$, $P_{22}^{(3)}$, $P_{23}^{(3)}$ из состояния а2, необходимо построить аналогичное дерево с исходным состоянием а2. Аналогично определяются $P_{31}^{(3)}$, $P_{32}^{(3)}$, $P_{33}^{(3)}$. Результаты можно записать в виде матрицы $P^{(3)}$.

$P^{(3)} =$

	a1	**a2**	**a3**
a1	1/6	1/4	7/12
a2	7/36	7/24	37/72
a3	4/27	7/18	25/54

Сумма элементов каждой строки по-прежнему равна единице. Это соответствует тому факту, что из какого бы состояния ни начинать, через три шага мы обязательно достигаем либо первого состояния, либо второго, либо третьего. Все элементы матрицы являются положительными числами.

Матрицу $P^{(3)}$ можно получить и другим способом. В [1] показано, что вероятности удовлетворяют следующим условиям:

$P_1^{(n)} = P_1^{(n-1)}\times p11 + P_2^{(n-1)}\times p21 + P_3^{(n-1)}\times p31,$

$P_2^{(n)} = P_1^{(n-1)}\times p12 + P_2^{(n-1)}\times p22 + P_3^{(n-1)}\times p32,$

$P_3^{(n)} = P_1^{(n-1)}\times p13 + P_2^{(n-1)}\times p23 + P_3^{(n-1)}\times p33.$

Это условие можно записать в виде произведения вектора на матрицу:

$P^{(n)} = P^{(n-1)}\times P.$

Используя матрицу P, представленную на рисунке 5.2, сначала находится матрица $P^{(2)}$, а затем $P^{(3)}$. Для этого необходимо перемножить матрицы P. $P^{(2)} = P\times P$. Например, для нахождения второй строки матрицы $P^{(2)}$ необходимо последовательно перемножить вторую строку матрицы P (см. рисунок 5.2) соответственно на первый, второй и третий столбцы этой матрицы.

$$P^{(2)} = \begin{vmatrix} 0 & 1 & 0 \\ 0 & 1/2 & 1/2 \\ 1/3 & 0 & 2/3 \end{vmatrix} \times \begin{vmatrix} 0 & 1 & 0 \\ 0 & 1/2 & 1/2 \\ 1/3 & 0 & 2/3 \end{vmatrix}$$

$P_{21}^{(2)} = 0\times0 + 1/2\times0 + 1/2\times1/3 = 1/6$

$P_{22}^{(2)} = 0\times1 + 1/2\times1/2 + 1/2\times0 = 1/4$

$P_{23}^{(2)} = 0\times0 + 1/2\times1/2 + 1/2\times2/3 = 7/12$

$P^{(2)} =$

	a1	**a2**	**a3**
a1	0	½	½
a2	**1/6**	**¼**	**7/12**
a3	2/9	1/3	4/9

Используя значения матрицы $P^{(2)}$ и исходной матрицы перехода P, приведенной на рисунке 5.2, можно получить значения матрицы $P^{(3)}$:

$$P^{(3)} = P^{(2)} \times P.$$

Упражнения

Задача 1. Составить диаграммы состояний для марковских цепей, вероятности перехода которых заданы следующими матрицами переходов P1 и P2:

P1=

	a1	a2	a3
a1	½	½	0
a2	0	1	0
a3	½	0	1/2

P2=

	a1	a2	a3
a1	1/3	1/3	1/3
a2	1/3	1/3	1/3
a3	1/3	1/3	1/2

Задача 2. Составить матрицу вероятностей перехода, соответствующие диаграммам состояний, изображенным на рисунке 5.4.

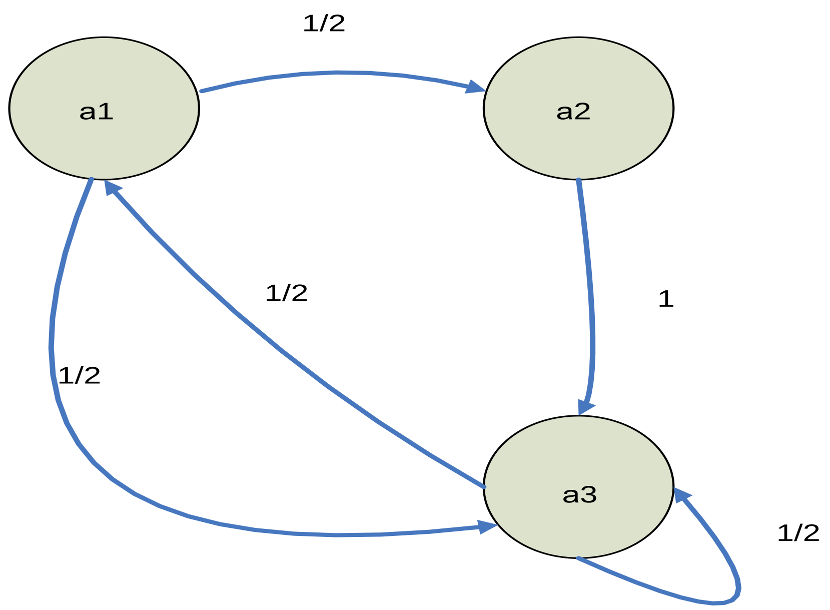


Рисунок 5.4

Задача 3. Найти матрицу перехода через 2 шага $P^{(2)}$ для марковской цепи, заданной матрицей вероятностей перехода P.

$$P = \begin{matrix} 1/2 & 1/2 \\ 1/3 & 2/3 \end{matrix} \qquad \text{Ответ: } P^{(2)} = \begin{vmatrix} 5/12 & 7/12 \\ 7/18 & 11/18 \end{vmatrix}$$

Задача 4. Найти матрицу перехода через 2 шага $P^{(2)}$ для марковской цепи, заданной матрицей вероятностей перехода P.

$$P = \begin{vmatrix} 0{,}4 & 0{,}6 \\ 0{,}3 & 0{,}7 \end{vmatrix} \qquad \text{Ответ: } P^{(2)} = \begin{vmatrix} 0{,}34 & 0{,}66 \\ 0{,}33 & 0{,}67 \end{vmatrix}$$

Литература

1. Кемени Дж. Конечные цепи Маркова/ Дж. Кемени. – М.: Наука, 1970.

6 Теория игр и принятие решений

6.1 Основные понятия теории игр

Формализованную модель конфликтной ситуации называют игрой, которая проводится по определенным правилам [1,2].

Если в игре можно выделить две антагонистические стороны, то игра называется парной. В противном случае игра называется множественной [1].

Игру называют с нулевой суммой, если одна из сторон выигрывает, а другая проигрывает.

Отдельные решения называют ходами. Ход называют личным, если выбор производится сознательно, и случайным, если действует случайный механизм выбора (бросание монеты).

Совокупность правил, определяющих однозначно выбор решения, называют стратегией.

Если у конфликтующих сторон имеется конечное число стратегий, то игру называют конечной.

Исход игры, в которой кроме личных ходов имеются случайные, определяется математическим ожиданием.

Если у одного игрока имеется n стратегий (**A1, A2, …,An**), а у другого m стратегий (**B1,B2,…,Bm**) и игра состоит только из личных ходов, выбор пары **AjBi** единственным образом определяет исход игры **aji**. Если известны значения aji для каждой пары стратегий, то можно составить платежную матрицу.

Рассмотрим пример, заимствованный из книги Е.С**.** Вентцель «Элементы теории игр». В нашем распоряжении имеется три вида вооружения **A1, A2, A3**; у противника три вида самолетов **B1, B2, B3**.

Известно, что при применении вооружения **A1** самолеты **B1, B2, B3** поражаются с вероятностями 0,9; 0,4; 0,2; при применении вооружения **A2** самолеты **B1, B2, B3** поражаются с вероятностями 0,3; 0,6; 0,8; при применении вооружения **A3** самолеты **B1, B2, B3** поражаются с вероятностями 0,5; 0,7; 0,2. Требуется сформулировать конфликтную ситуацию в терминах теории игр.

Решение. Конфликтную ситуацию можно рассматривать как игру 3 × 3 с двумя личными ходами и одним случайным. Личный ход противника – выбор типа самолета. Наш личный ход – выбор типа вооружения. Случайный ход – применение вооружения. Этот ход может закончится поражением или непоражением самолета. Нашими стратегиями являются три варианта применения типов вооружения. Стратегиями противника - три варианта применения типов самолета. Наш выигрыш равен единице, если самолет поражен, и нулю в противном случае. Среднее значение выигрыша при каждой заданной паре стратегий есть вероятность поражения самолета данным вооружением. Платежная матрица представлена в таблице 6.1.

Решить игру – значит найти для каждого игрока наилучшие стратегии, которые обеспечат ему наибольший выигрыш или наименьший проигрыш.

Таблица 6.1

	B1	**B2**	**B3**
A1	0,9	0,4	0,2
A2	0,3	0,6	0,8
A3	0,5	0,7	0,2

Пусть требуется найти оптимальную стратегию для игрока **А**. Проанализируем последовательно каждую из его стратегий, начиная с **А1**.

Если игрок **А** выбрал стратегию **Aj**, то он должен рассчитывать на то, что игрок **В** ответит на нее той из своих стратегий, для которой его (игрока **А**) выигрыш $\mathbf{a_{ji}}$ будет минимален. Выберем это значение выигрыша, т.е. минимальное из чисел в **j**-й строке:

$$a_j = \min_i a_{ji}.$$

Избегая всякого риска, игрок A должен выбрать из стратегий Aj, для которой значение $\mathbf{a_j}$ является максимальным:

$$A = \max_j a_j,$$

$$a = \max_j \min_i a_{ji}.$$

Выражение «**а**» определяет нижнюю цену игры (максиминный выигрыш) и тот гарантированный минимум, который получит игрок **А**, придерживаясь наиболее осторожной из своих стратегий. Для нашего примера (см. таблицу 6.1) вначале находятся минимальные значения в каждой строке (0,2; 0,3; 0,2), а затем среди них максимальное значение **а**=0,3. Это и есть нижняя цена игры. Стратегия, соответствующая нижней цене игры, называется максиминной стратегией.

Игрок **В** заинтересован в том, чтобы свести выигрыш игрока **А** к минимуму. Следовательно, он должен проанализировать каждую из своих стратегий с точки зрения максимального выигрыша при этой стратегии:

$$b_i = \max_j a_{ji},$$

$$b = \min_i b_i,$$

$$b = \min_i \max_j a_{ji}.$$

Выражение «**b**» определяет верхнюю цену игры (минимаксный выигрыш). Для нашего примера (см. таблицу 6.1) вначале находятся максимальные значения в каждом столбце (0,9; 0,7; 0,8), а затем среди них минимальное значение **а**=0,7. Это и есть верхняя цена игры. Стратегия, соответствующая верхней цене игры, называется минимаксной стратегией.

Анализируя платежную матрицу, можно сделать вывод о неустойчивости минимаксных стратегий. Это означает, что если игрок **А** применит свою наиболее осторожную стратегию **А2**, а игрок **В** – стратегию **В2**, то средний выигрыш равен 0,6 (см. таблицу 6.1). Как только игроку **В** становится известно, что игрок **А** применяет стратегию **А2**, он может ответить на нее стратегией **В1**, уменьшив выигрыш игрока **А** с 0,6 до 0,3.

Следовательно, выигрыш при использовании минимаксных стратегий является неустойчивым, поскольку зависит от сведений о стратегии антагонистической стороны.

Существуют игры, для которых минимаксные стратегии являются устойчивыми. Для таких игр нижняя цена игры равна верхней и это общее значение называется *чистой ценой игры.* Элемент матрицы, являющийся одновременно минимальным в своей строке и максимальным в своем столбце, называется *седловой точкой матрицы.* Игра, платежная матрица которой имеет седловую точку, называется игрой с седловой точкой.

Для игр с седловой точкой существует решение, определяющее пару оптимальных стратегий обеих из конфликтующих сторон. Если один из игроков будет придерживаться своей оптимальной стратегии, а другой отклоняться от нее, то он может только проиграть.

6.2 Применение смешанных стратегий

Для игр, не имеющих седловой точки, можно применять не одну чистую стратегию, а чередовать случайным образом несколько стратегий с целью гарантировать себе средний выигрыш, больший «а».

Смешанной стратегией называется набор вероятностей применения чистых стратегий.

Рассмотрим игру 2 × 2. Пусть игра задана матрицей, приведенной в таблице 6.2.

Таблица 6.2

	В1	**В2**
А1	a11	a12
А2	a21	a22

Оптимальные смешанные стратегии игроков **А** и **В** определяются вероятностями **Р**=(**p1,p2**) и **Q**=(**q1,q2**). Цена игры определяется следующими неравенствами:

$a\ 11 \times p\ 11 + a\ 21 \times\ p2 \geq R$

$a\ 12 \times p\ 11 + a\ 22 \times\ p2 \geq R$

$p1>=0;\ p2>=0;\ p1+p2=1$

$a\ 11 \times q\ 1 + a\ 12 \times\ q2 \leq R$

$a\ 21 \times q\ 1 + a\ 22 \times\ q2 \leq R$

$q1>=0; q2>=0; q1+q2=1$

Вероятности **p1,p2**, **q1,q2** определяются путем решения системы неравенств:

$p1= |a\ 22 – a\ 21| \ / [|a\ 12- a11| + |a22 –a21|]$
$p2= |a\ 12 – a\ 11| \ / [|a\ 12- a11| + |a22 –a21|]$
$p2=1-p1$
$q1 = |a\ 22 – a\ 12| \ / |a\ 11 + a22 –a12 – a21|$
$q2=1-q1$

Матрица игры приведена в таблице 6.3.

Таблица 6.3

	B1	B2
A1	2	3
A2	7	1

Решить игру – значит найти **p1,p2,q1,q2,R**, удовлетворяющие соотношениям:

$2 \times p1 + 7 \times p2 =R$ (6.1)
$3 \times p1 + p2= R$ $p2=1-p1$
$-5 \times p1+7=R$ (6.2)

$2 \times p1+1=R$ $p1= 6/7 \quad p2=1/7 \quad R=19/7$
$2 * q1+3 * q2=R$ (6.3)

$7 \times q1+q2=R$ $q2=1- q1$
$- q1+3 \quad =R$ (6.4)

$6 \times q1+1=R$ $q1=2/7 \quad q2=5/7 \quad R=19/7$

Графический метод решения игры приведен на рисунке 6.1. Из уравнения 6.2 при **p1**=0 получаем точку (0; 7), а при **p1**=1 получаем точку (1; 2). Соединяем отмеченные точки прямой линией.

Анализируя второе уравнение, получаем точки (0; 1) и (1; 3). Строим вторую линию. Точка пересечения этих прямых соответствует решению системы уравнений.

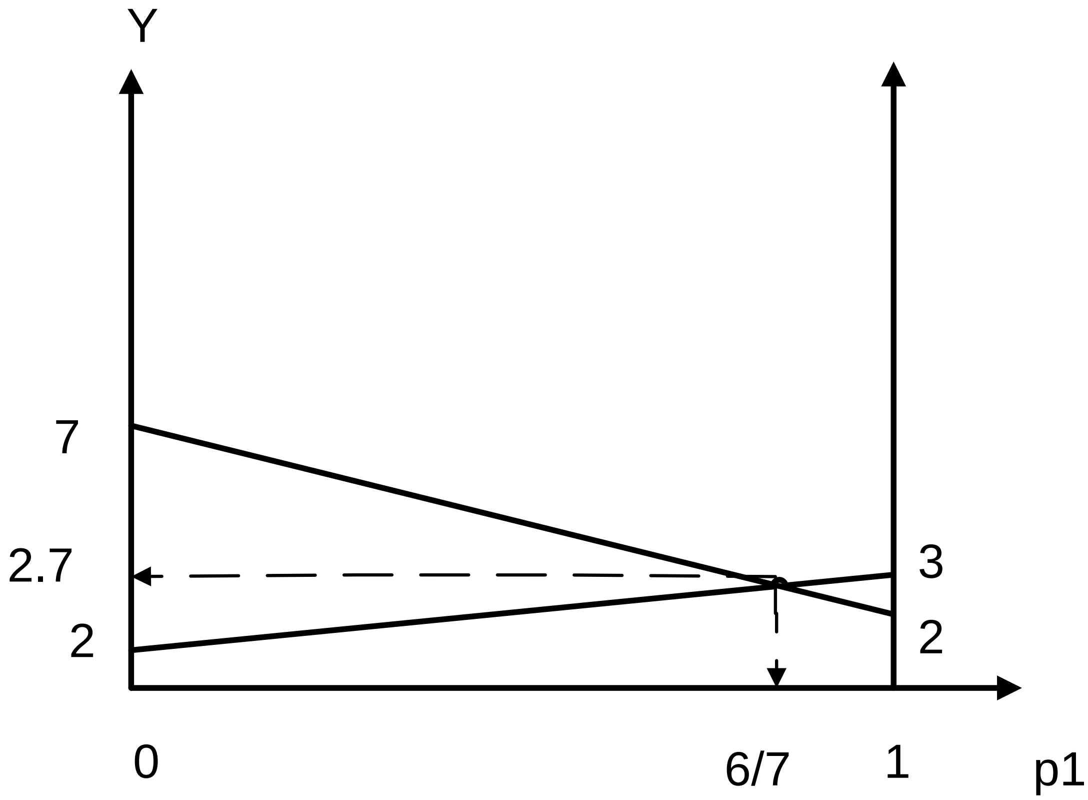


Рисунок 6.1

6.3 Игры n × 2 и 2 × m

Пусть имеется игра порядка **n × m**, где n обозначает число стратегий первого игрока (**A**); m – число стратегий второго игрока (**B**).

Если каждый элемент матрицы одной строки (столбца) больше соответствующего элемента другой строки (столбца) или равен ему, то говорят, что первая стратегия доминирует над второй.

Пример игры задан матрицей 3 × 3, приведенной в таблице 6.4.

Таблица 6.4

	B1	B2	B3	B4
A1	0	2	3	-1
A2	-3	3	4	9
A3	2	4	4	8

В этой игре стратегия **A3** доминирует над **A1**. Эта стратегия (**A3**) лучше, поэтому ее оставляют в матрице, удаляя доминируемую (**A1**).

Стратегия **B2** доминирует над **B1**; **B3** над **B2**; **B2** над **B1**. Здесь доминирующие стратегии (**B2** и **B3**) хуже, и их следует удалить.

Знание соотношения доминирования позволяет в ряде случаев свести игры $2 \times \mathbf{m}$ и $\mathbf{n} \times 2$ к играм 2×2 [1].

6.4 Графический метод решения игр 2 x n

Пример 1. Игра, задана платёжной матрицей. Первый игрок имеет стратегии A1, A2. Второй игрок имеет стратегии B1,B2,B3.

$$\begin{array}{cccc} & & \begin{array}{ccc} q1 & q2 & q3 \\ & 2 & \\ B_1 & B_2 & B_3 \end{array} & \\ 1 & \begin{array}{c} A_1 \\ A_2 \end{array} & \begin{pmatrix} 2 & 3 & 11 \\ 7 & 5 & 2 \end{pmatrix} & \begin{array}{c} \mathbf{P1} \\ \mathbf{P2} \end{array} \end{array}$$

На плоскости **R 0 p1** введём систему координат и на оси **0 p1** отложим отрезок единичной длины $\mathbf{A_1}$, $\mathbf{A_2}$, каждой точке которого поставим в соответствие некоторую смешанную стратегию игрока **1** (**p1, 1 –p1**). В частности, точке $\mathbf{A_1}$ (0;0) отвечает стратегия $\mathbf{A_1}$, точке $\mathbf{A_2}$ (1;0) – стратегия $\mathbf{A_2}$ и т.д.

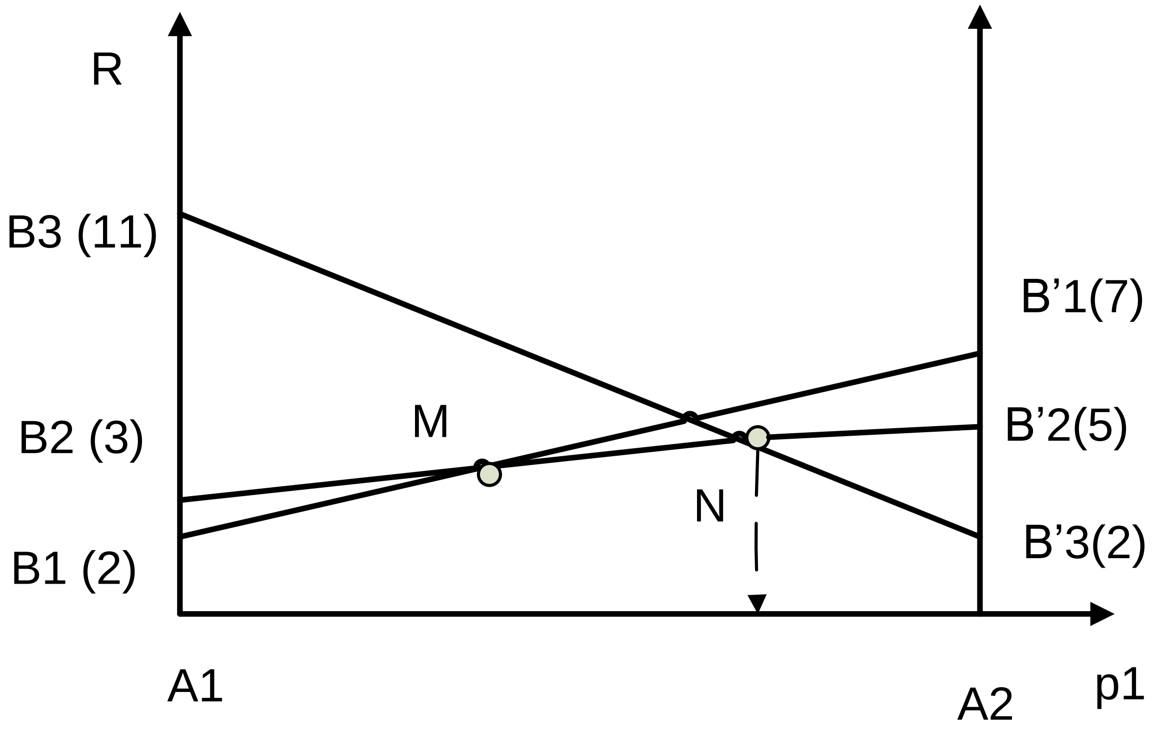


Рисунок 6.2

В точках $\mathbf{A_1}$ и $\mathbf{A_2}$ восстановим перпендикуляр и на полученных прямых будем откладывать выигрыш игроков (рисунок 6.2). На первом перпендикуляре (в данном случае он совпадает с осью $\boldsymbol{0R}$) отложим выигрыш игрока **1** при стратегии $\mathbf{A_1}$, а на втором – при стратегии $\mathbf{A_2}$. Если игрок **1** применит стратегию $\mathbf{A_1}$, то выиграет при стратегии $\mathbf{B_1}$ игрока **2** – 2, при стратегии $\mathbf{B_2}$ – 3, а при стратегии $\mathbf{B_3}$ – 11. Числам 2, 3, 11 на оси $\boldsymbol{0R}$ соответствуют точки $\mathbf{B_1}$, $\mathbf{B_2}$ и $\mathbf{B_3}$.

Если же игрок **1** применит стратегию $\mathbf{A_2}$, то его выигрыш при стратегии $\mathbf{B_1}$ равен 7, при $\mathbf{B_2}$ – 5, а при $\mathbf{B_3}$ – 2. Эти числа определяют точки $\mathbf{B'_1}$, $\mathbf{B_2'}$, $\mathbf{B_3'}$ на перпендикуляре, восстановленном в точке $\mathbf{A_2}$. Соединяя между собой точки $\mathbf{B_1}$ и $\mathbf{B'_1}$, $\mathbf{B_2}$ и $\mathbf{B_2'}$, $\mathbf{B_3}$ и $\mathbf{B_3'}$ получим три прямые, расстояние до которых от оси (**0 p1**) определяет средний выигрыш при любом сочетании соответствующих стратегий. Например, расстояние от любой точки отрезка $\mathbf{B_1\ B'_1}$ до оси **0p1** определяет средний выигрыш $\mathbf{R_1}$ при любом сочетании стратегий $\mathbf{A_1\ A_2}$ (с частотами **p1** и **1–p1**) и стратегией $\mathbf{B_1}$ игрока **2**. Это расстояние равно

$$2p1 + 6(1 - p2) = R1$$

(Вспомните планиметрию и рассмотрите трапецию $\mathbf{A_1\ B_1\ B'\ A_{2\ 1}}$

Таким образом, ординаты точек, принадлежащих ломанной $\mathbf{B_1\ M\ N\ B_3'}$ определяют минимальный выигрыш игрока **1** при применении им любых смешанных стратегий. Эта минимальная величина является максимальной в точке **N**.

