

**М.Б.Бровкова**

**СИСТЕМЫ  
ИСКУССТВЕННОГО  
ИНТЕЛЛЕКТА  
В МАШИНОСТРОЕНИИ**

**Учебное  
пособие**

**2004**

**Министерство образования и науки Российской Федерации**

**Саратовский государственный технический университет**

**М.Б.Бровкова**

## **СИСТЕМЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В МАШИНОСТРОЕНИИ**

*Допущено Учебно-методическим объединением вузов по образованию в области автоматизированного машиностроения (УМО АМ) в качестве учебного пособия для студентов высших учебных заведений, обучающихся по направлениям подготовки дипломированных специалистов – «Конструкторско-технологическое обеспечение машиностроительных производств»; «Автоматизированные технологии и производства»*

**Саратов 2004**

УДК 681.3.06

ББК 32.81

Б 88

Рецензенты:

Институт проблем точной  
механики и управления РАН (директор чл.-корр. РАН А.Ф.Резчиков);  
заведующий кафедрой «Станки» Московского государственного  
технологического университета «Станкин», д.т.н., проф. В.В.Бушуев

Одобрено  
редакционно-издательским советом  
Саратовского государственного  
технического университета

Бровкова М.Б.

Б 88 Системы искусственного интеллекта в машиностроении: Учеб.  
пособие. Саратов: Сарат. гос. техн. ун-т, 2004. 119 с.

*ISBN 5-7433-1384-9*

Содержит материалы, посвященные интеллектуализации  
машиностроительного производства, в том числе общие сведения о  
методах искусственного интеллекта, определении параметров  
качества детали в механообработке на основе нейронных сетей,  
применении генетических алгоритмов для оптимизации  
управляющих воздействий.

Предназначается для студентов специальностей 120200,  
120100, 210200 и может быть рекомендовано студентам других  
специальностей при изучении курсов, содержащих разделы по  
системам искусственного интеллекта машиностроительного  
профиля.

УДК 681.3.06

ББК 32.81

© Саратовский государственный  
технический университет, 2004  
© Бровкова М.Б., 2004

*ISBN 5-7433-1384-9*

## ВВЕДЕНИЕ

Высокие показатели качества, новые требуемые свойства современных машин и агрегатов в аэрокосмической, электронной, автомобильной и других областях машиностроения обуславливают применение прецизионных узлов и деталей. При этом можно отметить расширение номенклатуры изделий из специальных труднообрабатываемых сплавов и материалов, механообработка которых имеет специфические особенности. В таком сочетании задач технологические процессы автоматизированных производств должны быть обеспечены оборудованием высокого класса точности и с высокой стабильностью характеристик функционирования. На эти задачи направлены работы по совершенствованию конструкции оборудования и перспективные работы по использованию современных информационных технологий, обеспечивающих диагностирование, идентификацию состояния и оптимизацию режимов функционирования, в том числе на основе искусственного интеллекта.

В задачи предлагаемого учебного пособия входит освещение общих сведений о методах искусственного интеллекта, применение интеллектуальных методов для объектов машиностроения и изложение результатов авторских исследований в этой предметной области.

Технологическое оборудование с позиций теории управления является многоконтурной системой с нелинейными элементами и нестационарными свойствами. К этому следует добавить низкую «наблюдаемость» отдельных процессов и функциональных блоков. В процессе эксплуатации оборудования изменяется его динамическое состояние не только при переходе от одной операции к другой, но и при однотипных операциях технологического процесса. Априорно учесть эти изменения практически невозможно. Этим и определяется актуальность применения методов искусственного интеллекта для машиностроительного производства.

Предлагаемые методы могут быть использованы для задач обеспечения качества формообразования в процессе изнашивания оборудования, инструментов и других элементов технологической системы; для оптимизации режимов резания. Это является весьма актуальным в настоящее время из-за спада производства и ограниченных финансовых возможностей. С самого начала развития теории и практики резания исследователи и специалисты по эксплуатации оборудования стремились к увеличению периода стойкости режущего инструмента и формировали соответствующие критерии оптимизации на этой основе. По мере усложнения задач из разряда ограничений в разряд критериев оптимизации стали переводить такие параметры как качество микро- и макрорельефа вновь образованной поверхности. При этом в основном использовались эмпирические формулы, пригодные на практике лишь для отдельных реализаций технологических процессов, так как они не могли учитывать всего многообразия состояний заготовки, инструментов, технологической жидкости, станка и их сочетаний. При использовании гибких станочных модулей для всего спектра реализуемых технологических операций и их сочетаний решение задач оптимизации, улучшения качества деталей без учета каждый раз фактического состояния технологической системы, как было отмечено выше, вообще не реально.

В представленной работе обосновывается необходимость применения интеллектуальных методов для решения таких задач. Представлены результаты применения интеллектуального программного обеспечения для машиностроительного производства.

# 1. ВВЕДЕНИЕ В ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ

## 1.1. Краткая история развития искусственного интеллекта

Самыми первыми интеллектуальными задачами, которые стали решаться при помощи ЭВМ, были логические игры (шашки, шахматы), доказательство теорем. Хотя, правда, здесь надо отметить еще кибернетические игрушки типа «электронной мыши» Клода Шеннона, которая управлялась сложной релейной схемой. Эта мышка могла «исследовать» лабиринт и находить выход из него. Кроме того, помещенная в уже известный ей лабиринт, она не искала выход, а сразу же, не заглядывая в тупиковые ходы, выходила из лабиринта [62].

Американский кибернетик А. Самуэль составил для вычислительной машины программу, которая позволяет ей играть в шашки, причем в ходе игры машина обучается или, по крайней мере, создает впечатление, что обучается, улучшая свою игру на основе накопленного опыта. В 1962 г. эта программа сразилась с Р. Нили, сильнейшим шашкистом в США, и победила.

Каким образом машине удалось достичь столь высокого класса игры? Естественно, что в машину были программно заложены правила игры так, что выбор очередного хода был подчинен этим правилам. На каждой стадии игры машина выбирала очередной ход из множества возможных ходов согласно некоторому критерию качества игры. В шашках (как и в шахматах) обычно невыгодно терять свои фигуры, и, напротив, выгодно брать фигуры противника. Игрок (будь он человек или машина), который сохраняет подвижность своих фигур и право выбора ходов и в то же время держит под боем большое число полей на доске, обычно играет лучше своего противника, не придающего значения этим элементам игры. Описанные критерии хорошей игры сохраняют свою силу на протяжении всей

игры, но есть и другие критерии, которые относятся к отдельным ее стадиям — дебюту, миттэндшпилю, эндшпилю.

Разумно сочетая такие критерии (например, в виде линейной комбинации с экспериментально подбираемыми коэффициентами или более сложным образом), можно для оценки очередного хода машины получить некоторый числовой показатель эффективности — оценочную функцию. Тогда машина, сравнив между собой показатели эффективности очередных ходов, выберет ход, соответствующий наибольшему показателю. Подобная автоматизация выбора очередного хода не обязательно обеспечивает оптимальный выбор, но все же это какой-то выбор, и на его основе машина может продолжать игру, совершенствуя свою стратегию (образ действия) в процессе обучения на прошлом опыте. Формально обучение состоит в подстройке параметров (коэффициентов) оценочной функции на основе анализа проведенных ходов и игр с учетом их исхода.

По мнению А. Самуэля, машина, использующая этот вид обучения, может научиться играть лучше, чем средний игрок, за относительно короткий период времени.

Можно сказать, что все эти элементы интеллекта, продемонстрированные машиной в процессе игры в шашки, сообщены ей автором программы. Отчасти это так. Но не следует забывать, что программа эта не является «жесткой», заранее продуманной во всех деталях. Она совершенствует свою стратегию игры в процессе самообучения. И хотя процесс «мышления» у машины существенно отличен от того, что происходит в мозгу играющего в шашки человека, она способна у него выиграть.

Ярким примером сложной интеллектуальной игры до недавнего времени являлись шахматы. В 1974 г. состоялся международный шахматный турнир машин, снабженных соответствующими программами. Как известно, победу на этом турнире одержала советская машина с шахматной программой «Каисса».

Почему здесь употреблено «до недавнего времени»? Дело в том, что недавние события показали, что несмотря на довольно

большую сложность шахмат и невозможность, в связи с этим, произвести полный перебор ходов, возможность перебора их на большую глубину, чем обычно, очень увеличивает шансы на победу. К примеру, по сообщениям в печати, компьютер фирмы IBM, победивший Каспарова, имел 256 процессоров, каждый из которых имел 4 Гб дисковой памяти и 128 Мб оперативной. Весь этот комплекс мог просчитывать более 100'000'000 ходов в секунду. До недавнего времени редкостью был компьютер, могущий делать такое количество целочисленных операций в секунду, а здесь мы говорим о ходах, которые должны быть сгенерированы и для которых просчитаны оценочные функции. Хотя, с другой стороны, этот пример говорит о могуществе и универсальности переборных алгоритмов.

В настоящее время существуют и успешно применяются программы, позволяющие машинам играть в деловые или военные игры, имеющие большое прикладное значение. Здесь также чрезвычайно важно придать программам присущие человеку способности к обучению и адаптации. Одной из наиболее интересных интеллектуальных задач, также имеющей огромное прикладное значение, является задача обучения распознаванию образов и ситуаций. Решением ее занимались и продолжают заниматься представители различных наук — физиологи, психологи, математики, инженеры. Такой интерес к задаче стимулировался фантастическими перспективами широкого практического использования результатов теоретических исследований: читающие автоматы, системы искусственного интеллекта (ИИ), ставящие медицинские диагнозы, проводящие криминалистическую экспертизу и т. п., а также роботы, способные распознавать и анализировать сложные сенсорные ситуации.

В 1957 г. американский физиолог Ф. Розенблатт предложил модель зрительного восприятия и распознавания — перцептрон. Появление машины, способной обучаться понятиям и распознавать предъявляемые объекты, оказалось чрезвычайно интересным не только физиологам, но и представителям других областей знания, и



породило большой поток теоретических и экспериментальных исследований.

Перцептрон или любая программа, имитирующая процесс распознавания, работают в двух режимах: в режиме обучения и в режиме распознавания. В режиме обучения некто (человек, машина, робот или природа), играющий роль учителя, предъявляет машине объекты и о каждом из них сообщает, к какому понятию (классу) он принадлежит. По этим данным строится решающее правило, являющееся, по существу, формальным описанием понятий. В режиме распознавания машине предъявляются новые объекты (вообще говоря, отличные от ранее предъявленных), и она должна их классифицировать, по возможности, правильно.

Проблема обучения распознаванию тесно связана с другой интеллектуальной задачей — проблемой перевода с одного языка на другой, а также обучения машины языку. При достаточно формальной обработке и классификации основных грамматических правил и приемов пользования словарем можно создать вполне удовлетворительный алгоритм для перевода, скажем, научного или делового текста. Для некоторых языков такие системы были созданы еще в конце 60-х годов. Однако для того, чтобы связно перевести достаточно большой разговорный текст, необходимо понимать его смысл. Работы над такими программами ведутся уже давно, но до полного успеха еще далеко. Имеются также программы, обеспечивающие диалог между человеком и машиной на урезанном естественном языке.

Что же касается моделирования логического мышления, то хорошей модельной задачей здесь может служить задача автоматизации доказательства теорем. Начиная с 1960 г., был разработан ряд программ, способных находить доказательства теорем в исчислении предикатов первого порядка. Эти программы обладают, по словам американского специалиста в области ИИ Дж. Маккатти, «здоровым смыслом», то есть способностью делать дедуктивные заключения.

В программе К. Грина и других, реализующей вопросно-ответную систему, знания записываются на языке логики предикатов в виде набора аксиом, а вопросы, задаваемые машине, формулируются как подлежащие доказательству теоремы. Большой интерес представляет «интеллектуальная» программа американского математика Хао Ванга. Эта программа за 3 минуты работы IBM-704 вывела 220 относительно простых лемм и теорем из фундаментальной математической монографии, а затем за 8,5 мин выдала доказательства еще 130 более сложных теорем, часть из которых еще не была выведена математиками. Правда, до сих пор ни одна программа не вывела и не доказала ни одной теоремы, которая бы, что называется «позарез» была бы нужна математикам и была бы принципиально новой.

Очень большим направлением систем ИИ является робототехника. В чем основное отличие интеллекта робота от интеллекта универсальных вычислительных машин?

Для ответа на этот вопрос уместно вспомнить принадлежащее великому русскому физиологу И. М. Сеченову высказывание: «... все бесконечное разнообразие внешних проявлений мозговой деятельности сводится окончательно лишь к одному явлению — мышечному движению». Другими словами, вся интеллектуальная деятельность человека направлена в конечном счете на активное взаимодействие с внешним миром посредством движений. Точно так же элементы интеллекта робота служат прежде всего для организации его целенаправленных движений. В то же время основное назначение чисто компьютерных систем ИИ состоит в решении интеллектуальных задач, носящих абстрактный или вспомогательный характер, которые обычно не связаны ни с восприятием окружающей среды с помощью искусственных органов чувств, ни с организацией движений исполнительных механизмов.

Первых роботов трудно назвать интеллектуальными. Только в 60-х годах появились оживленные роботы, которые управлялись универсальными компьютерами. К примеру, в 1969 г. в Электротех-

нической лаборатории (Япония) началась разработка проекта «промышленный интеллектуальный робот». Цель этой разработки — создание оцувствленного манипуляционного робота с элементами искусственного интеллекта для выполнения сборочно-монтажных работ с визуальным контролем.

Манипулятор робота имеет шесть степеней свободы и управляется мини-ЭВМ NEAC-3100 (объем оперативной памяти 32000 слов, объем внешней памяти на магнитных дисках 273000 слов), формирующей требуемое программное движение, которое отрабатывается следящей электрогидравлической системой. Схват манипулятора оснащен тактильными датчиками.

В качестве системы зрительного восприятия используются две телевизионные камеры, снабженные красно-зелено-синими фильтрами для распознавания цвета предметов. Поле зрения телевизионной камеры разбито на 64\*64 ячеек. В результате обработки полученной информации грубо определяется область, занимаемая интересующим робота предметом. Далее, с целью детального изучения этого предмета, выявленная область вновь делится на 4096 ячеек. В том случае, когда предмет не помещается в выбранное «окошко», оно автоматически перемещается, подобно тому, как человек скользит взглядом по предмету. Робот Электротехнической лаборатории был способен распознавать простые предметы, ограниченные плоскостями и цилиндрическими поверхностями при специальном освещении. Стоимость данного экспериментального образца составляла примерно 400000 долларов.

Постепенно характеристики роботов монотонно улучшались. Но до сих пор они еще далеки по понятливости от человека, хотя некоторые операции уже выполняют на уровне лучших жонглеров - к примеру, удерживают на лезвии ножа шарик от настольного тенниса.

Еще, пожалуй, здесь можно выделить работы Киевского института кибернетики, где под руководством Н. М. Амосова и В. М. Глушкова (ныне покойного) проводился комплекс исследований, направ-

ленных на разработку элементов интеллекта роботов. Особое внимание в этих исследованиях уделяется проблемам распознавания изображений и речи, логического вывода (автоматического доказательства теорем) и управления с помощью нейроподобных сетей.

К примеру, можно рассмотреть созданный еще в 70-х годах макет транспортного автономного интегрального робота (ТАИР). Конструктивно ТАИР представляет собой трехколесное шасси, на котором смонтированы сенсорная система и блок управления. Сенсорная система включает в себя следующие средства очувствления: оптический дальномер, навигационную систему с двумя радиомаяками и компасом, контактные датчики, датчики углов наклона тележки, таймер и другие. И особенность, которая отличает ТАИР от многих других систем, созданных у нас и за рубежом, это то, что в его составе нет компьютера в том виде, к которому мы привыкли. Основу системы управления составляет бортовая нейроподобная сеть, на которой реализуются различные алгоритмы обработки сенсорной информации, планирования поведения и управления движением робота.

## **1.2. Основные направления искусственного интеллекта**

Имеется, по крайней мере, две точки зрения на то, что следовало бы назвать ИИ. Первую можно назвать нейробионической. Ее сторонники ставят перед собой цель воспроизвести искусственным образом те процессы, которые протекают в мозгу человека. Это путь изучения естественного мозга, выявления принципов его работы, создания технических средств для повторения биологических структур и протекающих в них процессов [25, 40].

Вторая точка зрения, доминирующая в ИИ, может быть названа информационной. Сторонники информационного подхода считают, что основной целью работ в ИИ является не построение технического аналога биологической системы, а создание средств для решения задач, традиционно считающихся интеллектуальными.

Информационная точка зрения, в свою очередь, неоднородна. В ней можно выделить три направления.

1. Часть специалистов считает: можно найти свой способ решения задачи на ЭВМ, который даст либо результат, подобный человеческому, либо даже лучший. Специалисты такого типа неоднократно демонстрировали свое искусство по созданию программ такого рода. Достаточно назвать, например, программы для игры в шахматы, благодаря которым ЭВМ играют в эту игру лучше подавляющего большинства людей, проводящих время за шахматной доской.

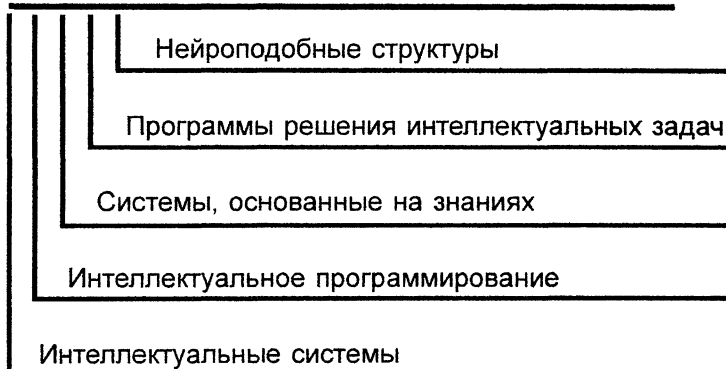
2. Другая часть специалистов считает, что ИИ должен имитировать не решение отдельных (пусть весьма творческих) задач. Ибо, естественный интеллект человека – это его способность при необходимости обучаться тому или иному виду творческой деятельности. Значит, и программы, создаваемые в ИИ, должны быть ориентированы не на решение конкретных задач, а на создание для автоматического построения необходимых программ решения конкретных задач, когда в этом возникает необходимость. Именно эта группа исследователей сейчас определяет лицо ИИ, составляя основную массу специалистов этого профиля.

3. Третья часть специалистов – это программисты, чьими руками делаются программы для решения задач ИИ. Они склонны рассматривать область своей деятельности как новый виток развития программирования. Они считают, что средства, разрабатываемые для написания программ решения интеллектуальных задач, в конце концов есть средства, позволяющие по описанию задачи на профессиональном естественном языке построить нужную программу на основании тех стандартных программных модулей, которые хранятся в памяти машины. Все метасредства, которые предлагают те, кто рассматривает ИИ как способ определения на информационном уровне, функции, которые реализует естественный интеллект, когда он решает задачу, программисты видят сквозь призму своей цели — создание интеллектуального программного обес-

печения (по существу, комплекса средств, автоматизирующих деятельность самого программиста).

На рис.1.1 показана схема строения ИИ, связанная с различными точками зрения на него.

## **ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ**



***Рис.1.1. Схема строения ИИ***

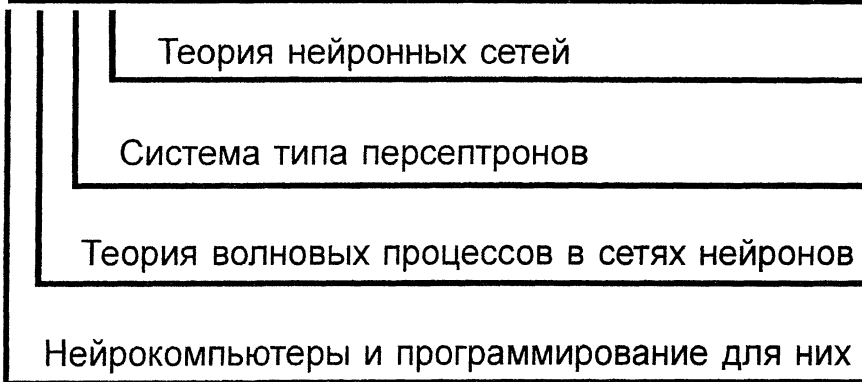
### **1.2.1. Нейроподобные структуры**

В последние десятилетия в мире бурно развивается новая прикладная область математики, специализирующаяся на искусственных нейронных сетях (НС). Актуальность исследований в этом направлении подтверждается массой различных применений НС. Это автоматизация процессов распознавания образов, адаптивное управление, аппроксимация функционалов, прогнозирование, создание экспертных систем, организация ассоциативной памяти и многие другие приложения. С помощью НС можно, например, предсказывать показатели биржевого рынка, выполнять распознавание оптических или звуковых сигналов, создавать самообучающиеся системы, способные управлять автомашиной при парковке или синтезировать речь по тексту. В то время как на Западе применение НС уже достаточно обширно, у нас это еще в некоторой степени экзоти-

ка – российские фирмы, использующие НС в практических целях, наперечет [37, 26, 63].

На рис.1.2 подробно представлена схема этого направления.

## НЕЙРОПОДОБНЫЕ СТРУКТУРЫ

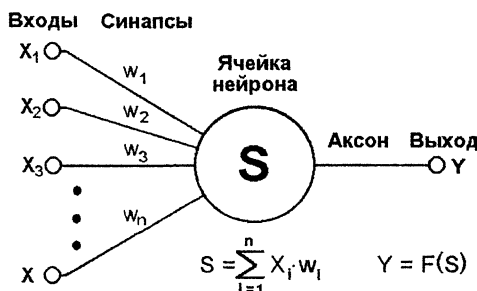


***Рис.1.2. Схема строения нейробионического направления***

Несмотря на существенные различия, отдельные типы НС обладают несколькими общими чертами [62].

Во-первых, основу каждой НС составляют относительно простые, в большинстве случаев – однотипные, элементы (ячейки), имитирующие работу нейронов мозга. Далее под нейроном будет подразумеваться искусственный нейрон, то есть ячейка НС. Каждый нейрон характеризуется своим текущим состоянием по аналогии с нервными клетками головного мозга, которые могут быть возбуждены или заторможены. Он обладает группой синапсов – однонаправленных входных связей, соединенных с выходами других нейронов, а также имеет аксон – выходную связь данного нейрона, с которой сигнал (возбуждения или торможения) поступает на синапсы следующих нейронов. Общий вид нейрона приведен на рис.1.3. Каждый синапс характеризуется величиной синаптической связи или ее ве-

сом  $w_i$ , который по физическому смыслу эквивалентен электрической проводимости.