Следовательно, этой точке соответствует оптимальная стратегия **P* = (p1, 1–p1**), а её ордината равна цене игры **R0**. Координаты точки **N** находим как точку пересечения прямых $\mathbf{B_2\ B'_2}$ и $\mathbf{B_3\ B'_3}$.

Соответствующие два уравнения имеют вид

$$3*p1+5*(1-p1)=R$$
$$11*p1+2*(1-p1)=R$$

p1=3/11: p2=8/11: R=49/11

Следовательно, P = (3/11; 8/11), при цене игры $R = \dfrac{49}{11}$.

Таким образом мы можем найти оптимальную стратегию при помощи матрицы (B1 не входит в оптимальную смешанную стратегию)

$$\begin{pmatrix} 3 & 11 \\ 5 & 2 \end{pmatrix}$$

Оптимальные стратегии для игрока 2 можно найти из системы

$$3*q2+11*q3=R$$
$$5*q2+2*q3=R$$

q2=9/11: q3=2/11 R=49/11

и, следовательно, **Q** = (0; $\frac{9}{11}$; $\frac{2}{11}$). (Из рисунка видно, что стратегия **B_1** не войдёт в оптимальную стратегию.

6.5 Графический метод решения игр m x 2

Пример 1. Игра, задана платёжной матрицей. Первый игрок имеет стратегии **A_1, A_2, A_3,A_4**. Второй игрок имеет стратегии B_1,B_2.
Найти решение игры, заданной матрицей.

$$\begin{array}{ccc} & & 2 \\ & & \begin{array}{cc} B_1 & B_2 \end{array} \\ 1 & \begin{array}{c} A_1 \\ A_2 \\ A_3 \\ A_4 \end{array} & \begin{pmatrix} 6 & 5 \\ 4 & 6 \\ 2 & 7 \\ 1 & 8 \end{pmatrix} \end{array}$$

Решение. Матрица имеет размерность 2 x 4. Строятся прямые, соответствующие стратегиям игрока **1**. Ломанная **A_1 K A'_4** соответствует верхней границе выигрыша игрока **1**, а отрезок **N K** –цене игры (рисунок 6.3).

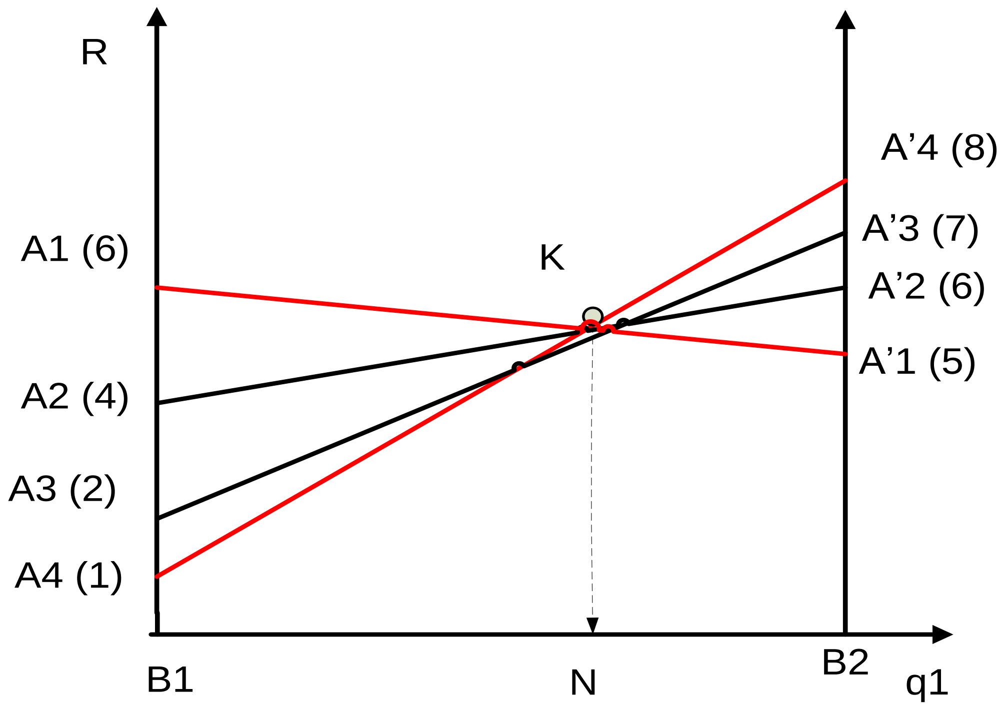


Рисунок 6.3

Решение игры таково

$$Q = \left(\frac{3}{8};\ \frac{5}{8}\right); \quad \mathrm{P} = \left(\frac{7}{8};\ 0;\ 0;\ \frac{1}{8}\right); \quad R = \frac{43}{8}.$$

Пример 2. В магазине имеются 2 зоны **A** и **B**. В зоне **A** большая толпа покупателей, а в зоне **B** - меньше. Администрация имеет 2 полицейских в штатском, и видеокамеру в месте **T** для наблюдения за зонами **A** и **B**. Полицейские имеют 6 стратегий поведения: **TT**, **AB**, **AA**, **BB**, **TA**, **TB**, которые обозначим как **p1**, **p2**, …,**p6**. В нашей привычной терминологии это игрок А, а вор – это игрок В.

Вор имеет 2 стратегии красть в зоне **A** (**B1**) или красть в зоне **B** (**B2**). Полицейские ищут такие смешанные стратегии **A1**, **A2**, …,**A6**, чтобы обеспечить задержание вора с вероятностью **R** независимо от поведения воров. Вор ищет такие стратегии (**B1, B2**), чтобы вероятность его задержания не превышала **R**. Задана матрица игры

Стратегия полицейских	Характеристика стратегии полицейских	Стратегия B1 (зона A)	Стратегия B2 (зона B)
A1	TT	0,51	0,75
A2	AA	0,64	0,36
A3	BB	0,19	0,91
A4	TA	0,58	0,60
A5	TB	0,37	0,85
A6	AB	0,46	0,76

Решение

В графическом методе решения задачи строятся прямые, соответствующие стратегиям полицейских.

B1 B2

A1 $0{,}51*q1+0{,}75*q2 \le R$

A2 $0.64*q1+ 0{,}36*q2 \le R$

A3 $0.19*q1+0.91*q2 \le R$

A4 $0.58*q1+0.60*q2 \le R$

A5 $0.37*q1+0.85*q2 \le R$

A6 $0{,}46*q1+0{,}76*q2 \le R$

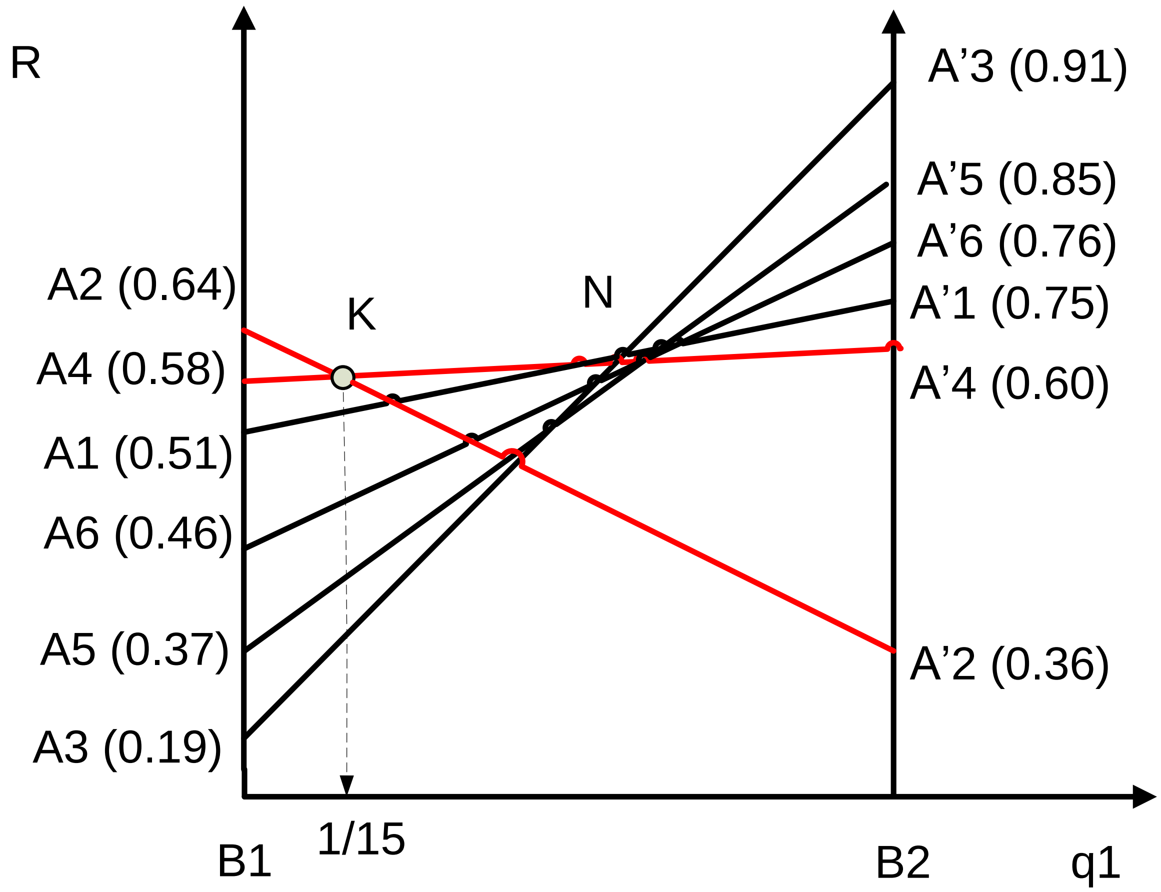


Рисунок 6.4

Ломанная A2 K N A′3 соответствует верхней границе выигрыша полицейских, а ордината точки **K** –цене игры (рисунок 6.4). В точке К полицейские имеют гарантированный выигрыш, равный 1/15.

Ответ:
P=(0; 1/15; 0; 14/15; 0;0) x2=1/15 (стратегия АА) и x4=14/15 (стратегия ТА)
Q= (0.8; 0.2) R= 0.584

6.6 Принятие решений в условиях риска

6.6.1 Критерий ожидаемого значения

Использование критерия ожидаемого значения обусловлено стремлением максимизировать ожидаемую прибыль (или минимизировать ожидаемые затраты). Использование ожидаемых величин предполагает возможность многократного решения одной и той же задачи, пока не будут получены достаточно точные расчётные формулы. Математически это выглядит так: пусть $\boldsymbol{X}$ – случайная величина с математическим ожиданием $\boldsymbol{MX}$ и дисперсией $\boldsymbol{DX}$. Если $\boldsymbol{x_1, x_2, ..., x_n}$ – значения случайной величины (с.в.) $\boldsymbol{X}$, то среднее арифметическое их (выборочное среднее) значений $\bar{x} = \frac{x_1 + x_2 + ... + x_n}{n}$ имеет дисперсию $\frac{DX}{n}$. Таким образом, при $\boldsymbol{n} \to \infty$

$\frac{DX}{n} \to 0$ и $\bar{x} \to MX$.

Другими словами при достаточно большом объёме выборки разница между средним арифметическим и математическим ожиданием стремится к нулю (так называемая предельная теорема теории вероятности). Следовательно, использование критерия ожидаемого значения справедливо только в случае, когда одно и то же решение приходится применять достаточно большое число раз. Верно и обратное: ориентация на ожидания будет приводить к неверным результатам для решений, которые приходится принимать небольшое число раз.

Пример 1. Требуется принять решение о том, когда необходимо проводить профилактический ремонт ПЭВМ, чтобы минимизировать потери из-за неисправности. В случае, если ремонт будет производиться слишком часто, затраты на обслуживание будут большими при малых потерях из-за случайных поломок.

Так как невозможно предсказать заранее, когда возникнет неисправность, необходимо найти вероятность того, что ПЭВМ выйдет из строя в период времени $\boldsymbol{t}$. В этом и состоит элемент ”риска”.

Математически это выглядит так: ПЭВМ ремонтируется индивидуально, если она остановилась из-за поломки. Через $\boldsymbol{T}$ интервалов времени выполняется профилактический ремонт всех n ПЭВМ. Необходимо определить оптимальное значение $\boldsymbol{T}$, при котором минимизируются общие затраты на ремонт неисправных ПЭВМ и на проведение профилактического ремонта в расчёте на один интервал времени.

Пусть $\boldsymbol{p_t}$ – вероятность выхода из строя одной ПЭВМ в момент $\boldsymbol{t}$, а $\boldsymbol{n_t}$ – случайная величина, равная числу всех вышедших из строя ПЭВМ в тот же момент. Пусть далее $\boldsymbol{C_1}$ – затраты на ремонт неисправной ПЭВМ и $\boldsymbol{C_2}$ – затраты на профилактический ремонт одной машины.

Применение критерия ожидаемого значения в данном случае оправдано, если ПЭВМ работают в течение большого периода времени. При этом ожидаемые затраты на один интервал составят

$$ОЗ = \frac{C_1 \sum_{t=1}^{T-1} M(n_t) + C_2 n}{T},$$

где $M(n_t)$ – математическое ожидание числа вышедших из строя ПЭВМ в момент t. Так как n_t имеет биномиальное распределение с параметрами (n, p_t), то $M(n_t) = np_t$. Таким образом,

$$ОЗ = \frac{n(C_1 \sum_{t=1}^{T-1} p_t + C_2)}{T}.$$

Необходимые условия оптимальности T^* имеют вид

$ОЗ\ (T^*-1) \geq ОЗ\ (T^*)$,

$ОЗ\ (T^*+1) \geq ОЗ\ (T^*)$.

Следовательно, начиная с малого значения T, вычисляют $ОЗ(T)$, пока не будут удовлетворены необходимые условия оптимальности.
Пусть $C_1 = 100$; $C_2 = 10$; $n = 50$. Значения p_t приведены в таблице 6.5.

Таблица 6.5

T	p_t	$\sum_{t=1}^{T-1} p_t$	$ОЗ(T)$
1	0.05	0	$\frac{50\ (100 \cdot 0 + 10)}{1} = 500$
2	0.07	0.05	375
3	0.10	0.12	366.7
4	0.13	0.22	400
5	0.18	0.35	450

$$T^* \rightarrow 3\ ,\quad ОЗ(T^*) \rightarrow 366.7$$

Следовательно, профилактический ремонт необходимо делать через $T^*=3$ интервала времени, так как этому интервалу соответствуют минимальные ожидаемые затраты 366.7.

6.7 Критерий «ожидаемое значение – дисперсия»

Критерий ожидаемого значения можно модифицировать так, что его можно будет применить и для редко повторяющихся ситуаций.
Если x – случайная величина с дисперсией DX, то среднее арифметическое $\bar{x}$ имеет дисперсию DX/n, где n – число слагаемых в $\bar{x}$. Следовательно, если DX уменьшается, и вероятность того, что $\bar{x}$ близко к MX, увеличивается.

Следовательно, целесообразно ввести критерий, в котором максимизация ожидаемого значения прибыли сочетается с минимизацией её дисперсии.

Пример 2. Применим критерий "ожидаемое значение – дисперсия" для примера 1. Для этого необходимо найти дисперсию затрат за один интервал времени, т.е. дисперсию

$$з_T = \frac{C_1\sum_{t=1}^{T-1} n_t + C_2 n}{T}.$$

Так как n_t, $t = \overline{1, T-1}$ – случайная величина, то $з_T$ также случайная величина. Параметр n_t имеет биномиальное распределение с $M(n_t) = np_t$ и $D(n_t) = np_t(1–p_t)$. Следовательно,

$$D(з_T) = D\left(\frac{C_1\sum_{t=1}^{T-1} n_t + C_2 n}{T}\right) = \left(\frac{C_1}{T}\right)^2 D\left(\sum_{t=1}^{T-1} n_t\right) =$$

$$= \left(\frac{C_1}{T}\right)^2 \sum_{t=1}^{T-1} Dn_t = \left(\frac{C_1}{T}\right)^2 \sum_{t=1}^{T-1} np_t(1-p_t) = n\left(\frac{C_1}{T}\right)^2 \left\{\sum_{t=1}^{T-1} p_t - \sum_{t=1}^{T-1} p_t^2\right\},$$

где $C_2 n = const$.

Из примера 1 следует, что

$M(з_T) = M(з(T))$.

Следовательно, искомым критерием будет минимум выражения

$M(з(T)) + к\, D(з_T)$.

Константу "*к*" можно рассматривать как уровень несклонности к риску, так как константа "*к*" определяет "степень возможности" дисперсии $Д(з_T)$ по отношению к математическому ожиданию. Например, если предприниматель особенно остро реагирует на большие отрицательные отклонения прибыли вниз от $M(з(T))$, то он может выбрать "*к*" много больше единицы. Это придаёт больший вес дисперсии и приводит к решению, уменьшающему вероятность больших потерь прибыли.

При $к = 1$ получаем задачу

$$M(З(T)) + D(З(T)) = n\left\{\left(\frac{C_1}{T} + \frac{C_1^2}{T^2}\right)\sum_{t=1}^{T-1} p_t - \left(\frac{C_1}{T}\right)^2 \sum_{t=1}^{T-1} p_t^2 + \frac{C^2}{T}\right\}$$

По данным примера 1 можно составить таблицу 6.6.

Таблица 6.6

T	p_t	p_t^2	$\sum_{t=1}^{T-1} p_t$	$\sum_{t=1}^{T-1} p_t^2$	$M(з(T))+D(З(T))$
1	0.05	0.0025	0	0	500.00
2	0.07	0.0049	0.05	0.0025	6312.50
3	0.10	0.0100	0.12	0.0074	6622.22
4	0.13	0.0169	0.22	0.0174	6731.25
5	0.18	0.0324	0.35	0.0343	6764.00

Из таблицы видно, что профилактический ремонт необходимо делать в течение каждого интервала $T^*=1$.

6.8 Принятие решений в условиях неопределённости

Данные, необходимые для принятия решения в условиях неопределенности, обычно задаются в форме матрицы, строки которой соответствуют возможным действиям, а столбцы – возможным состояниям системы.

Пусть, например, из некоторого материала требуется изготовить изделие, долговечность которого при допустимых затратах невозможно определить. Нагрузки считаются известными. Требуется решить, какие размеры должно иметь изделие из данного материала.

Варианты решения таковы:

E_1 – выбор размеров из соображений максимальной долговечности;

E_m– выбор размеров из соображений минимальной долговечности;

E_i – промежуточные решения.

Условия требующие рассмотрения таковы :

F_1 – условия, обеспечивающие максимальной долговечность;

F_n – условия, обеспечивающие минимум долговечности;

F_i – промежуточные условия.

Под результатом решения $e_{ij} = e(E_i\ ;\ F_j\ )$ здесь можно понимать оценку, соответствующую варианту E_i и условиям F_j и характеризующую прибыль, полезность или надёжность. Обычно мы будем называть такой результат полезностью решения. Тогда семейство (матрица) решений $\|e_{ij}\|$ имеет вид, приведенный на рисунке 6.5.

	F_1	F_2	. . .	F_n
E_1	e_{11}	e_{12}	. . .	e_{1n}
E_2	e_{21}	e_{22}	. . .	e_{2n}
. . .	. . . .			
E_m	e_{m1}	e_{m2}	. . .	e_{mn}

Рисунок 6.5

Чтобы прийти к однозначному и по возможности наивыгоднейшему варианту решению необходимо ввести оценочную (целевую) функцию. При этом матрица решений $\|e_{ij}\|$ сводится к одному столбцу.

Таким образом, каждому варианту E_i приписывается некоторый результат e_{ir}, характеризующий, последствия этого решения.

6.8.1 Минимаксный критерий

Правило выбора решения в соответствии с минимаксным критерием (ММ-критерием) можно интерпретировать следующим образом:
матрица решений дополняется ещё одним столбцом из наименьших результатов e_{ir} *каждой строки. Необходимо выбрать те варианты, в строках которых стоят наибольшие значения* e_{ir} *этого столбца.*

Выбранные варианты полностью исключают риск. Это означает, что принимающий решение не может столкнуться с худшим результатом, чем тот, на который он ориентируется.

Применение ММ-критерия бывает оправдано, если ситуация, в которой принимается решение следующая:

- о возможности появления внешних состояний F_j ничего не известно;
- приходится считаться с появлением различных внешних состояний F_j;
- решение реализуется только один раз;
- необходимо исключить какой бы то ни было риск.

6.8.2 Критерий Байеса – Лапласа

Обозначим через q_i вероятность появления внешнего состояния F_j. Соответствующее правило выбора можно интерпретировать следующим образом:
матрица решений $\|e_{ij}\|$ *дополняется ещё одним столбцом, содержащим математическое ожидание значений каждой из строк. Выбираются те варианты, в строках которых стоит наибольшее значение* e_{ir} *этого столбца.*

При этом предполагается, что ситуация, в которой принимается решение, характеризуется следующими обстоятельствами:

- вероятности появления состояния F_j известны и не зависят от времени;
- решение реализуется (теоретически) бесконечно много раз;
- для малого числа реализаций решения допускается некоторый риск.

При достаточно большом количестве реализаций среднее значение постепенно стабилизируется. Поэтому при полной (бесконечной) реализации какой-либо риск практически исключён.

Таким образом, критерий Байеса-Лапласа (B-L-критерий) более оптимистичен, чем минимаксный критерий, однако он предполагает большую информированность и достаточно длительную реализацию.

6.8.3. Критерий Сэвиджа

$$a_{ij} := \max_i e_{ij} - e_{ij}$$

$$e_{ir} := \max_i a_{ij} = \max_j (\max_i e_{ij} - e_{ij})$$

Величину a_{ij} можно трактовать как максимальный дополнительный выигрыш, который достигается, если в состоянии F_j вместо варианта E_i выбирать другой, оптимальный для этого внешнего состояния вариант. Величину a_{ij} можно интерпретировать и как потери (штрафы), возникающие в состоянии F_j при замене оптимального для него варианта на вариант E_i. В последнем случае e_{ir} представляет собой максимально возможные (по всем внешним состояниям F_j, $j = \overline{1, n}$) потери в случае выбора варианта E_i.

Соответствующее критерию Сэвиджа правило выбора трактуется так: *каждый элемент матрицы решений* $\| e_{ij} \|$ *вычитается из наибольшего результата max* e_{ij} *соответствующего столбца.*

Разности a_{ij} *образуют матрицу остатков* $\| e_{ij} \|$. *Эта матрица пополняется столбцом наибольших разностей* e_{ir}. *Выбирают те варианты, в строках которых стоит наименьшее для этого столбца значение.*

Требования, предъявляемые к ситуации, в которой принимается решение, совпадают с требованием к ММ-критерию.

6.8.4 Пример и выводы

Из требований, предъявляемых к рассмотренным критериям, становится ясно, что вследствие их жёстких исходных позиций они применимы только для идеализированных практических решений. В случае, когда возможна слишком сильная идеализация, можно применять одновременно поочерёдно различные критерии. После этого среди нескольких вариантов ЛПР выбирает окончательное решение. Такой подход позволяет, во-первых, лучше проникнуть во все внутренние связи проблемы принятия решений и, во-вторых, ослабляет влияние субъективного фактора.

Пример. При работе ЭВМ необходимо периодически приостанавливать обработку информации и проверять ЭВМ на наличие в ней вирусов. Приостановка в обработке информации приводит к определённым экономическим издержкам. В случае же, если вирус не будет вовремя обнаружен, возможна потеря и некоторой части информации, что приведёт и ещё к большим убыткам.

Варианты решения таковы:

E_1– полная проверка;

E_2– минимальная проверка;

E_3– отказ от проверки.

ЭВМ может находиться в следующих состояниях:

F_1– вирус отсутствует;

F_2– вирус есть, но он не успел повредить информацию;

F_3– есть файлы, нуждающиеся в восстановлении.

Результаты, включающие затраты на поиск вируса и его ликвидацию, а также затраты, связанные с восстановлением информации, приведены в таблице 6.7.

Таблица 6.7

	F_1	F_2	F_3	ММ-критерий		критерий B-L	
				$e_{ir}=\min\limits_j e_i$	$\max\limits_i e_{ir}$	$e_{ir}=\sum\limits_j e_{ij}$	$\max\limits_i e_{ir}$
E_1	-20.0	-22.0	-25.0	-25.0	$\underline{\underline{-25.0}}$	-22.33	
E_2	-14.0	-23.0	-31.0	-31.0		-22.67	
E_3	0	-24.0	-40.0	-40.0		-21.33	$\underline{\underline{-21.33}}$

Согласно ММ-критерию, следует проводить полную проверку. По критерию Байеса-Лапласа, в предположении, что все состояния машины равновероятны: $P(F_j) = q_j = 0.33$, рекомендуется отказаться от проверки.

В соответствии с критерием Сэвиджа матрица остатков для этого примера приведена в таблице 6.8.

Таблица 6.8

	F_1	F_2	F_3	Критерий Сэвиджа	
				$e_{ir}=\min\limits_j a$	$\min\limits_j e_{ir}$
E_1	+20.0	0	0	+20.0	
E_2	+14.0	+1.0	+6.0	+14.0	$\underline{\underline{+14.0}}$
E_3	0	+2.0	+15.0	+15.0	

Пример специально подобран так, что каждый критерий предлагает новое решение. Неопределённость состояния, в котором проверка застаёт ЭВМ, превращается в неясность, какому критерию следовать, поскольку различные критерии связаны с различными условиями, в которых принимается решение. В частности, если принимаемое решение относится к сотням машин с одинаковыми параметрами, то рекомендуется применять критерий Байеса-Лапласа. Если же число машин невелико, лучше пользоваться критериями минимакса или Севиджа.

6.9 Производные критерии

Стараясь занять наиболее уравновешенную позицию, Гурвиц предположил оценочную функцию, которая находится где-то между точкой зрения крайнего оптимизма и крайнего пессимизма:

$\max\limits_{i} e_{ir} = \{C \min\limits_{j} e_{ij} + (1\text{- } C) \max\limits_{j} e_{ij} \}$,

где *C*– весовой множитель.

Правило выбора согласно критерию Гурвица формируется следующим образом:

матрица решений $\| e_{ij} \|$ *дополняется столбцом, содержащим среднее взвешенное наименьшего и наибольшего результатов для каждой строки. Выбираются только те варианты, в строках которых стоят наибольшие элементы* e_{ir} *этого столбца.*

При *C*=1 критерий Гурвица превращается в ММ-критерий. При *C* = 0 он превращается в критерий «азартного игрока» $\max\limits_{i} e_{ir} = \max\limits_{i} \max\limits_{j} e_{ij}$, т.е. мы становимся на точку зрения азартного игрока, делающего ставку на то, что «выпадет» выгоднейший случай.

В технических приложениях сложно выбрать весовой параметр *C*, так как трудно найти количественную характеристику для тех долей оптимизма и пессимизма, которые присутствуют при принятии решения. Поэтому чаще используют параметр *C,* равный 0,5.

Критерий Гурвица применяется в следующих случаях:

- о вероятностях появления состояния F_j ничего не известно;
- с появлением состояния F_j необходимо считаться;
- реализуется только малое количество решений;
- допускается некоторый риск.

6.10 Статические игры с полной информацией

Задача дилемма заключённого. Два заключённых подозреваются как сообщники в совершении некоторого преступления. Они помещены в разные камеры и не имеют никакой возможности обмениваться информацией. Каждому по отдельности предлагается сознаться С) к определенному сроку, но можно и молчать (М). Если один сознался, другой молчит, то сознавшегося освобождают, а молчун получает максимальный срок (9 лет). Если сознаются оба, то обоим срок снижается до 6 лет. Если оба молчат, то вину по основному преступлению доказать невозможно, и они получат по 1 году за незаконное владение оружием (рисунок 6.6).

	M	C
M	-1 . -1	-9, 0
C	0, -9	-6, -6

Рисунок 6.6

В дилемме заключенного у каждого игрока есть строго доминирующая стратегия. Сознаться лучше, чем молчать, что бы ни делал другой игрок:
0>-1;   -6>-9.

Казалось бы исход игры с выбором стратегий очевиден (С,С). Это и есть равновесие Нэша. Однако вектор (-1;-1) превосходит (-6,-6).

Проблема противоречия

Проблема противоречия индивидуального и коллективного выбора состоит в том, что каждый может навязать выгодное для себя равновесие.

Парето-оптимальные решения обладают тем свойством, что улучшать значения целевой функции одного игрока можно только за счёт других игроков.

Устойчивость приобретает важное значение. Другой принцип принятия решений, связан с понятием устойчивости. При выборе устойчивого решения говорят, что достигнута ситуация равновесия.

Неустойчивость какой-либо ситуации проявляется в том, что в случае её возникновения ей грозит распад, обусловленный возможностями одного из игроков путём изменения только своей стратегии улучшить свое положение за счёт других. На этом основании возник принцип устойчивости Нэша (американский математик Джон Нэш).

Он гласит, что выбор рациональной стратегии должен производиться среди множества точек равновесия. Равновесные решения называются оптимальными по Нэшу. В дилемме заключённого равновесию Неша соответствует стратегия (С, С).

6.11 Приближенные методы решения игр

При исследовании игровых ситуаций часто нет необходимости в точном решении игры. Для этих целей может быть применен численный метод итераций.

Идея метода сводится к следующему. Мысленно игру проигрывают много раз и вычисляют математическое ожидание обоих выигрышей стратегов, и их среднее арифметическое принимают за цену игры.
К недостаткам итерационного алгоритма можно отнести сравнительно медленную сходимость.

Пример. Решить игру методом Брауна (метод итераций), Выполнить 3 итерации. Игра задана платёжной матрицей, приведённой в таблице 6.9

Таблица 6.9

	B1	B2	B3
A1	8	2	4
A2	4	5	6
A3	1	7	3

n – номер исследуемой пары ходов;
I – номер выбранной стратегии игрока А;
B1- выигрыш накопленный за первые n игр игроком В при стратегии B1;
B2 -выигрыш накопленный за первые n игр игроком В при стратегии B2;
B3- выигрыш накопленный за первые n игр игроком В при стратегии B3;
j – номер стратегии, выбранной противником;
A1- накопленный выигрыш за n игр при стратегии A1;
A2- накопленный выигрыш за n игр при стратегии A2;
A3- накопленный выигрыш за n игр при стратегии A3;
Rmin –минимальный средний выигрыш, который равен минимальному накопленному выигрышу, деленному на число игр;
Rmax –максимальный средний выигрыш;
Rcp=(Rmin+Rmax)/2.
На второй итерации минимальный накопленный выигрыш равен 7 (подчёркивание снизу в таблице 6.10), а максимальный накопленный

выигрыш равен 12. Разделив на число игр получаются следующие значения: Rmin=7/2=3,5; Rmax=12/2=6; Rср=(3,5+6)/2=4,75.

Этот метод даёт возможность найти приближённое значение оптимальных смешанных стратегий обоих игроков. Для этого надо подсчитать частоту применения каждой чистой стратегии (i, j) и принять её за приближенное значение вероятности использования этой чистой стратегии в оптимальной смешанной стратегии соответствующего игрока.

Таблица 6.10

n	i	B1	B2	B3	j	A1	A2	A3	Rmin	Rmax	Rср
1	3	1	7	3	1	8	4	1	1	8	4,5
2	1	9	9	7	3	12	10	4	3,5	6	4,75
3	1	17	11	11	2	14	15	11	3,67	5,0	4,33
4	2	21	16	17	2	16	20	18	4,0	5,0	4,5
5	2	25	21	23	2	18	25	25	4,2	5,0	4,6

При увеличении числа n все три величины Rср, Rmin, Rmax будут приближаться к цене игры, причём Rср наиболее быстро.

Упражнения

Задача 1. Решить игру 2 × 3, заданную матрицей, используя аналитический и графический методы. Указать доминирующие стратегии и упростить игру.

	B1	B2	B3
A1	2	4	1
A2	5	1	4

Задача 2. Решить игру 3 × 3, заданную матрицей, используя аналитический и графический методы. Указать доминирующие стратегии и упростить игру.

	B1	B2	B3
A1	7	3	5
A2	4	5	6
A3	2	7	4

Задача 3. Выбрать рациональный вариант профилактики ЭВМ. Затраты на поиск вируса и его ликвидацию, а также затраты, связанные с восстановлением информации, приведены ниже.

	F1	F2	F3
E1	-23	-21	-26
E2	-12	-20	-24
E3	-5	-21	-45

Литература

1. Лапа В.Г. Математические основы кибернетики/ В.Г. Лапа.-Киев: Вища шк.,1974,- 450с.
2. Фомин Г.П. Математические методы и модели в коммерческой деятельности: Учеб./Г.П. Фомин.- М.: Финансы и статистика, 2001.- 544с.

7 Задача многокритериального выбора

Многокритериальные задачи выбора вариантов из заданного конечного множества $X = (x_1, \ldots, x_2)$.

Предполагается, что каждый из вариантов x_i оценивается по m частным критериям F_k. Частные критерии могут иметь как числовые, так и порядковые шкалы. Примером числового критерия в задаче выбора автомобиля из заданного множества (например, при покупке подержанного автомобиля) может служить пробег автомобиля, выраженный в километрах — чем он меньше, тем лучше. Цвет автомобиля естественно оценивать в порядковой шкале, например, вида: черный—белый—коричневый—...—желтый. *Здесь предполагается, что степень предпочтительности убывает слева направо.*

Для решения задачи необходимо принять следующие допущения:

• *множество исходных вариантов конечно, и ЛПР (лицо, принимающее решение, пользователь) может перечислить элементы этого множества;*

• *ЛПР может определить цель по каждому критерию, а также задать некоторые параметры критериев, описанные ниже;*

• *ЛПР может ранжировать критерии по важности и указывать равноценные критерии.*

Кроме того, предполагается, что ЛПР может принимать или отвергать предлагаемые ему альтернативы, а также указывать не устраивающие его по некоторым критериям оценки отвергаемой альтернативы.

В такой системе должен поддерживаться непрерывный диалог с пользователем, в процессе которого у ЛПР запрашивается вся необходимая информация о задаче и создается отчет о ней.

7.1 Решение задачи многокритериального выбора

Решением многокритериальной задачи, является соответствующее множество Парето — ***множество недоминируемых по Парето альтернатив***. Это множество может оказаться достаточно обширным, а пользователя соответствующей системы принятия решений обычно интересует выбор какого-то одного "наилучшего" варианта или небольшого их числа. Если какая-либо дополнительная информация о задаче отсутствует, то множество Парето — это лучшее, что можно предложить. Однако при наличии дополнительной информации о системе предпочтений пользователя могут быть применены различные методы сужения исходного множества альтернатив — более сильные, чем методы, основанные на доминировании по Парето. Весьма часто исходной информацией для таких методов выступает само множество Парето и ставится задача его сужения с целью выбора одной или нескольких альтернатив в качестве окончательного результата. Далее рассмотрен подход к решению этой проблемы.

7.2 Метод ограничений

Простым и часто применяемым методом сжатия множества Парето является метод ограничений. Решается стандартная многокритериальная задача

$$\begin{gathered} f_i(x) \to \max, \quad x \in D, \\ f_i : D \to R, \quad i = 1, \dots, m, \end{gathered} \tag{7.1}$$

где D — произвольное абстрактное множество.

На первом этапе одним из известных методов строится множество Парето $P(D)$. Далее метод ограничений реализуется в соответствии со следующей последовательностью шагов.

Шаг 1. Пользователю предлагается назначить нижние допустимые границы t_i для всех m критериальных функций:

$$f_i(x) \geq t_i, \quad i = 1, \ldots, m \qquad (7.2)$$

Предполагается, что указанные значения критериальных функций дают удовлетворяющие пользователя варианты.

Шаг 2. Строится подмножество множества Парето, состоящее из точек, удовлетворяющих неравенствам:

$$D_i = \{x \in P(D) \mid f_i(x) \geq t_i, i = 1, \ldots, m\}, D_i \subset P(D) \qquad (7.3)$$

Шаг 3. Если D_i — пустое множество, то пользователю предлагается ослабить требования с помощью уменьшения какого-то из чисел t_i. Далее переходим к *шагу 2*. Если D_i не пусто — переходим к *шагу 4*.

Шаг 4. Выбирается $\forall x \in D_i$ и предъявляется пользователю в качестве кандидата на "решение" задачи. Если решение удовлетворяет пользователя, то процесс завершается. В противном случае переходим к *шагу 5*.

Шаг 5. Пользователю предлагается назначить новую (увеличенную) нижнюю границу по одному из критериев и осуществляется переход к *шагу 2*.

Метод ограничений целесообразно использовать на завершающей стадии процесса выбора, например, после применения метода *t*-упорядочения.

Тогда в качестве исходного для метода ограничений множества будет использоваться не *P(D),* а некоторое его подмножество, что сократит наиболее трудоемкую диалоговую часть процедуры выбора.

Непосредственно в методе ограничений какая-либо ординальная информация не используется, а сокращение исходного множества альтернатив производится в процессе поступления дополнительной информации от пользователя в виде последовательности наборов нижних границ $\{t_i\}$.

7.3 Выбор на основе метода t-упорядочения

Пусть решается детерминистская многокритериальная задача

$$f_i(x) \to \max, \; f_i : X \to R, \quad i = \overline{1,m} \qquad (7.4)$$

где X— произвольное абстрактное множество. Будем предполагать, что все критериальные функции f_i отражают "полезность" объекта с позиций различных критериев и являются соизмеримыми в том смысле, что значения каждой критериальной функции изменяются в одних и тех же пределах [a, b]:

$$\forall x \in X : 0 \le a \le fi(x) \le b \quad i = \overline{1,m} \qquad (7.5)$$

Таким образом, мы предполагаем, что оценочные шкалы критериев являются числовыми и одинаковыми. Отметим, что предположение о соизмеримости критериев является существенным и требует для каждой решаемой задачи отдельного обоснования и исследования. Требование, связанное с необходимостью приведения всех числовых шкал к единому промежутку, выглядит весьма невинно и формально достигается с помощью, например, следующих простых преобразований:

$$f_i(x) = a + (b-1)\frac{f_i(x) - \min f_i}{\max f_i - \min f_i} \qquad (7.6)$$

Здесь maxf_i minf_i— максимальные и минимальные значения f_i соответственно.

Новые оценочные функции $\overline{f_i}$ будут изменяться уже в пределах заданного промежутка [a, b]. При этом наименее предпочтительный по любому из частных критериев вариант получит оценку *a,* а наиболее предпочтительный — оценку *b.* Часто полагают *a = 0, b = 1.* Могут использоваться и другие (может быть, нелинейные) формулы нормировки.

Итак, пусть критериальные функции (или просто — критерии) соизмеримы и удовлетворяют условиям (7.5). В качестве примера изначально соизмеримых критериев можно привести систему школьных оценок по нескольким предметам.

Определение 1

Нормированные критерии f_i , f_j называются ***равноценными*** (что записывается в виде $f_i = f_j$), если всякие две векторные оценки Z, *W,* где

$$Z = (z_1;\ldots;z_i;\ldots,z_j;\ldots;z_m),\quad Z = f(x), x \in X$$

$$W = (z_1;\ldots;z_i+\delta;\ldots,z_j-\delta;\ldots;z_m) \qquad (7.7)$$

одинаковы по предпочтительности при любом δ (большем или меньшем нуля), удовлетворяющем неравенствам:

$$a \le z_i + \delta \le b,\; a \le z_i - \delta \le b$$

Легко видеть, что суммы частных оценок в позициях *i,j*, у векторных оценок *Z, W* совпадают.

Таким образом, если, например, два школьника оцениваются по четырем предметам и имеют оценки (которые необходимо максимизировать)

$$Z=(5,4,4,3),\; W=(5,5,3,3), \qquad (7.8)$$

то при условии равноценности критериев f_2, f_3 приведенные векторные оценки будут одинаковы по предпочтительности, т. к. 4 + 4 = 5 + 3, а остальные оценки совпадают.

Следовательно, если критерии f_i , f_j равноценны, то можно "забрать" δ единиц у частной оценки z_j в (7) и "передать" их частной оценке z_i. При этом получим векторную оценку, одинаковую с исходной по предпочтительности. Если в приведенном примере (8) считать, что оценка Z предпочтительнее, чем *W,* то естественно предположить, что критерий f_3 важнее критерия f_2. Дадим соответствующее определение.

Определение 2

Критерий f_i более важен, чем критерий f_j (что записывается в виде $f_i > f_j$), если векторная оценка

$$Z = (z_1;\ldots;z_i;\ldots,z_j;\ldots;z_m)$$

менее предпочтительна, чем оценка

$$W = (z_1;\ldots;z_i+\delta;\ldots,z_j-\delta;\ldots;z_m)$$

где

$$\delta \in \{\delta > 0 | z_i + \delta \le b, a \le z_j - \delta\}$$

Таким образом, перенос δ единиц (δ > 0) с частной оценки z_j на частную оценку z_i приводит к улучшению ситуации, если $f_i > f_j$. Определения 1, 2 показывают, как может интерпретироваться дополнительная ординальная (порядковая) информация пользователя (ЛПР) об относительной важности частных критериев, на основе которой и происходит сокращение множества Парето решаемой многокритериальной задачи. Приведенные в определениях 1, 2 интерпретации ординальной информации ЛПР позволяют строить отношения доминирования более сильные, чем отношение Парето.

Алгоритм, реализующий метод t-упорядочения.

В качестве исходной информации для алгоритма t-упорядочения принимается множество *S* высказываний пользователя (ЛПР) об относительной важности частных критериев вида:

$$S = \{ f_i = f_j ; \ldots ; f_q > f_p \}$$

Мы будем предполагать, что множество *S* скорректировано следующим образом. Во-первых, необходимо проверить и при необходимости обеспечить непротиворечивость высказываний из *S,* может быть, путем проведения дополнительного диалога с пользователем для уточнения его системы предпочтений. Во-вторых, необходимо расширить множество *S* за счет добавления новых высказываний, являющихся транзитивными следствиями уже имеющихся (выполнить операцию транзитивного замыкания).

Пусть теперь $Z = (z_i, ..., z_m)$; $W= (w_j, ..., w_m)$ — две векторные оценки, которые необходимо сравнить с учетом дополнительной скорректированной информации *S*.

Если эти оценки сравнимы по Парето, то задача решена. В противном случае вектор Z фиксируется, а по вектору *W* формируются следующие два множества (может быть, бесконечные, даже если исходные множества оценок конечны).

1. *WE*— множество *W*-эквивалентных векторов (включающее сам вектор *W)*, полученное из W с помощью всех возможных переносов δ между парами равноценных критериев. Следовательно, множество *WE* строится с учетом всех данных типа $f_i = f_j$ из *S*.

2. *WI*— множество *W*-улучшенных векторов, каждый из которых получен с помощью возможных переносов δ с учетом всех данных $f_i = f_j$, $f_q > f_p$

При этом предполагается, что переносы согласно информации $f_q > f_p$, производятся только с целью улучшения вновь полученного вектора и, по крайней мере, одна такая улучшающая передача выполнена для любого вектора из *WI*, Далее новое отношение предпочтения $\overset{t}{\succ}$ строится следующим образом: $$Z \overset{t}{\succ} W \leftrightarrow [\exists W' \in WE : Z^P \succ W'] \vee [\exists W'' \in WI : Z^P \underset{\sim}{\succ} W''] \quad (7.9)$$

Здесь через $\left(\underset{\sim}{\overset{p}{\succ}}\right)$ обозначено нестрогое предпочтение вида

$$Z \underset{\sim}{\overset{P}{\succ}} W \leftrightarrow \forall i \in [1:m]: z_i \geq w_i \qquad (7.10)$$

Определение (9) имеет весьма простой смысл. Если вектор Z строго лучше, чем некоторый вектор *W`*, эквивалентный *W*, или нестрого лучше, чем некоторый вектор *W``* который, в свою очередь, строго лучше, чем *W*, то полагаем, что Z строго лучше *W*.

7.4 Метод линейной свертки

Этот метод свертки , позволяет изменить векторный критерий ($f_1, f_2, \ldots, f_n$) на скалярный критерий и он основан на линейном объединении всех частных функционалов.

Теорема 2.2. Пусть $a \in A$. Тогда решение задачи

$$F_1(x,a) \overset{\Delta}{=} \sum_{i=1}^{m} a_i f_i(x) \to \max_{x \in D} \qquad (7.11)$$

есть эффективный вектор.

Доказательство. Пусть $x^0 \in D$ есть решение задачи (7.11) и существует такой $x' \in D$, что $f_i(x') \geq f_i(x^0)$, а для $i = i_0$ имеем $f_{i_0}(x') > f_{i_0}(x^0)$. Тогда

$$\sum_{i=1}^{m} a_i f_i(x) > \sum_{i=1}^{m} a_i f_i(x^0)$$

И, следовательно, $x^0 \in D$ не максимизирует функционал F_1. Полученное противоречие доказывает, что точки x' со сформулированными выше свойствами не существуют и поэтому x^0 — эффективный вектор. Теорема доказана.

Замечание *2.2*

Обратное утверждение без дополнительных предположений неверно. Существуют эффективные векторы, не являющиеся решениями задачи (7.11). Для доказательства этого утверждения достаточно привести хотя бы один опровергающий пример, что и будет сделано далее. Таким образом, согласно доказанной теореме

$$\bigcup_{a \in A} X(a) \subseteq P(D)$$

здесь

$$X(a) \overset{\Delta}{=} Arg \max_{x \in D} F_1(x, a)$$

Обратимся к геометрическим иллюстрациям для $m = 2$.*В* этом случае

$$F_1(x, a) = a_1 f_1(x) + a_2 f_2(x) = \Phi(f_1, f_2)$$

где функция Φ_1 считается определенной в пространстве критериев(*f1,f2)*. Построим линии уровня функции Φ_1.

$$a_1 f_1 + a_2 f_2 = const \qquad (7.12)$$

Набор коэффициентов a = ($a1,a2$) считается фиксированным (неизменным в процессе всего рассмотрения). Графики прямых (7.12) для различных констант в правой части и различных весовых коэффициентов показаны на рисунок 7.1.

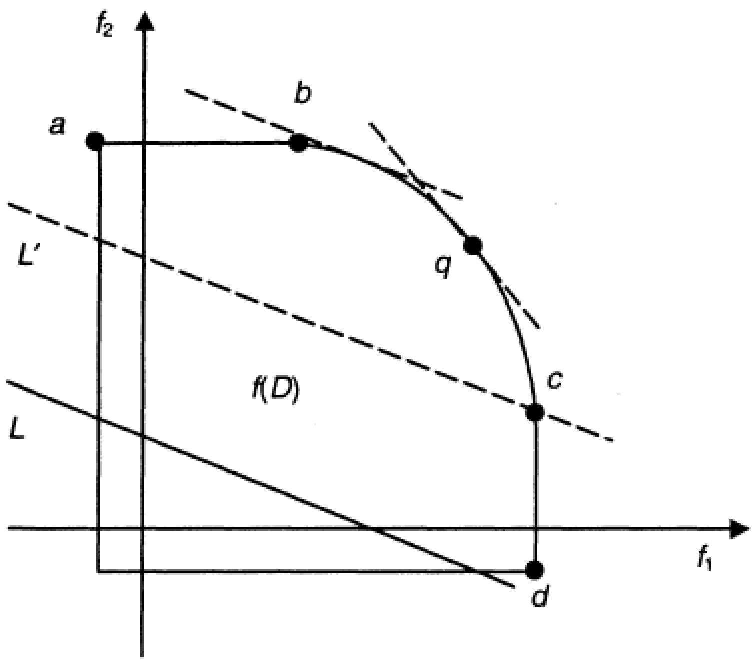


Рисунок 7.1 - Линии уровня функции Φ_1

Угловой коэффициент наклона прямой *L* определяется числами a_1, a_2 и равен ($-a_1/a_2$). При увеличении константы в правой части уравнения (7.12), прямая перемещается вверх параллельно *L* (занимая положение *L*\`). Таким образом, мы имеем целое семейство линий уровня, и максимум функции Φ_1, а вместе с ней и F_1, достигается в точках плоскости (*f1,f2),* соответствующих точкам касания (но не пересечения) самой "верхней" линии уровня и множества достижимости *f(D).* На рисунке 7.2 точка *b* с координатами (f_1^b, f_2^b)реализует найденную рассмотренным методом эффективную оценку. Легко видеть, что ни одна из точек интервалов [*a, b), (c, d*], соответствующих слабо эффективным, но не эффективным решениям, не может являться точкой касания *f(D)* и какой-либо линии уровня функции Φ_1 (угловой коэффициент ($-a_1/a_2$) не может равняться нулю или бесконечности, т. к. все $a_i > 0$ и их величина ограничена сверху единицей).