**Рис.1.3. Искусственный нейрон**

Текущее состояние нейрона определяется как взвешенная сумма его входов:

$$s = \sum_{i=1}^n x_i \cdot w_i . \quad (1.1)$$

Выход нейрона есть функция его состояния:

$$y = f(s). \quad (1.2)$$

Нелинейная функция  $f$  называется активационной и может иметь различный вид, как показано на рис.1.4. Одной из наиболее распространенных является нелинейная функция с насыщением, так называемая логистическая функция или сигмоид (т.е. функция S-образного вида)[2]:

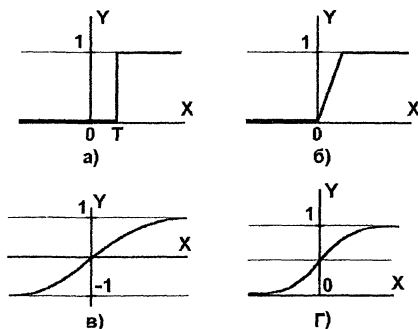
$$f(x) = \frac{1}{1 + e^{-\alpha x}} . \quad (1.3)$$

При уменьшении  $\alpha$  сигмоид становится более пологим, в пределе при  $\alpha=0$  вырождаясь в горизонтальную линию на уровне 0,5, при увеличении  $\alpha$  сигмоид приближается по внешнему виду к функции единичного скачка с порогом  $T$  в точке  $x=0$ . Из выражения для сигмоида очевидно, что выходное значение нейрона лежит в диапазоне  $[0,1]$ . Одно из ценных свойств сигмоидной функции – простое



выражение для ее производной, применение которого будет рассмотрено в дальнейшем.

$$f'(x) = \alpha \cdot f(x) \cdot (1 - f(x)). \quad (1.4)$$



**Рис.1.4. Нелинейная функция активации  $f$**

**а - функция единичного скачка;**

**б - линейный порог (гистерезис);**

**в - сигмоид – гиперболический тангенс;**

**г - сигмоид – формула (1.3)**

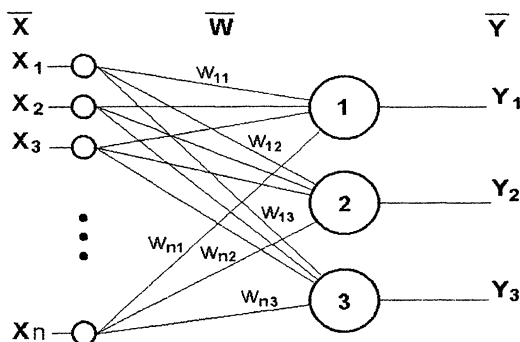
Следует отметить, что сигмоидная функция дифференцируема на всей оси абсцисс, что используется в некоторых алгоритмах обучения. Кроме того, она обладает свойством усиливать слабые сигналы лучше, чем большие, и предотвращает насыщение от больших сигналов, так как они соответствуют областям аргументов, где сигмоид имеет пологий наклон.

Возвращаясь к общим чертам, присущим всем НС, отметим, во-вторых, принцип параллельной обработки сигналов, который достигается путем объединения большого числа нейронов в так называемые слои и соединения определенным образом нейронов различных слоев, а также, в некоторых конфигурациях, и нейронов одного слоя между собой, причем обработка взаимодействия всех нейронов ведется послойно.

В качестве примера простейшей НС рассмотрим трехнейронный перцептрон (рис.1.5), то есть такую сеть, нейроны которой имеют активационную функцию в виде единичного скачка. На  $n$  входов

поступают некие сигналы, проходящие по синапсам на 3 нейрона, образующие единственный слой этой НС и выдающие три выходных сигнала:

$$y_j = f \left[ \sum_{i=1}^n x_i \cdot w_{ij} \right], \quad j=1 \dots 3. \quad (1.5)$$



**Рис.1.5. Однослойный перцептрон**

Очевидно, что все весовые коэффициенты синапсов одного слоя нейронов можно свести в матрицу  $W$ , в которой каждый элемент  $w_{ij}$  задает величину  $i$ -й синаптической связи  $j$ -го нейрона. Таким образом, процесс, происходящий в НС, может быть записан в матричной форме:

$$Y = F(XW), \quad (1.6)$$

где  $X$  и  $Y$  – соответственно входной и выходной сигнальные векторы;

$F(V)$  – активационная функция, применяемая поэлементно к компонентам вектора  $V$ .

Теоретически число слоев и число нейронов в каждом слое может быть произвольным, однако фактически оно ограничено ресурсами компьютера или специализированной микросхемы, на которых обычно реализуется НС. Чем сложнее НС, тем масштабнее задачи, подвластные ей.

Выбор структуры НС осуществляется в соответствии с особенностями и сложностью задачи. Для решения некоторых отдельных типов задач уже существуют оптимальные, на сегодняшний день, конфигурации. Если же задача не может быть сведена ни к одному из известных типов, разработчику приходится решать сложную проблему синтеза новой конфигурации. При этом он руководствуется несколькими основополагающими принципами: возможности сети возрастают с увеличением числа ячеек сети, плотности связей между ними и числом выделенных слоев; введение обратных связей наряду с увеличением возможностей сети поднимает вопрос о динамической устойчивости сети; сложность алгоритмов функционирования сети (в том числе, например, введение нескольких типов синапсов – возбуждающих, тормозящих и др.) также способствует усилению мощи НС. Вопрос о необходимых и достаточных свойствах сети для решения того или иного рода задач представляет собой целое направление нейрокомпьютерной науки. Так как проблема синтеза НС сильно зависит от решаемой задачи, дать общие подробные рекомендации затруднительно. В большинстве случаев оптимальный вариант получается на основе интуитивного подбора.

Очевидно, что процесс функционирования НС, то есть сущность действий, которые она способна выполнять, зависит от величин синаптических связей, поэтому, задавшись определенной структурой НС, отвечающей какой-либо задаче, разработчик сети должен найти оптимальные значения всех переменных весовых коэффициентов (некоторые синаптические связи могут быть постоянными).

Этот этап называется обучением НС, и от того, насколько качественно он будет выполнен, зависит способность сети решать поставленные перед ней задачи во время эксплуатации. На этапе обучения кроме параметра качества подбора весов важную роль играет время обучения. Как правило, эти два параметра связаны обратной зависимостью и их приходится выбирать на основе компромисса.

Обучение НС может вестись с учителем или без него. В первом случае сети предъявляются значения как входных, так и желательных выходных сигналов, и она по некоторому внутреннему ал-

горитму подстраивает веса своих синаптических связей. Во втором случае выходы НС формируются самостоятельно, а веса изменяются по алгоритму, учитывающему только входные и производные от них сигналы.

Существует великое множество различных алгоритмов обучения, которые, однако, делятся на два больших класса: детерминистские и стохастические. В первом из них подстройка весов представляет собой жесткую последовательность действий, во втором – она производится на основе действий, подчиняющихся некоторому случайному процессу.

Развивая дальше вопрос о возможной классификации НС, важно отметить существование бинарных и аналоговых сетей. Первые из них оперируют с двоичными сигналами, и выход каждого нейрона может принимать только два значения: логический ноль («заторможенное» состояние) и логическая единица («возбужденное» состояние). К этому классу сетей относится и рассмотренный выше перцептрон, так как выходы его нейронов, формируемые функцией единичного скачка, равны либо 0, либо 1. В аналоговых сетях выходные значения нейронов способны принимать непрерывные значения, что могло бы иметь место после замены активационной функции нейронов перцептрона на сигмоид.

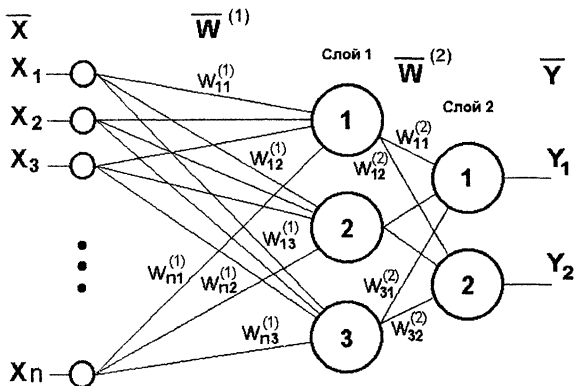
Еще одна классификация делит НС на синхронные и асинхронные [62]. В первом случае в каждый момент времени свое состояние меняет лишь один нейрон. Во втором – состояние меняется сразу у целой группы нейронов, как правило, у всего слоя. Алгоритмически ход времени в НС задается итерационным выполнением однотипных действий над нейронами. Далее будут рассматриваться только синхронные НС.

Сети также можно классифицировать по числу слоев. На рис.1.6 представлен двухслойный перцептрон, полученный из перцептрона с рис.1.5 путем добавления второго слоя, состоящего из двух нейронов. Здесь уместно отметить важную роль нелинейности активационной функции, так как, если бы она не обладала данным свойством или не входила в алгоритм работы каждого нейрона, результат функционирования любой р-слойной НС с весовыми матри-

цами  $W^{(i)}$ ,  $i=1,2,\dots,p$  для каждого слоя  $i$  сводился бы к перемножению входного вектора сигналов  $X$  на матрицу

$$W^{(\Sigma)} = W^{(1)} \cdot W^{(2)} \cdot \dots \cdot W^{(p)}, \quad (1.7)$$

то есть фактически такая  $p$ -слойная НС эквивалентна однослойной НС с весовой матрицей единственного слоя  $W^{(\Sigma)}$ :



**Рис.1.6. Двухслойный перцептрон**

$$Y = XW^{(\Sigma)}. \quad (1.8)$$

Продолжая разговор о нелинейности, можно отметить, что она иногда вводится и в синаптические связи. Большинство известных на сегодняшний день НС используют для нахождения взвешенной суммы входов нейрона формулу (1), однако в некоторых приложениях НС полезно ввести другую запись, например:

$$s = \sum_{i=1}^n x_i^2 \cdot w_i \quad (1.9)$$

или даже

$$s = \sum_{i=1}^n x_i \cdot x_{((i+1) \bmod n)} \cdot w_i. \quad (1.10)$$

Вопрос в том, чтобы разработчик НС четко понимал, для чего он это делает, какими ценными свойствами он тем самым дополнительно наделяет нейрон и каких лишает. Введение такого рода нелинейности, вообще говоря, увеличивает вычислительную мощь се-

ти, то есть позволяет из меньшего числа нейронов с «нелинейными» синапсами сконструировать НС, выполняющую работу обычной НС с большим числом стандартных нейронов и более сложной конфигурации [62].

#### 1.2.2. Информационное направление искусственного интеллекта

Программы для решения интеллектуальных задач могут быть разделены на несколько групп, определяемых типом задач, решаемых этими программами.

С самого начала появления ЭВМ стали создаваться программы для машинного перевода и автоматического реферирования текстов. Создание этих программ оказало значительное влияние на развитие ИИ, заложило основы тех работ, которые были непосредственно связаны с естественно-языковым общением пользователей с интеллектуальными системами. В системах машинного перевода были разработаны модели и методы, позволяющие автоматически проводить морфологический, синтаксический и во многом семантический анализ фраз естественного языка, нащупаны приемы анализа связного текста. Все эти результаты активно используются при обработке естественно-языковых текстов в интеллектуальных системах. В работах по автоматическому реферированию были заложены основы понимания общей структуры текста, как целого. От идеи «что говорится» был сделан переход к идее «о чем говорится». Это позволило на более высоком уровне создавать программы генерации текстов. Основные задачи данного направления представлены на рис.1.7.

Третье основное направление в ИИ образует его фундамент. Именно здесь создается теория данного научного направления, решаются основные проблемы, связанные с центральным объектом изучения ИИ – знаниями.

На рис.1.8 показана структура этого направления.

На рис.1.9 представлены основные направления интеллектуального программирования.

На рис.1.10 отражены прикладные интеллектуальные системы.

# ПРОГРАММЫ РЕШЕНИЯ ИНТЕЛЛЕКТУАЛЬНЫХ ЗАДАЧ

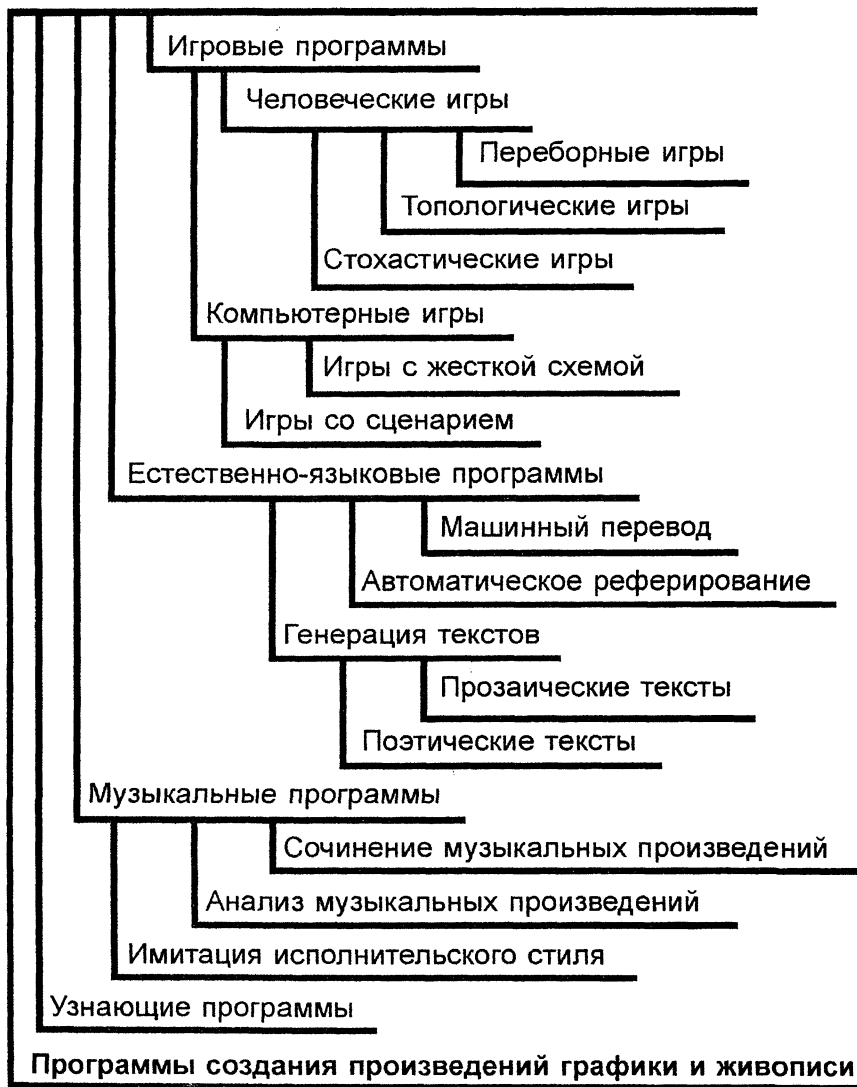


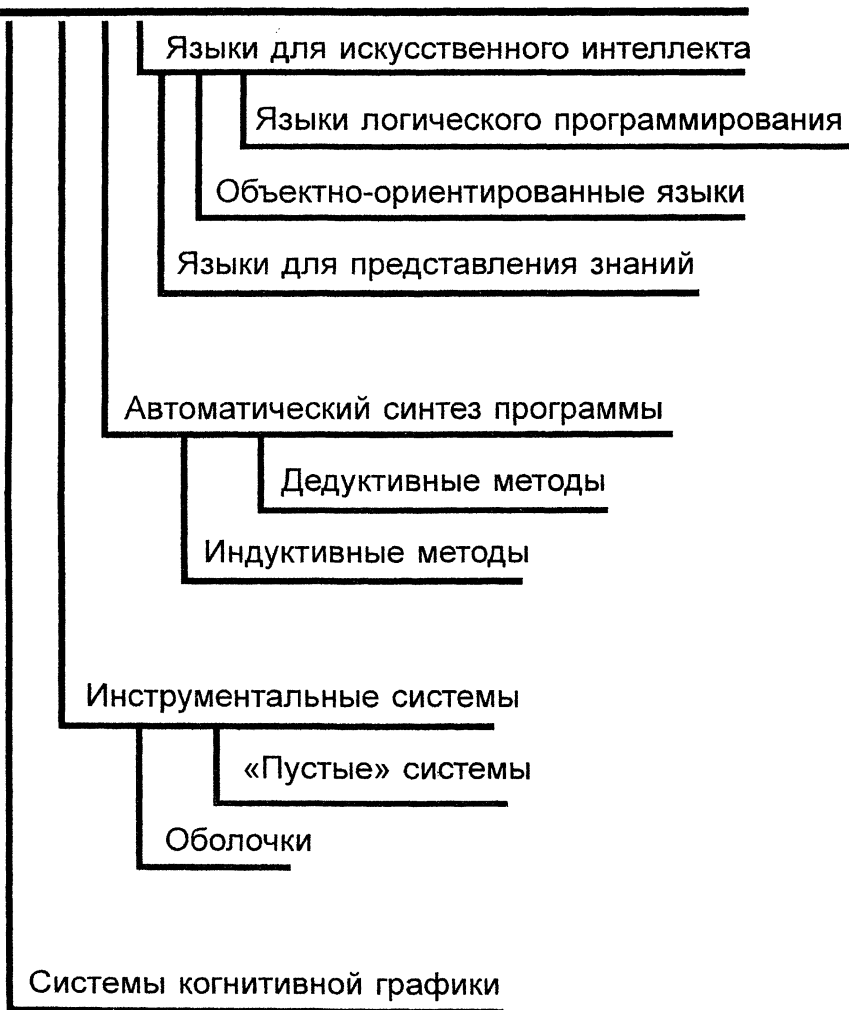
Рис. 1.7. Программы решения интеллектуальных задач

# СИСТЕМЫ, ОСНОВАННЫЕ НА ЗНАНИЯХ

**Рис.1.8. Системы, основанные на знаниях**

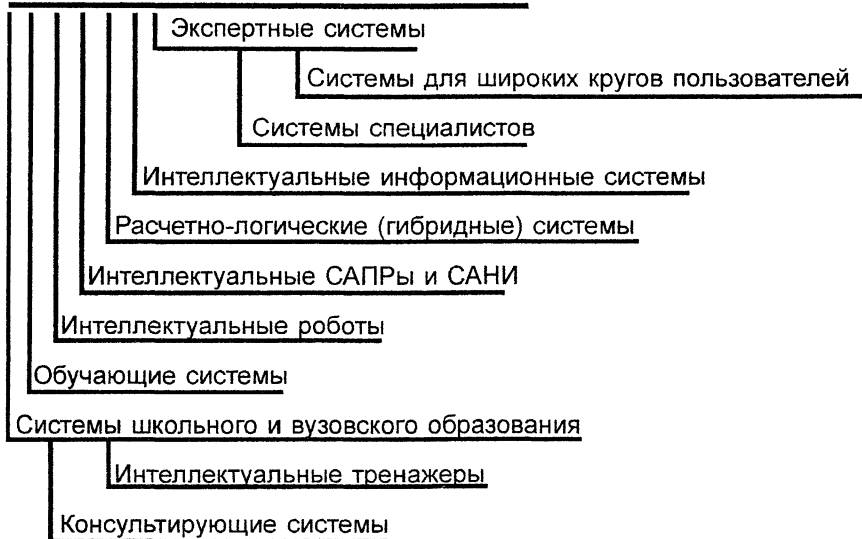


# ИНТЕЛЛЕКТУАЛЬНОЕ ПРОГРАММИРОВАНИЕ



*Рис.1.9. Интеллектуальное программирование*

# ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ



**Рис.1.10. Интеллектуальные системы**

## 1.3. Нечеткая логика

Нечеткая логика (fuzzy logic) - одно из немногих научных направлений, созданных в США, развитых в Японии и вновь признанных американцами уже после безнадёжной утраты стратегической инициативы [47].

Созданная в 60-х годах профессором Лотфи Заде (выходцем из Баку), развитая Бартом Коско (поклонником буддизма и обладателем черного пояса по карате) и воплощенная в первые коммерческие системы американскими фирмами Aptronix (основана Wei Xu, выходцем из Китая) и Togai InfraLogic (первый миллион долларов заработала в Японии) - нечеткая логика демонстрирует явное тяготение к Востоку.

В основе нечеткой логики лежит теория нечетких множеств, где функция принадлежности элемента множеству не бинарна (да/нет), а может принимать любое значение в диапазоне 0-1. Это

дает возможность определять понятия, нечеткие по самой своей природе: «высокий», «быстрый», «ажиотажный» и т.д. Соответственно, сделав еще лишь один шаг - научившись обрабатывать нечеткие импликации (типа «Если А принадлежит нечеткому множеству а, то В принадлежит нечеткому множеству b»), вы получаете возможность строить базы знаний и экспертные системы нового поколения, способные хранить и обрабатывать неточную информацию. Впрочем, на этот шаг наука потратила почти пятнадцать лет...

Первые же опыты применения новых интеллектуальных систем показали их необычайно широкие возможности. Адаптивное управление роботами и системами вооружений; высокодоходная игра практически на всех финансовых рынках мира; интеллектуальные пылесосы, видеокамеры и швейные машины - послужной список успехов нечеткой логики нарастал как снежный ком. И каждый раз, начиная новую разработку с использованием нечеткой логики, инженеры сталкивались с необходимостью отрешиться от привычной бинарной логики и вступить в зыбкую область нечетких рассуждений. Жизнь показала, что лучше всего это удастся приверженцам восточных философских систем, в первую очередь - японцам.

В 1988 году одна из научных лабораторий технологического института MITI (Япония), в поисках спонсоров для своего проекта в области нечеткой логики провела мини-исследование крупных промышленных фирм на предмет использования нечеткой логики в их разработках. К изумлению исследователей, вместо ожидаемых 7-10 фирм они обнаружили «нечеткий след» в изделиях сорока девяти компаний! Первой была Matsushita, отметившая 70-летие фирмы выпуском самоходного пылесоса. Управляемое нечеткой логикой устройство, похожее на большого жука, самостоятельно чистило комнату за комнатой, легко приспосабливаясь к любой расстановке мебели. Впрочем, это был скорее рекламный, чем коммерческий продукт. А первой на массовом рынке стала стиральная машина той же фирмы. Снабженная всего одной кнопкой и способная автоматически выбирать оптимальный режим стирки из более чем четырех-

сот вариантов, машина получила поэтическое название, которое в приближенном русском переводе значит что-то вроде «День моей любимой жены». Любимым женам понравилось - и уже через месяц объем продаж стиральных машин вырос вдвое, превысил рубеж в 35000 штук в месяц и обрек на стахановский режим все заводы компании. Разумеется, конкуренты не пожелали отставать и к 90-му году крупные заголовки «FUZZY» украшали уже целые прилавки торгового района Akihabara, своеобразного «Митинского рынка» Токио. Более двадцати (!) видов бытовых изделий обрели «интеллектуальную начинку» в виде управляющих кристаллов на основе нечеткой логики.

В числе основных промышленных применений теории нечеткой логики можно указать экономическое управление, распознавание образов и обработку изображений, принятие решений, анализ надежности и т.д. В настоящее время наметилась тенденция применения нечетких множеств в гуманитарных науках, лингвистике, психологии и в социологии. Вообще применение нечетких множеств более характерно для гуманитарных наук, т.к. там чаще приходится сталкиваться с нечеткими, субъективными данными. А теория нечетких множеств прежде всего ставит перед собой задачи непосредственного анализа и обработки именно таких неопределенных, неясных данных. В 1972 г. Заде предложил теоретико-множественную интерпретацию лингвистических переменных и ограничений, которая отражала лингвистические аспекты отношения принадлежности в нечетких множествах.

Определение «нечеткого множества» чаще всего интерпретируют как величину  $M_A(x)$ , которая обозначает субъективную оценку степени принадлежности  $x$  множеству  $A$ , например  $M_A(x) = 80\%$  означает, что  $x$  на 80% принадлежит  $A$ . Следовательно, должны существовать «моя функция принадлежности», «ваша функция принадлежности», «еще чья-нибудь функция принадлежности» и т.д.

Несколько позже были сформулированы понятия нечеткой логики с лингвистическим, а не числовым значением истинности. Со-

гласно такой логике, высказывание может принимать истинное значение типа: истинно, ложно, абсолютно истинно, совсем ложно и т.п. - каждое такое значение представляет собой нечеткое подмножество единичного интервала.