Перебирая различные *a*, можно получить достаточно точную аппроксимацию множеств *P(f)* и *P(D).*

Ситуация, связанная с существованием эффективных решений, не являющихся решениями задачи (7.11) ни при каких *a*, проиллюстрирована на рисунок 7.2.

Все точки дуги *a, b* являются эффективными оценками, но ни одна из них (кроме самих точек *a и b)* не может являться точкой касания линий уровня функции Φ_1 к множеству *f(D)* ни при каком наборе коэффициентов *a.*

Таким образом, ясно, что отсутствие выпуклости множества *f(D)* приводит к принципиальным затруднениям при применении метода линейной свертки. Аналогично, наличие строго прямолинейных участков "северо-восточной" границы множества *f(D)* может приводить к похожим трудностям из-за приближенного характера вычислений (точки внутри таких прямолинейных участков оказываются "неустойчивыми" решениями задачи (7.11).

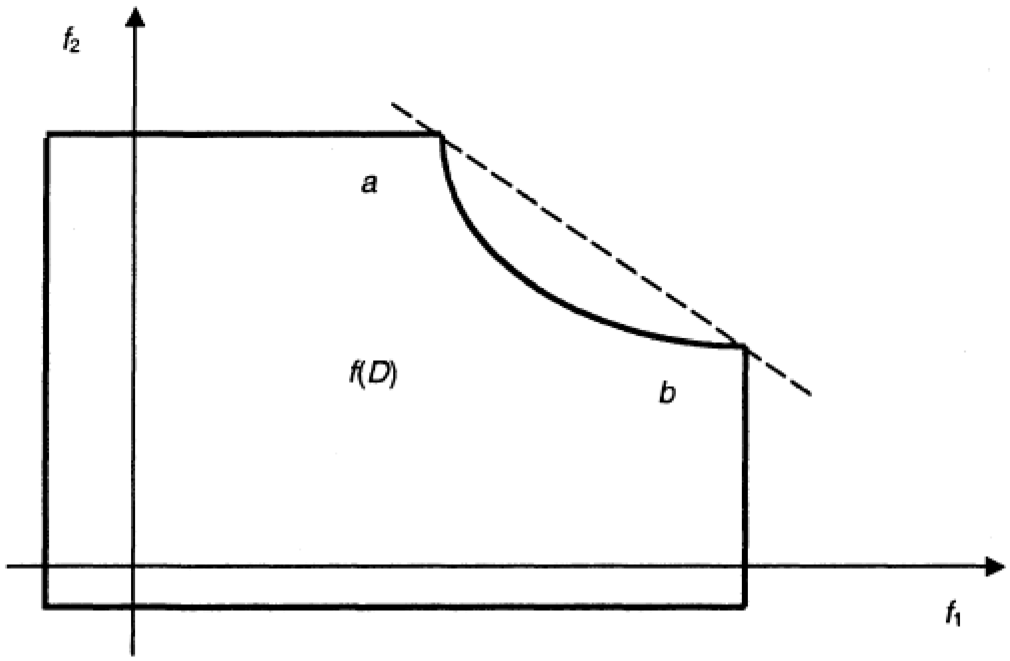


Рисунок 7.2 - Случай невыпуклой границы множества достижимости

7.5 Quick Choice – система многокритериальногго выбора вариантов

Система Quick Choice (QCH) предназначена для решения многокритериальных задач выбора вариантов из заданного конечного множества$X = (x_1, \ldots, x_n)$ [1].

Предполагается, что каждый из вариантов x_1 оценивается по m частным критериям F_k. Частные критерии могут иметь как числовые, так и порядковые шкалы. Примером числового критерия в задаче выбора автомобиля из заданного множества (например, при покупке подержанного автомобиля) может служить пробег автомобиля, выраженный в километрах – чем он меньше, тем лучше. Цвет автомобиля естественно оценивать в порядковой шкале, например, вида: черный – белый – коричневый - … - желтый. Здесь предполагается, что степень предпочтительности убывает слева направо.

В данной системе приняты следующие предположения об исходной задаче:

- множество исходных вариантов конечно, и ЛПР (лицо, принимающее решение, пользователь) может перечислить элементы этого множества;
- ЛПР может определить цель по каждому критерию, а также задать некоторые параметры критериев, описанные ниже;
- ЛПР может ранжировать критерии по важности и указывать равноценные критерии.

Кроме того, предполагается, что ЛПР может принимать или отвергать предлагаемые ему альтернативы, а также указывать не устраивающие его по некоторым критериям оценки отвергаемой альтернативы.

Система QCH поддерживает непрерывный диалог с пользователем, в процессе которого у ЛПР запрашивается вся необходимая информация о задаче и создается отчет о ней. Для запуска этого диалога нужно в меню «Запуск» выбрать пункт «Диалог».

Исходные данные

Для работы с системой необходимо задать следующие исходные данные об альтернативах и критериях:

- количество критериев – число m;
- тип каждого критерия (порядковый или числовой);
- цель по каждому критерию (максимум или минимум);
- все принимаемые значения для перечислимых критериев;
- набор ординальной (порядковой) информации вида «критерий P важнее критерия Q»;
- набор ординальной информации вида «критерий P равноценен критерию Q»;
- количество альтернатив (вариантов), среди которых производится выбор – число n;
- весь выбор альтернатив;
- значения всех частных критериев для каждой альтернативы;
- минимальные требования к выбираемой альтернативе по каждому из частных критериев.

Типы критериев

Для каждого критерия необходимо задать некоторые свойства. Эти свойства описываются далее. Свойство «цель» описывает намерения ЛПР по данному критерию: минимум или максимум. Если целью данному критерию является максимум, то система предпочитает альтернативы с большим значением данного критерия, если же минимум, то наоборот.

Структура задания типа критерия изображена на рисунке 7.3. Для непрерывного критерия ЛПР должен задать максимальное и минимальное значения. Непрерывный критерий принимает значения на заданном промежутке. Перечислимый критерий принимает значения из фиксированного набора. Значения перечислимых критериев могут задаваться численно, например:

- отлично – 5;
- хорошо – 4;
- удовлетворительно – 3;
- плохо – 2.

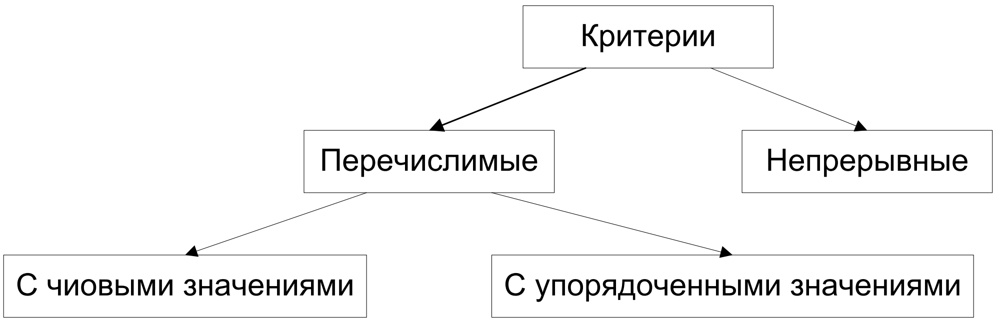


Рисунок 7.3 - Типы критериев

В таком случае пользователь задает текстовые значения и в системе они заменяются на численные.

Система не требует от ЛПР задания численных значений, если задать тип критерия как «Перечислимый с упорядоченными значениями». В этом случае имеется в виду порядковая шкала, и достаточно задать лишь набор значений данного критерия, а также упорядочить их в порядке предпочтения.

Функции, реализованные в системе

Описываемая система позволяет выполнять следующие функции:

- задание информации о задаче в диалоге с пользователем;

- построение множества оптимальных вариантов в соответствии с методом t-упорядочения;
- реализация метода ограничений, позволяющего пользователю выбрать одну альтернативу из t- оптимального множества;
- импорт из базы данных;
- импорт данных из текстового файла;
- проверка корректности ординальной информации;
- создание отчета.

Обобщённая структура системы приведена на рисунке 7.4. Исходная информация о задаче задается в диалоге с пользователем. После завершения диалога система имеет информацию об альтернативах, критериях, и ординальную информацию о критериях. В системе также предусмотрено получение данных об альтернативах из текстового файла или баз данных dBase и paradox. После этого выполняется нормализация исходных данных и запуск метода t-упорядочения.

Выбор результирующей альтернативы осуществляется на основе метода ограничений.

Для более эффективного использования метода ограничений выполняется предварительное ранжирование альтернатив, по результатам которого определяется порядок предложения альтернатив на рассмотрение пользователя. В качестве метода ранжирования альтернатив применяется метод аддитивной свертки. При этом учитывается ординарная информация о критериях, полученная от пользователя.

При работе метода ограничений помимо текущего варианта пользователю предлагается «идеальная» альтернатива (с наилучшими возможными оценками по всем критериям) и установленные ограничения. Эти данные предоставляются пользователю как в численном виде, так и в виде диаграммы. При таком представлении пользователь может наилучшим образом оценить достоинства и недостатки текущей оцениваемой альтернативы.

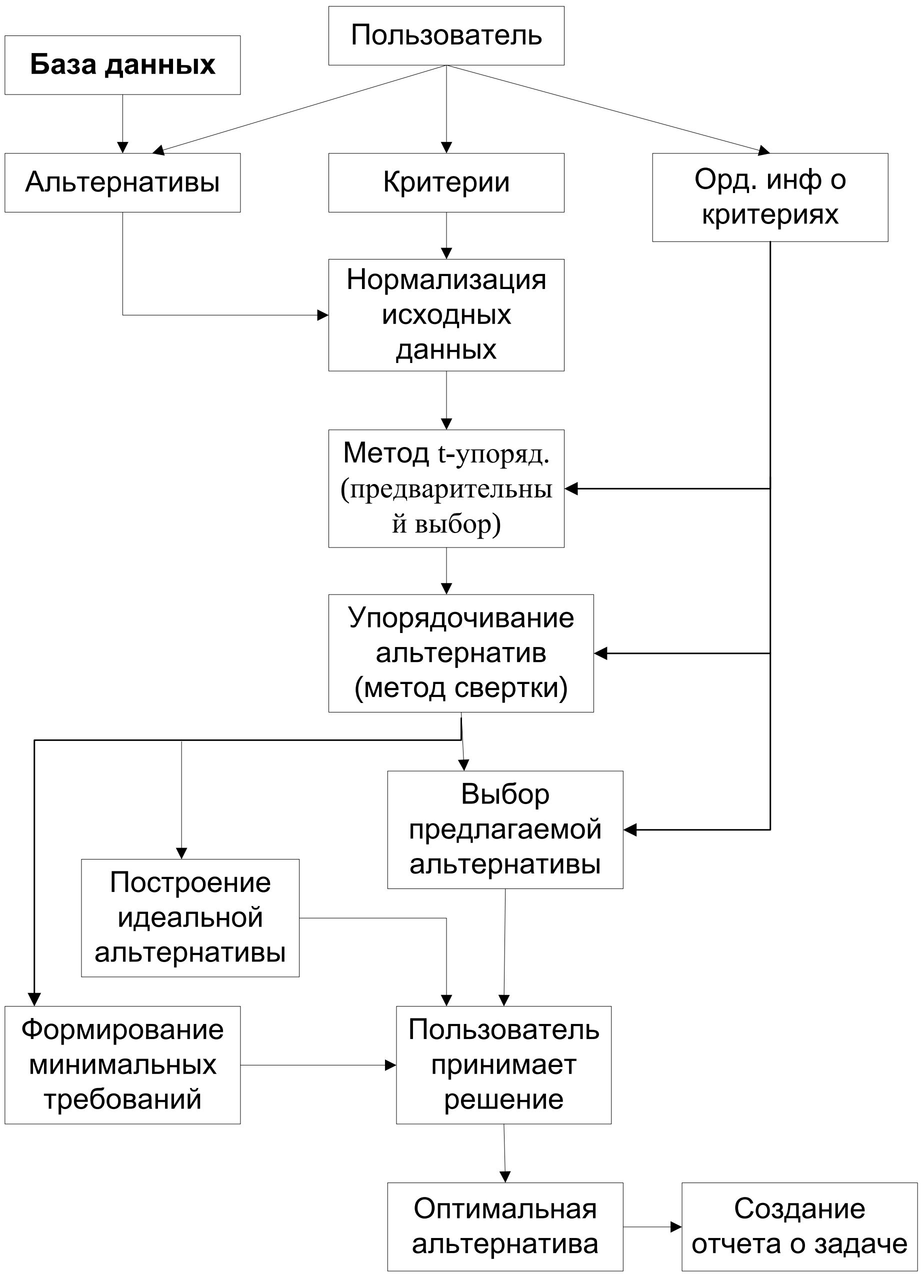


Рисунок 7.4- Обобщённая структура системы

После получения решения система формирует отчёт о задаче.

Принятие решений в диалоге с пользователем

В данной системе реализован диалог с пользователем, позволяющий пользователю описать интересующую его задачу и получить решение, отвечая на вопросы, задаваемые системой. Такой подход может расширить круг пользователей системы, потому что от них не требуется специальных знаний в области принятия решений. Процедура принятия решения состоит из пяти этапов:

1. Задание критериев.
2. Задание альтернатив.
3. Задание дополнительной информации о критериях.
4. Предварительный выбор альтернатив (метод t-упорядочения).
5. Реализация метода ограничений.

На первых двух этапах пользователь непосредственно задает данные о задаче, описывая множества критериев и альтернатив. Далее в диалоге с пользователем определяются предпочтения ЛПР. Исходя из заданных предпочтений, система отбирает наиболее предпочтительные варианты. На пятом этапе система предлагает пользователю выбрать одну альтернативу в соответствии с методом ограничений.

Метод ограничений

Для выбора одной альтернативы из выбранного множества альтернатив в данной системе используется метод ограничений. Суть этого метода заключается в том, что ЛПР предоставляется один из вариантов, который он сможет принять или отвергнуть. В случае если ЛПР не устраивает данный вариант, ему предоставляется возможность задать минимальные требования по какому-нибудь критерию с целью предложить ему более приемлемую альтернативу. Если же в исходном множестве не имеется альтернатив, удовлетворяющих требованию ЛПР, то система попросит ослабить требования по какому-нибудь критерию.

7.6 Применение теории генетических алгоритмов для поиска многоэкстремальных значений

7.6.1 Основные понятия теории генетических алгоритмов

Для решения многоэкстремальных задач могут быть использованы генетические алгоритмы (ГА).

Генетические алгоритмы представляют собой алгоритмы поиска, построенные на принципах, сходных с принципами естественного отбора и генетики. Если говорить обобщенно, они объединяют в себе принцип выживания наиболее перспективных особей – решений и структуризированный обмен информацией, в котором присутствует элемент случайности, который моделирует природные процессы наследования и мутации[2,3]. Дополнительным свойством этих алгоритмов является невмешательство человека в развивающийся процесс поиска.

Генетические алгоритмы имеют целью нахождение лучшего, а не оптимального решения задачи. Это связанно с тем, что для сложной системы часто требуется найти удовлетворительное решение, а проблема достижения оптимума отходит на второй план. При этом другие методы, ориентированные на поиск именно оптимального решения, вследствие чрезвычайной сложности задачи становятся вообще неприменимыми. В этом кроется причина появления, развития и роста популярности генетических алгоритмов. Хотя, как и всякий другой метод поиска, этот подход не является оптимальным методом решения любых задач.

Преимущества генетических алгоритмов становятся еще более прозрачным, если рассмотреть основные их отличия от традиционных методов. Основных отличий четыре.

1. Генетические алгоритмы работают с кодами, в которых представлен набор параметров, напрямую зависящих от аргументов целевой функции. Причем интерпретация этих кадров происходит только перед началом работы алгоритма и после завершения его работы для получения результата. В процессе работы манипуляции с кодами происходят совершенно независимо от их интерпретации. Код рассматривается просто как битовая строка .

2. Для поиска генетический алгоритм использует несколько точек поискового пространства одновременно, а не переходит от точки к точке, как это делается в традиционных методах. Это позволяет преодолеть один из их недостатков – опасность попадания в локальный экстремум целевой функции, если она не является унимодальной, т.е. имеет несколько таких

экстремумов. Использование нескольких точек одновременно значительно снижает такую возможность.

3. Генетические алгоритмы в процессе работы не используют никакой дополнительной информации, что повышает скорость работы. Единственной используемой информацией может быть область допустимых значений параметров и целевой функции в произвольной точке.

4. Генетический алгоритм использует как вероятностные правила для порождения новых точек анализа, так и детерминированные правила для перехода от одних точек к другим. Выше уже говорилось, что одновременное использование элементов случайности и детерминированности дает значительно больший эффект, чем раздельное.

Прежде чем рассматривать непосредственно работу генетического алгоритма, введем ряд терминов, которые широко используются в данной области.

Выше было показано, что генетический алгоритм работает с кодами безотносительно их смысловой интерпретации. Поэтому сам код и его структура описываются понятием генотип, а представляет, по сути, точку пространства поиска. С целью максимального приближения к биологическим терминам экземпляр кода называют хромосомой, особью или индивидуумом. Далее для обозначения строки будет использоваться термин “особь”.

На каждом шаге работы генетический алгоритм использует несколько точек поиска одновременно. Совокупность этих точек является набором особей, который называется популяцией. Количество особей в популяции называют размером популяции. Поскольку в данном разделе рассматриваются классические генетические алгоритмы, то можно сказать, что размер популяции является фиксированным и представляет одну из характеристик генетического алгоритма. На каждом шаге работы генетический алгоритм обновляет популяцию путем создания новых особей и уничтожения старых. Чтобы отличать популяцию на каждом из шагов и сами эти шаги, их называют поколениями и обычно идентифицируют по номеру. Например, популяция, полученная из исходной популяции после первого шага работы алгоритма, будет первым поколением, после следующего шага - вторым, и т. д.

В процессе работы алгоритма генерация новых особей происходит на основе моделирования процесса размножения. При этом, естественно, порождающие особи называют родителями, а порожденные – потомками. Родительская пара, как правило, порождает пару потомков. Непосредственная генерация новых кодовых строк из двух выбранных происходит за счет работы оператора скрещивания, который также называют кроссинговером (от англ. crossover). При порождении новой популяции оператор скрещивания может применяться и ко всем парам родителей. Часть этих пар может переходить в популяцию следующего поколения непосредственно. Насколько часто будет возникать такая ситуация, зависит от значения вероятности применения оператора скрещивания, который является одним из основных параметров генетического алгоритма.

Моделирование процесса мутации новых особей осуществляется за счет работы оператора мутации. Основным параметром оператора мутации также является вероятность мутации.

Поскольку размер популяции фиксирован, то порождение потомков должно сопровождаться уничтожением других особей. Выбор пар родителей из популяции для порождения потомков производится оператором отбора, а выбор особей для уничтожения – оператором редукции.

Таким образом, можно перечислить основные понятия и термины, используемые в области генетических алгоритмов:

- генотип и фенотип, особь и качество особи;
- популяция и размер популяции;
- поколение;
- родители и потомки.

К характеристикам генетического алгоритма относятся:

- размер популяции;
- оператор скрещивания и вероятность его использования;
- оператор мутации и вероятность мутации;
- оператор отбора;
- оператор редукции;
- критерий останова.

Операторы отбора, скрещивания, мутации и редукции называют еще генетическими операторами.

Критерием останова работы генетического алгоритма может быть одно из трех событий:

- сформировано заданное пользователем число поколений;
- популяция достигла заданного пользователем качества (например, значение качества всех особей превысило заданный порог);
- достигается некоторый уровень сходимости, т.е. особи в популяции стали настолько подобными, что дальнейшее их улучшение происходит чрезвычайно медленно.

Характеристики генетического алгоритма выбираются таким образом, чтобы обеспечить малое время работы, с одной стороны, и поиск как можно лучшего решения, с другой.

Рассмотрим теперь непосредственно работу генетического алгоритма.

1. Создание исходной популяции (задается случайным образом).
2. Выбор родителей для процесса размножения (работает оператор отбора).
3. Создание потомков выбранных пар родителей (работает оператор скрещивания).
4. Мутация новых особей (работает оператор мутации).
5. Расширение популяции за счет добавления новых, только что порожденных особей.
6. Сокращение расширенной популяции до исходного размера (работает оператор редукции).

7. Останов работы генетического алгоритма по заданному критерию.
8. Поиск лучшей достигнутой особи в конечной популяции является результатом работы генетического алгоритма.

В основе оператора отбора, который служит для выбора родительских пар и уничтожения особей, лежит принцип “выживает сильнейший”. В качестве примера можно привести следующий оператор. Выбор особи для размножения производится случайно. Вероятность участия особи в процессе размножения вычисляется по формуле

$$\mathrm{Pi} = fi \Big/ \sum_{j=1}^{n} fj,$$

где n – размер популяции; i – номер особи; Pi – вероятность участия особи в процессе размножения; Fi – значение целевой функции для i-й особи. Очевидно, что одна особь может быть задействована в нескольких родительских парах.

Аналогично может быть решен вопрос уничтожения особей. Только вероятность уничтожения, соответственно, должна быть обратно пропорциональна качеству особей. Однако обычно происходит просто уничтожение особей с наихудшим качеством. Таким образом, выбирая для размножения наиболее качественные особи и уничтожая наиболее слабые, генетический алгоритм постоянно улучшает популяцию, приводя к нахождению все лучших решений.

Оператор скрещивания призван моделировать природный процесс наследования, т.е. обеспечивать передачу свойств родителей потомкам.

Опишем простейший оператор скрещивания. Он исполняется в два этапа. Пусть особь представляет собой строку из n элементов. На первом этапе равновероятно выбирается число k от 1 до n-1. Это число называется точкой разбиения. В соответствии с ним обе исходные строки разбиваются на две подстроки. На втором этапе строки обмениваются своими подстроками, лежащими после точки разбиения, то есть элементами с k+1-го по n -й. Так получается две новые строки, которые наследовали частично свойства обоих родителей. Этот процесс проиллюстрирован ниже.

Стока 1	X1X2. . .XkXk+1. . .Xn	$\Rightarrow$	X1X2. . .XkYk+1. . .Yn
Строка 2	Y1Y2. . .YkYk+1. . .Yk		Y1Y2. . .YkXk+1. . .Xn

Вероятность применения оператора скрещивания обычно выбирается достаточно большой, в пределах от 0,9 до 1, чтобы обеспечить постоянное появление новых особей, расширяющих пространство поиска. При значении вероятности менее единицы часто используют элитизм. Это особая стратегия, которая предполагает переход в популяцию следующего поколения элиты, т.е. лучших особей текущей популяции, без всяких изменений. Применение элитизма способствует сохранению общего качества популяции на высоком уровне. При этом элитные особи участвуют еще и в процессе отбора родителей для последующего скрещивания. Количество элитных особей определяется обычно по формуле

$K=(1-P)\times N$,

где K- количество элитных особей; P – вероятность применения оператора скрещивания; N- размер популяции.

В случае использования элитизма все выбранные родительские пары подвергаются скрещиванию, несмотря на то, что вероятность применения оператора скрещивания меньше единицы. Это позволяет сохранять размер популяции постоянным.

Оператор мутации служит для моделирования природного процесса мутации. Его применение в генетических алгоритмах обусловлено следующими соображениями. Исходная популяция, какой бы большой она и была, охватывает ограниченную область пространства поиска. Оператор скрещивания, безусловно, расширяет эту область, но все же до определенной степени, поскольку использует ограниченный набор знаний, заданный исходной популяцией. Внесение случайных изменений в особи позволяет преодолеть это ограничение и иногда значительно сократить время поиска или улучшить качество результата.

Как правило, вероятность мутации, в отличие от вероятности скрещивания, выбирается достаточно малой. Сам процесс мутации заключается в замене одного из элементов строки на другое значение. Это может быть перестановка двух элементов в строке, замена элемента строки значением элемента из другой строки, в случае битовой строки может применяться инверсия одного из битов и т. д.

В процессе работы алгоритма все указанные выше операторы применяются многократно и ведут к постепенному изменению исходной популяции. Поскольку операторы отбора, скрещивания, мутации и редукции по своей сути направлены на улучшения каждой отдельной особи, то результатом их работы является постепенное улучшение популяции. В этом и заключается основной смысл работы генетического алгоритма - улучшить популяцию решений по сравнению с исходной.

После завершения работы генетического алгоритма из конечной популяции выбирается та особь, которая дает максимальное (минимальное) значение целевой функции и является, таким образом , работы генетического алгоритма. За счет того, что конечная популяция лучше исходной, полученный результат представляет собой улучшенное решение.

7.6.2 Пример решения задачи поиска максимума одномерной функции

Пример поиска максимума одномерной функции. Пусть имеется набор натуральных чисел от 1 до 31 и функция, определенная на этом наборе чисел, f(x)=x. Требуется найти максимальное значение функции. Задача тривиальна и не требует применения столь изощренных методов поиска, но ее решение необходимо для иллюстрации функционирования генетического алгоритма [1].

В качестве кода будем использовать двоичное представление аргументов функции. Это положение представляет собой фенотип нашего алгоритма. Сам код будет представлять собой двоичную строку из 5 бит. Это генотип алгоритма. Целевой функцией будет непосредственно сама рассматриваемая функция, аргументом которой является число, чье двоичное представление использует алгоритм.

Определим некоторые характеристики генетического алгоритма. Пусть размер популяции 4, вероятность мутации 0,001, сам процесс мутации заключается в инверсии одного из битов строки, выбираемого случайно по равномерному закону. Оператор скрещивания и отбора аналогичны описанным выше. Поскольку задача простейшая, будем считать, что алгоритм не использует элитизм.

Пусть на основе равномерного распределения создана исходная популяция из четырех особей, представленная в таблице 7.1

Таблица 7.1

Номер строки	Код генотипа	Значение целевой функции	Вероятность участия в размножении
1	01011	11	11/43
2	10010	18	18/43
3	00010	2	2/43
4	01100	12	12/43
		Сумма 43	

Предположим, что оператор отбора выбрал для производства потомков две пары строк (1, 2) и (2, 4). Работа оператора скрещивания проиллюстрирована в таблице 7.2. При этом в каждой паре разбиение на подстроки происходит независимо. Точка разбиения задана звездочкой.

Пропорциональный простейший отбор (рулетка) выбирает n особей после n запусков. Колесо рулетки содержит по одному сектору для каждого гена популяции. Размер сектора пропорционален вероятности участия.

Таблица 7. 2

№ строки	Родители	Потомки	Значение целевой функции для потомков
1	0*1011	00010	2
2	1*0010	11011	27
3	100*10	10000	16
4	011*00	01110	14

Пусть оператор мутации, несмотря на низкую вероятность, срабатывает для младшего бита потомка в строке 3 и данный код изменил свое значение с 10000 на 10001.

Таким образом, популяция за счет порожденных потомков расширилась до восьми особей, представленных в таблице 7.3.

Таблица 7.3

№ строки	код	Значение целевой функции
Исходная популяция		
1	01011	11
2	10010	18
3	00010	2
4	01100	12
Порожденные потомки		
5	00010	2
6	11011	27
7	10001	17
8	01110	14

Оператор редукции далее сокращает популяцию до исходного числа особей, исключив из нее особи с минимальным значением целевой функции. То есть будут исключены строки 1, 3, 4, 5, и популяция первого поколения примет вид, представленный в таблице 7.4.

Таблица 7.4

№ строки	Код	Значение целевой функции	Вероятность участия в процессе размножения
1	10010	18	18/ 76
2	11011	27	27/ 76
3	10001	17	17/ 76
4	01110	14	14/ 76

На этом шаг работы генетического алгоритма закончится. Очевидно, что даже за эту одну итерацию качество популяции значительно возросло. Если в исходной популяции среднее значение целевой функции было 10, 75, а ее минимальное значение составляло 2, то в популяции первого поколения среднее значение увеличилось до 19, а минимальное значение составило 14. Лучшее же решение увеличилось с 18 до 27 при оптимальном решении 31.

Таким образом, данный пример наглядно иллюстрирует процесс улучшения как популяции в целом, так и поиск наилучшего решения в результате работы генетического алгоритма [1,2].

7.6.3 Пример решения задачи коммивояжера

Задача коммивояжера является классической оптимизационной задачей. Суть ее заключается в следующем. Дано множество из n городов и матрица расстояний между ними или стоимостей переезда (в зависимости от интерпретации). Цель коммивояжера – объехать все эти города по кратчайшему пути или с наименьшими затратами на поездку. Причем в каждом городе он должен побывать один раз и свой путь закончить в том же городе, откуда начал.

Для решения предлагается следующая задача: имеется пять городов, стоимость переезда между которыми представлена матрицей (таблица 7.5).

Таблица 7.5

	Г1	Г2	Г3	Г4	Г5
Г 1	0	4	6	2	9
Г 2	4	0	3	2	9
Г 3	6	3	0	5	9
Г 4	2	2	5	0	8
Г 5	9	9	9	8	0

Для решения задачи применим следующий генетический алгоритм. Решение представим в виде перестановки чисел от 1 до 5, отображающей последовательность посещения городов. Значение целевой функции будет равно стоимости всей поездки, вычислений в соответствии с вышеприведенной матрицей. Одним из оптимальных решений задачи является последовательность 51435 стоимостью 25. Из пятого города в первый стоимость равна 9, из первого в четвертый – 2, из четвертого в третий - 5, из третьего в пятый – 9, из пятого в пятый - 0. Таким образом, 9+2+5+9+0=25.

В качестве оператора отбора будем использовать традиционный оператор, применявшийся в предыдущем примере. При этом заметим, что, чем меньше значение целевой функции, тем лучше.

В качестве оператора скрещивания выберем более изощренную процедуру, похожую на двухточечный оператор скрещивания. Пусть есть две родительские перестановки (1/234/5) и (3/452/1). Случайно и равновероятно выбираются две точки разрыва. Для примера возьмем ситуацию, когда первая точка разрыва находится между первым и вторым элементами перестановки, а вторая точка – между четвертым и пятым: (1/234/5), (3/452/1). На первом этапе перестановки обмениваются фрагментами, заключенными между точками разрыва: (*/452/*), (*/234/*). <u>На втором этапе вместо звездочек вставляются соответствующие числа из исходной родительской перестановки, начиная со второго числа.</u> В данном случае в первой перестановке (1/234/5) таким начальным числом является 3, а за ним идет 4, которое есть в новой перестановке, и мы его пропускаем, также пропускаем число 5, переходим на начало перестановки и выбираем число 1. В итоге вместо (*/452/*) получаем (34521), аналогично из (/3/452/1) и (*/234/*) получаем (52341).

Оператор мутации будет представлять собой случайную перестановку двух чисел в хромосоме, также выбранных случайно по равномерному закону. Например, вероятность мутации равна 0,01. Размер популяции выберем равным 4. Исходная популяция представлена в таблице 7.6.

Таблица 7.6

№ строки	Код	Значение целевой функции	Вероятность участия в процессе размножения
1	12345	29	32/122 (0,26)
2	21435	29	32/122
3	54312	32	29/122 (0,23)
4	43125	32	29/122

Пусть для скрещивания были выбраны следующие пары: (1,3) и (2,4). В результате были получены потомки, представленные в таблице 7.7.

Таблица 7.7

№ строки	Родители	Потомки	Значение целевой функции для потомков
1 (1)	1 | 23 | 45	5 | 43 | 12	32
2 (3)	5 | 43 | 12	1 | 23 | 54	28
3 (2)	2 | 143 | 5	4 | 312 | 5 мутация 13254	32
4 (4)	4 | 312 | 5	2 | 143 | 5	29

Пусть для потомка (12345) сработал оператор мутации, и обменялись местами числа 2 и 3. В данном случае строка (12345) изменилась и приняла значение (13254). Популяция первого поколения после отсечения «худших» особей в результате работы оператора редукции приняла вид, представленный в таблице 7.8.

Таблица 7.8

№ строки	Код	Значение целевой функции	Вероятность участия в процессе размножения
1	12345	29	28/111 29/115
2	21435	29	28/111 28/115
3	13254	28	29/111 29/115
4	21435	29	28/111 28/115

Пусть для получения второго поколения были выбраны следующие пары строк: (1, 4) и (2, 3). В результате были получены потомки, показанные в таблице 7.9.

Таблица 7.9

№ строки	Родители	Потомки	Значение целевой функции для потомков
1	∣ 123 ∣ 45	∣ 214 ∣ 35	29
2	∣ 214 ∣ 35	∣ 123 ∣ 45	29
3	∣ 21 ∣ 435	∣ 13 ∣ 452	32
4	∣ 13 ∣ 254	∣ 21 ∣ 354	29

Популяции второго поколения после отсечения худших особей приняла вид, показанный в таблице 7.10.

Таблица 7.10

№ строки	Код	Значение целевой функции	Вероятность участия в процессе размножения
1 (1)	12345	28	29/114
2 (2)	21435	29	28/114
3 (3)	13254	28	29/114
4 (н)	21354	29	28/114

Таким образом, после двух итераций значение целевой функции для лучшего решения изменилось с 29 на 28, среднее значение изменилось с 30,5 до 28,75, а общее качество с 122 до 114. Наблюдается улучшение популяции.

Упражнения

Задача 1. Провести поиск минимума одномерной функции f(x)= x^2-16x +5 с помощью генетических алгоритмов. Описать 2 популяции. Генотип алгоритма представляет собой строку из 5 бит. Размер популяции - 4. Например, строка 01011 соответствует числу x=11, а f(x)= -50. Использовать одноточечный кроссовер. Мутация заключается в инверсии одного из битов строки, выбираемого случайно. Элитизм необходимо использовать.

Задача 2. Решить задачу коммивояжера с использованием генетических алгоритмов. Коммивояжер должен объехать все города и вернуться в тот же город, откуда выехал первый раз, потратив при этом минимум средств. Затраты на поездки представлены матрицей.

	Г1	Г2	Г3	Г4	Г5
Г1	0	3	4	4	5
Г2	3	0	6	5	7
Г3	4	6	0	3	5
Г4	4	5	3	0	4
Г5	**5**	**7**	**5**	**4**	**0**

Выполнить формирование двух поколений. Использовать необходимо элитизм. Мутацию выполнить методом перестановки соседних позиций для случайно выбранного числа. Применить изощренный оператор кроссовера для размножения особей.

Библиографический список

1. Черноруцкий И.Г. Методы принятия решения. – СПб.: БХВ-Петербург, 2005. -416с.
2. Кернеев В.В. Базы данных. Интеллектуальная обработка информации/ В.В. Корнеев и др. – М.: Нолидж, 2000. – 352 с.
3. Романов В.П. Интеллектуальные информационные системы в экономике: Учеб. пособие/ В.П. Романов.; Под ред. д.э.н., проф. Н.П. Тихомирова.-М.: Издательство Экзамен, 2003.-496 с.

8 Коллективные решения

8.1 Парадокс Кондорсе

В 1996 г. перед первым туром президентских выборов в России по московскому радио передавали выступление избирателя, недовольного системой голосования. Он предлагал разрешить каждому избирателю не только голосовать за одного кандидата, но и упорядочивать всех кандидатов по своему предпочтению от лучшего к худшему. Только после этого, утверждал выступавший, будет ясно истинное отношение населения России к кандидатам в президенты.

Интересно, что большой интерес к разным системам голосования наблюдался примерно за 200 лет до этого во Франции. При этом ситуации в двух странах были близкими: и тут, и там происходил переход от тоталитаризма к новой системе, позволяющей каждому избирателю голосовать свободно и тайно.

Одним из первых, кто заинтересовался системами голосования, был французский ученый маркиз де Кондорсе (1743— 1794). Он сформулировал принцип или критерий, позволяющий определить победителя в демократических выборах. Принцип де Кондорсе состоит в следующем: *кандидат, который побеждает при сравнении один на один с любым из других кандидатов, является победителем на выборах.*

Система голосования, предложенная де Кондорсе, совпадала с системой, которую предлагал 200 лет спустя избиратель в России. Каждый из голосующих упорядочивал кандидатов по степени своего желания видеть его победителем. Согласно де Кондорсе, справедливое определение победителя возможно путем попарного сравнения кандидатов по числу голосов, поданных за них. Принцип де Кондорсе предлагался как рациональный и демократический. Однако вскоре маркиз де Кондорсе столкнулся с парадоксом, получившим впоследствии его имя. Рассмотрим пример голосования в собрании представителей из 60 чел. [1]. Пусть на голосование поставлены три кандидата: А, В и С, и голоса распределились, как в таблице 8.1.

Таблиц а 8.1 - *Распределение голосов (парадокс Кондорсе)*

Число голосующих	Предпочтения
23	A → B → C
17	B → C → A
2	B → A → C
10	C → A → B
8	C → B → A

Сравним предпочтения по отношению к парам кандидатов. Берем А и С: тогда А предпочитают 23+2=25; С по сравнению с А предпочитают: 17+10+8=35. Следовательно, С предпочтительнее А (С —> А) по воле большинства.

Сравнивая попарно аналогичным образом А и В, В и С, получаем: В->С (42 против 18), С-»А (35 против 25) и А -» В (33 против 27). Следовательно, мы пришли к противоречию, к нетранзитивному отношению А -> В -> С -> А .

Столкнувшись с этим парадоксом, Кондорсе выбрал «наименьшее зло», а именно то мнение, которое поддерживается большинством голосов (избранным следует считать А).

8.2 Правило большинства голосов

Изменим несколько результаты голосования, чтобы избежать парадокса Кондорсе. Предположим, что голоса распределились так, как показано в таблице 8.2. Нетрудно подсчитать, что при этих новых результатах голосования, в соответствии с принципом Кондорсе, избранным будет кандидат С, который при попарном сравнении побеждает двух других кандидатов.

Таблица 8.2- *Распределение голосов (правило большинства)*

Число голосующих	Предпочтения
23	$A \rightarrow C \rightarrow B$
17	$B \rightarrow C \rightarrow A$
2	$C \rightarrow B \rightarrow A$
10	$C \rightarrow A \rightarrow B$

Однако если мы используем другой принцип выбора: *большинство голосующих,* которые назвали данного кандидата лучшим, то победителем оказывается кандидат А. Но при этом кандидат А не набрал абсолютного большинства голосов.

Мы видим, что способ определения победителя при демократической системе голосования (один человек — один голос) зависит от процедуры голосования.

8.3 Метод Борда

Отметим еще одну процедуру голосования из множества предложенных: метод Борда [2]. Согласно этому методу, результаты голосования выражаются в виде числа баллов, набранных каждым из кандидатов. Пусть число кандидатов равно п. Тогда за первое место присуждается п баллов, за второе — п-1, за последнее — один балл.

Применим метод Борда к приведенному выше примеру. Подсчитаем число баллов для каждого из кандидатов:

А: 23-3 + 19-1 + 16-1 + 2-2 = 108;

В: 23-1 + 19-3 + 16-2 + 2-1 = 114;

С: 23-2 + 19-2 + 16-2 + 2-3 = 138.

В соответствии с методом Борда мы должны объявить победителем кандидата С.

Однако с методом Борда, как и с принципом Кондорсе, возникают проблемы. Предположим, что результаты голосования в выборном органе представлены в таблице 8.3 18. Подсчитав баллы в соответствии с методом Борда, получим: А — 124, В — 103, С — 137. В соответствии с методом Борда победителем следует объявить кандидата С. Однако в данном случае явным победителем является кандидат А, набравший абсолютное большинство голосов: 31 из 60.

Таблица 8.3 - *Распределение голосов (метод Борда)*

Число голосующих	Предпочтения
31	*A* → *C* → *B*
12	*B* → *C* → *A*
17	*C* → *B* → *A*
2	*C* → *A* → *B*

Приведенные примеры позволяют понять, что парадоксы при голосовании не возникают лишь в случае, когда победитель определяется по принципу абсолютного большинства голосов. Однако такой случай нетипичен для большинства выборов в демократических странах. Обычно число кандидатов больше, чем два, и редки случаи, когда кто-то из них сразу же получает поддержку абсолютного большинства избирателей.

Интересно, что парадоксы голосования сохраняются и при введении двух туров и условии, что во второй тур выходят два кандидата, набравшие большинство голосов. Обратимся к таблице 8.1, составленной Кондорсе. В соответствии с предпочтениями во второй, тур выходят А (23 голоса) и В (19 голосов), после чего побеждает А. Однако при небольшом усилении первоначальной позиции А: предпочтения двух избирателей (3-я строка) выглядят как А -> В -> С, во второй тур выходят А (25 голосов) и С (20 голосов), после чего побеждает С. Ясно, что такой результат голосования противоречит здравому смыслу.

8.4 Аксиомы Эрроу

Выше мы привели примеры нескольких различных систем голосования. Возможны и другие системы. В качестве примеров можно указать на систему многотурового выбора с вычеркиванием кандидатов, набравших наименьшее число голосов [2], на систему вычеркивания нежелаемых кандидатов (approval voting) [3] и т.д.

Систематическое исследование всех возможных систем голосования провел в 1951г. Кеннет Эрроу из Стенфордского университета [4]. Он поставил вопрос в наиболее общем виде: можно ли создать такую систему голосования, чтобы она была одновременно рациональной (без противоречий), демократической (один человек — один голос) и решающей (позволяла осуществить выбор)? Вместо попыток изобретения такой системы Эрроу предложил набор требований, аксиом, которым эта система должна удовлетворять. Эти аксиомы были интуитивно понятны, приемлемы с точки зрения здравого смысла и допускали математическое выражение в виде некоторых условий. На основе этих аксиом Эрроу попытался в общем виде доказать существование системы голосования, удовлетворяющей одновременно трем перечисленным выше принципам: рациональная, демократическая и решающая [4,5].

Первая аксиома Эрроу требует, чтобы система голосования была достаточно общей для того, чтобы учитывать все возможные распределения голосов избирателей. Интуитивно это требование вполне очевидно. Заранее нельзя предсказать распределение голосов. Совершенно необходимо, чтобы система была действенной при любых предпочтениях избирателей. Эта аксиома получила название *аксиомы универсальности.*

Еще более очевидной с точки зрения здравого смысла является вторая аксиома Эрроу: *аксиома единогласия.* В соответствии с ней необходимо, чтобы коллективный выбор повторял в точности единогласное мнение всех голосующих. Если, например, каждый из голосующих считает, что кандидат А лучше кандидата В, то и система голосования должна приводить к этому результату.

Третья аксиома Эрроу получила название *независимости от несвязанных альтернатив.* Пусть избиратель считает, что из пары кандидатов А и В лучшим является А. Это предпочтение не должно зависеть от отношения избирателя к прочим кандидатам. Третья аксиома достаточно привлекательна, но не столь очевидна с точки зрения каждодневного человеческого поведения. Так, в [6] приводится убедительный пример нарушения этой аксиомы. Посетитель ресторана первоначально сравнивает блюдо А и В и хочет заказать А, потому что приготовление блюда В требует высокой квалификации повара, а , по его мнению, такой повар вряд ли есть в данном ресторане. Вдруг он замечает в меню блюдо С - очень дорогое и также требующее высокого искусства приготовления. Тогда он выбирает блюдо В, считая, что повар умеет хорошо готовить.

Часто третья аксиома Эрроу нарушается судьями в фигурном катании. Давая сравнительные оценки двум сильным фигуристам в одиночном катании, они стараются учесть возможность хорошего выступления третьего сильного кандидата, оставляя ему шансы стать победителем. Отличное выступление в произвольном катании фигуриста С, имевшего ранее не очень высокий результат в обязательной программе, может повлиять на оценки фигуристов А и В. Если А имел отличный результат в обязательной программе, судьи иногда ставят его ниже фигуриста В при примерно равном выступлении, чтобы повысить шансы фигуриста С.

Тем не менее сама возможность предъявления требования независимости к системе голосования в качестве обязательного не вызывает сомнения.

Четвертая аксиома Эрроу носит название *аксиомы полноты:* система голосования должна сравнить любую пару кандидатов, определив, кто из них лучше. При этом имеется возможность объявить двух кандидатов равнопривлекательными. Требование полноты не кажется слишком строгим для системы голосования.

Пятая аксиома Эрроу является уже знакомым *условием транзитивности:* если в соответствии с мнением избирателей кандидат B не лучше кандидата A (хуже или эквивалентен), кандидат C не лучше кандидата B, то кандидат C не лучше кандидата A. Считается, что система голосования, не допускающая нарушения транзитивности, ведет себя рациональным образом.

Определив пять аксиом — желательных свойств системы голосования, Эрроу доказал, что системы, удовлетворяющие этим аксиомам, обладают недопустимым с точки зрения демократических свобод недостатком: каждая из них является правилом диктатора — личности, навязывающей всем остальным избирателям свои предпочтения.

Результаты, выявленные Эрроу, получили широкую известность. Они развеяли надежды многих экономистов, социологов, математиков найти совершенную систему голосования.

Требование исключения диктатора приводит к невозможности создания системы голосования, удовлетворяющей всем аксиомам Эрроу. Поэтому результат Эрроу называют «теоремой невозможности».

8.5 Попытки пересмотра аксиом

Более 70 лет математики и экономисты предпринимают попытки изменить требования Эрроу, «смягчить» аксиомы, чтобы избежать вывода, столь неприятного для демократической системы голосования.

Очень интересное изменение первой аксиомы предложил Д. Блейк [7]. Если каждый избиратель упорядочивает кандидатов в соответствии со своей политической позицией, вывода Эрроу можно избежать. На практике это означает, что каждый избиратель должен упорядочить кандидатов в соответствии с их политическими взглядами. Если он сторонник рынка и монетаризма и считает, что А лучше В, В лучше С, то это означает, что А ближе всех к его позиции, а С — дальше всех.

Однако на практике при оценке кандидата избиратели чаще всего руководствуются многими критериями. Далеко не все избиратели понимают свою политическую позицию. Результаты голосований, основанных на эмоциях, широко известны.

Другим интересным изменением аксиом Эрроу является правило консенсуса, сформулированное А.Сеном. Он предложил изменить аксиому транзитивности, сохранив правило транзитивности только для случая строгого предпочтения между кандидатами. Согласно правилу А.Сена, если хотя бы один избиратель по-иному сравнивает кандидатов А и В, чем все остальные, то система голосования объявляет кандидатов эквивалентными. Ясно, что такое правило приводит к коллективному безразличию.

8. 6 Теорема невозможности и реальная жизнь

Итак, серьезность результатов К.Эрроу безусловна. Нельзя отказаться от требования рациональности: система голосования не должна приводить к нетранзитивности. Нельзя не потребовать, чтобы система голосования была решающей: коллективное безразличие, неумение сделать выбор ведет в тупик. Нельзя отказаться от требования демократичности выборов: человечество заплатило (и продолжает платить) высокую цену за право каждого человека выражать свое мнение. Кроме того, демократичность в решении социальных проблем особенно важна в наше время, когда меньшинство имеет массу возможностей защищать свою позицию перед большинством.

С точки зрения реальной жизни важно знать, насколько часто нарушаются все эти три условия одновременно. Исследования французских ученых показали, что при моделировании всех возможных распределений голосов избирателей и сохранении условий демократичной и решающей системы голосования рациональность нарушается примерно в 6—9% случаев [1].

Конечно, каждый раз неизвестны ни распределение голосов избирателей, ни возможности нарушения рациональности. Однако в реальных процедурах выбора есть и многие, не менее существенные недостатки. Известны ситуации манипулирования в процессе выборов, когда преднамеренное искажение предпочтений группой избирателей приводит к желаемому для этой группы результату (см. пример с двумя турами голосования, приведенный выше).