В том случае, когда речь идет о логике, то представляется некоторая четкая и жесткая система, которая позволяет все разделить на «Да» и «Нет» (истину и ложь). Это представление соответствует компьютерной логике или двузначной булевой алгебре. Однако в реальной действительности очень трудно все разделить на черное и белое. В работе [22] Л.Заде двузначная оценка 0 или 1 расширена до неограниченной многозначной оценки выше 0 и ниже 1, т.е. впервые было введено понятие «нечеткое множество». Теоретико-вероятностное понятие случайности уже давно отнесено к категории объективных понятий и рассматривается как дополнительное к понятию причинности; такое восприятие подкрепляется концепцией воспроизводимых элементов, которая согласуется с наблюдениями в области естественных наук и в технике. По-видимому, и к субъективной вероятности можно относиться как к шкале неоднозначности. Подобно объективной вероятности - самой популярной и неопределенной из концепций неопределенности - субъективная вероятность удовлетворяет аксиоме вероятностной меры и оказывается положительной и воспроизводимой. Однако широкое распространение разнообразия неясных, неопределенных и неточных явлений, событий и фактов, а также связей между объектами и операциями, показывает, что существуют различные классы неясности или неопределенности, которые не всегда будут связаны со случайностью или нечеткостью.

#### 1.4. Нечеткие множества

Пусть  $E$  - универсальное множество,  $x$  - элемент  $E$ , а  $R$  - некоторое свойство. Обычное (четкое) подмножество  $A$  универсального множества  $E$ , элементы которого удовлетворяют свойству  $R$ , опре-

деляется как множество упорядоченных пар  $A = \{\mu_A(x)/x\}$ , где  $\mu_A(x)$  - характеристическая функция, принимающая значение 1, если  $x$  удовлетворяет свойству  $R$ , и 0 - в противном случае [13, 10, 29, 32, 25, 64].

Нечеткое подмножество отличается от обычного тем, что для элементов  $x$  из  $E$  нет однозначного ответа «да-нет» относительно свойства  $R$ . В связи с этим, нечеткое подмножество  $A$  универсального множества  $E$  определяется как множество упорядоченных пар  $A = \{\mu_A(x)/x\}$ , где  $\mu_A(x)$  - характеристическая функция принадлежности (или просто функция принадлежности), принимающая значения в некотором вполне упорядоченном множестве  $M$  (например,  $M = [0,1]$ ). Функция принадлежности указывает степень (или уровень) принадлежности элемента  $x$  подмножеству  $A$ . Множество  $M$  называют множеством принадлежностей. Если  $M = \{0,1\}$ , то нечеткое подмножество  $A$  может рассматриваться как обычное или четкое множество.

#### 1.4.1. Примеры записи нечетких множеств

Пусть  $E = \{x_1, x_2, x_3, x_4, x_5\}$ ,  $M = [0,1]$ ;  $A$  - нечеткое множество, для которого

$$\mu_A(x_1) = 0,3;$$

$$\mu_A(x_2) = 0;$$

$$\mu_A(x_3) = 1;$$

$$\mu_A(x_4) = 0,5;$$

$$\mu_A(x_5) = 0,9.$$

Тогда  $A$  можно представить в виде:  $A = \{0,3/x_1; 0/x_2; 1/x_3; 0,5/x_4; 0,9/x_5\}$  или  $A = 0,3/x_1 + 0/x_2 + 1/x_3 + 0,5/x_4 + 0,9/x_5$ .

Замечание. Здесь знак «+» не является обозначением операции сложения, а имеет смысл объединения.

#### Основные характеристики нечетких множеств

Пусть  $M = [0,1]$  и  $A$  - нечеткое множество с элементами из универсального множества  $E$  и множеством принадлежностей  $M$ .

- Величина  $\sup_{x \in E} \mu_A(x)$  называется высотой нечеткого множества

А. Нечеткое множество А нормально, если его высота равна 1, т.е. верхняя граница его функции принадлежности равна 1

( $\sup_{x \in E} \mu_A(x)=1$ ). При  $\sup_{x \in E} \mu_A(x)<1$  нечеткое множество называется суб-

нормальным.

• Нечеткое множество пусто, если  $\forall x \in E \mu_A(x)=0$ . Непустое субнормальное множество можно нормализовать по формуле

$$\mu_A(x) := \frac{\mu_A(x)}{\sup_{x \in E} \mu_A(x)}.$$

• Нечеткое множество унимодально,  $\mu_A(x)=1$  только на одном  $x$  из  $E$ .

• Носителем нечеткого множества А является обычное подмножество со свойством  $\mu_A(x)>0$ , т.е. носитель  $A = \{x/\mu_A(x)>0\} \forall x \in E$ .

• Элементы  $x \in E$ , для которых  $\mu_A(x)=0,5$ , называются **точками перехода** множества А.

#### 1.4.2. Примеры нечетких множеств

1. Пусть  $E = \{0,1,2,...,10\}$ ,  $M = [0,1]$ . Нечеткое множество «несколько» можно определить следующим образом: «несколько» =  $0,5/3+0,8/4+1/5+1/6+0,8/7+0,5/8$ ; его характеристики: *высота* = 1, *носитель* =  $\{3,4,5,6,7,8\}$ , *точки перехода* -  $\{3,8\}$ .

2. Пусть  $E = \{0,1,2,3,...,n,...\}$ . Нечеткое множество «малый» можно определить:

$$\text{«малый»} = \mu_{\text{«малый»}}(n) = \frac{1}{1 + \frac{n}{10}} \cdot \frac{1}{n}.$$

#### 1.4.3. Операции над нечеткими множествами

##### **Включение.**

Пусть А и В - нечеткие множества на универсальном множестве Е. Говорят, что А содержится в В, если  $\forall x \in E \mu_A(x) \leq \mu_B(x)$ .

Обозначение:  $A \subset B$ .

Иногда используют термин «доминирование», т.е. в случае, когда  $A \subset B$ , говорят, что  $B$  доминирует  $A$ .

**Равенство.**

$A$  и  $B$  равны, если  $\forall x \in E \mu_A(x) = \mu_B(x)$ .

Обозначение:  $A = B$ .

**Дополнение.**

Пусть  $M = [0, 1]$ ,  $A$  и  $B$  - нечеткие множества, заданные на  $E$ .  $A$  и  $B$  дополняют друг друга, если  $\forall x \in E \mu_A(x) = 1 - \mu_B(x)$ .

Обозначение:  $B = \bar{A}$  или  $A = \bar{B}$ . Очевидно, что  $(\bar{\bar{A}}) = A$ . (Дополнение определено для  $M = [0, 1]$ , но очевидно, что его можно определить для любого упорядоченного  $M$ ).

**Пересечение.**

$A \cap B$  - наибольшее нечеткое подмножество, содержащееся одновременно в  $A$  и  $B$ .

$$\mu_{A \cap B}(x) = \min(\mu_A(x), \mu_B(x)).$$

**Объединение.**

$A \cup B$  - наименьшее нечеткое подмножество, включающее как  $A$ , так и  $B$ , с функцией принадлежности:  $\mu_{A \cup B}(x) = \max(\mu_A(x), \mu_B(x))$ .

**Разность.**

$A - B = A \cap \bar{B}$  с функцией принадлежности:

$$\mu_{A-B}(x) = \mu_{A \cap \bar{B}}(x) = \min(\mu_A(x), 1 - \mu_B(x)).$$

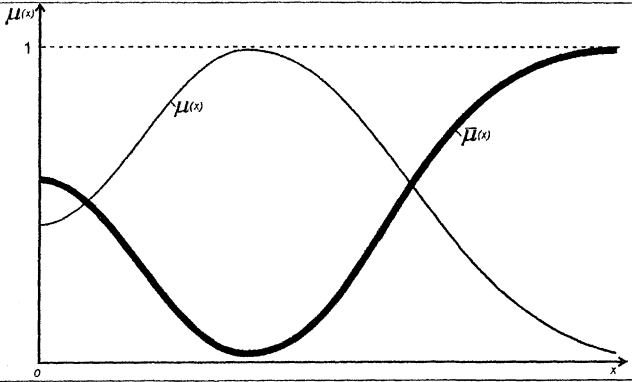
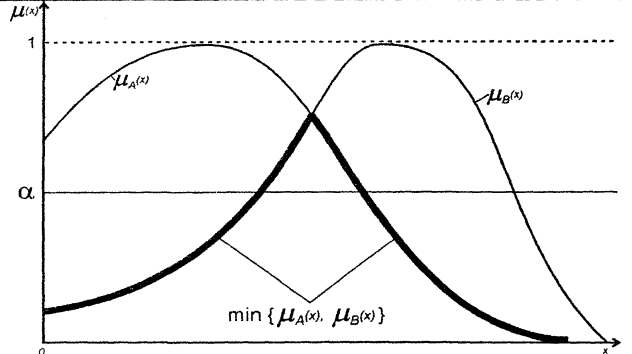
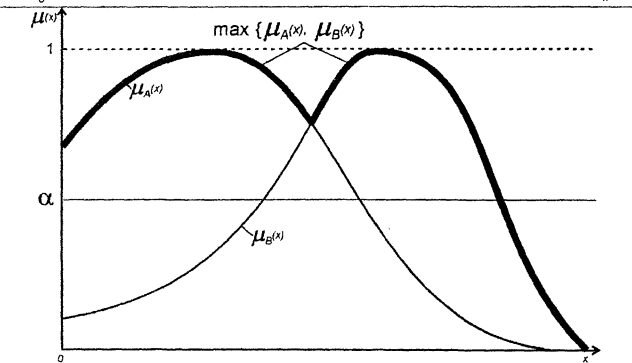
**Дизъюнктивная сумма.**

$A \oplus B = (A - B) \cup (B - A) = (A \cap \bar{B}) \cup (\bar{A} \cap B)$  с функцией принадлежности:

$$\mu_{A \oplus B}(x) = \max\{\min\{\mu_A(x), 1 - \mu_B(x)\}; \min\{1 - \mu_A(x), \mu_B(x)\}\}$$



# Графическое представление основных операций

Название операции	Графическое представление
Дополнение	 <p>The graph shows a coordinate system with a vertical axis labeled <math>\mu(x)</math> and a horizontal axis labeled <math>x</math>. A dashed horizontal line is drawn at <math>\mu(x) = 1</math>. Two curves are plotted: <math>\mu_A(x)</math>, a bell-shaped curve peaking at 1, and <math>\mu_B(x)</math>, an S-shaped curve starting at 0 and approaching 1. The resulting membership function <math>\mu(x)</math> is the complement of <math>\mu_A(x)</math>, which is a curve starting at 1 and approaching 0.</p>
Пересечение I (минимум: не взаимодействующие переменные)	 <p>The graph shows a coordinate system with a vertical axis labeled <math>\mu(x)</math> and a horizontal axis labeled <math>x</math>. A dashed horizontal line is drawn at <math>\mu(x) = 1</math>, and a solid horizontal line is drawn at <math>\mu(x) = \alpha</math>. Two curves are plotted: <math>\mu_A(x)</math>, a bell-shaped curve peaking at 1, and <math>\mu_B(x)</math>, a bell-shaped curve peaking at 1. The resulting membership function <math>\mu(x)</math> is the minimum of <math>\mu_A(x)</math> and <math>\mu_B(x)</math>, shown as a thick line. A label <math>\min \{ \mu_A(x), \mu_B(x) \}</math> points to the thick line.</p>
Объединение I (максимум: не взаимодействующие переменные)	 <p>The graph shows a coordinate system with a vertical axis labeled <math>\mu(x)</math> and a horizontal axis labeled <math>x</math>. A dashed horizontal line is drawn at <math>\mu(x) = 1</math>, and a solid horizontal line is drawn at <math>\mu(x) = \alpha</math>. Two curves are plotted: <math>\mu_A(x)</math>, a bell-shaped curve peaking at 1, and <math>\mu_B(x)</math>, a bell-shaped curve peaking at 1. The resulting membership function <math>\mu(x)</math> is the maximum of <math>\mu_A(x)</math> and <math>\mu_B(x)</math>, shown as a thick line. A label <math>\max \{ \mu_A(x), \mu_B(x) \}</math> points to the thick line.</p>

Название операции	Графическое представление
Разность	

### Свойства операций $\cup$ и $\cap$ .

Пусть  $A, B, C$  - нечеткие множества, тогда выполняются следующие свойства:

- $A \cap B = B \cap A$   
 $A \cup B = B \cup A$  } - коммутативность;
- $(A \cap B) \cap C = (A \cap B) \cap C$   
 $(A \cup B) \cup C = A \cup (B \cup C)$  } - ассоциативность;
- $A \cap A = A$   
 $A \cup A = A$  } - идемпотентность;
- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$   
 $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$  } - дистрибутивность;
- $A \cup \emptyset = A$ , где  $\emptyset$  - пустое множество, т.е.  $\mu_{\emptyset}(x) = 0$

$\forall x \in E$ ;

- $A \cap \emptyset = \emptyset$ ;
- $A \cap E = A$ , где  $E$  - универсальное множество;
- $A \cup E = E$ .

В отличие от четких множеств, для нечетких множеств в общем случае:

$$A \cap \overline{A} \neq \emptyset,$$

$$A \cup \overline{A} \neq E.$$

Замечание. Введенные выше операции над нечеткими множествами основаны на использовании операций **max** и **min**. В теории нечетких множеств разрабатываются вопросы построения обобщенных, параметризованных операторов пересечения, объединения и дополнения, позволяющих учесть разнообразные смысловые оттенки соответствующих им связей «и», «или», «не».

Один из подходов к операторам пересечения и объединения заключается в их определении в классе *треугольных норм и конорм*.

**Треугольной нормой** (*t-нормой*) называется двухместная действительная функция  $T:[0,1] \times [0,1] \rightarrow [0,1]$ , удовлетворяющая следующим условиям:

1.  $T(0,0)=0$ ;  $T(\mu_A, 1) = \mu_A$ ;  $T(1, \mu_A) = \mu_A$  - *ограниченность*;
2.  $T(\mu_A, \mu_B) \leq T(\mu_C, \mu_D)$ , если  $\mu_A \leq \mu_C$ ,  $\mu_B \leq \mu_D$  - *монотонность*;
3.  $T(\mu_A, \mu_B) = T(\mu_B, \mu_A)$  - *коммутативность*;
4.  $T(\mu_A, T(\mu_B, \mu_C)) = T(T(\mu_A, \mu_B), \mu_C)$  - *ассоциативность*.

Простым случаем треугольных норм являются:

$$\min(\mu_A, \mu_B)$$

$$\text{произведение } \mu_A \cdot \mu_B$$

$$\max(0, \mu_A + \mu_B - 1).$$

**Треугольной конормой** (*t-конормой*) называется двухместная действительная функция  $\perp:[0,1] \times [0,1] \rightarrow [0,1]$ , со свойствами:

1.  $T(1,1) = 1$ ;  $T(\mu_A, 0) = \mu_A$ ;  $T(0, \mu_A) = \mu_A$  - *ограниченность*;
2.  $T(\mu_A, \mu_B) \geq T(\mu_C, \mu_D)$ , если  $\mu_A \geq \mu_C$ ,  $\mu_B \geq \mu_D$  - *монотонность*;
3.  $T(\mu_A, \mu_B) = T(\mu_B, \mu_A)$  - *коммутативность*;
4.  $T(\mu_A, T(\mu_B, \mu_C)) = T(T(\mu_A, \mu_B), \mu_C)$  - *ассоциативность*.

Примеры *t-конорм*:

$$\max(\mu_A, \mu_B)$$

$$\mu_A + \mu_B - \mu_A \cdot \mu_B$$

$$\min(1, \mu_A + \mu_B).$$

#### 1.4.4. Алгебраические операции над нечеткими множествами

Алгебраическое произведение **A** и **B** обозначается **A**·**B** и определяется так:

$$\forall x \in E \quad \mu_{A \cdot B}(x) = \mu_A(x) \mu_B(x).$$

Алгебраическая сумма этих множеств обозначается  $A \mp B$  и определяется так:

$$\forall x \in E \quad \mu_{A \mp B} = \mu_A(x) + \mu_B(x) - \mu_A(x)\mu_B(x).$$

Для операций  $\{ \cdot, \mp \}$  выполняются свойства:

- $\left. \begin{array}{l} A \cdot B = B \cdot A \\ A \mp B = B \mp A \end{array} \right\}$  - коммутативность;
- $\left. \begin{array}{l} (A \cdot B) \cdot C = A \cdot (B \cdot C) \\ (A \mp B) \mp C = A \mp (B \mp C) \end{array} \right\}$  - ассоциативность;
- $A \cdot \emptyset = \emptyset, A \mp \emptyset = A, A \cdot E = A, A \mp E = E$
- $\left. \begin{array}{l} \overline{A \cdot B} = \overline{B \cdot A} \\ \overline{A \mp B} = \overline{B \mp A} \end{array} \right\}$  - теоремы де Моргана.

Не выполняются:

- $\left. \begin{array}{l} A \cdot A = A \\ A \mp A = A \end{array} \right\}$  - идемпотентность;
- $\left. \begin{array}{l} A(B \mp C) = (A \cdot B) \mp (A \cdot C) \\ A \mp (BC) = (A \mp B) \cdot (A \mp C) \end{array} \right\}$  - дистрибутивность;
- а также  $A \cdot \overline{A} = \emptyset, A \mp \overline{A} = E$ .

Продолжим обзор основных операций над нечеткими множествами. На основе операции алгебраического произведения (по крайней мере для целых  $\alpha$  эта основа очевидна) определяется операция возведения в степень  $\alpha$  нечеткого множества  $A$ , где  $\alpha$  - положительное число. Нечеткое множество  $A^\alpha$  определяется функцией принадлежности  $\mu_{A^\alpha} = \mu_A^\alpha(x)$ . Частным случаем возведения в степень являются:

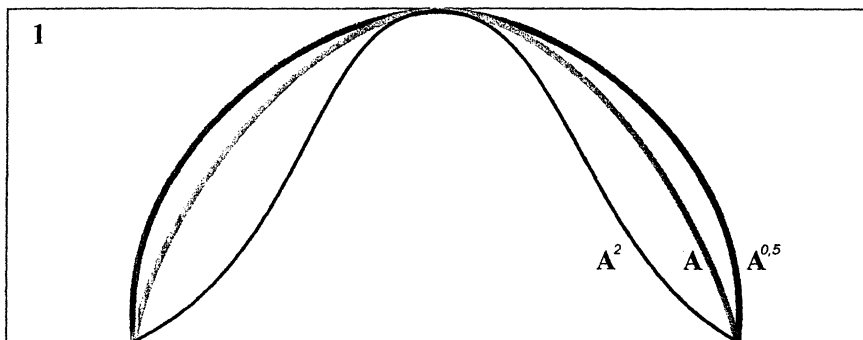
- $\text{CON}(A) = A^2$  - операция *концентрирования*,
- $\text{DIL}(A) = A^{0,5}$  - операция *растяжения*,

которые используются при работе с лингвистическими неопределенностями.

Выпуклая комбинация нечетких множеств. Пусть  $A_1, A_2, \dots, A_n$  - нечеткие множества универсального множества  $E$ , а  $\omega_1, \omega_2, \dots, \omega_n$  - неотрицательные числа, сумма которых равна 1.

Выпуклой комбинацией  $A_1, A_2, \dots, A_n$  называется нечеткое множество  $A$  с функцией принадлежности:

$$\forall x \in E \quad \mu_A(x_1, x_1, \dots, x_n) = \omega_1 \mu_{A_1}(x) + \omega_2 \mu_{A_2}(x) + \dots + \omega_n \mu_{A_n}(x).$$



**Рис.1.11. Операции концентрирования и растяжения**

Декартово произведение нечетких множеств. Пусть  $A_1, A_2, \dots, A_n$  - нечеткие подмножества универсальных множеств  $E_1, E_2, \dots, E_n$  соответственно. Декартово произведение  $A = A_1 \times A_2 \times \dots \times A_n$  является нечетким подмножеством множества  $E = E_1 \times E_2 \times \dots \times E_n$  с функцией принадлежности:

$$\mu_A(x_1, x_2, \dots, x_n) = \min\{\mu_{A_1}(x_1), \mu_{A_2}(x_2), \dots, \mu_{A_n}(x_n)\}.$$

Оператор увеличения нечеткости используется для преобразования четких множеств в нечеткие и для увеличения нечеткости нечеткого множества.

Пусть  $A$  - нечеткое множество,  $E$  - универсальное множество и для всех  $x \in E$  определены нечеткие множества  $K(x)$ . Совокупность всех  $K(x)$  называется ядром оператора увеличения нечеткости  $\Phi$ . Результатом действия оператора  $\Phi$  на нечеткое множество  $A$  является нечеткое множество вида:

$$\Phi(A, K) = \bigcup_{x \in E} \mu_A(x) K(x),$$

где  $\mu_A(x)K(x)$  - произведение числа на нечеткое множество.

Пример:

$$E = \{1, 2, 3, 4\};$$

$$A = 0,8/1 + 0,6/2 + 0,3/3 + 0,4/4;$$

$$K(1) = 1/1 + 0,4/2;$$

$$K(2) = 1/2 + 0,4/1 + 0,4/3;$$

$$K(3) = 1/3 + 0,5/4;$$

$$K(4) = 1/4.$$

Тогда

$$\begin{aligned}\Phi(A,K) &= \mu_A(1) K(1) \cup \mu_A(2) K(2) \cup \mu_A(3) K(3) \cup \mu_A(4) K(4) = \\ &= 0,8(1/1+0,4/2) \cup 0,6(1/2+0,4/1+0,4/3) = \\ &= 0,8/1+0,6/2+0,24/3.\end{aligned}$$

### 1.5. Нечеткая и лингвистическая переменные

Понятие нечеткой и лингвистической переменных используется при описании объектов и явлений с помощью нечетких множеств.

**Нечеткая переменная** характеризуется тройкой  $\langle \alpha, X, A \rangle$ ,

где  $\alpha$  - наименование переменной,

$X$  - универсальное множество (область определения  $\alpha$ ),

$A$  - нечеткое множество на  $X$ , описывающее ограничения (т.е.  $\mu_A(x)$ ) на значения нечеткой переменной  $\alpha$ .

**Лингвистической переменной** называется набор  $\langle \beta, T, X, G, M \rangle$ , где  $\beta$  - наименование лингвистической переменной;  $T$  - множество ее значений (терм-множество), представляющих собой наименования нечетких переменных, областью определения каждой из которых является множество  $X$ . Множество  $T$  называется базовым терм-множеством лингвистической переменной;

$G$  - синтаксическая процедура, позволяющая оперировать элементами терм-множества  $T$ , в частности, генерировать новые термы (значения). Множество  $T \cup G(T)$ , где  $G(T)$  - множество сгенерированных термов, называется расширенным терм-множеством лингвистической переменной;

$M$  - семантическая процедура, позволяющая превратить каждое новое значение лингвистической переменной, образуемое процедурой  $G$ , в нечеткую переменную, т.е. сформировать соответствующее нечеткое множество.

Замечание. Чтобы избежать большого количества символов

- символ  $\beta$  используют как для названия самой переменной, так и для всех ее значений;
- пользуются одним и тем же символом для обозначения нечеткого множества и его названия, например терм «*молодой*», являющийся значением лингвистической переменной  $\beta$  = «*возраст*», одновременно есть и нечеткое множество  $M$  («*молодой*»).

Присвоение нескольких значений символам предполагает, что контекст позволяет разрешить возможные неопределенности.

Пример: Пусть эксперт определяет толщину выпускаемого изделия с помощью понятий «малая толщина», «средняя толщина» и «большая толщина», при этом минимальная толщина равна 10 мм, а максимальная - 80 мм.