Исключительно сильное воздействие на умы избирателей оказывает так называемая «промывка мозгов» — целенаправленные кампании в пользу какого-то кандидата с искажением фактов, подтасовкой и т. д. Для стран, не имеющих опыта демократических выборов, такие явления приводят к разочарованию избирателей в демократических институтах власти. Как каждый человек, так и народы в целом должны учиться делать свой выбор, различая слова и дела политиков, трезво оценивая обещания, используя разные и независимые источники информации.

Вернемся к парадоксальному результату Эрроу. Примириться с фактом его существования помогут известные слова У. Черчилля о том, что демократия является плохой формой правления, но человечество пока не придумало ничего лучшего.

8.7 Принятие коллективных решений в малых группах

Принятие коллективных .решений не сводится только к голосованию избирателей на выборах. Решения принимаются в комиссиях, жюри, коллегиях, словом, в небольших группах. В роли ЛПР в этом случае выступает группа, принимающая решения (ГПР). Как организовать работу ГПР? Где гарантии, что люди, имеющие различные предпочтения, могут прийти к соглашению?

Традиционным способом решения этих проблем является организация совещаний (заседаний), на которых члены коллективного органа, принимающего решения, выступают как эксперты, оценивая различные варианты решений и убеждая других членов присоединиться к их мнению. Во многих случаях эти обсуждения позволяют прийти к единому мнению, которое иногда отражает компромисс между членами коллективного органа, принимающего решения.

Несомненными преимуществами такого способа принятия коллективных решений являются:

- возможность для каждого из членов ГПР высказать свое мнение и обосновать его;
- возможность для каждого из членов ГПР выслушать мнение всех других членов.

Наряду с указанными достоинствами применение традиционного способа решения задач коллективного выбора в ГПР в ряде случаев сопровождается следующими отрицательными явлениями:

- чрезмерно сильное влияние на ГПР доводов одного или нескольких членов (коалиции), направленных на выпячивание положительных особенностей предпочитаемых ими вариантов решений;
- большая и зачастую неэффективная трата времени членами ГПР, особенно при сильном расхождении мнений у некоторых из них;
- поспешное применение правила большинства, не позволяющего учесть мнения всех членов ГПР. С попытками преодолеть те или иные отрицательные черты традиционных способов принятия коллективных решений связаны различные направления исследований. Мы остановимся далее на проблемах выбора, решаемых ГПР: даны какие-то объекты (варианты капиталовложений, акции, предприятия и т.д.). Необходимо выделить из них лучший, разделить на классы, упорядочить, т. е. решить задачи, типичные для принятия индивидуальных решений.

Принятие решений в ГПР разительно отличается от принятия индивидуальных решений. У каждого из членов группы имеется, как правило, свой взгляд на проблему. Если эти взгляды полностью совпадают либо если в группе есть диктатор, навязывающий свои предпочтения, то задача принятия коллективных решений не возникает. В общем же случае основной для ГПР является проблема поиска компромисса, приемлемого для всех членов группы.

Можно выделить три направления исследований по принятию решений в малых группах.

1. Неантагонистические игры. Одно из направлений в теории игр [8], ориентированное на разработку математических моделей, описывающих процесс выработки компромисса — поиск точек равновесия. Работы в данном направлении имеют, как правило, чисто теоретический характер.

2. Групповые системы поддержки принятия решений. Разрабатываются локальные сети для членов ГПР, а также формальные алгоритмы сравнения предпочтений на заданном множестве объектов. Как правило, системы поддержки принятия решений предназначены для ознакомления каждого из членов ГПР с мнениями других. Задача согласования мнений членов ГПР либо не ставится, либо сводится к усреднению мнений. С практической точки зрения данный подход не соответствует задачам принятия ответственных решений.

3. *Организация работы ГПР с помощью посредника* (аналитика, консультанта). Это направление исследований с практической точки зрения является наиболее перспективным. Ярким примером могут служить так называемые конференции по принятию решений (decision conference). Организация и проведение конференций по принятию решений связаны с именами С. Ка-мерера (США) и Л.Филипса (Англия). Они первыми разработали методологические основы организации конференций по принятию решений и получили хороший практический результат.

8.8 Организация и проведение конференций по принятию решений

Предположим, что ГПР представляет собой совет директоров крупной фирмы. Фирма переживает трудности: уменьшилась доля рынка изделий, выпускаемых компанией, упали прибыли. Возникла необходимость что-то предпринять. Одни из директоров предлагают перейти к выпуску нового изделия, другие — более активно вести рекламу, третьи — модернизировать выпускаемые изделия.

Для выработки согласованной стратегии президент фирмы решил пригласить консультанта (facilitator), специализирующегося на проведении конференций по принятию решений. После получения заказа консультант посещает фирму, беседует с некоторыми из директоров, знакомится с положением дел. Конечно, он не может знать так же глубоко, как директора, все промышленные и финансовые проблемы фирмы. Но он стремится освоить деловой язык, на котором идут обсуждения, понять распределение власти и влияния среди руководителей фирмы.

Через пару недель назначается конференция по принятию решений. Она проводится в специальной комнате, технически подготовленной для этой цели: есть компьютер, есть экраны, на которые проецируются данные, обрабатываемые компьютером, есть проектор, доски, фломастеры и т.д. Конференция продолжается один—два дня (часто суббота и воскресенье), во время которых руководители заняты не текущими задачами, а только разработкой стратегии. Конференцию проводит консультант. Он дает слово участникам, представляющим различные точки зрения на обсуждаемую проблему. Он регулирует процесс обсуждения, направляя его по конструктивному пути, он сам ставит вопросы, пытаясь выяснить сильные и слабые стороны обсуждаемых вариантов решений. Он старается показать участникам конференции то общее, что объединяет различные варианты, подчеркивая элементы согласия. Он пытается помочь участникам конференции найти смешанные стратегии (если это возможно), достойно, не теряя лица, отказаться от вариантов, недостатки которых стали очевидными. Консультант предлагает совместно оценить варианты решений по различным критериям. Если есть расхождения в оценках, он проверяет чувствительность выбора к этим расхождениям и т. д.

Ясно, что успех конференций по принятию решений в значительной степени зависит от квалификации консультанта. Он должен обладать необходимыми личностными характеристиками: умением быстро схватывать суть обсуждаемой проблемы и имеющихся разногласий, быстро «стать своим» во время обсуждения и т.д. Он должен обладать большим практическим опытом, чтобы правильно вести дискуссию. Как опытный дирижер оркестра, он должен вовремя «взмахнуть палочкой», предоставляя слово участнику, который в данный момент наиболее продуктивно влияет на ход дискуссии. Такому набору умений нельзя выучиться по учебнику, да и таких учебников не существует. Есть отдельные публикации [9], в которых на примерах представляют основные идеи конференций по принятию решений.

Успех этих конференций приводит к впечатляющим результатам в деятельности компаний. Даже относительный успех может повлиять на дальнейшую выработку стратегии. Но и 1 неудача оказывает сильное отрицательное воздействие, закрепляя расхождения во взглядах.

8.9 Метод организации работы в группе принятия решений

В 1975 г. был разработан и применен один из возможных методов организации работы группы принятия решений (ГПР) при проведении конференции по принятию решений [10]. В основу метода были положены следующие соображения:

- для эффективного обсуждения проблемы в ГПР полезен предварительный анализ, вскрывающий совпадение или расхождение мнений членов ГПР о достоинствах и недостатках рассматриваемых объектов;

• такой анализ должен определить те конкретные вопросы, по которым совпадают или расходятся мнения членов ГПР, выявить намечающееся единство мнений, а также образующиеся коалиции.

Практический опыт организации и проведения экспертных опросов привел к выводу об исключительной важности разбиения обсуждаемого вопроса на части. При сравнении объектов, при оценке их качества очень важно выделить составляющие этого качества — отдельные критерии, и оценить объекты по каждому из них. Такой подход позволяет:

1. Достичь большей объективности получаемой информации. При оценке объектов в целом сильнее проявляется возможность субъективных искажений. Определив положительные качества объектов, члены ГПР часто (невольно или преднамеренно) не принимают во внимание его отрицательные качества. При определении оценок объекта по многим критериям член ГПР в значительной степени выступает как квалифицированный специалист, своего рода «измеритель», определяющий оценки объекта по шкалам, общим для группы. Сознательное завышение или занижение этих оценок связано с возможной утратой репутации квалифицированного, знающего специалиста (при проверке и обсуждении этих оценок другими специалистами), в то время как давая оценку в целом, член ГПР почти всегда может объяснить ее учетом определенных качеств объекта.

2. Сделать более конкретным и четким обсуждение полученной информации. При обсуждении вопроса об оценке объекта по одному из критериев расхождения во мнениях членов ГПР, как правило, меньше, чем при обсуждении оценки объекта в целом. По узкому вопросу легче получить дополнительную информацию, легче собрать достаточно бесспорные факты.

8.9.1 Предварительные этапы

Представляется целесообразной следующая последовательность действий.

1. Определение списка критериев. Для рассматриваемого множества объектов ставится вопрос о выделении группы критериев, которые необходимо принимать во внимание при оценке объектов. Список этих критериев образуется в результате опроса каждого из членов ГПР. Перечень критериев, полученных от всех членов ГПР, согласовывается с каждым из ее членов в отдельности. Иногда консультант может предложить предварительный список критериев, которые могли бы учитываться при рассмотрении подобных вопросов.

2. Разработка шкал оценки по критериям. На основе предварительного знакомства с терминологией, применяемой обычно в данной ГПР, консультант разрабатывает для каждого критерия шкалу из нескольких словесных качественных оценок, расположенных от лучшей к худшей. Эти шкалы согласовываются с каждым из членов ГПР. Оценки на шкалах должны быть понятными и исключать неоднозначное толкование.

3. *Сбор информации.* Членам ГПР раздаются формы, содержащие список критериев со шкалами. Число форм равно числу объектов. Каждый из членов ГПР оценивает рассматриваемые объекты по каждому из критериев — определяет одну из оценок по шкале критерия, характеризующую данный объект. При необходимости каждый из членов ГПР может через консультанта (или самостоятельно) запросить дополнительную информацию, необходимую ему для оценки объектов.

Данные, полученные при индивидуальном опросе членов ГПР, содержат информацию о степени совпадения или расхождения их мнений. Для выявления этой информации требуется специальный анализ.

8.9.2 Анализ собранной информации

Желательно иметь такой способ анализа информации, который позволял бы оценивать степень согласия между членами ГПР. Необходимо, чтобы этот способ выявлял объекты, оценки по которым наиболее противоречивы, а также критерии, оценки по которым показывают наибольшую несогласованность мнений между членами ГПР.

На наш взгляд, удовлетворяет этим требованиям способ понижения размерности данных - способ «проецирования» [11, 12]. Методы понижения размерности данных являются одной из групп методов факторного анализа [13]. Основная идея методов проецирования заключается в следующем. Пусть в r-мерном пространстве задана конфигурация из n точек. Необходимо спроецировать эти точки в подпространство меньшей размерности (как правило, двух- или трехмерное), передав при этом с наименьшими (в смысле определенного критерия) возможными искажениями расстояния между объектами в исходном пространстве.

Как и другие методы факторного анализа, методы проецирования основаны на гипотезе о том, что существует небольшое число обобщенных факторов, которыми характеризуются рассматриваемые объекты. Справедливость этой гипотезы проверяется с помощью применения метода и изучения полученных результатов (критерий практики).

В рассматриваемом случае каждый объект может быть представлен точкой в пространстве критериев. В этом же пространстве можно введением определенной метрики расположить точки, характеризующие отношение членов ГПР к рассматриваемым объектам. Далее аналитику предоставляется возможность, применяя методы проецирования, попытаться представить на плоскости объекты и членов ГПР так, чтобы [10]:

- расстояние между членами ГПР характеризовало степень согласия между ними, малые расстояния соответствовали бы совпадению мнений;
- расстояние между членом ГПР и объектом характеризовало ценность этого объекта для данного члена ГПР; малые расстояния соответствовали бы наиболее ценным объектам;
- расстояния между объектами соответствовали бы степени их сходства по оценкам членов ГПР.

Полученное при этом распределение на плоскости точек, соответствующих объектам и экспертам, дает общее представление о степени совпадения мнений членов ГПР в целом и по каждому из объектов.

Анализ информации заключается в последовательном применении метода проецирования при «загрублении» первоначально используемой информации последовательным удалением ряда критериев, объектов, оценок и наблюдении при этом за степенью совпадения мнений членов ГПР при оставшейся информации. Таким образом можно определить те спорные вопросы (критерии, оценки, объекты), на которых необходимо сосредоточить внимание членов ГПР при последующем обсуждении.

Хотя совместное проецирование объектов и экспертов применялось ранее, рассматриваемая задача обладает существенными особенностями: учёт оценок объектов по многим критериям резко увеличивает число связей, накладываемых на объекты и экспертов при проецировании. Поэтому появляется реальная опасность внесения больших искажений в первоначальные расстояния между объектами в многомерном пространстве.

В связи с этим принята следующая процедура.

1. Первоначально вводятся лишь расстояния между точками, соответствующими членам ГПР. Они характеризуют степень совпадения мнений членов ГПР при оценке ими объектов по многим критериям.

2. Задача проецирования решается только для точек, представляющих членов ГПР, в соответствии с введенными расстояниями. Для эффективного отображения используется критерий минимизации суммы относительных разностей исходных расстояний между точками и расстояний на плоскости. Минимизация осуществляется методом сопряженного градиента.

3. Поэтапно производится «загрубление» исходной информации до достижения удовлетворительного совпадения мнений между членами ГПР либо до появления одной или нескольких коалиций. Степень совпадения мнений устанавливается при анализе результата проецирования на соответствующих рисунках.

4. При достижении совпадения мнений между членами ГПР или членами коалиций (ценой удаления части информации) для них вводятся расстояния между экспертом и объектом и решается заново задача проецирования совместно для всех объектов и всех членов ГПР (или членов коалиции). При этом предполагается, что предыдущие этапы (удаление части информации) обеспечили достаточную близость точек, представляющих членов ГПР, и поэтому проецирование не связано с чрезмерными искажениями.

При совместном проецировании используется критерий, аналогичный приведенному выше.

Среднее расстояние объекта от группы членов ГПР с совпадающими мнениями определяет его место в окончательном упорядочении объектов по мнению этой группы.

Определим перечень этапов анализа первоначальной информации («загрубления» информации).

1. Первоначально вводится простое метрическое (евклидово) расстояние между точками, соответствующими экспертам, как относительная сумма разностей между получаемыми от экспертов оценками объектов по многим критериям.

2. Вводится расстояние, учитывающее только совпадение точек зрения экспертов о превосходстве одного объекта над другим по каждому из критериев.

3. Если на шкалах критериев имеется по несколько оценок, то производится операция «объединения оценок», т.е. делается попытка объединить эти оценки в две, характеризующие наличие и отсутствие качества по данному критерию.

4. Определяются критерии, по которым противоречия между экспертами проявляются в наибольшей степени. Разногласия подсчитываются для каждой пары экспертов, а затем определяется коэффициент Tk, характеризующий степень несогласия экспертов по k-му критерию. Критерии ранжируются в соответствии с знанием коэффициента T_k.

На каждом из перечисленных этапов анализируется последовательность рисунков, характеризующих результат проецирования. При установлении достаточного сближения группы точек, представляющих членов ГПР, процесс удаления части информации прекращается. Если такое сближение не достигается после удаления на этапе 4 половины критериев, то результатом анализа является факт глубоких противоречий между членами ГПР. Можно предположить, что такой случай на практике встречается достаточно редко. Если анализ не выявляет общего согласия между членами ГПР, то он позволяет выявить коалиции.

После выделения согласных между собой членов ГПР для этой группы вводятся метрические расстояния между объектом и членом ГПР.

Полученная при проецировании ранжировка объектов характеризует среднее мнение членов данной группы. Степень совпадения взглядов внутри группы можно оценить с помощью коэффициента ранговой корреляции Кэндалла [16], сравнивая ранжировки членов группы и общую ранжировку.

8.9.3. Проведение конференции по принятию решений

Результаты анализа сообщаются всем членам ГПР при первом обсуждении рассматриваемого вопроса. Эти результаты способствуют более эффективной организации работы ГПР: в первую очередь обсуждаются расхождения мнений по оценкам отдельных критериев, в случае необходимости запрашивается дополнительная информация и т.д. Результаты анализа, представленные в наглядном виде, направляют дискуссию членов ГПР.

8.9.4 Практический пример

Разработанный метод был использован для организации деятельности комиссии, в задачу которой входило разделение представленных на конкурс результатов научно-исследовательских работ на две группы —лучшие (первая премия) и худшие (вторая премия) [10]. Использовались следующие критерии оценок результатов работ: новизна, важность, оригинальность, широта применения, практическое значение, уровень по отношению к другим работам.

На рисунок 31,а представлены первоначальные результаты «проецирования» на плоскость точек, характеризующих членов > комиссии, при использовании расстояния между их оценками объектов. Из рисунка видно намечающееся единство среди большой группы членов комиссии (семь из восьми), причем три из них (№ 3, 6, 7) образуют плотную группу. Далее было осуществлено «слияние» оценок объектов по критериям — результат представлен на рисунок 31,6. При этом было достигнуто практически полное совпадение мнений большинства членов комиссии. Этот результат был доложен консультантом перед началом обсуждения работ, представленных на конкурс. Было сообщено также, что лишь один член ГПР (№ 5) расходится во мнении с остальными в оценке четырех объектов по трем критериям. При обсуждении удалось убедить его пересмотреть первоначальные оценки. Обсуждение продолжалось значительно меньшее время, чем аналогичные заседания подобных комиссий, причем было принято намечавшееся решение большинства.

Выводы

1. Существует *множество* систем голосования. Наиболее известные из них - системы Кондорсе, Борда, правило большинства голосующих - кажутся справедливыми и убедительными с точки зрения здравого смысла. Однако

они приводят к нарушению рациональности. Парадоксы не возникают, если один из кандидатов набрал абсолютное большинство голосов.

2. Результаты К. Эрроу показывают, что в принципе невозможно найти систему голосования, которая была бы одновременно решающей, рациональной и демократической. Это означает, что применение демократических процедур голосования требует внимания и тщательного анализа результатов.
3. Широко распространенной на практике является задача принятия коллективных решений в малых группах: комиссиях, жюри и т. д. Эффективность работы таких групп могут существенно повысить методы анализа позиций членов группы и рациональная организация работы группы.

Список литературы

1. Гюйбо Д.Т. Теории общего интереса и логическая проблема агрегирования // Математические методы в социальных науках / Под ред. П. Лазерпс-фельда, Н. Генри. М.: Прогресс, 1973.
2. Tannenbaum P., Arnold R. Excursions in Modern Mathematics, NJ: Prentis-Hall Inc.1992.
3. Regenwetter M., Grofman B. Approval voting, Borda Winners and Condorcet Winners: Evidence from seven Elections // Management Science. 1998. V. 44. № 4.
4. Arrow K. J. Social Choice and individual values. N.Y.: Wiley, 1951.
5. Миркин Б.Г. Проблема группового выбора. М.: Наука, 1974.
6. Льюс Р. Д., Райфа Г. Игры и решения. М.: ИЛ, 1961.
7. Блэйр Д. Х., Поллак Р. Э. Рациональный коллективный выбор // В мире науки . 1983. №10.
8. Нейман Д. фон, Моргенштерн О. Теория игр и экономическое поведение. М.: Наука, 1970.
9. Philips L.D. Theory of Requisite Decision Models // K. Borcherding, B.Brehmer, C.Vlek, W. Wagenaar. Research perspectives on Decision Making under Uncertainty. North Holland, Amsterdam, 1984.

10. Ларичев О.И., Терехина А.Ю., Павельев В.В. Метод организации работы коллективного органа, принимающего решения // Перспективное планирование научных исследований и разработок / Под. ред. С. В.Емельянова. М.: Наука, 1974.
11.'Терехина А.Ю. Анализ данных методами многомерного шкалирования. М.: Наука, 1986.
12. Айвазян С.А., Бежаева З.И., Староверов О.В. Классификация многомерных наблюдений. М.: Статистика, 1974.
13. Иберла К. Факторный анализ.М., 1980.

Контрольные вопросы

Дайте определения следующих ключевых понятий:

1. Принцип Кондорсе и Парадокс Кондорсе
2. Правило большинства голосов
3. Метод Борда
4. Аксиомы Эрроу и теорема о невозможности
5. Попытки пересмотра аксиом
6. Принятие решений в малых группах
7. Конференции по принятию решений

9 Принятие решений и сети Петри

9.1 Основные понятия теории сетей Петри

Сети Петри предназначены для моделирования систем, которые состоят из множества взаимодействующих друг с другом компонент. При этом компонента сама может быть системой. Действиям различных компонент системы присущ параллелизм. Примерами таких систем могут служить вычислительные системы, в том числе и параллельные, компьютерные сети, программные системы, обеспечивающие их функционирование, а также экономические системы, системы управления дорожным движением, химические системы, и т. д.

Впервые сети Петри предложил немецкий математик Карл Адам Петри. Одной из областей применения сетей Петри является моделирование параллельных вычислений и многопроцессорных систем. Успешная работа в этом направлении требует хорошего знания сетей Петри и методов сетей Петри. Сети Петри — это математический аппарат моделирования динамических дискретных систем.

Сеть Петри - инструмент для моделирования динамических систем. Теория сетей Петри делает возможным моделирование системы математическим представлением ее в виде сети Петри, анализ которой помогает получить важную информацию о структуре и динамическом поведении моделируемой системы.

Возможно несколько путей практического применения сетей Петри при проектировании и анализе систем. В одном из подходов сети Петри рассматриваются как вспомогательный инструмент анализа. Здесь для построения системы используются общепринятые методы проектирования, затем построенная система моделируется сетью Петри, и построенная модель анализируется.

В другом подходе весь процесс проектирования и определения характеристик проводится в терминах сетей Петри. В этом случае задача заключается в преобразовании представления сети Петри в реальную информационную систему.

Несомненным достоинством сетей Петри является математически строгое описание модели. Это позволяет проводить их анализ с помощью современной вычислительной техники (в том числе с массово-параллельной архитектурой).

9.2 Структура сети Петри

Сеть Петри представляет собой двухдольный граф, состоящий из двух типов вершин - позиций и переходов, соединённых между собой дугами, вершины одного типа не могут быть соединены непосредственно. В позициях могут размещаться метки (маркеры), способные перемещаться по сети.

Событием называют срабатывание перехода, при котором метки из входных позиций этого перехода перемещаются в выходные позиции. События происходят мгновенно, разновременно при выполнении некоторых условий.

Сеть Петри N состоит из пяти компонентов: множества *позиций* P, множества *переходов* T, множества *входных функций* I(t) и множества *выходных функций* O(t) и начальной *маркировки* M(P). Входная функция $I(t_i)$ отображает связи от множество позиций в переход t_i, а выходная функция $O(t_i)$ отображает связи от перехода t_i в множество позиций [1].

Сеть Петри является пятёркой, N = <P, T, I, O, M>,

где **$P = \langle p_1, p_2, \ldots, p_n \rangle$** - множество позиций;

$T = \langle t_1, t_2, \ldots, t_m \rangle$ - множество переходов;

I(t) - множество входных функций;

O(t) - множество выходных функций;

M(P) – начальная маркировка.

9.3 Графическое представление сетей Петри

Сеть Петри может быть представлена в аналитическом или графическом виде.

Пример задания сети Петри в аналитическом виде представлен ниже.

$P = \{p_1, p_2, p_3, p_4, p_5, p_6\}$

$T = \{t_1, t_2, t_3, t_4, t_5\}$

$I(t_1) = p_1$

$I(t_2) = p_3$

$I(t_3) = p_2, p_3$

$I(t_4) = p_4, p_5, p_5, p_5$

$I(t_5) = p_2$

$O(t_1) = p_2, p_3$

$O(t_2) = p_3, p_5, p_5$

$O(t_3) = p_4, p_5$

$O(t_4) = p_4$

$O(t_5) = p_6$

$M(P) = \{3\ 1\ 2\ 2\ 4\ 0\}$

Эта же сеть Петри, представленная в графическом виде, приведена на рисунке 9.1. Граф сети Петри обладает двумя типами узлов. Кружок является позицией, а планка – переходом. Ориентированные дуги (стрелки) соединяют позиции и переходы, при этом некоторые дуги направлены от позиций к переходам, а другие – от переходов к позициям. Дуга, направленная от позиции P_i к переходу t_j, определяет позицию, которая является входом для перехода. Кратные входы в переход обозначаются перечёркиванием наклонной чертой соответствующей дуги и указанием количества этих дуг.