Формализация такого описания может быть проведена с помощью следующей лингвистической переменной  $\langle \beta, T, X, G, M \rangle$ , где  $\beta$  - толщина изделия;

$T$  - {«малая толщина», «средняя толщина», «большая толщина»};

$X$  - [10, 80];

$G$  - процедура образования новых термов с помощью связок «и», «или» и модификаторов типа «очень», «не», «слегка» и др. Например: «малая или средняя толщина», «очень малая толщина» и др.;  $M$  - процедура задания на  $X = [10, 80]$  нечетких подмножеств  $A_1$  = «малая толщина»,  $A_2$  = «средняя толщина»,  $A_3$  = «большая толщина», а также нечетких множеств для термов из  $G(T)$  в соответствии с правилами трансляции нечетких связок и модификаторов «и», «или», «не», «очень», «слегка» и других операций над нечеткими множествами вида:  $A \cap B$ ,  $A \cup B$  и др.

Замечание. Наряду с рассмотренными выше базовыми значениями лингвистической переменной «толщина» ( $T = \{\text{«малая толщина»}, \text{«средняя толщина»}, \text{«большая толщина»}\}$ ) возможны значения, зависящие от области определения  $X$ . В данном случае значения лингвистической переменной «толщина изделия» могут быть определены как «около 20 мм», «около 50 мм», «около 70 мм», т.е. в виде нечетких чисел.

## Нечеткие числа

**Нечеткие числа** - нечеткие переменные, определенные на числовой оси, т.е. нечеткое число определяется как нечеткое множество  $A$  на множестве действительных чисел  $R$  с функцией принадлежности  $\mu_A(x) \in [0, 1]$ , где  $x$  - действительное число, т.е.  $x \in R$ .

Нечеткое число  $A$  **нормально**, если  $\mu_A(x) = 1$ , **выпуклое**, если для любых  $x \leq y \leq z$  выполняется  $\mu_A(x) \geq \mu_A(y) \wedge \mu_A(z)$ .

**Множество  $\alpha$  - уровня** нечеткого числа  $A$  определяется как  $A_\alpha = \{x/\mu_A(x) \geq \alpha\}$ .

Подмножество  $S_A \subset R$  называется **носителем** нечеткого числа  $A$ , если  $S = \{x/\mu_A(x) > 0\}$ .

Нечеткое число  $A$  **положительно**, если  $\forall x \in S_A, x > 0$  и **отрицательно**, если  $\forall x \in S_A, x < 0$ .

## 1.6. Нечеткие высказывания и нечеткие модели систем

Нечеткими высказываниями будем называть высказывания следующего вида:

1. Высказывание  $\langle \beta \text{ есть } \beta' \rangle$ , где  $\beta$  - наименование лингвистической переменной,  $\beta'$  - ее значение, которому соответствует нечеткое множество на универсальном множестве  $X$ .

Например, высказывание  $\langle \text{давление большое} \rangle$  предполагает, что лингвистической переменной «давление» придается значение «большое», для которого на универсальном множестве  $X$  переменной «давление» определено соответствующее данному значению «большое» нечеткое множество.

2. Высказывание  $\langle \beta \text{ есть } m\beta' \rangle$ , где  $m$  - модификатор, которому соответствуют слова «очень», «более или менее», «много больше» и др.

Например:  $\langle \text{давление очень большое} \rangle$ ,  $\langle \text{скорость много больше средней} \rangle$  и др.

3. Составные высказывания, образованные из высказываний видов 1 и 2 и союзов «и», «или», «если..., то...», «если..., то..., иначе».

Высказывания на множестве значений фиксированной лингвистической переменной.

То, что значения фиксированной лингвистической переменной соответствуют нечетким множествам одного и того же универсального множества  $X$ , позволяет отождествлять модификаторы «очень» или «не» с операциями «CON» и «дополнение», а союзы «и», «или» с операциями «пересечение» и «объединение» над нечеткими множествами.

Для иллюстрации понятия лингвистической переменной мы в качестве примера рассматривали лингвистическую переменную «толщина изделия» с базовым терм-множеством  $T = \{\text{«малая»},$



«средняя», «большая»}. При этом на  $X = [10, 80]$  мы определили нечеткие множества  $A_1, A_2, A_3$ , соответствующие базовым значениям: «малая», «средняя», «большая». В этом случае высказыванию <толщина изделия очень малая> соответствует нечеткое множество  $CONA = A^2$ ; высказыванию <толщина изделия не большая или средняя> - нечеткое множество  $A_2 \cup \overline{A_3}$ ; высказыванию <толщина изделия не малая и не большая> -  $A_1 \cap \overline{A_3}$ .

Высказывания <толщина изделия много больше средней> или <толщина изделия близка к средней> требуют использования нечетких отношений  $R$  («много больше, чем») и  $R$  («близко к»), заданных на  $X \times X$ . Тогда этим высказываниям будут соответствовать нечеткие множества  $A \bullet R_1$  и  $A \bullet R_2$ , индуцированные нечеткими отношениями  $R_1$  и  $R_2$ .

### Случай двух и более лингвистических переменных

Пусть  $\langle \alpha, T_\alpha, X, G_\alpha, M_\alpha \rangle$  и  $\langle \beta, T_\beta, Y, G_\beta, M_\beta \rangle$  - лингвистические переменные, и высказываниям  $\langle \alpha \text{ есть } \alpha' \rangle$ ,  $\langle \beta \text{ есть } \beta' \rangle$  соответствуют нечеткие множества  $A$  и  $B$ , заданные на  $X$  и  $Y$ .

Составные нечеткие высказывания вида 3, связывающие значения лингвистических переменных  $\alpha$  и  $\beta$ , можно привести к высказываниям вида 1, введя лингвистическую переменную  $(\alpha, \beta)$ , значениям которой будут соответствовать нечеткие множества на  $X \times Y$ .

Напомним, что нечеткие множества  $A$  и  $B$ , заданные на  $X$  и  $Y$ , порождают на  $X \times Y$  нечеткие множества, называемые цилиндрическими продолжениями, с функциями принадлежности:

$$\mu_A(x, y) = \mu_A(x) \text{ при любом } y,$$

$$\mu_B(x, y) = \mu_B(y) \text{ при любом } x,$$

где  $(x, y) \in X \times Y$ .

Нечеткие множества, соответствующие составным высказываниям

$\langle \alpha \text{ есть } \alpha' \text{ и } \beta \text{ есть } \beta' \rangle$  и

$\langle \alpha \text{ есть } \alpha' \text{ или } \beta \text{ есть } \beta' \rangle$ ,

определяются по следующим правилам (преобразования к виду 1), справедливым при условии не взаимодействия переменных, т.е. множества  $X$  и  $Y$  таковы, что их элементы не связаны какой-либо функциональной зависимостью.

## Правила преобразований нечетких высказываний

### Правило преобразования конъюнктивной формы

Справедливо выражение:

$$\langle \alpha \text{ есть } \alpha' \text{ и } \beta \text{ есть } \beta' \rangle \Rightarrow \langle (\alpha, \beta) \text{ есть } (\alpha' \cap \beta') \rangle.$$

Здесь  $\Rightarrow$  - знак подстановки,  $\alpha' \cap \beta'$  - значение лингвистической переменной  $(\alpha, \beta)$ , соответствующее исходному высказыванию  $\langle \alpha \text{ есть } \alpha' \text{ и } \beta \text{ есть } \beta' \rangle$ , которому на  $X \times Y$  ставится в соответствие нечеткое множество  $\hat{A} \cap \hat{B}$  с функцией принадлежности

$$\mu_{\hat{A} \cap \hat{B}}(x, y) = \mu_{\hat{A}}(x, y) \wedge \mu_{\hat{B}}(x, y) = \mu_{\hat{A}}(x) \wedge \mu_{\hat{B}}(y)$$

### Правило преобразования дизъюнктивной формы

Справедливо выражение:

$\langle \alpha \text{ есть } \alpha' \text{ или } \beta \text{ есть } \beta' \rangle \Rightarrow \langle (\alpha, \beta) \text{ есть } (\alpha' \cup \beta') \rangle$ , где значению  $(\alpha' \cup \beta')$  лингвистической переменной  $(\alpha, \beta)$  соответствует нечеткое множество  $\hat{A} \cup \hat{B}$ , с функцией принадлежности

$$\mu_{\hat{A} \cup \hat{B}}(x, y) = \mu_{\hat{A}}(x, y) \vee \mu_{\hat{B}}(x, y) = \mu_{\hat{A}}(x) \vee \mu_{\hat{B}}(y)$$

Замечание 1. Правила справедливы также для переменных вида  $\langle \alpha, T_1, X, G_1, M_1 \rangle$  и  $\langle \alpha, T_2, Y, G_2, M_2 \rangle$ , когда в форме значений лингвистических переменных формализованы не взаимодействующие характеристики одного и того же объекта.

Замечание 2. Если задана совокупность лингвистических переменных  $\{ \langle \alpha_i, T_i, X_i, G_i, M_i \rangle \}$ ,  $i = 1, 2, \dots, n$ , то любое составное высказывание, полученное из высказываний  $\langle \alpha \text{ есть } \alpha' \rangle$  с использованием модификаторов «очень», «не», «более или менее» и др. и связок «и», «или», можно привести к виду  $\langle \alpha \text{ есть } \alpha' \rangle$ , где  $\alpha$  - составная лингвистическая переменная  $(\alpha_1, \alpha_2, \dots, \alpha_n)$ ,  $\alpha'$  - ее значение, определяемое (как и функция принадлежности) в соответствии с вышеуказанными правилами.

## Правило преобразования высказываний импликативной формы

Справедливо выражение:

$\langle \text{если } \alpha \text{ есть } \alpha', \text{ то } \beta \text{ есть } \beta' \rangle \Rightarrow \langle (\alpha, \beta) \text{ есть } (\alpha' \rightarrow \beta') \rangle$ , где значению  $(\alpha' \rightarrow \beta')$  лингвистической переменной  $(\alpha, \beta)$  соответствует нечеткое отношение  $XRY$  на  $X \times Y$ . Функция принадлежности  $\mu_R(x, y)$  зависит от выбранного способа задания нечеткой импликации.

### **1.7. Логико-лингвистическое описание систем, нечеткие модели**

Логико-лингвистические методы описания систем основаны на том, что поведение исследуемой системы описывается на естественном (или близком к естественному) языке в терминах лингвистических переменных.

Входные и выходные параметры системы рассматриваются как лингвистические переменные, а качественное описание процесса задается совокупностью высказываний следующего вида:

$L_1$  : если  $\langle A_1 \rangle$  то  $\langle B_1 \rangle$ ,

$L_2$  : если  $\langle A_2 \rangle$  то  $\langle B_2 \rangle$ ,

.....

$L_k$  : если  $\langle A_k \rangle$  то  $\langle B_k \rangle$ ,

где  $\langle A_i \rangle$ ,  $i=1,2,...,k$  - составные нечеткие высказывания, определенные на значениях входных лингвистических переменных, а  $\langle B_i \rangle$ ,  $i=1,2,...,k$  - высказывания, определенные на значениях выходных лингвистических переменных.

С помощью правил преобразования дизъюнктивной и конъюнктивной формы описание системы можно привести к виду:

$L_1$  : если  $\langle A_1 \rangle$  то  $\langle B_1 \rangle$ ,

$L_2$  : если  $\langle A_2 \rangle$  то  $\langle B_2 \rangle$ ,

.....

$L_k$  : если  $\langle A_k \rangle$  то  $\langle B_k \rangle$ ,

где  $A_1, A_2, ..., A_k$  - нечеткие множества, заданные на декартовом произведении  $X$  универсальных множеств входных лингвистических переменных, а  $B_1, B_2, ..., B_k$  - нечеткие множества, заданные на декар-

товом произведении  $Y$  универсальных множеств выходных лингвистических переменных.

Совокупность импликаций  $\{L_1, L_2, \dots, L_k\}$  отражает функциональную взаимосвязь входных и выходных переменных и является основой построения нечеткого отношения  $XRY$ , заданного на произведении  $X \times Y$  универсальных множеств входных и выходных переменных. Если на множестве  $X$  задано нечеткое множество  $A$ , то композиционное правило вывода  $B = A \bullet R$  определяет на  $Y$  нечеткое множество  $B$  с функцией принадлежности

$$\mu_B(y) = (\mu_A(x) \wedge \mu_R(x, y)).$$

Таким образом, композиционное правило вывода в этом случае задает закон функционирования нечеткой модели системы.

### **1.8. Основные аспекты интеллектуализации автоматизированных систем**

Создание и развитие интеллектуальной системы (ИС) представляет собой сложный многоэтапный процесс, характеризующийся значительными капиталовложениями, длительным сроком реализации и существенной неопределённостью, связанной с возможными изменениями как целей проектирования и применения, так и воздействий различного рода возмущений внешней среды на ИС на различных этапах жизненного цикла, имеющих как объективный, так и субъективный характер.

Центральной задачей при разработке и модернизации крупномасштабных ИС является выбор её структуры, под которой понимаются состав элементов, соответствующая система взаимосвязей (отношений) между элементами и распределение функций, выполняемых ИС. Применительно к ИС принято выделять следующие базовые структуры: структуру целей, функций и задач ИС, технологи-

ческую структуру, техническую, организационную, топологическую структуры, структуры информационного, математического и программного обеспечения ИС.

Перечисленные шесть видов структур являются инвариантными для любого класса ИС и должны рассматриваться в первую очередь при системном исследовании и комплексном моделировании соответствующей ИС.

Важность задачи синтеза вышеперечисленных структур применительно к ИС состоит в том, что от того, какой окончательный вариант структуры ИС выбран, во многом зависит эффективность применения указанной ИС по целевому назначению. Дополнительную особенность рассматриваемая задача синтеза приобретает в том случае, когда учитывается такой важный аспект, как динамика развития структур проектируемой ИС, которая, в свою очередь, обусловлена большой продолжительностью общего периода создания ИС, необходимостью учёта изменения и уточнения технических требований к параметрам и характеристикам ИС по этапам развития; расширением круга решаемых задач; необходимостью учёта пространственно-временных, технических, технологических ограничений, связанных с проектированием, производством, испытанием, поставкой, внедрением и развитием основных элементов и подсистем ИС. Говоря о процессах создания и применения ИС, необходимо подчеркнуть, что данные процессы, как правило, имеют эволюционный характер, в ходе которого новые поколения аппаратно-программных средств (АПС) (серверы, рабочие станции, новые версии операционных систем и т.п.) совместно используются с существующими АПС, входящими в состав уже действующих ИС. В этих условиях особую актуальность приобретают совместная постановка и решение задач синтеза облика ИС (выбора оптимального вариан-

та её облика) и планирования развития перечисленных видов структур.

На первом шаге решения предложенной обобщённой задачи синтеза и планирования развития структур ИС необходимо прежде всего рассмотреть вопрос формирования методологических основ её исследования. Масштабность и сложность рассматриваемой в этой связи проблемы требует выбора соответствующей методологии, в качестве которой должна быть выбрана методология современного обобщённого системного анализа, представляющего собой одно из главных направлений реализации системного подхода, в рамках которого на основе гармоничного сочетания формально-математических и логико-эвристических методов осуществляется конструктивное решение разнородных и разноуровневых задач анализа и синтеза ИС на различных этапах их жизненного цикла. Применительно к современным ИС в качестве основных этапов проведения обобщённого системного анализа указанной проблемы синтеза можно выделить:

- этап ретроспективного критического сравнительного анализа существующих отечественных и зарубежных разработок в области создания и применения ИС;
- этап оценивания эффективности существующей ИС;
- этап постановки обобщённой задачи синтеза и планирования развития ИС;
- этап анализа целей и задач, которые требуется решать ИС на новом этапе её развития, формирования системы показателей эффективности функционирования создаваемой ИС на различных этапах её жизненного цикла;
- этап анализа основных пространственно-временных, технических, технологических, стоимостных и ресурсных ограничений, связанных с процессом создания и применения ИС;

- этап анализа (конструирования) альтернативных вариантов структур перспективной ИС;
- этап многокритериального оценивания указанных вариантов структуры ИС и выбора из них наиболее предпочтительных;
- этап формирования, оценивания и выбора оптимальных эволюционных планов перехода от существующей ИС к создаваемой («новой», перспективной) ИС без снижения эффективности их применения.

Конструктивное решение рассматриваемой проблемы поиска и выбора наилучших вариантов создания и развития ИС предполагает, во-первых, построение соответствующего полимодельного комплекса, описывающего различные аспекты жизненного цикла существующей и создаваемой ИС, во-вторых, разработку методов, алгоритмов и методик многокритериального синтеза структуры перспективной ИС, и, в-третьих, разработку многоэтапной интерактивной процедуры поиска решения задачи синтеза ИС и программ их развития.

## **2. ИСПОЛЬЗОВАНИЕ СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ ОПРЕДЕЛЕНИЯ ПАРАМЕТРОВ КАЧЕСТВА ДЕТАЛИ В МЕХАНООБРАБОТКЕ**

### **2.1. Оценка качества и определения оптимальной настройки технологического объекта в реальном времени**

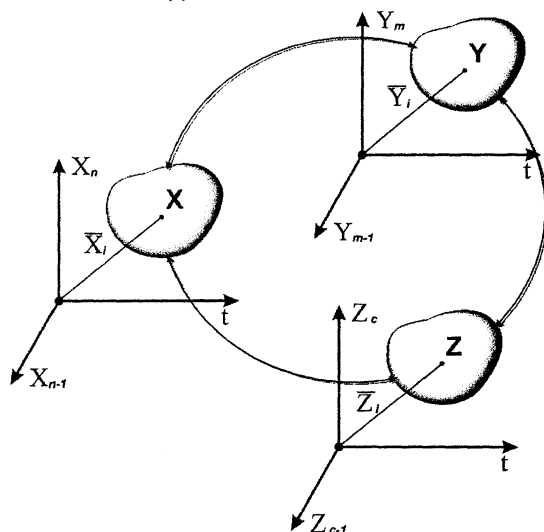
Рассматриваемая проблема является многоцелевой как непосредственно для оценки качества изделий в реальном времени, так и для диагностирования оптимизации адаптивного управления и других задач при различных укрупненных этапах жизненного цикла изделий: изготовление, исследования, эксплуатация в производственных условиях, ремонтные и наладочные процессы. Для каждой из рассмотренных задач имеются особые требования к идентификации качества (точность, локализация параметров, оперативность) рассматриваемых объектов [14].

Следует отметить, что к этой задаче примыкает мониторинг. Организация мониторинга состояния рассматриваемого оборудования сама по себе является крупной научно-технической проблемой, так как для всей гаммы технологического оборудования практически отсутствуют однотипные критерии, технически доступные для «наблюдения», и соответствующие правила принятия решений для целей мониторинга. Аналогичная проблематика возникает и для других технологических систем.

Технологическое металлорежущее оборудование с системных позиций представляется в виде сложных технических систем и несмотря на наличие многочисленных работ в области теоретического моделирования и экспериментальных исследований станков априорное прогнозирование его текущего состояния практически невозможно. В связи с этим, чаще всего приходится лишь констатировать качество готовых деталей после обработки. Практически все серийное технологическое оборудование обеспечено лишь локальными



обратными связями, что обуславливает низкую «управляемость» и «наблюдаемость» по выходным критериям. Главная обратная связь по выходным параметрам практически реализуется лишь в редких случаях. Это в целом не обеспечивает условия инвариантности системы к управляющим, возмущающим воздействиям и к изменению параметров объектов управления. В результате приобретает огромное значение косвенная оценка технологического процесса, поскольку он является практически единственной информационной базой автоматизированной подналадки оборудования в процессе его эксплуатации с изменяющимися воздействиями и параметрами и определяет качество изделий.



**Рис. 2.1. Модель функциональных связей состояния**

Объект рассматривался во временной и частотной областях, для которых и были выполнены исследования. Это связано с тем, что с позиций теории множеств для целей мониторинга должны быть введены множества состояний оборудования  $Y$  и готовых деталей  $Z$ , которые определенным образом соотносятся с первичными показателями качества  $X$ , заданными чертежом детали, и соответствуют друг другу в пространстве возможных состояний. Рассматри-

ваемые множества имеют разную физическую природу и являются непересекающимися. Характеристики множеств  $X$ ,  $Y$ ,  $Z$  являются нечеткими и могут определяться в пространстве состояний как совокупность упорядоченных пар, составленных из элементов  $X$ ,  $Y$ ,  $Z$ , универсальных множеств  $X$ ,  $Y$ ,  $Z$  и соответствующих степеней принадлежности  $\mu_a(x)$ ,  $\mu_b(y)$ ,  $\mu_c(z)$ .

На рис.2.1 приведена графическая интерпретация рассматриваемой проблемы.

В общем случае обеспечение требуемого соответствия между представленными множествами возможно двумя путями: оператор за счет интуитивных и априорно накопленных знаний и автоматизированной системой с искусственным интеллектом. Обращение к системе, обеспеченной искусственным интеллектом, связано с «нечеткостью» рассматриваемых объектов. При этом нечеткость образов, представлений и понятий человека вводится в формальные модели различными способами, которые можно классифицировать по виду представления нечеткой субъективной оценки какой-либо величины (нечеткого множества); по виду области значений функции принадлежности; по виду области определения функции принадлежности; по виду соответствия между областью определения и областью значений (однозначное, многозначное); по признаку однородности или неоднородности области значений функции принадлежности. Оценку качества можно производить по этим характеристикам множеств  $Y$  и  $Z$ , сопоставляя (отображая) его результаты со свойствами множества  $X$ . С этих позиций и рассмотрим формирование базовых принципов оценивания.

Во-первых, выбранные множества и их элементы должны быть состоятельными оценками оборудования или технологических процессов. Во-вторых, они должны однозначно характеризовать состояние рассматриваемых объектов в пространстве состояний. В-третьих, они должны быть доступны для измерений. В-четвертых, динамический мониторинг эффективен в том случае, если без большой вычислительной работы можно определить динамическое

состояние оборудование или процессов по прямым или косвенным характеристикам.

Часто для подобного анализа в теории автоматического управления используют временные, частотные и коэффициентные методы. В последнем случае качество системы управления оценивается, например, по коэффициентам передаточной функции без построения переходного процесса. В нестационарных системах с нелинейными свойствами это возможно для мгновенных значений указанных оценок без интерпретации или прогнозирования состояния в других интервалах времени и при смене режимов функционирования. В противном случае возникают сложные проблемы. Задачи анализа качества, если их необходимо решать для большого числа систем, становятся трудоемкими. Такие задачи возникают всякий раз, когда, например, удастся перейти от одной и (или) нестационарной системы к множеству линейных стационарных систем, эквивалентных одной многорежимной системе, т.е. когда одна система с переменными параметрами заменяется множеством систем с постоянными «замороженными» по тому или иному принципу параметрами. Оценки сводятся к тому, что если «замороженные» системы устойчивы, то и исходная нестационарная система устойчива, что для нелинейных систем доказать практически невозможно. Для количественных показателей качества задача еще более усложняется и для целей динамического мониторинга практически не применима.

Можно обратиться к опыту исследования узкого класса нелинейных систем, имеющих единственное равновесное состояние (или единственную зону равновесного состояния и показателей качества переходной функции). Эти системы характеризуются двухтемповыми движениями, где «быстрые» парциальные движения описываются нелинейными, а «медленные» парциальные движения – линейными стационарными уравнениями.

Если отношения соседних коэффициентов первого уравнения для всех возможных значений коэффициентов будут существенно

больше отношений соседних коэффициентов второго уравнения и решения этих уравнений будут асимптотически устойчивы, то показатели качества рассматриваемой модели будут определяться вторым звеном. Оценить априорно этот факт напрямую из-за отсутствия моделей не является возможным. Однако в проводимых нами экспериментальных исследованиях при разделении частотной области на две части адаптация к указанным условиям возможна. Нелинейные свойства, в свою очередь, приводят к многочастотным колебаниям, основные режимы которых в соответствии с коэффициентами исходного дифференциального уравнения так же лежат в разделимых областях Фурье-образов. Следовательно, является достоверным и перспективным динамический мониторинг по двум разделенным областям множества Фурье-образов с последующим переходом при необходимости во временные пространства состояния обратным преобразованием Фурье.

По результатам оценки возможностей использования приведенных выше классификаций и с учетом особенностей рассматриваемых объектов для целей динамического мониторинга предложены обобщенные показатели качества их состояния в реальном времени. При этом учитывались изложенные факторы и общепринятые в теории управления показатели качества и методы моделирования и идентификации сложных нелинейных систем с нестационарными параметрами. Кроме того, было принято во внимание следующее. Строгие аналитические выражения в целом для нелинейных систем, а особенно для сложных конфигураций системы и числа нелинейных элементов свыше одного, отсутствуют и тем более отсутствуют количественные оценки показателей качества с параметрами передаточной функции. Особые нелинейные стационарные системы имеют указанные выше приближенные оценки, но для рассматриваемой предметной области эти реализации редки и априорно установить их невозможно. В связи с этим можно использовать некоторые модельные представления на ЭВМ с целью отработки принципов создания мониторинга и выбора приближенных границ критери-

ев качества рассматриваемого класса нелинейных систем. Структурный анализ различного металлообрабатывающего оборудования определил модели, включающие нелинейные и линейные элементы, причем приведение нелинейностей к процессу резания не представляется возможным, так как в общем случае нелинейные элементы разделены линейными. Для имитационных исследований нами были использованы модели в виде комбинаций линейных элементов как колебательных звеньев второго порядка и нелинейных элементов с существенно нелинейной характеристикой типа «завор» («сухое трение»). Это является наиболее общим представлением элементов и дает возможность, осуществляя вариацию параметров, адаптировать модель к реальным характеристикам звеньев технологической системы оборудования различного назначения, в котором рассмотренные выше существенно нелинейные характеристики проявляются по-разному. Вычисляемые с помощью моделей обобщенные критерии качества желательно иметь однотипными для станков различной реализации и типов, поскольку по существу оценивают похожие свойства системы управления в низкочастотной области, достаточно точно описываемые вырожденным дифференциальным уравнением, решением которого являются медленные парциальные движения. В высокочастотной области содержится результирующая информация, но о более тонких процессах. Правда, в нелинейных системах она связана даже в двухтемповом представлении управления.

Априорное моделирование с указанными ограничениями для мониторинга, и, соответственно, оценки качества позволяет принципиально создать с определенной «грубостью» эталонную модель, использование которой может быть связано с определением общих свойств идеальной системы, структурной моделью для адаптивной настройки и для определения диапазонов основных частот при, например, модальном анализе или для оценки узкополосной фильтрации в узких областях при наличии базы данных в процессе. В некоторых случаях это является необходимым и очень эффективно

реализуется при оценке качества изделий в реальном времени по косвенным оценкам, например, по виброакустическим сигналам. В предлагаемой работе предстоит решить представленные выше задачи, на основе систем искусственного интеллекта (нечеткие множества и нейронные сети) и интегральных показателей во временной и частотной областях.

С достаточно высокими метрологическими характеристиками прямая оценка качества (микро- и макрогеометрия) изделий непосредственно при резании в производственных условиях практически не реализуема. В данном учебном пособии мы рассмотрим косвенную оценку качества по относительным колебаниям инструмента и детали регистрируемыми пьезоакселерометрами. За многие годы исследований в этом направлении получены положительные результаты, но внедрение их в автоматизированном производстве требует дальнейших научных исследований в рамках предлагаемых информационных технологий. Одной из актуальных задач повышения точности металлообработки, находящейся в контексте сертификации, является получение достоверной оценки качества и надежности станков на этапе приемо-сдаточных испытаний. Для этого создаются гибкие автоматизированные системы и комплексы с применением ЭВМ, которые позволяют в условиях действующего производства оценить техническое состояние станка по обобщенному технологическому критерию и способны выполнять множество конкретных операций:

- *измерять в автоматическом режиме выходные характеристики точности станков* в виде параметров точности траекторий движения их основных формообразующих узлов;
- регистрировать характеристики и параметры формообразующих узлов: параметры точности траекторий опорных точек узлов, характеристики точности их взаимного расположения, кинематические и силовые параметры, тепловые характеристики и характеристики износа и т.д.;
- осуществлять *цифровую фильтрацию входных сигналов*;

- производить *статистическую обработку результатов испытаний*;

- осуществлять *оценку точности* и параметрической надежности испытуемых узлов станка.

Среди первоочередных задач, которые требуют решения для успешной практической реализации этих систем, находятся задачи разработки:

1) *комплекса показателей, пригодных для их автоматизированной оценки с помощью ЭВМ*, т.к. существующие, регламентированные соответствующими ГОСТами, с этой целью использованы быть не могут;

2) математических моделей, алгоритмов и программ для оценки и прогнозирования таких выходных характеристик станков, как параметры точности, показатели надежности, динамические, тепловые, кинематические и другие характеристики.

Целесообразность внедрения подобных систем диктуется и тем обстоятельством, что в современном машиностроении 70-80% объема выпускаемой продукции приходится на многономенклатурные и мелкосерийные производства. Поэтому становятся актуальными вопросы обеспечения высокой надежности и быстрой адаптации станков к постоянно изменяющимся режимам функционирования и номенклатуре изготавливаемых деталей. Решение этой проблемы возможно только в комплексе и находит свое практическое воплощение в создании интегрированного автоматизированного (интегрированного программируемого) производства, характеризующегося в отличие от традиционного рядом специфических особенностей, в числе которых:

- многоуровневая структура, первый уровень которой представляет собой само производство, второй предназначен для его технологической подготовки, третий - для автоматизированного проектирования конструкции изготавливаемой детали, четвертый - для планирования и полного автоматического управления им;

- программирование всей функциональной деятельности на базе широкого использования ЭВМ и микропроцессоров;
- совмещение программного и адаптивного управления для выработки оптимальной стратегии функционирования;
- наличие высокоточных *автоматических измерительных систем* и систем диагностирования.

## **2.2. Распознавание качества формообразования как объект применения нейросетевых методов**

Работы по реализации систем диагностики процессов резания традиционно принято проводить в два этапа. Первый этап заключается в установлении корреляционных связей между состоянием процесса резания и качеством поверхности; далее разрабатывается алгоритм обработки электрического сигнала, регистрируемого датчиком. Второй этап заключается в разработке программной и аппаратной части.

Станок, как и любое другое оборудование, имеющее подвижные части, создаёт виброакустические колебания. Протекание процессов, порождающих вибрации, определяется большим числом факторов: макро- и микрогеометрией инструмента, физико-механическими свойствами материалов инструмента и детали, динамическими характеристиками технологической системы, внешними условиями нагружения и применения технологических сред и др. Изменение хотя бы одного из этих факторов приводит к изменению характеристик виброакустического сигнала, его интенсивности и частотного спектра, что в свою очередь отражается на качестве поверхности детали. Зависимость параметров качества обрабатываемой детали от виброакустических колебаний, создаваемых станком, является существенно нелинейной, поэтому трудно определить те частоты, которые отвечают конкретно за тот или иной узел станка. В области построения диагностических моделей процесса резания известно много работ. Их основной целью является определение наиболее информативного набора диагностирующих признаков и установление строгой корреляции между состоянием процесса резания и результатами косвенных измерений. Подав-



ляющее большинство известных моделей построено на основе регрессионных или корреляционных зависимостей, полученных путем статистической обработки экспериментальных данных, т.е. состояние процесса идентифицируется статистическими методами. При этом помимо установления качественного проявления зависимостей сигналов стараются получить соответствующее математическое описание, чтобы найти количественную форму выражения зависимостей.

Как показали теоретические и экспериментальные исследования, наиболее логично решать задачи, связанные с идентификацией состояния процесса резания, специализированными методами теории распознавания образов. Эффективным математическим механизмом обобщения и распознавания образов могут служить искусственные нейронные сети (ИНС). В любом алгоритме распознавания всегда присутствует процедура сравнения входного сигнала с имеющимися в памяти эталонами. Вне зависимости от наличия или отсутствия предварительной обработки сигнала (выделение основных признаков, преобразование в другую форму в новом параметрическом пространстве и т.д.) сигнал будет представлять собой вектор в каком-либо параметрическом пространстве, и этот вектор сравнивается с векторами, используемыми на стадии обучения. Как раз эту операцию и выполняет большинство нейросетевых моделей.

Одним из приложений теории распознавания образов является распознавание параметров качества обработки детали на металлорежущем станке по виду спектра входного виброакустического сигнала.

В связи с вышесказанным можно выделить ряд задач, которые необходимо решить при построении системы мониторинга, основанного на применении математического аппарата нейронных сетей:

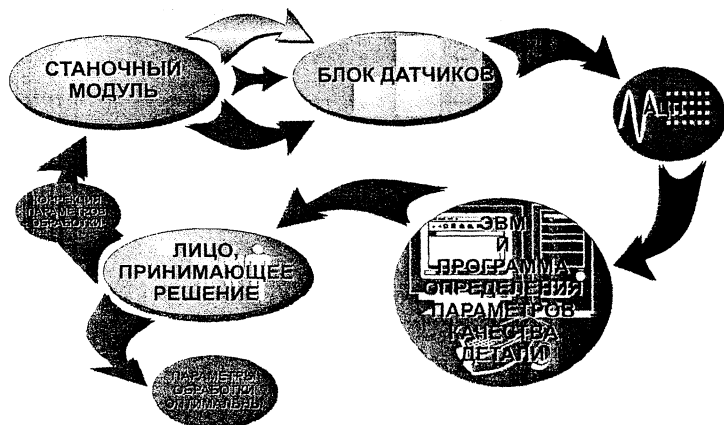
- исследование методов ввода, обработки и анализа акустических сигналов при помощи компьютера;
- изучение специфики акустических сигналов, полученных через АЦП от пьезодатчика, установленного на металлорежущем станке, определение их характерных свойств и построение на основе этих

знаний модели распознавания параметров качества обработки детали;

– решение проблем, возникающих при практическом применении нейросетевых алгоритмов.

Место программного обеспечения для контроля параметров качества в технологическом процессе показано на рис. 2.2.

При обработке детали акустические сигналы от станочного модуля фиксируются блоком датчиков и через АЦП поступают на ЭВМ, где программа определения параметров качества детали по виду спектра акустического сигнала показывает параметры качества обрабатываемой детали в реальном времени. Далее лицо, принимающее решение (рабочий или специалист по наладке станочного модуля) принимает решение о необходимости коррекции параметров обработки.



**Рис. 2.2. Функциональная структура ПО**

### **2.3. Выбор нейросетевого алгоритма в контексте решаемой задачи**

В общем случае искусственная нейросеть не является универсальной по своей структуре. Для ряда прикладных задач существует множество различных структур реализации, алгоритмов, методов обу-

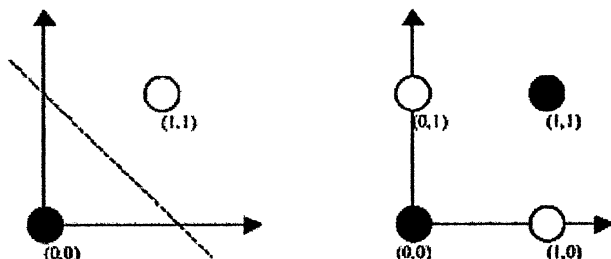
чения и т.п., описанных, например, в [28, 57] и других изданиях. Манипулируя этими различными видами, можно получить, в конечном итоге, НС с оптимальными свойствами и параметрами. Наибольшее использование для решения задач распознавания образов находят следующие нейросетевые методы: нейросетевой гауссов классификатор (Neural Gaussian Classifier), Сеть Хопфилда (hopfield net), Сеть Хемминга (hamming net), Сеть Кохонена (kohonen's neural network), Сеть поиска максимума, Классификатор Карпентера/Гроссберга, Сеть встречного распространения (counterpropagation network), Сеть поиска максимума с прямыми связями (Feed-forward maxnet), Входная звезда (instar).

При выборе структуры нейросети для построения мониторинга качества формообразования необходимо решить ряд задач. В связи с разделением системы анализа параметров качества на несколько уровней (ввода/вывода, уровень распознавания, уровень синтеза), применение в создаваемом программном комплексе какого-то определенного вида НС и алгоритма обучения не представляется возможным. Поэтому необходимо использовать несколько различных видов НС с разными выходными параметрами и моделировать нейросеть для каждого уровня, в соответствии с требованиями по обработке информации на нём.

На уровне ввода/вывода происходит выделение из сигнала значимых системе образцов и представление их одним нейроном или нейронным ансамблем на следующих уровнях. Как при обучении, так и при распознавании входные векторы являются нечеткими, т.е. имеется небольшой разброс векторов, принадлежащих к одному классу. В связи с этим нейросеть, осуществляющая эту операцию, должна обладать определенной способностью к статистическому усреднению. Напротив, может оказаться, что группа векторов находится в непосредственной близости друг к другу, но все они представляют разные классы. Тогда нейросеть должна определять тонкие различия между векторами. Ещё одно требование к нейросети низкого уровня обработки сигнала – способность разделять входные сигналы на классы.

Большое количество нейросетевых алгоритмов выполняют функцию разделения входного сигнала на классы. Известно 3 математических модели этого разделения:

1. Разделение входных сигналов гиперплоскостями (простой персептрон). Применение этого алгоритма оправдано только для задач, обладающих высокой линейностью. Например, можно построить нейросеть, разбивающую точки  $(0,0)$  и  $(1,1)$  на два класса для двумерного сигнала, но невозможно решить задачу по разбиению точек  $(0,0)$ ,  $(1,1)$  – первый класс, и  $(0,1)$ ,  $(1,0)$  – второй. Это широко известный пример неспособности простого персептрона решить задачу «исключающее или» или теорему Минского (см. рис. 2.3).



**Рис. 2.3. Теорема Минского**

2. Разделение входного сигнала гиперповерхностями (много-слойные персептроны). При последовательном соединении слоев, подобных простому персептрону, появляется возможность комбинировать гиперплоскости и получать гиперповерхности довольно сложной формы, в том числе и замкнутые. Такая нейросеть в принципе при достаточном числе нейронов способна разделять сигналы на классы практически любой сложности. Но применение таких нейросетей ограничено сложностью их обучения.

3. Поиск наибольшего соответствия (углового или линейного состояния). При нормализованных векторах входного сигнала все они располагаются на поверхности гипersферы.

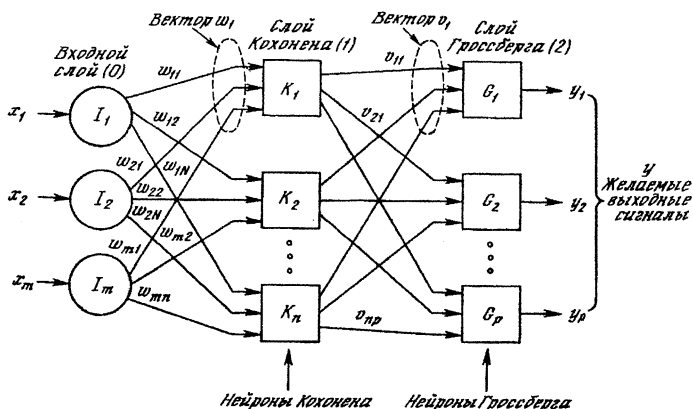
Проведенный анализ сетевых структур и исследования ряда алгоритмов, представленных в литературе, определили выбор нейросети встречного распространения, как наиболее подходящей, для при-

менения в системе мониторинга. Такой выбор был сделан по ряду причин.

1. Возможности сети встречного распространения превосходят возможности однослойных сетей. Время обучения по сравнению с обратным распространением может уменьшаться в сто раз.
2. Встречное распространение не столь общо, как обратное распространение, но оно может давать решение в тех приложениях, где долгая обучающая процедура невозможна.
3. Во встречном распространении объединены два хорошо известных алгоритма: самоорганизующаяся карта Кохонена и звезда Гроссберга [104]. Их объединение ведет к свойствам, которых нет ни у одного из них в отдельности. Самоорганизующаяся карта Кохонена выполняет классификацию входных векторов в группы схожих. Задача классификации представляет собой задачу отнесения образца к одному из нескольких попарно не пересекающихся множеств. Примером таких задач является, например, задача определения информативного частотного диапазона при анализе вибросигналов, возникающих в процессе резания. При решении задач классификации необходимо отнести имеющиеся статические образцы (характеристики виброакустического сигнала, например) к определенным классам. Возможно несколько способов представления данных. Наиболее распространенным является способ, при котором образец представляется вектором. Компоненты этого вектора представляют собой различные характеристики образца, которые влияют на принятие решения о том, к какому классу можно отнести данный образец. Например, для целей вибродиагностики в качестве компонентов этого вектора используются данные о частотном составе сигнала.
4. Обобщающая способность сети встречного распространения позволяет получать правильный выход даже при приложении входного вектора, который является неполным или слегка неверным. Также сеть обладает замечательным свойством: успешно обрабатывать зашумленную, искаженную или частично поврежденную информацию.

## 2.4. Структура сети

На рис.2.4 показана упрощенная структура сети встречного распространения пространства.



**Рис.2.4. Сеть встречного распространения без обратных связей**

Нейроны слоя 0 (показанные кружками) служат лишь точками разветвления и не выполняют вычислений. Каждый нейрон слоя 0 соединен с каждым нейроном слоя 1 (называемого слоем Кохонена) отдельным весом  $w_{mn}$ . Эти веса в целом рассматриваются как матрица весов  $\mathbf{W}$ . Аналогично, каждый нейрон в слое Кохонена (слое 1) соединен с каждым нейроном в слое Гроссберга (слое 2) весом  $v_{nr}$ . Эти веса образуют матрицу весов  $\mathbf{V}$ . Все это весьма напоминает другие сети, различие, однако, состоит в операциях, выполняемых нейронами Кохонена и Гроссберга.

Как и многие другие сети, встречное распространение функционировать в двух режимах: в нормальном режиме, при котором принимается входной вектор  $\mathbf{X}$  и выдается выходной вектор  $\mathbf{Y}$ , и в режиме обучения, при котором подается входной вектор и веса корректируются, чтобы дать требуемый выходной вектор.

После того как определена общая структура сети, необходимо уточнить ее параметры. К параметрам сети относятся количество нейронов входа и выхода, число нейронов слоев Кохонена и Гроссберга, а также величины синаптических связей.

#### 2.4.1. Входной слой

Для построения классификатора необходимо определить, какие параметры влияют на принятие решения о том, к какому классу принадлежит образец. При этом возникают две проблемы. Во-первых, если количество параметров мало, то может возникнуть ситуация, при которой один и тот же набор исходных данных соответствует примерам, находящимся в разных классах. Тогда невозможно обучить нейронную сеть, и система не будет корректно работать (невозможно найти минимум, который соответствует такому набору исходных данных). Исходные данные обязательно должны быть непротиворечивы. Для решения этой проблемы необходимо увеличить размерность пространства признаков (количество компонент входного вектора, соответствующего образцу). Но при увеличении размерности пространства признаков может возникнуть ситуация, когда число примеров станет недостаточным для обучения сети, и она вместо обобщения просто запомнит примеры из обучающей выборки и не сможет корректно функционировать. Таким образом, при определении признаков необходимо найти компромисс с их количеством. Количество нейронов входного слоя будем определять исходя из особенностей входного сигнала.

Ввод виброакустических колебаний осуществляется через аналого-цифровой преобразователь, переводится в цифровую форму и сохраняется в виде файлов формата WAVE в кодировке PCM (разрядность 8 бит, частота дискретизации 22050 Гц). Работа с файлами позволяет многократно повторять процессы их обработки нейросетью, что особенно важно при обучении.

При вводе звука определяющими являются частота дискретизации ( $f_d$ ) и разрядность преобразования (сколько единиц информации кодирует один отсчет). Частота дискретизации определяет макси-

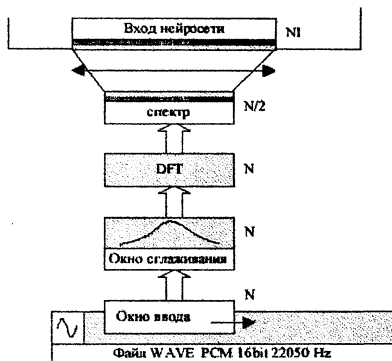
мальную частоту сигнала, которую можно закодировать. Как следует из теоремы Колесникова,

$$f_{\max} < f_d / 2, \quad (2.1)$$

где  $f_d$  – частота дискретизации.

Процесс ввода информации в нейросеть изображен на рис. 2.5.

Очевидно, что представление виброакустического сигнала во временной форме неэффективно. Оно не отражает характерных особенностей акустического сигнала. Для того, чтобы виброакустические колебания можно было подать на вход нейросети, необходимо представить их в виде энергетического спектра.



**Рис.2.5. Ввод информации на вход нейросети**

Спектр непериодического сигнала  $x(t)$  вычисляется с помощью комплексного преобразования (интеграла) Фурье:

$$F(\omega) = \int x(t) \exp(-j\omega t) dt \quad (2.2)$$

Интегрирование выполняется по всему временному интервалу действия сигнала. Вычисление производится для выбираемого диапазона частот: 128, 256, 512, ..., 65536 (по степеням двойки). Непосредственное вычисление ДПФ (дискретного преобразования Фурье) требует выполнения  $N$  умножений комплексных чисел и  $N-1$  сложений комплексных чисел для каждого значения  $X(k)$ . Так как нужно вычислить  $N$  значений, то общее число вычислений сводится к  $N*N$  умножениям и  $N*(N-1)$  сложениям. Для больших  $N$  ( $N > 1000$ ) число операций непомерно большое. Поэтому применяются так называемые алгоритмы быстрого преобразования Фурье (БПФ), применяемые для случая



$N = 2^r$ ,  $r$  - целое число. Алгоритмы основаны на итерациях, каждая из которых сводит предыдущее значение ДПФ к ДПФ степени  $N/2$ , т.е. исходная последовательность из  $N$  точек разбивается на две подпоследовательности из  $N/2$  точек, что позволяет сократить число операций умножения в два раза.

Пусть

$$W_N = e^{-\frac{2i\pi}{N}}, \quad (2.3)$$

$$X(k) = \sum_{n=0}^{N-1} W_N^{nk} x(n)$$

( $N$  – период точечной последовательности)

Разобьем исходную последовательность:

$$X_1 = x(2n), X_2 = x(2n+1), n = 0, 1, 2, \dots, (N/2 - 1). \quad (2.4)$$

Тогда получим следующее:

$$\begin{aligned} X(k) &= \sum_{n=0}^{N-1} W_N^{nk} x(n) = \sum_{n=0}^{\frac{N}{2}-1} W_N^{nk} x_1(n) + \sum_{n=0}^{\frac{N}{2}-1} W_N^{nk} x_2(n) = \\ &= \sum_{n=0}^{\frac{N}{2}-1} W_{N/2}^{2nk} x(2n) + \sum_{n=0}^{\frac{N}{2}-1} W_{N/2}^{(2n+1)k} x(2n+1) = \\ &= X_1(k) + X_2(k) \cdot W_N^k. \end{aligned} \quad (2.5)$$

$$W_N^{k+N/2} = e^{-i\frac{2\pi}{N}(k+N/2)} = W_N^k (\cos \pi - i \sin \pi) = -W_N^k. \quad (2.6)$$

$$X(k + \frac{N}{2}) = X_1(k) - X_2(k) \cdot W_N^k, \quad (2.7)$$

$$X(k) = X_1(k) + X_2(k) \cdot W_N^k.$$

Произведя расчеты, определяем количество входных нейронов, которое составит 11025. После проведенных преобразований спектр вибросигнала накладывается на вход нейросети. Вход нейросети — это линейно упорядоченный массив элементов, которым будем присваивать уровни соответствующих частот в спектре. Эти элементы не выполняют никаких решающих функций, а только передают сигналы дальше в нейросеть.