Маркировка М есть присвоение фишек (точек внутри позиций) позициям сети. Фишка изображается точкой в кружке, которая представляет позицию сети Петри. Количество и положение фишек при выполнении сетей Петри могут изменяться. Они используются для определения выполнения сети Петри.

Маркировка М сети Петри есть функция, отображающая множество позиций в множество неотрицательных целых чисел N.

$M: P \rightarrow N$.

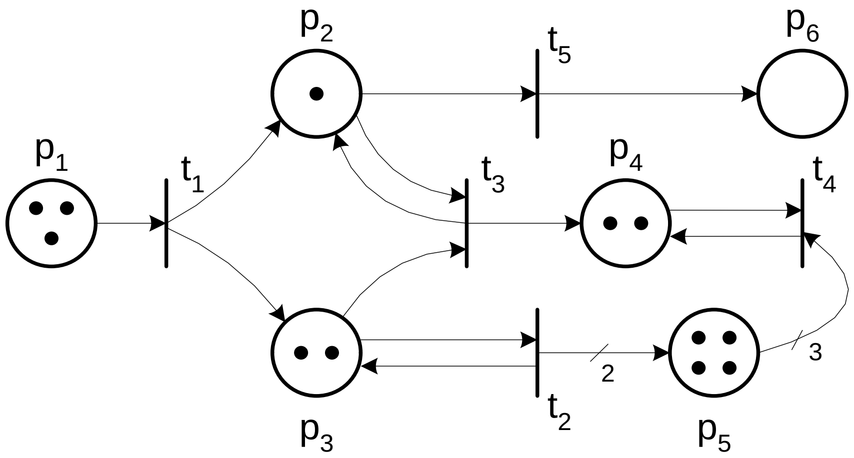
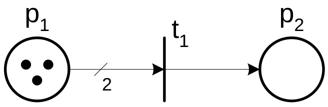


Рисунок 9.1- Графическое представление сети Петри

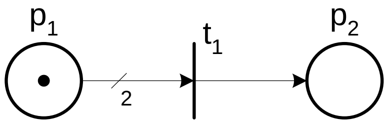
9.4 Выполнение сетей Петри и правила запуска

Под выполнением понимается последовательность путей запусков переходов, в результате которой фишки изымаются из входных позиций и перемещаются в выходные. Переход может запускаться только в том случае, когда он разрешен.

Переход называется *разрешенным к запуску*, если каждая из его входных позиций имеет число фишек по крайней мере равное числу входных дуг из каждой позиции в данный переход.



- запуск перехода **t1** разрешен, так как число фишек в позиции Р1 равно трём, а количество входных дуг две.



- запуск перехода **t1** запрещен, так как в позиции Р1 имеется одна фишка, а количество входных дуг две.

Переход tj сети Петри с маркировкой М может быть запущен всякий раз, когда он разрешен. В результате запуска разрешенного перехода tj образуется новая маркировка М’.

Правила запуска:

1) запуск переходов носит случайный характер;

2) при запуске перехода из его входных позиций изымается число фишек равное числу входных дуг;

3) фишки перемещаются в выходные позиции перехода в количестве равном числу дуг, идущих от перехода в выходные позиции.

Например, при запуске перехода t_1 (рисунок 9.1) из позиции p_1 изымается одна фишка и добавляется по одной фишке в позиции p_2 и p_3. При запуске перехода t_2 в полученной маркировке из позиции p_2 изымается одна фишка и добавляется две фишки в позицию p_5 и одна фишка в позицию p_3;

При запуске перехода t_3 изымается по одной фишке из позиций p_2 и p_3 и добавляется по одной фишке в позиции p_2 и p_4.

Переход, не имеющий ни одной входной позиции, называется *истоком*. Переход-исток всегда разрешен (рисунок 9.2).

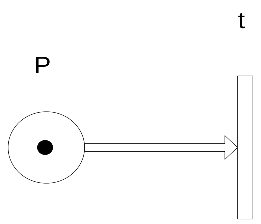


Рисунок 9.2.

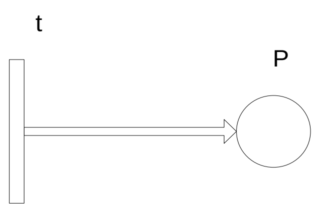


Рисунок 9.3.

Переход, не имеющий ни одной выходной позиции, называется *стоком*. Переход-сток всегда разрешен (рисунок 9.3).

Запуск стока приводит к изъятию фишек без порождения новых.

Пара, состоящая из позиции p и перехода t, называется *петлей*, если позиция P является входной и выходной для перехода t (рисунок 9.4).

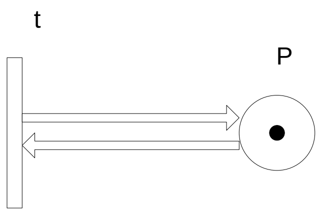


Рисунок 9.4.

Сеть Петри называется *однородной*, если в ней отсутствуют петли.

Сеть Петри называется *простой*, если в ней используется по одной дуге между позициями и переходами.

9.5 Примеры использования сетей Петри

Сеть Петри, моделирующая работу торгового автомата, приведена на рисунок 9.5. Переходам поставлено в соответствие опускание монеты достоинством 5 руб., 10 руб. в автомат или нажатие кнопки выбора Nuts, Twix.

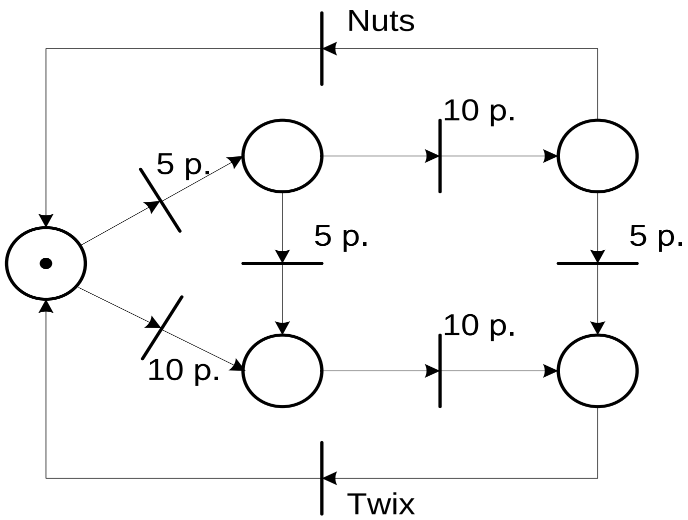


Рисунок 9.5. Модель торгового автомата

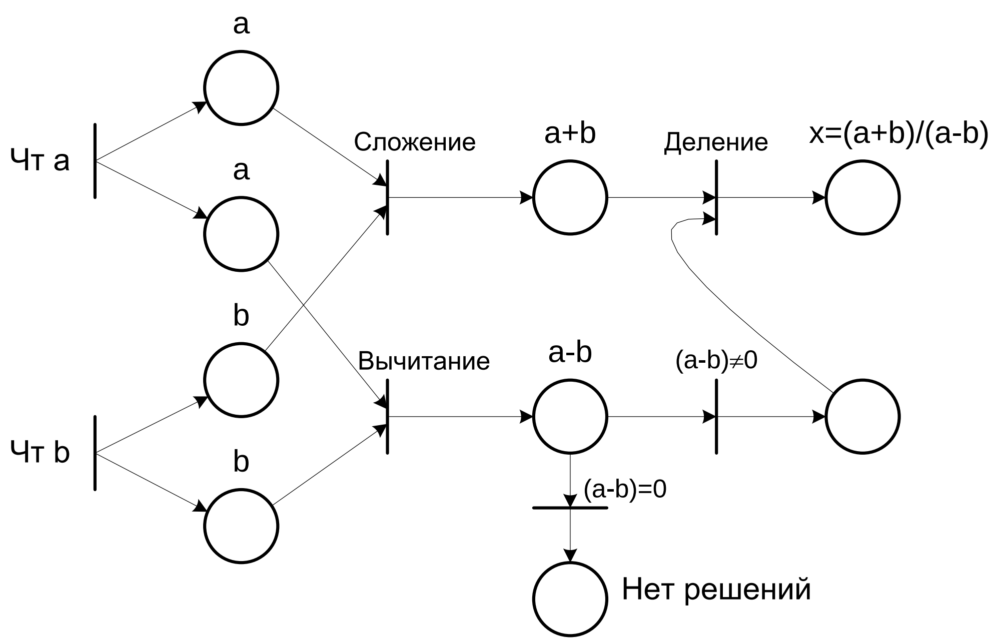


Рисунок 9.6. Модель решения алгебраического уравнения

Пример моделирования алгебраического уравнения представлен на рисунок 9.6.

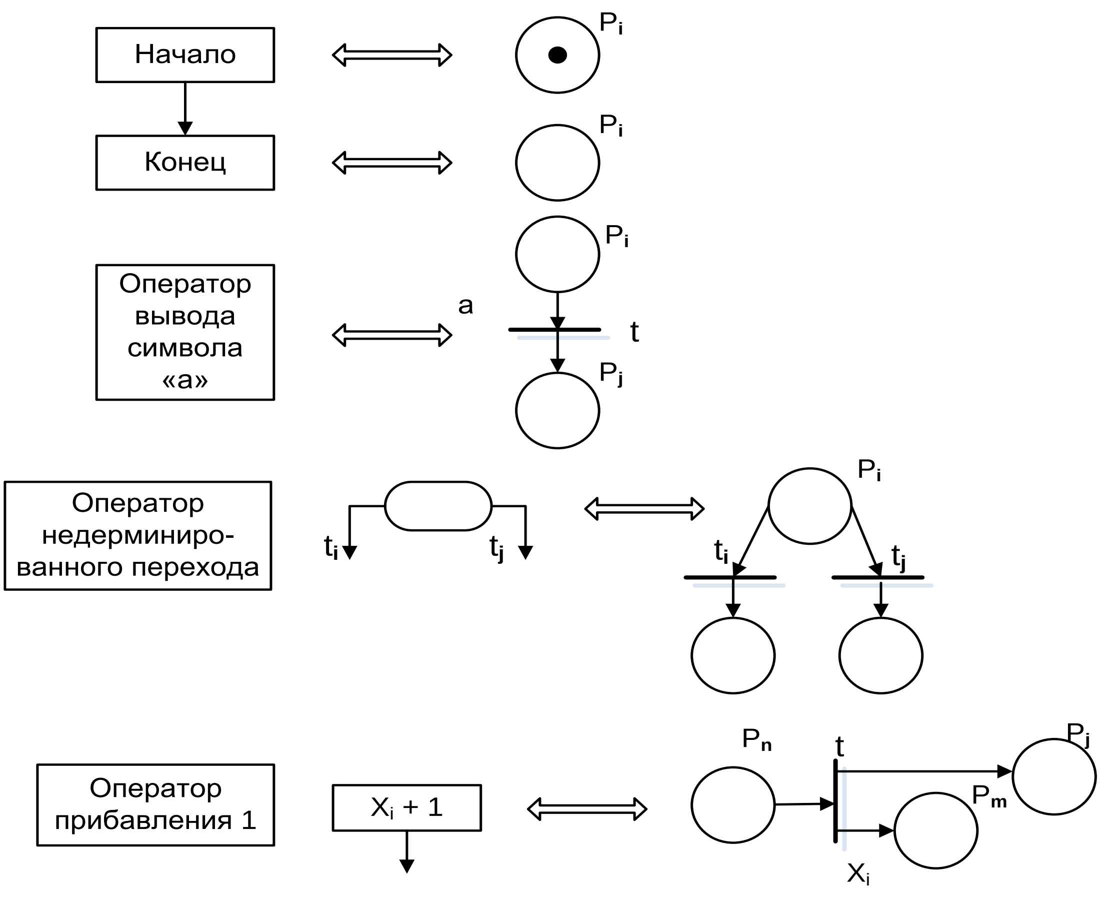


Рисунок 9.7- Модели основных операторов

Основные операторы алгоритмов решения задачи могут быть реализованы с помощью сети Петри. На рисунок 9.7 сопоставлены фрагменты сетей Петри и основных операторов.

Оператору начало соответствует позиция без входных дуг, а оператору конец – без выходных дуг.

Оператору вывода символа «а» соответствует переход, помеченный этим символом. Оператор недетерминированного перехода моделируется двумя переходами с общим конфликтным входом и разными выходами [2].

Оператор прибавления единицы $x_i{:=}x_i{+}1$ может быть представлен переходом с одной входной позицией и двумя выходными позициями. Позиция x_i соответствует счётчику, в котором накапливаются фишки.

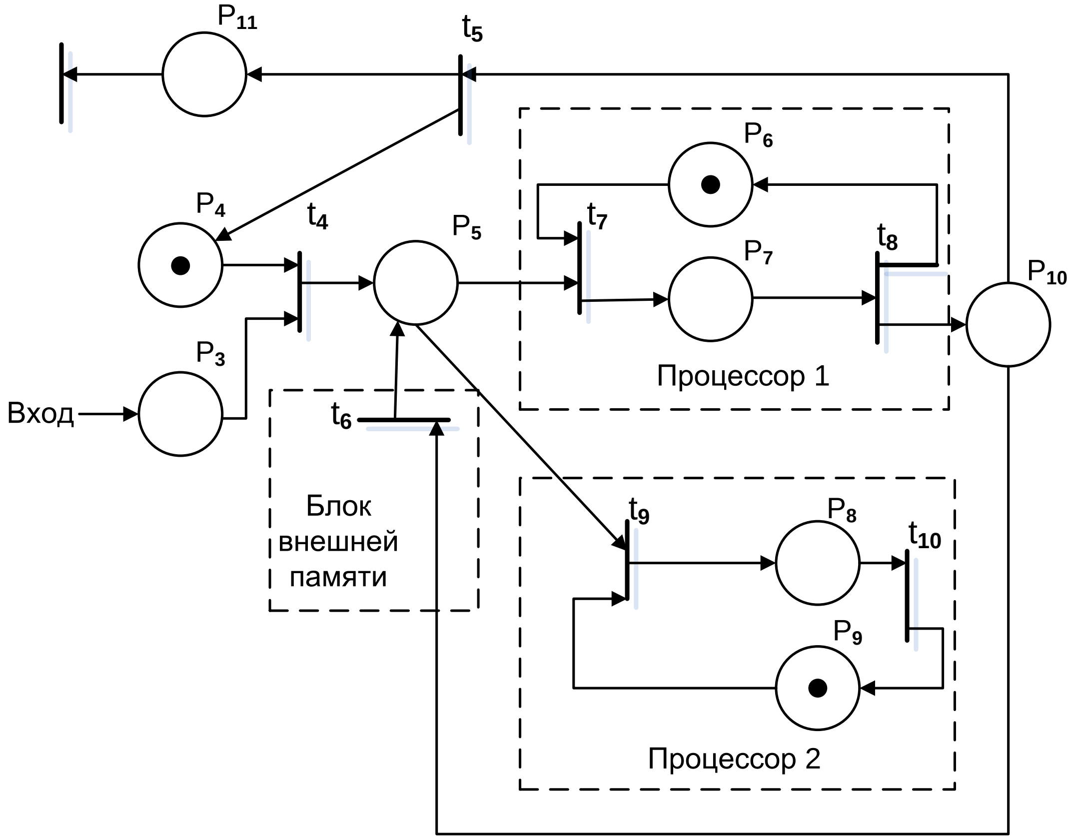


Рисунок 9.8 - Модель двухпроцессорной вычислительной системы

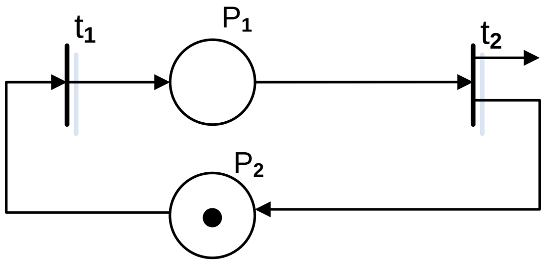


Рисунок 9.9. Модель блока формирования заданий

Каждой задаче, поступающей в систему, должен выделяться один блок памяти (позиция P4). Количество фишек в позиции P4 соответствует количеству блоков в системе. Например, при наличии 8 блоков памяти, в позиции P4 должно находиться 8 фишек. Во время решения задачи возможно обращение к блоку внешней памяти. При моделировании задаются законы распределения времени обработки порций информации в процессорах и вероятность q выхода задачи из системы между обращениями к к блоку внешней памяти. Задержки на обслуживание в процессорах 1 и 2, а так же в блоке внешней памяти реализуются переходами t8, t10, t6. Вероятность *q* учитывается при выходах из позиции P10. Формирование задачи, подаваемой на вход (рисунок 9.8) представлено моделью сети Петри на рисунок 9.9.

9.6 Свойства сетей Петри

9.6.1 Ограниченность и безопасность

Сеть Петри называется *k-ограниченной*, если для любой маркировки, достижимой из M_0, число фишек в любой позиции не превышает число k.

При k=1 сеть Петри называется безопасной, т.е. число фишек в любой позиции не превышает единицу. Пример безопасной сети приведён на рисунок 9.10.

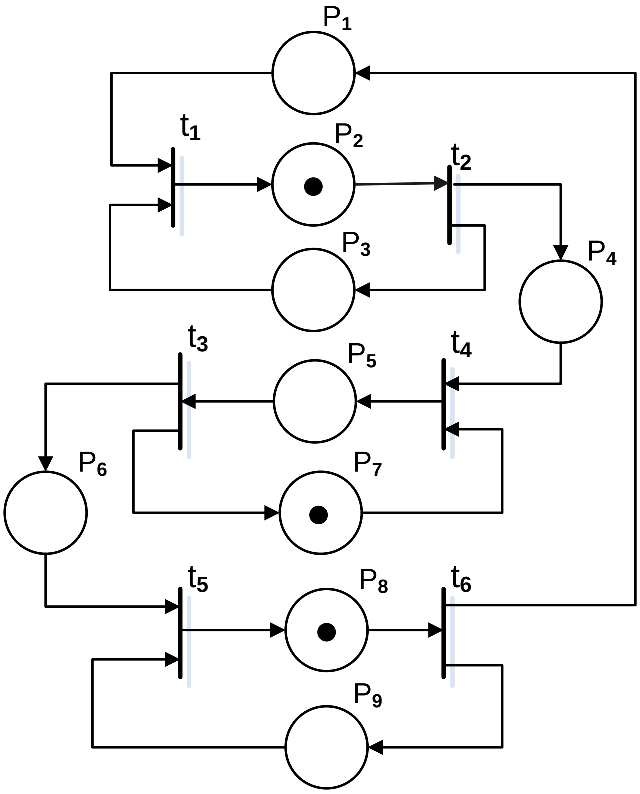
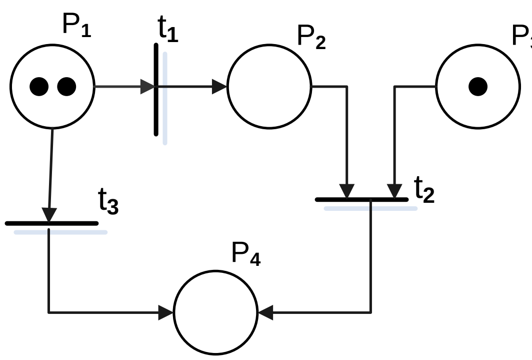


Рисунок 9.10 - Пример безопасной сети Петри

9.6.2 Достижимость

Проблема достижимости является центральной в теории сетей Петри. Маркировка M_n достижима из маркировки M_0, если существует последовательность запусков $\delta = M_0\ t_1\ M_1\ t_2 \ldots M_n$, приводящих от M_0 к M_n.



$M_0 = (2010)$

$M_1(t_1) = (1110)$

$M_2(t_2) = (1001)$

$M_3(t_3) = (0002)$

Рисунок 9.11.

В примере (рисунок 9.11) маркировка $M_3(t_3)$ = (0002) достижима из начальной маркировки M_0 = (2010) путём последовательного запуска переходов t1, t2, t3.

Это свойство можно ослабить. В результате проводится поиск не полной маркировки, а её подмножества (подмаркировки).

9.6.3 Активность

Сеть Петри называется *активной*, если независимо от достигнутой из M_0 маркировки для любого перехода существует последовательность запусков, приводящая к изменению маркировки.

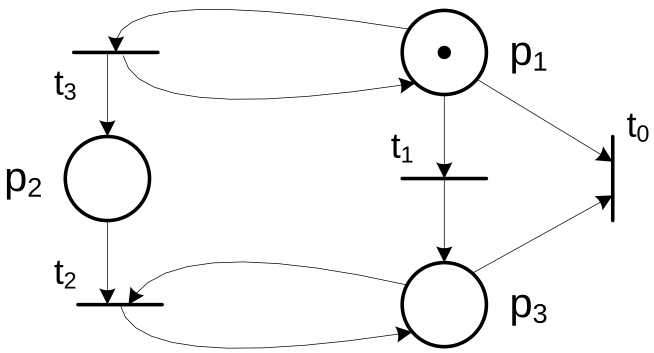


Рисунок 9.12.

Это жесткое свойство может быть ослаблено. Для этого вводится 5 свойств L – активности: L_0, L_1, L_2, L_3, L_4.

L_0 – активным называется переход, который не может срабатывать в сети Петри ни в какой последовательности запусков (на рисунок9.12 переход t_0).

L_1 – активным называется переход, который в любой последовательности запусков может сработать только один раз (на рисунок 9.12 переход t_1).

Переход называется *L_2 – активным*, если существует последовательность запусков из M_0, в которой он может сработать k раз (k > 1) (на рисунок 9.12 переход t_2).

Переход называется L_3 – *активным,* если существует последовательности запусков из M_0, в которой он может сработать бесконечное число раз (на рисунок 9.12 переход t_3).

Переход называется L_4 – *активным*, если это L_1 – активный переход для любой маркировки из M_0.

В ряде источников литературы указанное свойство называется *живостью* (возможность срабатывания любого перехода при функционировании моделируемого объекта).

9.6.4 Обратимость и базовое состояние

Сеть Петри *обратима*, если существует последовательность запусков, по которой можно вернуться к M_0.

Во многих задачах необязательно возвращаться в M_0, а достаточно вернуться в некоторое базовое состояние M*.

Маркировка M* называется *базовым состоянием*, если она достижима из маркировки M_n.

9.6.5 Покрываемость

Маркировка M в сети Петри *покрываема*, если существует маркировка M*, достижимая из M_0, такая что выполняется условие $M^*(P) \geq M(P)$ для любой позиции P.

9.6.6 Устойчивость

Сеть Петри называется *устойчивой*, если для двух любых разрешенных переходов запуск одного из них не приводит к запрещению другого.

9.6.7 Сохраняемость

При моделировании систем распределения ресурсов сетями Петри фишки представляют ресурсы. В результате запуска переходов фишки, представляющие ресурсы, не уничтожаются и не создаются, т.е. наблюдается своего рода сохранение.

Сеть Петри с начальной маркировкой M_0 называется *сохраняющей*[3] по отношению к вектору взвешивания позиций $W=(w_1, w_2, \ldots, w_n)$, если для всех маркировок M' выполняется выражение

$$\sum_{i=1}^{т} w_i M'(p_i) = \sum_{i=1}^{n} w_i M_0(p_i)$$

Здесь проведено взвешивание фишек путём присвоения веса позициям и вектор $W=(w_1, w_2, \ldots, w_n)$ определяет вес w_i для каждой позиции [3].

Другими словами – сеть называется строго сохраняющей, если в любой маркировке сумма фишек по всем позициям постоянна; более общим свойством является сохранение по отношению к некоторому ненулевому (и неединичному) вектору, на который умножается всякий вектор маркирования в целях проверки постоянства этого произведения.

9.7 Методы анализа сетей Петри

Методы анализа сетей Петри можно разбить на три основные группы:

- методы на основе дерева достижимости;
- применение матричного подхода к решению задачи достижимости;
- методы преобразования и декомпозиции.