## 2.4.2. Слой Кохонена

Слой Кохонена классифицирует входные векторы в группы схожих, что достигается такой подстройкой весов слоя Кохонена, когда близкие входные векторы активируют один и тот же нейрон данного слоя.

Ассоциированное с каждым нейроном Кохонена множество весов соединяет его с каждым входом. Например, на рис.2.4 нейрон Кохонена  $K_1$  имеет веса  $w_{11}, w_{21}, \dots, w_{m1}$ , составляющие весовой вектор  $\mathbf{W}_1$ . Они соединяются через входной слой с входными сигналами  $x_1, x_2, \dots, x_m$ , составляющими входной вектор  $\mathbf{X}$ . Подобно нейронам большинства сетей выход NET каждого нейрона Кохонена является просто суммой взвешенных входов. Это может быть выражено следующим образом:

$$NET_j = w_{1j}x_1 + w_{2j}x_2 + \dots + w_{mj}x_m \quad (2.8)$$

где  $NET_j$  – это выход NET нейрона Кохонена  $j$ ,

$$NET_j = \sum_i k_i w_{ij} \quad (2.9)$$

или в векторной записи

$$\mathbf{N} = \mathbf{XW}, \quad (2.10)$$

где  $\mathbf{N}$  – вектор выходов NET слоя Кохонена.

Нейрон Кохонена с максимальным значением NET является «победителем». Его выход равен единице, у остальных он равен нулю.

Чтобы получить наилучшие результаты от применения НС для предварительной классификации, например амплитудно-частотных спектров вибросигналов, необходимо изначально знать, сколько классов присутствует в обрабатываемых данных. Если количество нейронов окажется недостаточным, то НС отнесет спектры «смежных» классов к более представительным классам. Если перезадаваться числом нейронов, то «смежные» спектры НС отнесет к несуществующим на самом деле классам. В связи с этим, в работе определим количе-

ство нейронов слоя Кохонена экспериментальным путем. Необходимым и достаточным будем считать такое количество, при котором нейросеть будет правильно классифицировать предъявляемые на ее вход образы.

### 2.4.3 Слой Гроссберга

Слой Гроссберга функционирует в сходной манере. Его выход NET является взвешенной суммой выходов  $k_1, k_2, \dots, k_n$  слоя Кохонена, образующих вектор  $\mathbf{K}$ . Вектор соединяющих весов, обозначенный через  $\mathbf{V}$ , состоит из весов  $v_{11}, v_{21}, \dots, v_{np}$ . Тогда выход NET каждого нейрона Гроссберга есть

$$NET_j = \sum_i k_i w_{ij}, \quad (2.11)$$

где  $NET_j$  – выход  $j$ -го нейрона Гроссберга, или в векторной форме

$$\mathbf{Y} = \mathbf{KV}, \quad (2.12)$$

где  $\mathbf{Y}$  – выходной вектор слоя Гроссберга;

$\mathbf{K}$  – выходной вектор слоя Кохонена;

$\mathbf{V}$  – матрица весов слоя Гроссберга.

Если слой Кохонена функционирует таким образом, что лишь у одного нейрона величина NET равна единице, а у остальных равна нулю, то лишь один элемент вектора  $\mathbf{K}$  отличен от нуля, и вычисления очень просты. Фактически каждый нейрон слоя Гроссберга лишь выдает величину веса, который связывает этот нейрон с единственным ненулевым нейроном Кохонена.

## 2.5. Обучение нейросети

Обучение состоит в многократном предъявлении характерных примеров, а также адаптивной модификации весовых коэффициентов до тех пор, пока НС на своем выходе не станет выдавать желаемый отклик. Например, для предсказания качества поверхности это значения параметров качества (шероховатость, некруглость, волнистость). Обучение может вестись с учителем или без него. В первом случае сети предъявляются значения как входных, так и желательных выходных сигналов, и она по некоторому внутреннему алгоритму подстраивает

веса своих синаптических связей. Существует и другой принцип самодиагностики весовых коэффициентов. Суть его в том, что модификации подвергаются весовые коэффициенты, соответствующие определенным, получившим приоритет нейронам. Сначала определяются согласно выбранному критерию выигравший нейрон и нейроны из его окружения. На следующем этапе обучения область вокруг выигравшего нейрона сжимается, а шаг адаптации уменьшается. Обучение прекращается после многократных повторов, когда выигрывать начнет один и тот же нейрон.

### 2.5.1. Обучение слоя Кохонена

Так как в конце обучения вектора весов будут располагаться на единичной гиперсфере, то желательно, чтобы вектора весов формировались случайным образом на гиперсфере единичного радиуса. Входной вектор, содержащий информацию об энергетическом спектре и параметрах качества детали нормируется на 1.0 по формуле:

$$x_i' = \frac{x_i}{\sqrt{x_1^2 + x_2^2 + \dots + x_n^2}}. \quad (2.13)$$

Это превращает входной вектор в единичный вектор с тем же направлением в  $n$ -мерном пространстве. Далее нормированный на 1.0 входной вектор подается на вход, который распределяет его дальше через матрицу весов  $\mathbf{W}$  на нейроны слоя Кохонена. В отличие от классической сети встречного распространения каждый нейрон в слое Кохонена вычисляет не скалярное произведение весового вектора на входной, а *среднеквадратическое отклонение* по формуле:

$$NET_j = \sum_i \sqrt{(X_i - W_{ij})^2}. \quad (2.14)$$

Нейрон с наименьшим среднеквадратическим отклонением становится активным, и его выход устанавливается в 1. Так как отработка этого механизма требует значительных вычислительных ресурсов, его лучше заменить нахождением нейрона с максимальной активностью и присвоением ему активности 1, а всем остальным нейронам 0. Таким образом, срабатывает нейрон, для которого входной вектор ближе всего к вектору весов связей.

Когда сеть находится в режиме обучения, для выигравшего нейрона происходит коррекция весов матрицы связи по формуле

$$w_H = w_c + \alpha(x - w_c), \quad (2.15)$$

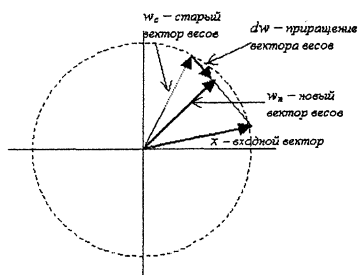
где  $w_H$  — новое значение веса;

$w_c$  — старое значение;

$\alpha$  — скорость обучения;

$x$  — величина входа.

Геометрически это правило проиллюстрировано на рис. 2.6 (двумерный случай).



**Рис.2.6. Иллюстрация механизма обучения**

Метод обучения Кохонена обладает полезной и интересной способностью извлекать статистические свойства из множества входных данных. Для полностью обученной сети вероятность того, что случайно выбранный входной вектор будет ближайшим к любому заданному весовому вектору, равна  $1/k$ , где  $k$  — число нейронов Кохонена. Это является оптимальным распределением весов на гиперсфере.

### 2.5.2. Обучение слоя Гроссберга

Обучение слоя Гроссберга — это обучение с учителем, алгоритм располагает желаемым выходом, по которому он обучается. Основная задача слоя Гроссберга при обучении — *статистическое усреднение сходных входных векторов*. Его вход NET является взвешенной суммой выходов  $k_1, k_2, \dots, k_m$  слоя Кохонена, образующих бинарный вектор

**К.** Вектор соединяющих весов, обозначенный через **V**, состоит из весов  $v_{11}, v_{21}, \dots, v_{nm}$ . Тогда выход NET каждого нейрона Гроссберга есть

$$NET_i = \sum_j k_j w_{ji}, \quad (2.16)$$

где  $NET_i$  – выход  $i$ -го нейрона Гроссберга, или в векторной форме

$$Y = KV, \quad (2.17)$$

где **Y** – выходной вектор слоя Гроссберга;

**K** – выходной вектор слоя Кохонена;

**V** – матрица весов слоя Гроссберга.

Если слой Кохонена функционирует таким образом, что лишь у одного нейрона величина NET равна единице, а у остальных равна нулю, то лишь один элемент вектора **K** отличен от нуля и вычисления очень просты. Фактически каждый нейрон слоя Гроссберга лишь выдает величину веса, который связывает этот нейрон с единственным ненулевым нейроном Кохонена.

При обучении слоя Гроссберга входной вектор, являющийся выходом слоя Кохонена, подается на слой нейронов Гроссберга, и выходы слоя Гроссберга вычисляются, как при нормальном функционировании. Далее, каждый вес корректируется лишь в том случае, если он соединен с нейроном Кохонена, имеющим ненулевой выход. Величина коррекции веса пропорциональна разности между весом и требуемым выходом нейрона Гроссберга, с которым он соединен. В символической записи:

$$v_{ijn} = v_{ijc} + \beta (y_j - v_{ijc})k_i, \quad (2.18)$$

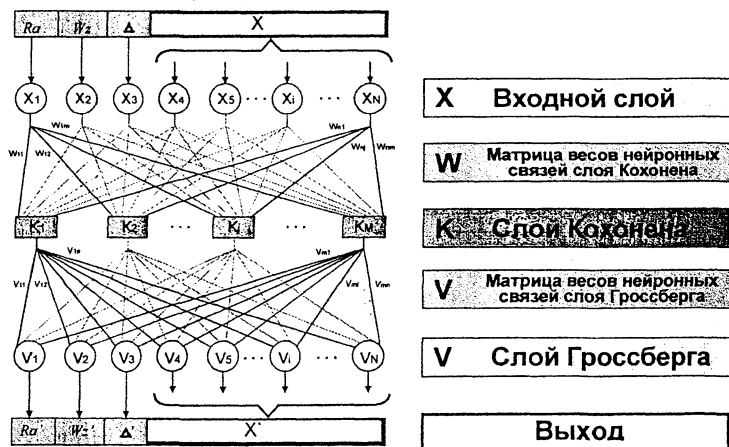
где  $k_i$  – выход  $i$ -го нейрона Кохонена (только для одного нейрона Кохонена он отличен от нуля);

$y_j$  –  $j$ -я компонента вектора желаемых выходов.

Первоначально  $\beta$  берется равным  $\sim 0,1$  и затем постепенно уменьшается в процессе обучения.

## 2.6. Применение ИНС встречного распространения в контексте задачи определения параметров качества

В контексте задачи определения параметров качества детали топология классической сети встречного распространения претерпевает ряд изменений (см. рис. 2.7).



**Рис.2.7. Топология сети встречного распространения в контексте задачи**

Эти изменения связаны с наличием в обучающей выборке помимо спектрального представления акустического сигнала от станка ещё и значений параметров качества обрабатываемой детали, которые добавляются в виде ещё трёх компонент входного вектора. При обучении вектор, содержащий спектр акустического сигнала и значения параметров качества детали, подаётся и как входной вектор сети и как желаемый выходной вектор. При использовании обученной нейросети на вход вместо полного входного вектора подаётся входной вектор, содержащий только спектр акустического сигнала. Сеть встречного распространения благодаря свойству восстановления информации на выходе дает полный вектор, содержащий помимо усреднённого энергетического спектра акустического сигнала ещё и искомые параметры качества детали.

Обучение разработанной ИНС для задач распознавания качества, проводилось в несколько этапов:

1. Составление базы данных, характерных для данной задачи.
2. Разбиение всей совокупности данных на два подмножества: обучающее и тестовое.
3. Оптимизация архитектуры сети с целью выбора наилучшей для решения задачи классификации виброакустических сигналов.
4. Оценивание качества работы сети на основе подтверждающего множества.

На основе изложенного подхода был создан программный продукт, работающий под управлением Microsoft Windows для IBM-совместимого компьютера. Программа содержит все необходимые для проведения исследований инструменты, обладает удобным интерфейсом пользователя, средствами для визуализации входных сигналов и их спектрального представления, возможностью сохранения результатов обучения на жесткий диск и документирования. Интерфейс пользователя программы представлен на рис. 2.8.

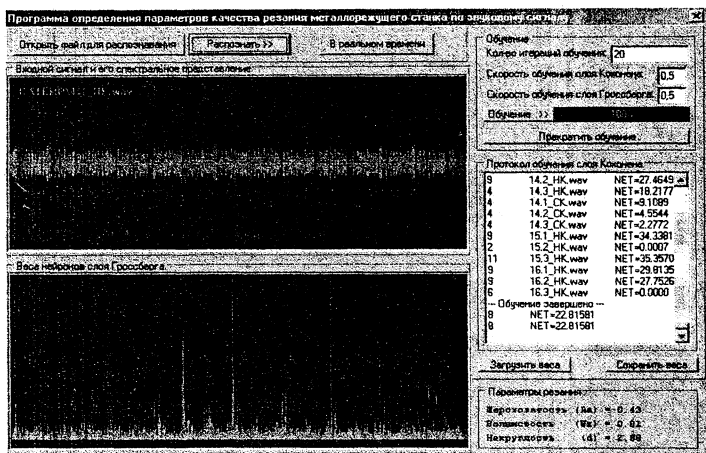
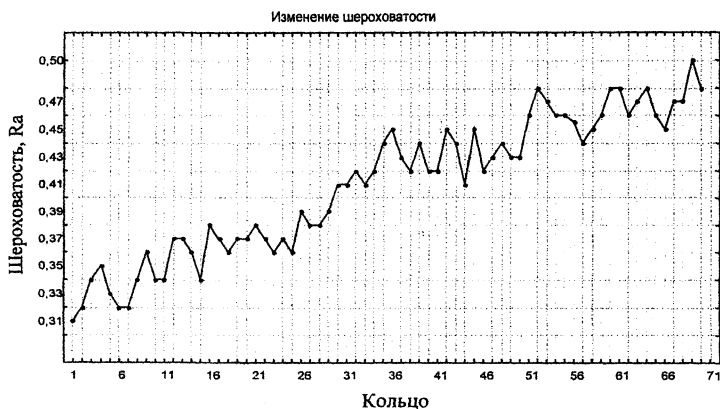


Рис.2.8. Окно программы определения показателей качества

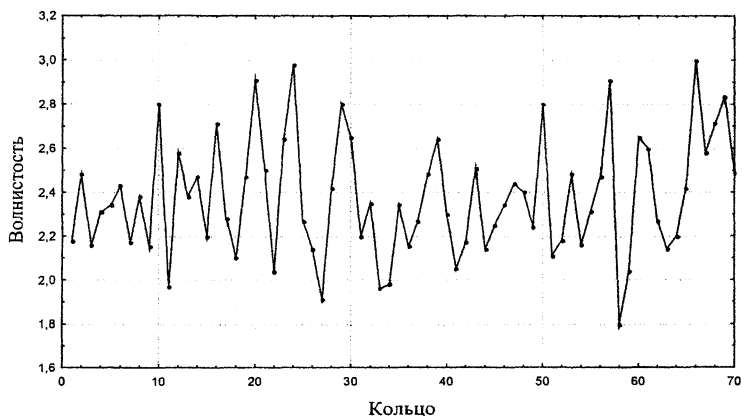


### 2.6.1. Формирование обучающей и тестовой выборки

Для формирования обучающей выборки был проведен эксперимент на внутришлифовальном станке SIW4 при обработке колец подшипников на операции окончательного шлифования с кругами различной структуры. Показатели качества колец замерялись в лаборатории с применением аттестованных измерительных приборов Talysond 73 System. Сигналы о виброколебаниях станка были записаны в полосе частот 2...1000 Гц пьезодатчиком, установленным на магнитной опоре кольца по горизонтальной координате. Общее количество исходных данных для обучения составило порядка 200 выборок. После анализа полученных данных они были разбиты на два подмножества: обучающее и тестовое. На рис.2.9 и 2.10 представлены значения параметра шероховатости и волнистости для 70 колец из обучающей выборки. Шероховатость колец на графике представлена по мере ее возрастания.



**Рис.2.9. Показатель шероховатости колец в обучающей выборке**



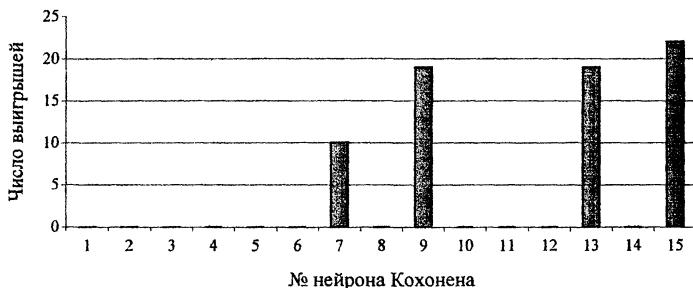
**Рис.2.10. Показатель волнистости колец в обучающей выборке**

## 2.6.2. Оптимизация архитектуры сети

Правильный выбор архитектуры сети имеет значение не меньшее, чем выбор исходных данных. Обработка результатов моделирования структуры сети выполнялась с использованием нейронно-сетевого пакета *STATISTICA Neural Networks*, представляющего собой реализацию всего набора нейросетевых методов анализа.

При выборе архитектуры сети исследовалось две конфигурации с различным количеством нейронов слоя Кохонена. На первом этапе моделирования было произведено исследование НС с количеством нейронов слоя Кохонена, равным 18. На рис.2.11 представлены частоты выигрышей отдельных нейронов.

Из рис.2.11 видно, что нейроны слоя Кохонена при классификации входных сигналов задействованы неэффективно. Это привело к недостаточно точному распределению показателей качества по нейронам Кохонена (рис.2.12).



**Рис.2.11. Частота выигрышей отдельных нейронов слоя Кохонена**

В силу особенностей входного сигнала основная часть входных векторов не распределена равномерно по всей поверхности гиперсферы, а сосредоточена в некоторых небольших областях. При этом лишь небольшое количество весовых векторов способно выделить входные векторы, и в этих областях возникает дефицит нейронов, тогда как в областях, где плотность входных векторов намного ниже, число нейронов избыточно.

Для решения проблемы неоднородного распределения входных векторов существует несколько методов. Часто применяется правило «нахождения центра масс», т.е. небольшое стремление всех весовых векторов на начальном этапе обучения к входным векторам. В результате в местах с большой плотностью входного сигнала оказывается много весовых векторов. Это правило записывается так:

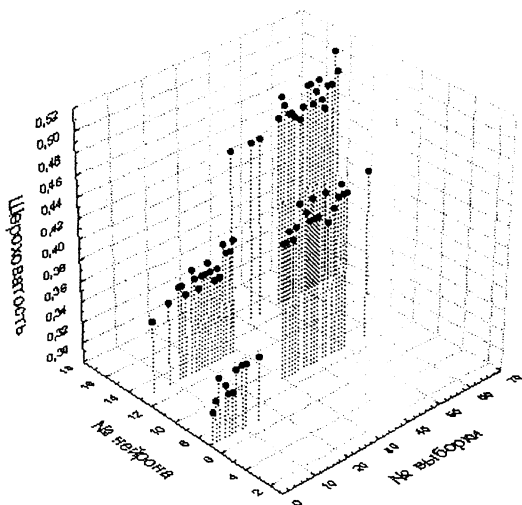
$$w_n = w_c + \alpha(x - w_c), \quad (2.19)$$

где  $w_n$  — новое значение веса;

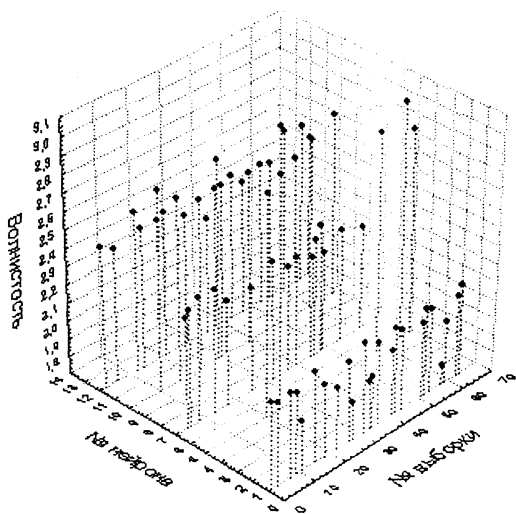
$w_c$  — старое значение;

$\alpha$  — скорость модификации;

$x$  — входной вектор.



a)



б)

Рис.2.12. Классификация показателей качества по нейронам слоя Кохонена: а - шероховатости, б - волнистости

Если весовой вектор окажется далеко от области входных сигналов, он никогда не даст наилучшего соответствия, всегда будет иметь нулевой выход, следовательно, не будет корректироваться и окажется бесполезным. Оставшихся же нейронов может не хватить для разделения входного пространства сигналов на классы. Для решения этой проблемы применяется правило *«желания работать»*: если какой-либо нейрон долго не находится в активном состоянии, он повышает веса связей до тех пор, пока не станет активным и не начнет подвергаться обучению. Правило *«желание работать»* записывается в следующей форме:

$$w_n = w_c + w_c \beta (1 - \alpha), \quad (2.20)$$

где  $w_n$  – новое значение веса;

$w_c$  – старое значение;

$\beta$  – скорость модификации;

$\alpha$  – активность нейрона.

Этот метод позволяет также решить проблему тонкой классификации: если образуется группа входных сигналов, расположенных близко друг к другу, с этой группой ассоциируется и большое число нейронов Кохонена, которые разбивают её на классы. Чем меньше активность нейрона, тем больше увеличиваются веса связей.

Самым эффективным решением является более точное моделирование механизма латерального торможения. То есть находится нейрон с максимальной активностью, затем искусственно при помощи латеральных связей устанавливается активность окружающих его нейронов по правилу:

$$a_j = e^{\frac{-(i-j)^2}{\theta}}, \quad (2.21)$$

где  $a_j$  – активность нейрона;

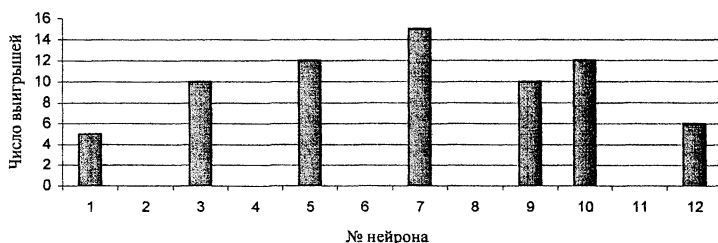
$i$  – выигравший нейрон;

$j$  – индекс нейрона;

$\theta$  – определяет радиус действия латеральных связей (уменьшается в процессе обучения).

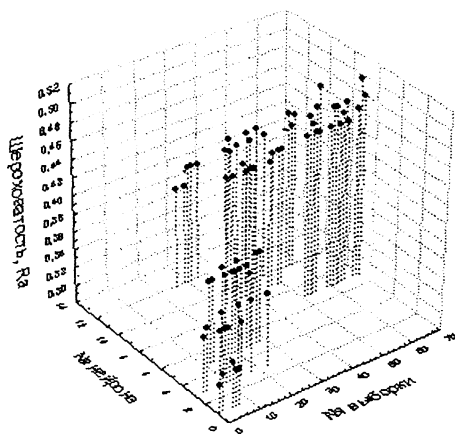
При этом предполагается, что все нейроны имеют определённую позицию по отношению к другим нейронам. Это топологическое отношение одномерно и линейно, позиция каждого нейрона определяется его индексом. Возбуждается не один нейрон, а группа топологически близких нейронов. В результате обучения образуется упорядоченная одномерная карта признаков. Упорядоченность означает, что ближайшие два нейрона в ней соответствуют двум ближайшим векторам в пространстве сигнала, но не наоборот (так как невозможно непрерывно отобразить многомерное пространство на одномерное). Сначала радиус действия латеральных связей достаточно большой, и в обучении участвуют практически все нейроны. При этом они находят «центр масс» всей обучающей выборки. В процессе обучения коэффициент  $\theta$  уменьшается, нейроны разделяются на группы, соответствующие локальным центрам масс. В конце концов, радиус латеральных связей снижается настолько, что нейроны функционируют независимо друг от друга и могут разделять очень близкие векторы.

В свете вышесказанного алгоритм обучения нейросети был доработан. Количество нейронов на втором этапе было сокращено до 12. Анализ протокола обучения показал, что использование вышеназванных правил позволяет избежать «паралича» сети, который проявился в постоянном выигрыше лишь ограниченного числа нейронов слоя Кохонена. На рис.2.13 представлено распределение входных векторов по нейронам после модернизации алгоритма обучения.

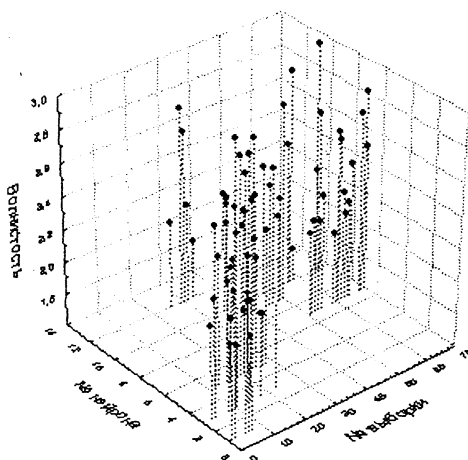


**Рис.2.13. Частота выигрышей отдельных нейронов слоя Кохонена после модернизации алгоритма обучения**

На рис.2.14 представлены распределения показателей качества по нейронам после модернизации алгоритма обучения.



а)



б

**Рис.2.14. Классификация показателей качества по нейронам слоя Кохонена после модернизации алгоритма обучения:**  
**а - шероховатости, б - волнистости**

Анализируя полученные результаты, можно прийти к выводу, что между параметрами качества и нейронами, которые классифицируют входной сигнал, существует явная зависимость, выражаемая в соответствии определенным нейронам колец с близкой шероховатостью  $R_a$  (рис.2.14, а) и волнистостью (рис.2.14, б). Наличие четкой кластеризации в рамках данной обучающей выборки доказывает возможность идентифицировать показатели шероховатости и волнистости по виброакустическим колебаниям. Зависимости между входными сигналами и параметрами некруглости  $W$ , в рамках данной обучающей выборки, отмечено не было. Это связано с основным формированием некруглости детали на предварительных операциях. К тому же на операции окончательного шлифования дорожки качения базирование кольца происходит по наружному диаметру, то есть воспроизводится его некруглость.

Слой Гроссберга обучался на основе априорной информации, представляющей собой измеренные с помощью приборов Talysond 73 и Surtronic 3+ параметры качества деталей. При обучении вектор, содержащий помимо спектра акустического сигнала еще и значения параметров качества детали, подаётся и как входной вектор сети и как желаемый выходной вектор. При поступлении каждого входного образа слой Кохонена (в соответствии с информацией, сформированной в нем во время самообучения) возбуждает один из своих нейронов и передает возбуждение на входы слоя Гроссберга. Далее производится подстройка синаптических весов нейронов этого слоя таким образом, чтобы выходные образы ИНС соответствовали выходным образам, поступившим в процессе обучения.

### 2.6.3 Корректность распознавания параметров качества сетью

Корректность распознавания параметров качества обученной нейросетью была проверена на тестовом множестве, состоящем из выборок, не вошедших в обучающее множество. Кольца подвергались обработке с различными вариациями параметров технологического режима.

Результаты измерений параметров качества колец представлены в табл. 2.1. Замер микрогеометрии проводился на приборах Talysond 73 и Surtronic 3+.



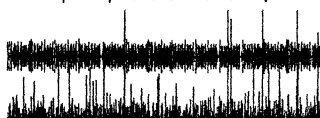
Таблица 2.1

Вариации параметров качества деталей при изменении параметров технологического режима

№ кольца	Параметры качества, мкм		
	$R_a$	$\Delta$	$w_z$
1	0,20	2,27	0,6
2	0,25	1,21	0,5
3	0,28	1,72	0,6
4	0,30	2,25	0,9
5	0,34	2,44	1,2
6	0,35	1,73	0,75
7	0,40	2,02	0,9
8	0,42	1,86	0,7
9	0,46	2,01	0,6
10	0,50	1,99	0,8

Результаты определения параметров качества деталей программой представлены на рис.2.15 и позволяют констатировать способность нейросети встречного распространения достаточно уверенно идентифицировать качество обработки по виду спектра вибросигнала станка, генерируемого в его упругой системе процессом резания.

Вход слоя Кохонена: спектры колебаний станка при обработке 5-го кольца



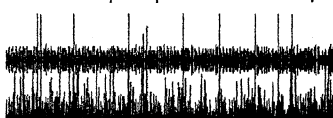
Спектральный выход слоя Гроссберга, соответствующий нейрону № 5 слоя Кохонена



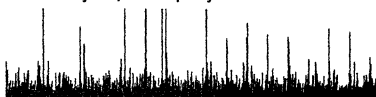
Параметрический выход слоя Гроссберга, соответствующий найденному спектру:

$R_a = 0,34$  мкм;  $w_z = 2,4$  мкм;  $\Delta = 1,1$  мкм

Вход слоя Кохонена: спектры колебаний станка при обработке 8-го кольца



Спектральный выход слоя Гроссберга, соответствующий нейрону № 12 слоя Кохонена



Параметрический выход слоя Гроссберга, соответствующий найденному спектру:

$R_a = 0,40$  мкм;  $w_z = 1,8$  мкм;  $\Delta = 0,8$  мкм

**Рис.2.15. Результаты программного распознавания параметров качества**

### 3. ПРИМЕНЕНИЕ СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ ОПТИМИЗАЦИИ В МЕХАНООБРАБОТКЕ

#### 3.1. Оптимизация в технике: общие вопросы

Технологическая система должна получать определенную оценку с помощью некоторых показателей качества, представляющих собой количественную характеристику системы. Эти показатели качества исследуемой системы (например, станка) должны находиться в тесной взаимосвязи с решаемыми системой задачами, а также средствами и способами их достижения и ее свойствами. Только в этом случае возможно целенаправленно управлять, например, настройкой станка по выбранным показателям качества. Очевидно, что в рассматриваемой нами предметной области эти показатели качества непосредственно связаны с процессом формообразования, реализуемым в соответствии с **требованиями к готовой детали и фактическим состоянием технологической системы в процессе эксплуатации.**

Состояние технологической системы характеризуется вектором состояния  $R^m \in M$  и соответствует желаемым требованиям к готовой детали  $R^n \in N$  в другой области функционального пространства состояний, причем, эти области различны и могут не пересекаться. К этому следует добавить, что одна и та же технологическая система может быть представлена различными векторами и областями определения, например, статическое и динамическое состояние; динамическое состояние, отображенное во временную или частотную Фурье - область. Размерности векторов состояния системы и их параметры могут отличаться друг от друга для каждой операции из всего множества реализуемой гибкой технологии. При этом оценка качества процессов управления должна выполняться с учетом требований, предъявляемых к реализациям объектов управления, связанных с их структурной перестройкой, и стратегии оптимизации. Однако и в этом случае оценка может получиться субъективной для

каждого станка и каждой реализации гибкой технологии. Таким образом, априорная реализация оптимизирующих воздействий может быть неточной. С учетом вышеизложенного рассмотрим критерии, принципы и примеры оптимизации в смежных задачах управления состоянием различных объектов [1, 2, 5, 6, 35].

Представленные в литературе критерии, характеризующие качество управления, дают возможность оценить систему на стадии проектирования по таким важным показателям, как качество работы, допустимые изменения параметров системы (например, с точки зрения запаса по фазе и амплитуде); расход энергии; быстродействие; условия достижения конечного состояния (например, прибытие объекта в некоторую область к моменту времени  $T_k$ ) и другим показателям.

В [45] отмечено, что число показателей качества  $J_1(t), \dots, J_l(t)$  обычно определяется требованиями, предъявляемыми к системе. Увеличение этого числа усложняет решение задач исследования и управления для рассматриваемых систем. Чтобы устранить эти трудности, в ряде случаев вводят обобщенный показатель качества систем  $J$ , являющийся функцией частных показателей  $J_1(t), \dots, J_l(t)$  и переходных функций состояний  $\varphi_i(t_0, t, z, x)$  ( $i=1, \dots, N$ ), где  $z$  - внутреннее состояние системы,  $X$  - входное воздействие. Характеристики систем являются функцией её структуры, представленной некоторым оператором  $W$ , значениями параметров (вектора параметров  $K$ ) и внешних воздействий  $X(t)$ . Таким образом, можно судить о качестве исследуемой системы на основании соотношения  $J=J\{J_1, \dots, J_l, Y_1, \dots, Y_n, X_1, \dots, X_m, t_1, t_{of}\}$ , где  $X$  - множество входных координат системы,  $Y$  - множество выходных координат системы. Знание такой общей взаимосвязи между показателем качества  $J$  и процессом изменения состояния системы позволяет решать задачи как анализа, так и синтеза систем.

Числовое значение обобщенного показателя  $J$ , которое соответствует определенному состоянию системы, и является критерием системы. В общем случае при использовании временной области

для различных операторов структуры системы значениям действительных чисел критерия соответствуют некоторые отображения множества процессов функционирования системы во времени. Данная функциональная связь позволяет трактовать критерий как функционал.

Все множество критериев можно разделить на две группы: регулярные и статистические. Системы, в которых все процессы являются детерминированными, исследуются с помощью регулярных критериев [45].

К регулярным, например, относятся критерии  $J$ , априорно выраженные через конструктивные параметры системы  $K$  без учета стохастической составляющей. Тогда  $J$  представляет  $J=J(K)$ .

В качестве критерия оптимальности могут быть приняты различные технические и экономические показатели: производительность, качество продукции, надежность затраты сырья или энергии. В зависимости от решаемой задачи необходимо достижение минимума либо максимума  $Q$ , например

$$Q(\bar{X}_{\text{вых}}, \bar{X}_{\text{вх}}, \bar{U}, \bar{F}, t) = \min, \quad (3.1)$$

где  $\bar{X}_{\text{вых}}$  - выходные координаты;

$\bar{X}_{\text{вх}}$  - входные координаты;

$\bar{U}$  - управляющие воздействия;

$\bar{F}$  - возмущающие воздействия.

Оптимальное управление системой определяется достижением максимального (минимального) значения функционала.

Если на систему воздействует случайный процесс  $X(t)$ , то векторы состояния  $Z$  и, следовательно, выходных координат  $Y$  представляют собой случайные процессы. Тогда критерий будет случайной величиной. Целесообразно ввести статистический критерий, являющийся неслучайной характеристикой случайного процесса. За наиболее общую форму статистического критерия может быть принято условное математическое ожидание [45, 58]

$$J = \int_{\chi} Q(\chi, K) f(\chi) d\chi, \quad (3.2)$$

где  $Q$  - функционал векторов  $\chi = \chi(X, Y, Z)$  и  $K$ ;  
 $f$  - закон распределения случайного процесса;  
 $X$  - входные координаты;  
 $Y$  - выходные координаты;  
 $Z$  - внутреннее состояние системы;  
 $K$  - конструктивные параметры системы.

Применительно к классу динамических систем управления существующие задачи анализа усложняются и обычно формулируются как задача определения статистического критерия системы, задаваемого в интегральной форме:

$$J = \int_{a_1}^{b_1} \dots \int_{a_m}^{b_m} (\Phi(\lambda_1, \dots, \lambda_m, t) f(\lambda_1, \dots, \lambda_m) d\lambda_1, \dots, \lambda_m, \quad (3.3)$$

где  $\Phi$  - критериальная функция, определяемая конкретным видом критерия и характеризующая реакцию исследуемой системы;

$\lambda_1, \dots, \lambda_m$  - система случайных величин;

$f$  - плотность распределения системы  $m$  случайных величин  $\lambda_1, \dots, \lambda_m$ .

При синтезе параметрически оптимизируемых систем для оценки качества управления чаще всего, как отмечено в работе [24], удобно использовать какой-либо единственный показатель. В частности, для непрерывных систем таким показателем может служить интегральный критерий качества (для дискретных систем вместо интеграла берется сумма). Следует отметить, что сумма квадратов ошибок управления предпочтительнее с математической точки зрения, кроме того, этот критерий может быть интерпретирован как средняя мощность и, в связи с этим, использоваться в других методах проектирования регуляторов.

Для параметрической оптимизации используются квадратичные критерии качества, представленные в следующем виде [24]

$$S_{lu}^2 = \sum_{k=0}^m (l^2(k) + r\Delta U^2(k)), \quad (3.4)$$

где  $l(k) = W(k) - Y(k)$  - ошибка управления;

$\Delta U(k) = U(k) - \bar{U}$  - «отклонение управляемой переменной» от установившегося значения  $\bar{U} = E\{U(k)\}$  для стохастических возмущений;

$r$  - весовой коэффициент при управляющей переменной.

В этом квадратичном критерии качества соотношение среднего квадрата ошибки управления

$$S_l^2 = \overline{l^2(k)} = \frac{1}{M+1} \sum_{k=0}^M \Delta U^2(k) \quad (3.5)$$

и усредненного квадратичного отклонения управляющей переменной или средней входной мощности

$$S_u^2 = \overline{\Delta U^2(k)} = \frac{1}{M+1} \sum_{k=0}^M \Delta U^2(k) \quad (3.6)$$

определяется выбором весового коэффициента  $r$ .

При оптимизации параметров регулятора параметры  $q^T = [q_0 q_1 \dots q_v]$  должны выбираться, так, чтобы обеспечить минимальное значение  $S_{lu}^2$ , т.е. выполнение условия  $\frac{dS_{lu}^2}{dq} = 0$ .

Для сравнения качества управления используются следующие достаточно простые показатели:

- среднеквадратичная ошибка управления

$$S_l' = \sqrt{\overline{l^2(k)}} = \sqrt{\frac{1}{M+1} \sum_{k=0}^M \Delta U^2(k)}; \quad (3.7)$$

- среднеквадратичное изменение управляющей переменной (затраты на управление)

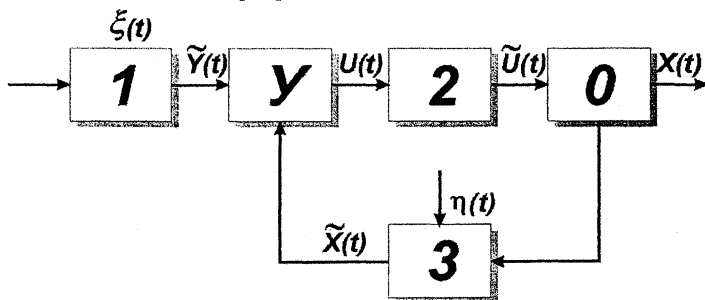
$$S_u' = \sqrt{\overline{U^2(k)}} = \sqrt{\frac{1}{M+1} \sum_{k=0}^M \Delta U^2(k)}; \quad (3.8)$$

- перерегулирование

$$Y_m = Y_{\max}(K) - W(K); \quad (3.9)$$

- время установления выходной координаты;
- начальное значение управляющей переменной  $U(0)$  при ступенчатом изменении сигнала  $W(0)$ .

Весьма распространенный способ задания критерия оптимальности основан на использовании функций штрафов, являющихся аналогами «расстояния» между элементами в некотором метрическом пространстве [27].



**Рис. 3.1. Функциональная схема системы оценки штрафов**

В момент времени  $t$  состояние системы (рис.3.1) характеризуется набором векторов  $X(t)$ ,  $Y(t)$ ,  $U(t)$ , штрафуются величиной  $C(X, Y, U)$ , где  $C(X, Y, U)$  – заданная неотрицательная функция своих аргументов. Если  $C(X, Y, U)$  имеет смысл удельных штрафов в единицу времени, то одним из возможных критериев оценки качества функционирования системы на отрезке времени  $[0, T]$  является интеграл вида

$$\int_0^T C(X(t), Y(t), U(t)) dt. \quad (3.10)$$

В силу случайного характера процессов  $X(t)$ ,  $Y(t)$  при любом фиксированном законе управления  $U(t)$ ,  $0 \leq t \leq T$  интеграл (3.10) является случайной величиной. Ее важной характеристикой является среднее значение, определяющее величину средних «затрат» или «потерь» на управление при многократном применении алгоритма управления. Эта величина

$$I_1[U] = M \int_0^T C(X(t), Y(t), U(t)) dt \quad (3.11)$$

может быть названа интегральным критерием оптимальности.

Часто рассматриваются задачи, в которых характер переходных процессов при  $0 \leq t \leq T$  несущественен, а важно лишь состояние системы в конечный момент времени  $T$  (терминальное управление). В этом случае, используя соответствующую функцию штрафа  $\phi[X(t), Y(t)]$ , можно получить терминальный критерий оптимальности

$$I_2[U] = M \Psi(X(T), Y(T)). \quad (3.12)$$

Расширением фазового вектора  $X$  интегральный критерий (3.11) может быть представлен в форме (3.12), и, таким образом, математически критерий (3.11) является частным случаем (3.12). Тем не менее, эти критерии в дальнейшем будут различаться, поскольку они имеют существенно различную техническую интерпретацию.

Помимо (3.11) и (3.12), часто используется критерий, являющийся их комбинацией:

$$J_3[U] = M \left[ \int_0^T C(X(t), Y(t), U(t)) dt + \psi(X(T), Y(T)) \right], \quad (3.13)$$

зависящей как от переходного процесса, так и от конечного состояния системы.

Если определяющим фактором является наихудшее (с точки зрения выбранной функции штрафа) состояние управляемой системы на фиксированном временном интервале  $[0, T]$ , то вместо интеграла (3.10) следует взять величину

$$\max_{0 \leq t \leq T} c(X(t), Y(t), U(t)). \quad (3.14)$$

Усредняя (3.14), получаем критерий оптимальности

$$J_4[U] = M \max C(X(t), Y(t), U(t)). \quad (3.15)$$

Оптимальная система, построенная из условия минимума критерия (3.15), обеспечивает наилучший результат лишь в наихудших режимах работы.



Качество работы с квадратичной функцией штрафа оценивает интегральная функция вида:

$$J(U) = M \int_t^T [qx^2(t) + hU^2(t)] dt. \quad (3.16)$$

Для ряда систем требуется рассчитать оптимальный регулятор (демпфер), который наилучшим образом, по критерию среднеквадратичной ошибки, успокаивает колебания, постоянно возникающие в системе из-за случайных возмущений  $\xi(t)$ . В качестве критерия оптимизации принимают [27]

$$J(U) = M \int_0^T [x^2(t) + \dot{x}^2(t) dt], \quad (3.17)$$

имеющий смысл средней энергии случайных колебаний в системе. Выбор среднеквадратичного критерия (3.17) для оценки качества системы является наиболее распространенным и, кроме того, он соответствует наиболее естественной постановке задачи оптимального демпфирования.

Представляют интерес в плане оптимизации качества «наблюдаемой» информационной модели статистического оценивания параметров. Предположим, что наблюдаемый сигнал  $\{Y\}$  имеет вид [164]

$$Y_t = S_t(\tau, k) + v_t, \quad (3.18)$$

где  $S_t$  - изменяющийся во времени полезный сигнал, зависящий известным образом от набора  $\tau$  информационных существенных параметров и набора  $K$  паразитных параметров;  $v_t$  - помеха наблюдения, изменяющаяся во времени  $t = 1, 2, \dots$

Требуется оценить информационные параметры  $\tau$ . В ряде случаев параметры  $\tau$ ,  $K$  предполагаются случайными с известными статистическими свойствами. В другом варианте набор параметров может быть неслучайным. Интерес представляет получение оптимальных оценок, соответствующих экстремизации функционала качества, например

$$W_t(\tau') = M \left\{ \sum_{i=1}^T |Y_i - S_i(\tau', k)|^2 / \tau \right\}, \quad (3.19)$$

где  $T$  - время наблюдения сигнала, а  $M$  означает операцию усреднения по ансамблю реализаций сигналов  $\{Y_i\}$ ,  $\{S_i\}$ , отвечающих фиксированному значению параметра  $\tau$ . Вычисление  $W_T(\tau')$  предполагает знание статистики сигналов  $\{Y_i\}$ ,  $\{S_i\}$ . Для этих целей возможны и другие разнообразные функционалы качества.

Близкой по смыслу к предыдущей задаче является задача о разладке, которая может произойти, например, при изнашивании инструмента

$$Y_t = S_t(\tau_t) + v_t, \quad (3.20)$$

где  $\tau_t$  — величина, определяющая статистические свойства полезного сигнала. Величина  $\tau_t$  изменяется во времени, оставаясь постоянной на интервалах значительной продолжительности. Требуется оценить моменты времени, когда происходит «переключение» величины  $\tau_t$ .

Представленный материал затрагивает не все аспекты техники, где применяются процессы оптимизации. Имеется много других работ по оптимизации. Из теоретических следует выделить работу [34], содержащую изложение теории оптимальных процессов, основным стержнем которой является принцип максимума. Этот принцип позволяет решать ряд задач математического и прикладного характера, которые являются вариационными, но не укладываются в классическую схему вариационного исчисления. Между тем к задачам такого неклассического типа приводят многие вопросы техники.

В заключение следует отметить, что существует большое количество критериев, но для всех них функционал оптимальных систем характеризует либо наилучшее поведение всей системы в динамике (при решении задач динамической оптимизации), либо наилучшие показатели в установившемся режиме (при решении задач статической оптимизации). В основном же, процесс оптимизации

производится на стадии конструирования систем или регуляторов. Информация по использованию оптимизации в реальном времени с учетом возможной структурной перестройки объекта при эксплуатации реальных технических систем с множеством нелинейных элементов и нестационарными свойствами в литературе практически не отражена, так как готовые «рецепты» в этом случае ввиду математической сложности решаемых задач вряд ли могут быть получены.

### **3.2. Оптимизация в металлообработке**

В настоящее время имеется множество работ по оптимизации качества формообразования. Здесь сознательно применен термин «качество формообразования», так как в нашем понимании современное представление об оптимальном управлении резанием включает в себя все наукоемкие разделы станкостроения: конструирование станков и инструментов; теорию резания; динамику станков; диагностику и идентификацию технологических систем в целом, инструментов и технологических жидкостей, а также исследования, как самих технологических процессов, так и возмущающих факторов различной природы и других явлений, определяющих действительное состояние технологической системы в процессе эксплуатации. Кроме того, представляет интерес как результат формообразования - качество готовой детали, так и сам процесс формообразования в реальном времени. С этих позиций и проведем анализ работ, посвященных оптимизации качества технологических процессов.

Еще в ранних исследованиях по резанию металлов нашли свое отражение соответствующие задачи оптимизации и качества. Основоположниками науки о резании металлов признаны русские ученые. Первый выдающийся труд о процессе резания металлов принадлежит талантливому русскому ученому профессору Петербургского горного института Ивану Августовичу Тиме (1838-1920 гг.). И.А.Тиме производил свои исследования в то время, когда научных

данных о процессе резания почти не было, возможности измерительной аппаратуры были ограничены, но, несмотря на это, установленные им закономерности в большей своей части не потеряли значения и в настоящее время. Теоретические положения, выдвинутые и обоснованные И.А.Тиме, в дальнейшем были уточнены и углублены русскими учеными А.Бриксом, проф. К.А.Зворыкиным, А.Н.Челюсткиным и другими. Более подробное изучение физической сущности процесса резания металлов было осуществлено самобытным ученым, мастером-механиком Петербургского политехнического института Я.Г.Усачевым. Особое значение имеют работы Я.Г.Усачева в области тепловых явлений, сопутствующих процессу резания металлов. Им разработан ряд конструкций термопар, применение которых дало возможность определить температурное поле резца и влияние скорости резания, подачи и глубины на температуру в зоне резания [11]. Эти работы нашли использование в [31] и в ряде других современных работах, что является хорошим примером преемственности науки в области резания металлов.