9.7.1 Построение дерева достижимости

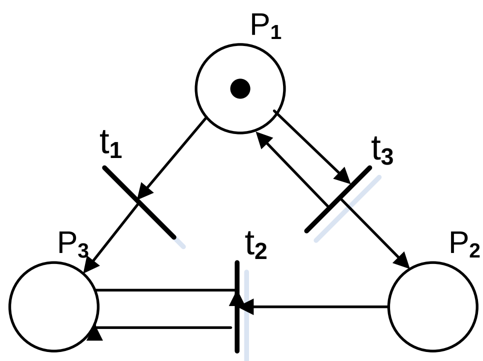


Рисунок 9.13.

Алгоритм построения дерева достижимости для графа сети Петри, приведённой на рисунок 9.13, указан ниже.

1. Обозначим начальную маркировку как корень дерева и присвоим ей метку «новая» (рисунок 9.14).

2. Зарисуем дуги, по которым могут быть произведены запуски переходов.

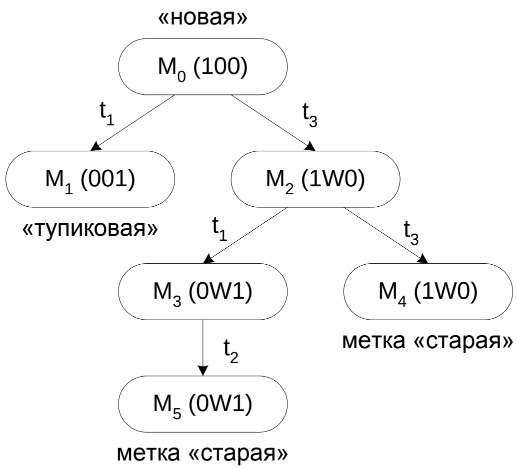


Рисунок 9.14. Дерево достижимости

Анализируем маркировки, полученные на втором ярусе:

M_1 – тупиковая, присваиваем метку «тупиковая»;

M_2 – при многократном запуске t_3 в p_2 накапливаются фишки, чтобы дерево не разрасталось, количество фишек в p_2 заменяется бесконечно большим числом W=W+1.

Анализ продолжается до тех, пока не будут просмотрены все возможные ситуации.

9.7.2 Алгебраические методы

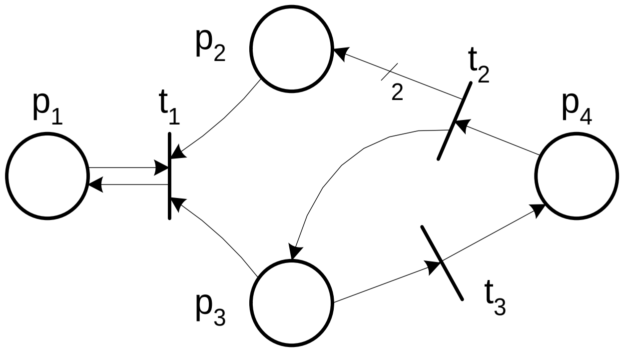


Рисунок 9.15

Для представления сети Петри в матричной форме используются матрицы входных и выходных функций и матрица инциденции.

Пример формирования матрицы входных и выходных функций для графа, представленного на рисунок 9.15, рассмотрен ниже.

Матрица входных функций:

$$D^+ = \begin{array}{c|cccc} & p_1 & p_2 & p_3 & p_4 \\ \hline t_1 & 1 & 1 & 1 & 0 \\ t_2 & 0 & 0 & 0 & 1 \\ t_3 & 0 & 0 & 1 & 0 \end{array}$$

Матрица выходных функций:

$$D^- = \begin{array}{c|cccc} & p_1 & p_2 & p_3 & p_4 \\ \hline t_1 & 1 & 0 & 0 & 0 \\ t_2 & 0 & 2 & 1 & 0 \\ t_3 & 0 & 0 & 0 & 1 \end{array}$$

Матрица инциденции вычисляется по формуле: $D = D^- - D^+$

$$D = \begin{array}{c|cccc} & p_1 & p_2 & p_3 & p_4 \\ \hline t_1 & 0 & -1 & -1 & 0 \\ t_2 & 0 & 2 & 1 & -1 \\ t_3 & 0 & 0 & -1 & 1 \end{array}$$

Физический смысл матрицы инциденции состоит в следующем.

Рассмотрим переход t_2.

-1 означает, что изымается одна фишка из позиции p_4.

1 – добавляется одна фишка в позицию p_3.

2 – добавляется две фишки в позицию p_2.

Для анализа достижимости маркировки M* из начальной маркировки M_0 требуется решить матричное уравнение вида

$$M^* = M_0 + x^*D, \qquad (9.1)$$

где x – последовательность запусков переходов сети Петри, приводящая к маркировке M* из маркировки M_0;

D – матрица инциденции.

<u>Пример:</u> Пусть *M* = (1, 8, 0, 1); M_0 = (1, 0, 1, 0).*

Подставив в (9.1) числовые значения получим

$$(1,8,0,1) = (1,0,1,0) + x * \begin{vmatrix} 0 & -1 & -1 & 0 \\ 0 & 2 & 1 & -1 \\ 0 & 0 & -1 & 1 \end{vmatrix}$$

После переноса значения $M_0 = (1, 0, 1, 0)$ в левую часть имеем

$$(0,8,-1,1) = x * \begin{vmatrix} 0 & -1 & -1 & 0 \\ 0 & 2 & 1 & -1 \\ 0 & 0 & -1 & 1 \end{vmatrix}$$

В результате решения матричного уравнения находим *x = (0, 4, 5).*

Это означает, что маркировка M* достижима из маркировки M_0; при этом переход t_1 запускается 0 раз, t_2 - 4 раза, t_3 - 5 раз.

Решение уравнения не дает ответа на вопрос: «в какой последовательности следует запускать переходы?». Если решения нет, то маркировка M* не достижима из маркировки M_0. Для графа сети Петри, приведенного на рисунок 1.14, к маркировке M* приводит следующая последовательность запусков: $\delta = t_3\ t_2\ t_3\ t_2\ t_3\ t_2\ t_3\ t_2\ t_3$.

9.8 Разновидности сетей Петри

9.8.1 Сети Петри с ингибиторными (запрещающими) дугами.

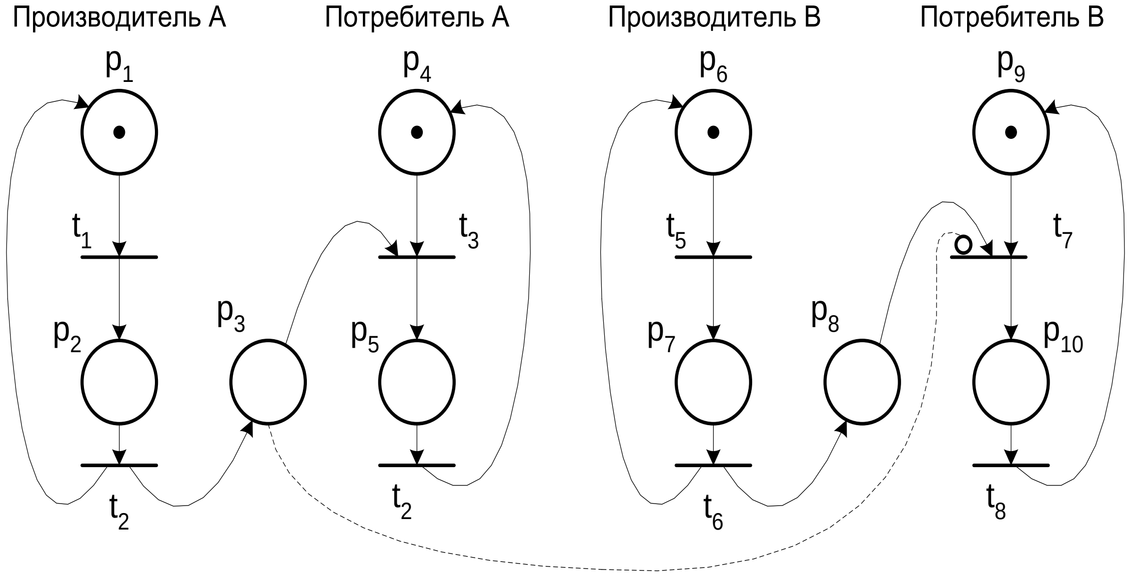


Рисунок 9.16 –Сеть Петри с ингибиторными дугами

В этой модели потребитель А имеет приоритет по сравнению с потребителем В. Если буфер А полон и в позиции p_3 фишка, она блокирует запуск перехода t_7. Запуск перехода t_7 возможен, если в позиции p_3 нет фишки (рисунок 9.16).

По ингибиторной дуге фишки не передаются, она служит только для управления.

Сети Петри, содержащие ингибиторные дуги, называются *расширенными* сетями Петри или *ингибиторными* сетями Петри.

9.8.2 Приоритетные сети Петри

При описании функционирования сетей Петри отмечался недетерминизм следующего рода: если несколько переходов могут сработать, то срабатывает любой из них [2]. В реальных дискретных системах имеют место ситуации, когда из двух готовых работать устройств требуется запустить сначала одно, а затем второе. Другими словами, одно из устройств имеет приоритет на запуск перед другими в том случае, если оба готовы работать. Такие ситуации не моделируются в сетях Петри в силу принятого в них правила срабатывания нескольких готовых к запуску переходов.

Эту задачу решают приоритетные сети Петри. Для этого вводится множество приоритетов PR, элементы которого частично упорядочены некоторым отношением (меньше или равно). Каждому переходу t сети Петри устанавливается его приоритет $pr(ti) \in PR$.

Правило срабатывания перехода модифицируется и дополняется условием: переход ti срабатывает в первую очередь, если для любого другого перехода tj этой сети приоритет срабатывания ниже, $pr(tj) \leq pr(ti)$.

Другими словами, если несколько переходов могут сработать, то срабатывает тот переход, приоритет которого выше. Приоритеты срабатывания могут динамически изменяться в процессе моделирования. Такую модификацию сети Петри называют сетью *с приоритетами или приоритетной сетью Петри.*

На рисунок 9.17 приведён пример простой сети с приоритетами. В качестве приоритетов можно выбрать множество целых чисел. Здесь могут сработать оба перехода, но поскольку $pr(t2) \leq pr(t1)$, то срабатывает только переход t1.

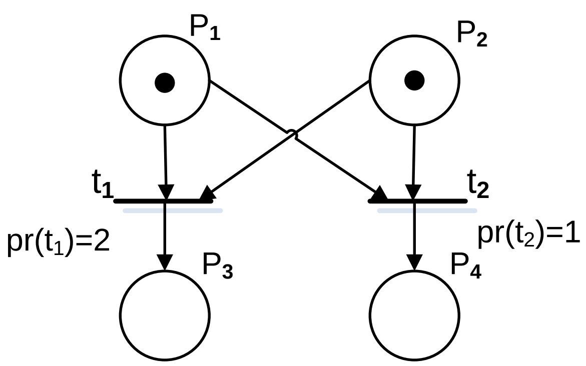


Рисунок 9.17- Сеть Петри с приоритетами

9.8.3 Временная сеть Петри

Время может быть сопоставлено различным сетевым элементам: местам, переходам, дугам, фишкам, шагам (множествам параллельных переходов). При этом различают временные ограничения, сопоставленные некоторому элементу временной сети Петри, и временные счетчики, введенные в модель для контроля локального или глобального времени. Временная информация может быть представлена как одним числом (что соответствует дискретному представлению времени), так и интервалом (что соответствует непрерывному представлению времени). В модели либо разрешается только одиночное срабатывание перехода, либо предполагается принцип максимального шага срабатывания (т.е. одновременного срабатывания максимально возможного количества параллельных переходов). Отдельным вопросом является наличие в модели автопараллелизма переходов (одновременного срабатывания нескольких экземпляров одного и того же перехода), а также возможность наличия в месте фишек с различными временными характеристиками. Кроме того, в разных моделях либо позволяют, либо запрещают "умирание" фишек по истечении заданного времени и т.п. Наиболее популярными стали модели, в которых временные характеристики связаны с переходами, а именно дискретно-временная модель Рамхандани-Штарке, и непрерывно-временная модель Мерлина. В модели Рамхандани с каждым переходом связывается некоторое (как правило, целочисленное) значение, соответствующее длительности срабатывания перехода. Модель Мерлина сопоставляет каждому переходу пару неотрицательных чисел, которые указывают на наиболее раннее и наиболее позднее времена, когда может сработать переход. По сравнению с другими временными сетевыми моделями временные сети Мерлина являются наиболее удобной и выразительной моделью для описания и изучения параллельных систем реального времени.

Временная сеть Петри определяется шестеркой:

$N = <P, T, I, O, M, \beta>$

Множество β связывает с каждой позицией сети Петри временные параметры. Фишка в позиции p_i становится готовой к запуску ее выходного перехода только после времени τ_i, заданного для данной позиции

$\tau_i = \beta(p_i)$.

Возможно задание временной сети Петри через задержку переходов

$\tau_i = \beta(t_i)$.

Переходы t_i обладают весом $\beta(t)_i$, определяющим продолжительность срабатывания (задержку).

9.8.4 Стохастическая сеть Петри характеризуется случайными задержками.

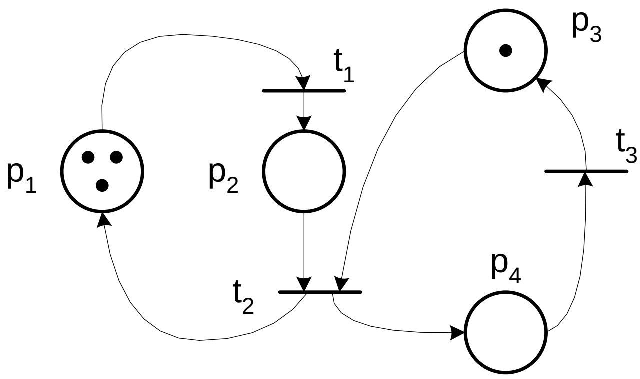


Рисунок 9.18

Переход t_1 на рисунок 9.18 имитирует отказ блока, t_2 – замену блока, t_3 – ремонт блока.

$M_0 = (3, 0, 1, 0)$ соответствует наличию в системе трех рабочих блоков и одного запасного. Переходам соответствуют случайные временные задержки:

t_1 – интервал времени между отказами;

t_2 – время замены;

t_3 – время восстановления отказавшего блока.

9.8.5 Функциональная сеть Петри

Такая сеть отражает не только последовательность событий, но и процессы обработки потока данных. Для этого в описание каждого перехода добавляется функция обработки данных.

С помощью функциональных сетей Петри можно моделировать разнообразные вычислительные системы, производить статистическую обработку результатов моделирования и отображать различные алгоритмы функционирования.

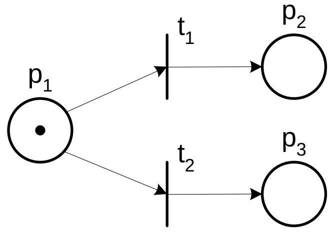
В функциональных сетях Петри можно использовать задержки, которые определяются как функции некоторых аргументов. Например, количества фишек в каких-либо позициях или состояния некоторых переходов.

Фишки могут быть различных типов, обозначаемых цветами. В этом случае тип метки может быть использован как аргумент в функциональных сетях.

9.8.6 Автоматная сеть Петри

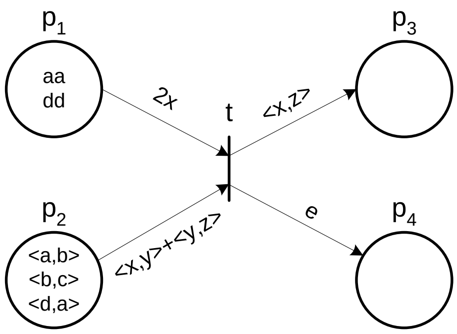
Каждый переход имеет только один вход и один выход. Число маркеров в такой сети постоянно. Такие сети по своим возможностям эквивалентны конечным автоматам.

***9.8.7* Цветная или раскрашенная сеть Петри** используется в тех случаях, когда требуется отличать друг от друга некоторые группы фишек.



Например, при моделировании гибких произвольных систем фишки могут отображать детали разных фишек.

Для разнотипных деталей вводятся разноцветные фишки со своими правилами запуска переходов. Например, красные фишки запускают переход t_1, а зеленые – t_2.



<u>Пример:</u>

Начальная маркировка:

- p_1 содержит четыре цветных фишки: две цвета ***a*** и две цвета ***d***;

- p_2 содержит три упорядоченных пары цветных фишек: <***a***,***b***>, <***b***,***c***>, <***d***,***a***>;

- p_3 и p_4 не имеют фишек.

Названия дуг определяют сколько фишек и какого цвета будут удалены и появятся вновь:

- p_1 теряет две фишки цвета ***x***;

- p_2 теряет две фишки цветов <***x***,***y***> и <***y***,***z***>;

- p_3 получает одну цветную фишку <***x***,***z***>;

- p_4 получает одну цветную фишку ***e*** (константу).

x, y, z – переменные

a, b, c, d, e - константы

Переход t называется разрешенным к запуску, если в каждой входной позиции существует достаточное число фишек правильного цвета. Под *правильным цветом* понимается такой цвет, при котором существуют непротиворечивые варианты подстановок констант в переменные.

Например, существует подстановка {***a*** | ***x***, ***b*** | ***y***, ***c*** | ***z***}. Тогда после запуска перехода из p_1 изымаются две фишки ***a***, из p_2 - <***a***,***b***> и <***b***,***c***>, в p_3 перемещается <***a***,***c***> и в p_4 – ***e***.

Другая подстановка: {***d*** | ***x***, ***a*** | ***y***, ***b*** | ***z***}.

9.8.8 WF-сети

WF-сети - подкласс сетей Петри называемый также сетями потоков работ. Формализм WF-сетей введён Вил Ван Дер Аальстом (англ. *Wil Van Der Aalst*) для моделирования потоков работ в workflow-системах.

Сеть Петри Pn = (P,T,F) называется сетью потоков работ (WF-сетью), если выполняются следующие условия:

- существует только одна исходная позиция P_{in} , такая что отсутствуют переходы, входящие в P_{in};
- существует только одна конечная позиция P_{out}, такая что отсутствуют переходы выходящие из P_{out};
- каждый узел данной сети расположен на пути от P_{in} к P_{out}.

WF-сети используются для проверки графов потоков работ на наличие таких структурных конфликтов, как «тупики» (англ. *deadlocks*) и «недостатки синхронизации» (англ. *lack of synchronization*).

Структурные конфликты отсутствуют, если WF-сеть является бездефектной. Свойство бездефектности или правильной завершаемости соответствует следующим требованиям:

- конечная позиция P_{out} достижима при любой последовательности запуска переходов от позиции P_{in};
- WF-сеть не содержит лишних позиций (которые никогда не будут выполнены);
- при достижении конечной позиции данной сети не должно оставаться фишек в промежуточных позициях.

Свойство бездефектности соответствует двум свойствам сетей Петри — живости и ограниченности, рассмотренным выше.

Контрольные вопросы и упражнения

1. Сеть Петри может ли быть представлена двудольным ориентированным графом и мультиграфом?
2. Правила запуска перехода-истока.
3. Правила запуска перехода-стока.
4. Является ли сеть Петри однородной, если содержит дугу, входящую в переход от позиции Pi, и дугу, выходящую из данного перехода в позицию Pi?
5. Может ли быть однородная сеть Петри простой?
6. Сохраняется ли суммарное число фишек в сети Петри после запуска одного, двух переходов?
7. После запуска перехода t3 (рисунок1.1) возможен ли запуск перехода t5?
8. После двукратного запуска перехода t3 (рисунок1.1) возможен ли запуск перехода t2?
9. Возможен ли запуск перехода t7 на рисунок1.7?
10. Возможен ли запуск перехода t5 на рисунок1.7?
11. Построить граф сети Петри по формальному описанию. P={p1,p2,p3,p4,p5} T={t1,t2,t3,t4} I(t1)=p1 I(t2)=p2, p4, p4 I(t3)=p3, p3, p2 I(t4)=p5, p5 O(t1)=p1, p2, p5, p5 O(t2)=p3,p3 O(t3)= p4,p4,p4 O(t4)=p4, p4, p4 M(p1, p2, p3, p4, p5)=21220.
12. Описать последовательный запуск переходов t1,t2,t3,t4 и перемещения фишек для графа сети Петри в упражнении 11 (ответ: M(t4)=20260).
13. Разработать граф сети Петри по формальному описанию. P={p1,p2,p3,p4,p5} T={t1,t2,t3,t4} I(t1)=p1 I(t2)=p2,p4,p4 I(t3)=p3,p3 I(t4)=p5,p5,p5 O(t1)=p2,p2,p5,p5,p5 O(t2)=p3,p3,p3,p3 O(t3)= p4,p4,p4 O(t4)=p4,p4,p4 M(p1,p2,p3,p4,p5)=23330

14. Описать последовательный запуск переходов t1,t2,t3,t4 и перемещения фишек для графа сети Петри в упражнении 13 (ответ: M(t4)=13543).

15. Разработать граф сети Петри по формальному описанию. P={p1,p2,p3,p4,p5} T={t1,t2,t3,t4} I(t1)=p1 I(t2)=p2,p4,p4 I(t3)=p3,p3 I(t4)=p5,p5,p5 O(t1)=p2,p2,p5,p5,p5 O(t2)=p3 O(t3)= p4,p4,p4 O(t4)=p4,p4,p4 M(p1,p2,p3,p4,p5)=13330. Какая будет маркировка после запуска перехода t2? (ответ: M(t2)=12410).

16. Описать последовательный запуск переходов t1,t2,t3,t4 и перемещения фишек для графа сети Петри в упражнении 15. (ответ: M(t4)=04270).

17. Может ли быть k-ограниченная сеть Петри безопасной?

18. Может ли быть обратимая сеть Петри безопасной?

19. Может ли быть устойчивая сеть Петри безопасной?

20. Построить 3-х ярусное дерево достижимости для графа сети Петри, представленной формальным описанием: P={p1,p2,p3,p4,p5} T={t1,t2,t3,t4} I(t1)=p1 I(t2)=p2,p4,p4 I(t3)=p3,p3,p2 I(t4)=p5,p5,p5 O(t1)=p1,p5,p5 O(t2)=p3,p3 O(t3)= p4,p4,p4 O(t4)=p4,p4,p4 M(p1,p2,p3,p4,p5)=21220.

21. Построить 3-х ярусное дерево достижимости для графа сети Петри, представленной формальным описанием: P={p1,p2,p3,p4,p5} T={t1,t2,t3,t4} I(t1)=p1 I(t2)=p2,p4,p4 I(t3)=p3,p3 I(t4)=p5,p5,p5 O(t1)=p2,p2,p5,p5,p5 O(t2)=p3,p3,p3,p3 O(t3)= p4,p4,p4 O(t4)=p4,p4,p4 M(p1,p2,p3,p4,p5)=23330.

22. Определить L-активность каждого перехода сети Петри, приведённой в п.20 и 21.

23. Является ли данная сеть Петри устойчивой, покрываемой и обратимой. Сеть Петри задана в п.21?

24. Составить матрицы входных, выходных функций и инциденции для сети Петри (см. п.20 «Контрольные вопросы и упражнения») и матричное уравнение достижимости маркировки $М_к$ (12221) из начальной маркировки (23330).

25. Решить матричное уравнение средствами MathCAD и определить достижимость маркировки $M_к$ (12221) из начальной маркировки (23330).

26. В каком случае может быть запущен переход t3 на графе сети Петри, представленном на рисунок 9.19.

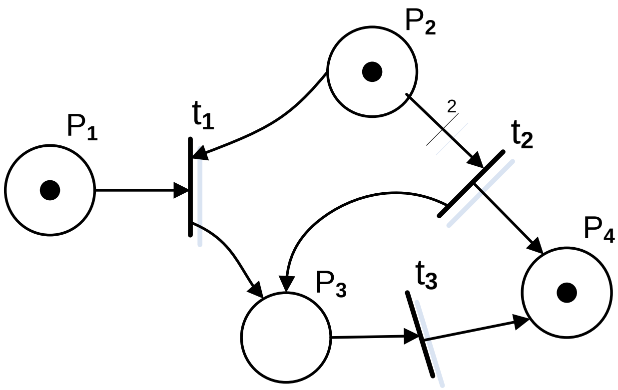


Рисунок 9.19 - Пример графа сети Петри

Список сокращений и обозначений

ГА- генетические алгоритмы

ГПР – группа принимающая решение

ЛПР –лицо принимающее решение