По мере развития науки о резании металлов все большее внимание уделялось рациональному (оптимальному) сочетанию режимов резания, геометрии режущего инструмента с учетом его взаимодействия с обрабатываемым материалом и наличия технологической жидкости для смазки и охлаждения в зоне резания. Таким образом, поиск рациональных структур, процессов и сочетаний возможных технологических, геометрических факторов, определяющих оптимальное по тому или иному критерию функционирование технологической системы, находился всегда в поле зрения проектировщиков и исследователей в области станкостроения.

Оптимизация на практике чаще всего обеспечивалась априорно и часто интуитивно за счет управления параметрами технологического режима ( $v$ ,  $S$ , и  $t$ ) по определенным критериям качества без учета фактического состояния технологического оборудования и возмущающих воздействий при эксплуатации. Были сформированы справочники, и выбор режимов резания в какой-то мере формализ-

зовался и производился по соответствующим нормативам. Нормативы по выбору режимов резания соответствуют средним (типовым) условиям работы и содержат сведения о рекомендуемых значениях  $v$ ,  $S$ , и  $t$  обычно в виде таблиц без учета возможностей адаптации к реальным условиям. Эти справочники в основном формировались на основе обобщения экспериментальных исследований по обрабатываемости различных металлов и ряда других факторов.

В начальной стадии процесса оптимизации формируют критерии оптимальности, статические и динамические ограничения. Критерии оптимальности необходимы для распознавания решений и, кроме того, составляют основу большинства используемых методов решения [48, 49].

Следует отметить, что наибольшее число опубликованных работ по оптимизации резания применяют критерии оптимизации, в той или иной степени связанные с экономической эффективностью. Это в первую очередь минимальная себестоимость, максимальная норма сменной выработки и так далее. К этим показателям часто сводят стойкость инструмента, выражая через нее какой-либо экономический показатель. Эти критерии оптимизации описаны во многих работах по резанию металлов.

Так, в [52] обосновывается динамика изменений затрат на обработку в условиях вариаций стойкости инструмента и предлагается оценка эффективности применения системы автоматического выбора оптимальной скорости резания только с точки зрения затрат. В работе [12] выделена переменная доля технологической себестоимости  $U$ , подсчитаны скорости резания и подача, обеспечивающие при периоде стойкости  $T_v$  максимально возможный объем срезаемого слоя. Разработанные в [6] алгоритмы позволяют априорно выбрать такие режимы резания, при которых производительность возрастает на 15-20%, а себестоимость снижается на 10-15% за счет обработки партии деталей за минимально возможное время без замены инструмента. В [53] для обработки корпусных деталей используется форсирование режимов резания путем увеличения минутной

подачи для повышения производительности и снижения себестоимости. Определены режимы резания, при дальнейшем ужесточении которых производительность обработки практически стабилизируется, а себестоимость возрастает из-за увеличивающегося расхода инструментов. Сведения по показателям качества обработки не приводятся. В исследованиях [52] отмечается улучшение поиска оптимального использования инструмента за счет введения в расчет верхней границы стойкости как функции износа режущей кромки. Оптимизация процессов обеспечивается таким образом, чтобы задача обработки была выполнена при минимуме производственных затрат в случае полной загрузки и при максимальном выигрыше во времени в случае перегрузки.

Наряду с этими примерами следует отметить работы, в которых критерии оптимизации связаны с техническим состоянием технологической системы и ее элементов. В работе [35] рассматривается математическое моделирование процесса резания для целей оптимизации, устанавливается статическая связь показателей качества с  $v$ ,  $S$ , и  $t$  и подробно оценивается влияние геометрии режущего инструмента на эти показатели. Приведены технологические ограничения.

В работе [60] усовершенствуется поиск оптимума использования инструмента введением в расчет верхней границы стойкости как функции износа режущей кромки.

В [9] предлагается методика определения оптимальных режимов резания, сущность которой заключается в выборе комплекса заданных параметров качества поверхностного слоя (высоту неровностей обработанной поверхности, величину остаточных напряжений в поверхностном слое, глубину наклепанного слоя) и определении оптимального сочетания скорости резания и подачи по разработанной авторами формуле.

В [51] предложен критерий максимальной технологической надежности, которая связана с потерей качества обработки. Рассмотрены особенности построения технических ограничений для про-

дольного точения и фрезерования. Предложена компромиссная целевая функция. Например, двухкритериальная задача сведена к минимизации функции  $F(v, S)$ .

Для повышения производительности обработки В.Н. Подураев в работе [43] предложил методы назначения оптимальных режимов резания при нестационарных методах обработки, за счет сокращения машинного времени.

В [19] используются статистические критерии точности и стабильности технологических операций.

В работе [56] для оценки эффективности вариантов конструкций сборного режущего инструмента (РИС) и условий его эксплуатации сформулирована система критериев, объединяющая: минимальное время на операцию механообработки; наибольшую стойкость РИС; минимальные стоимость нового инструмента, затраты на инструмент, себестоимость операций механообработки; наивысшую конкурентоспособность РИС.

Оптимизации процесса износа токарного инструмента при нестационарном резании, характеризующемся изменением режима обработки, посвящена работа [30]. Учитывается количество подналадок системы СПИД. Оптимизация технологических параметров позволила снизить себестоимость обработки на 20% .

В [33] реализуется структурно-параметрическая оптимизация с использованием многоцелевой функции (износ режущего элемента, сила, мощность и температура резания), учитывающая технологические ограничения  $v$ ,  $S$ , и  $t$  и прочность режущего инструмента. Исходными параметрами математической модели являются механические и теплофизические характеристики материала заготовки, технические требования к точности и качеству обработанной детали, жесткость элементов технологической системы.

В [3] предложен способ оптимизации процесса резания, по которому измеряют уровень ускорения звуковых колебаний, возникающих в зоне резания, и приращения активной мощности двигателя шпинделя, а выбор оптимальных технологических параметров

производят по первому отрицательному наименьшему приращению уровня ускорения звуковых колебаний и первому наибольшему приращению активной мощности двигателя.

В [55] в основу предлагаемых критериев оптимизации заложен принцип минимума энергии. С его помощью определяется энергия, накапливаемая поверхностным слоем деталей, а также энергетические затраты на процесс резания в целом. Приводится в качестве критерия оптимизации процесса резания его удельная энергоемкость.

В работе [59] удельная энергоемкость процесса резания рассчитывается как

$$\eta_1 = (N_1 - N_2) / (v \bullet S \bullet t) \quad , \quad (3.21)$$

где  $N_1$  - мощность станка, потребляемая под нагрузкой;

$N_2$  - мощность станка на вспомогательном ходу.

Скрытая энергия деформирования поверхностного слоя была предложена [54] в качестве критерия комплексной оценки физико-механического состояния поверхностного слоя детали, обработанной резанием. Его величина представляет собой плотность энергии, накопленной поверхностным слоем в результате пластической деформации при резании, и является энергией дислокаций. Проведены исследования связи указанных критериев со стойкостью инструментов. На эксплуатационные характеристики скрытая энергия оказывает более сильное влияние, чем шероховатость обработанной поверхности [54].

Метод акустической эмиссии (АЭ) [44] позволяет осуществить диагностирование износа инструмента, определить качество обработанной поверхности, разработать методики эспресс-оптимизации режимов резания и геометрии заточки инструмента, т.е. исследовать комплекс показателей обрабатываемости как существующих, так и перспективных материалов, выбрать марку рационального инструментального материала, оценить фактор технологической наследственности. Физические принципы метода связаны с основными



ми процессами структурообразования обрабатываемого материала и представляют собой высокочастотные волны упругой деформации, которые генерируются в зоне резания. Параметры данных волновых процессов непосредственно отражают энергетическую картину резания, зависят от ее изменения и могут быть зарегистрированы соответствующей измерительной аппаратурой.

В действующем производстве метод АЭ позволяет корректировать режимы обработки с целью их интенсификации, осуществлять непрерывный или активный дискретный контроль за ходом технологического процесса обработки, а также проводить входной контроль обработки и входной контроль обрабатываемости материала и режущих свойств инструмента. Экономическая эффективность применения АЭ определяется повышением надежности процессов обработки, увеличением их производительности.

Высокая информативность и разрешающая способность метода АЭ обуславливают его перспективное применение в комплексно-автоматизированных металлообрабатывающих производствах, например при создании адаптивных систем управления процессом резания от ЭВМ. В этом случае появляется возможность разработки моделей аварийных ситуаций – нерасчетных режимов функционирования технологического процесса и, как следствие, решения задачи их прогнозирования и устранения. Хорошие результаты дает обработка сигналов АЭ по критерию

$$W_u = A \dot{N}_{\Sigma} / V_{рез.}, \quad (3.22)$$

где  $A$  - амплитуда сигнала.

Однако использование метода АЭ в реальном времени для гибкой технологии требует обширной базы исходных данных.

В работе [31] рассмотрен целый комплекс вопросов оптимизации процессов резания конструкционных, нержавеющей и жаропрочных сталей и сплавов при условии постоянства оптимальной температуры резания. Указанное условие составляет основу метода ускоренного определения оптимальных сочетаний параметров про-

цесса резания, соответствующих максимальной размерной стойкости инструмента. При этом на одной из подач исследуют интенсивность износа для пяти – семи скоростей резания и ту скорость, на которой износ минимальный, принимают за оптимальную и измеряют на ней термоЭДС. Аналогичные испытания проводят на других подачах. Для применения метода в процессе эксплуатации оборудования необходимы апостериорные экспериментальные исследования.

В работе [50] выделяются две группы критериев оптимизации: термодинамические критерии (аналоги производства избыточной энтропии) и гидродинамические критерии (аналоги производства кинетической энергии от флуктуаций скорости и давления). Полученные экспериментальные результаты показывают, что тенденции изменения указанных критериев в зависимости от варьируемого параметра внешних воздействий на трибосистему (скорость, нагрузка), носят противоположный и экстремальный характер в критической точке, соответствующей образованию диссипативной структуры и минимизации изнашивания.

В работе [23] процесс резания рассматривается как нелинейная и нестационарная среда, с которой взаимодействует упруго-диссипативная система металлорежущего станка. Причем ее нестационарность учитывается силовой эмиссией.

В [20] сообщается о методах анализа и синтеза динамического качества фрезерных станков по критериям уровня вынужденных колебаний и предельной глубины резания, при которой нарушается устойчивость процесса фрезерования. В основе этого метода лежат априорно построенные математические модели, алгоритмы и программные комплексы многокритериальной многопараметрической оптимизации и функции чувствительности. Обнаружена взаимосвязь колебаний несущей системы с крутильными колебаниями привода главного движения. При этом увеличение (уменьшение) амплитуд колебаний элементов привода главного движения приводило к уменьшению (увеличению) амплитуд колебаний элементов несущей

системы, что свидетельствовало о перекачке энергии из одной подсистемы в другую.

В работе [18] допустимые (критические) значения амплитуд колебаний определяются исходя из требований к точности станка и обрабатываемой детали, качеству обработанной поверхности, стойкости инструмента и т.д. В [7] осуществляется расчет нелинейных колебаний станков на ЭВМ, результаты которого могут быть использованы для целей оптимизации.

В работе [17] установлено, что оптимизация параметров технологической системы механической обработки (ТСМО) по критерию устойчивости является одной из наиболее важных и часто встречающихся задач оптимизационного проектирования в технологических процессах механической обработки маложестких заготовок, которая решается как задача параметрической оптимизации с критериями эффективности, отражающими степень устойчивости динамической модели технологической системы в пространстве варьируемых параметров. При наличии у базовой модели в рабочем скоростном диапазоне резонансных режимов также решена задача частотной отстройки, которая реализуется в большей степени за счет варьирования скорости резания и в меньшей степени – упруго инерционных параметров подсистемы ТСМО. Для решения задачи обеспечения стабильности качества разработан обобщенный векторный критерий качества, представляющий собой совокупность критерия размерной точности и вектора состояния поверхностного слоя.

Имеется ряд других подобных работ, имеющих несколько иную направленность, например, по диагностированию изнашивания инструмента, специальных исследований его стойкости.

С учетом сформулированной в работе цели исследования оптимизация должна производиться в реальном времени, то есть в процессе эксплуатации станка в производственных условиях с учетом фактических состояний его динамической системы, которые, как

было отмечено выше, изменяются из-за множества внутренних и внешних программируемых и возмущающих факторов.

Анализируя рассмотренные выше работы, можно отметить, что в такой постановке задача оптимизации полностью не решена. Прежде всего, это связано с оперативностью получения первичной информации о состоянии всей динамической системы в целом, включая и процесс резания. В этом смысле, во-первых, анализировать отдельно разомкнутую технологическую систему без резания не имеет смысла, так как наличие нелинейных элементов и изменение в пространстве и во времени вектора силы резания может переводить систему из одного структурно-устойчивого состояния в другое или же приводить к структурно-неустойчивым состояниям. Во-вторых, идентификация системы для целей оптимизации должна осуществляться по свойствам информации, получаемой непосредственно в реальном времени при резании. Такую информацию, как следует из анализа опубликованных работ и наших исследований, можно получить реально только на основе анализа процессов, сопровождающих функционирование технологической системы и ее элементов: выделение тепла при резании, механические колебания в зоне резания и в других подсистемах, акустическая эмиссия и изменение параметров электромагнитного поля в зоне резания. Непосредственно же сам процесс резания генерирует (выделяет) тепловой поток, виброакустические колебания, акустическую эмиссию, электромагнитное поле и формирует с учетом временных воздействий состояние вектора силы в пространстве и во времени. Для всех этих процессов имеются в той или иной степени апробированные датчики и преобразователи для различных целей, в том числе и для оптимизации резания, диагностирования состояния инструмента и так далее. Однако количественная оценка и выделение требуемого структурно-устойчивого состояния для оптимизации качества формообразования в реальном времени требует специальной информационной технологии.

Рассмотрим один из методов, позволяющих оптимизировать процессы на основе итерационных процедур генетических алгоритмов, где целевая функция определяется параметрами качества формообразования, получаемыми на основе нейронных сетей, представленных выше.

### **3.3. Использование генетических алгоритмов для оптимизации**

Генетические алгоритмы (ГА) – адаптивные методы поиска, которые в последнее время часто используются для решения задач функциональной оптимизации. Они основаны на генетических процессах биологических организмов: биологические популяции развиваются в течение нескольких поколений, подчиняясь законам естественного отбора и по принципу «выживает наиболее приспособленный» (survival of the fittest), открытому Чарльзом Дарвином. Подражая этому процессу, генетические алгоритмы способны «развивать» решения реальных задач, если те соответствующим образом закодированы. Например, ГА могут использоваться, чтобы проектировать структуры моста, для поиска максимального отношения прочности/веса, или определять наименее расточительное размещение для нарезки форм из ткани. Они могут также использоваться для интерактивного управления процессом, например на химическом заводе, или балансировании загрузки на многопроцессорном компьютере. Вполне реальный пример: израильская компания Schema разработала программный продукт Channeling для оптимизации работы сотовой связи путем выбора оптимальной частоты, на которой будет вестись разговор. В основе этого программного продукта и используются генетические алгоритмы.

Основные принципы ГА были сформулированы Голландом (Holland, 1975) и хорошо описаны во многих работах. В отличие от эволюции, происходящей в природе, ГА только моделируют те процессы в популяциях, которые являются существенными для развития. Точный ответ на вопрос: какие биологические процессы суще-

ственные для развития и какие нет? – все еще открыт для исследователей.

В природе особи в популяции конкурируют друг с другом за различные ресурсы, такие, например, как пища или вода. Кроме того, члены популяции одного вида часто конкурируют за привлечение брачного партнера. Те особи, которые наиболее приспособлены к окружающим условиям, будут иметь относительно больше шансов воспроизвести потомков. Слабо приспособленные особи либо совсем не произведут потомства, либо их потомство будет очень немногочисленным. Это означает, что гены от высоко адаптированных или приспособленных особей будут распространяться в увеличивающемся количестве потомков на каждом последующем поколении. Комбинация хороших характеристик от различных родителей иногда может приводить к появлению «суперприспособленного» потомка, чья приспособленность больше, чем приспособленность любого из его родителей. Таким образом, вид развивается, лучше и лучше приспосабливаясь к среде обитания.

ГА используют прямую аналогию с таким механизмом. Они работают с совокупностью «особей» – популяцией, каждая из которых представляет возможное решение данной проблемы. Каждая особь оценивается мерой ее «приспособленности» согласно тому, насколько «хорошо» соответствующее ей решение задачи. Например, мерой приспособленности могло бы быть отношение силы/веса для данного проекта моста. В природе это эквивалентно оценке того, насколько эффективен организм при конкуренции за ресурсы. Наиболее приспособленные особи получают возможность «воспроизводить» потомство с помощью «перекрестного скрещивания» с другими особями популяции. Это приводит к появлению новых особей, которые сочетают в себе некоторые характеристики, наследуемые ими от родителей. Наименее приспособленные особи с меньшей вероятностью смогут воспроизвести потомков, так что те свойства, которыми они обладали, будут постепенно исчезать из популяции в процессе эволюции.

Так и воспроизводится вся новая популяция допустимых решений, выбирая лучших представителей предыдущего поколения, скрещивая их и получая множество новых особей. Это новое поколение содержит более высокое соотношение характеристик, которыми обладают хорошие члены предыдущего поколения. Таким образом, из поколения в поколение, хорошие характеристики распространяются по всей популяции. Скрещивание наиболее приспособленных особей приводит к тому, что исследуются наиболее перспективные участки пространства поиска. В конечном итоге, популяция будет сходиться к оптимальному решению задачи.

Имеются много способов реализации идеи биологической эволюции в рамках ГА. Традиционным считается ГА, представленный на схеме.

НАЧАЛО /\* генетический алгоритм \*/

Создать начальную популяцию

Оценить приспособленность каждой особи

останов := FALSE

ПОКА НЕ останов ВЫПОЛНЯТЬ

НАЧАЛО /\* создать популяцию нового поколения \*/

ПОВТОРИТЬ (размер\_популяции/2) РАЗ

НАЧАЛО /\* цикл воспроизводства \*/

Выбрать две особи с высокой приспособленностью из предыдущего поколения для скрещивания

Скрестить выбранные особи и получить двух потомков

Оценить приспособленности потомков

Поместить потомков в новое поколение

КОНЕЦ

ЕСЛИ популяция сошлась ТО останов := TRUE

КОНЕЦ

КОНЕЦ

В последние годы реализовано много генетических алгоритмов и в большинстве случаев они мало похожи на этот ГА. По этой причине в настоящее время под термином «генетические алгоритмы» скрывается не одна модель, а достаточно широкий класс алгоритмов, подчас мало похожих друг от друга. Исследователи экспериментировали с различными типами представлений, операторов кроссовера и мутации, специальных операторов и различных подходов к воспроизводству и отбору.

Хотя модель эволюционного развития, применяемая в ГА, сильно упрощена по сравнению со своим природным аналогом, тем не менее, ГА является достаточно мощным средством и может с успехом применяться для широкого класса прикладных задач, включая те, которые трудно, а иногда и вовсе невозможно решить другими методами. Однако ГА, как и другие методы эволюционных вычислений, не гарантирует обнаружения глобального решения за полиномиальное время. ГА не гарантируют и того, что глобальное решение будет найдено, но они хороши для поиска «достаточно хорошего» решения задачи «достаточно быстро». Там, где задача может быть решена специальными методами, почти всегда такие методы будут эффективнее ГА и в быстройдействии и в точности найденных решений. Главным же преимуществом ГА является то, что они могут применяться даже на сложных задачах, там, где не существует никаких специальных методов. Даже там, где хорошо работают существующие методики, можно достигнуть улучшения сочетанием их с ГА.

### **3.4. Символьная модель простого генетического алгоритма**

Цель в оптимизации с помощью ГА состоит в том, чтобы найти лучшее возможное решение или решение задачи по одному или нескольким критериям. Чтобы реализовать генетический алгоритм, нужно сначала выбрать подходящую структуру для представления этих решений. В постановке задачи поиска экземпляра этой структу-



ры данных представляет точку в пространстве поиска всех возможных решений.

Структура данных генетического алгоритма состоит из одной или большого количества хромосом (обычно из одной). Как правило, хромосома - это битовая строка, так что термин «строка» часто заменяет понятие «хромосома». В принципе ГА не ограничены бинарным представлением. Известны другие реализации, построенные исключительно на векторах вещественных чисел. Несмотря на то, что для многих реальных задач, видимо, больше подходят строки переменной длины, в настоящее время структуры фиксированной длины наиболее распространены и изучены.

Каждая хромосома (строка) представляет собой конкатенацию ряда подкомпонентов, называемых генами. Гены располагаются в различных позициях или локусах хромосомы и принимают значения, называемые аллелями. В представлениях с бинарными строками, ген - бит, локус - его позиция в строке, и аллель - его значение (0 или 1). Биологический термин «генотип» относится к полной генетической модели особи и соответствует структуре в ГА. Термин «фенотип» относится к внешним наблюдаемым признакам и соответствует вектору в пространстве параметров. Чрезвычайно простой, но иллюстративный пример - задача максимизации следующей функции двух переменных:

$$f(x_1, x_2) = \exp(x_1 x_2), \text{ где } 0 < x_1 < 1 \text{ и } 0 < x_2 < 1.$$

Обычно методика кодирования реальных переменных  $x_1$  и  $x_2$  состоит в их преобразовании в двоичные целочисленные строки достаточной длины - достаточной для того, чтобы обеспечить желаемую точность. Предположим, что 10-разрядное кодирование достаточно и для  $x_1$ , и  $x_2$ . Установить соответствие между генотипом и фенотипом закодированных особей можно, разделив соответствующее двоичное целое число - на  $2^{10}-1$ . Например, 0000000000 соответствует 0/1023 или 0, тогда как 1111111111 соответствует 1023/1023 или 1. Оптимизируемая структура данных - 20-битная строка, представляющая конкатенацию кодировок  $x_1$  и  $x_2$ . Перемен-

ная  $x_1$  размещается в крайних левых 10 разрядах, тогда как  $x_2$  размещается в правой части генотипа особи (20-битовой строке). Генотип – точка в 20-мерном хеммингговом пространстве, исследуемом ГА. Фенотип – точка в двумерном пространстве параметров.

Чтобы оптимизировать структуру, используя ГА, нужно задать некоторую меру качества для каждой структуры в пространстве поиска. Для этой цели используется функция приспособленности. В функциональной максимизации целевая функция часто сама выступает в качестве функции приспособленности (например, наш двумерный пример); для задач минимизации целевую функцию следует инвертировать и сместить затем в область положительных значений.

### **3.5. Работа простого генетического алгоритма**

Простой ГА случайным образом генерирует начальную популяцию структур. Работа ГА представляет собой итерационный процесс, который продолжается до тех пор, пока не выполнятся заданное число поколений или какой-либо иной критерий остановки. На каждом поколении ГА реализуются отбор пропорционально приспособленности, одноточечный кроссовер и мутация. Сначала отбор назначает каждой структуре вероятность  $P_s(i)$ , равную отношению ее приспособленности к суммарной приспособленности популяции. Затем происходит отбор (с замещением) всех  $n$  особей для дальнейшей генетической обработки, согласно величине  $P_s(i)$ . Простейший пропорциональный отбор – рулетка – отбирает особей с помощью  $n$  «запусков» рулетки. Колесо рулетки содержит по одному сектору для каждого члена популяции. Размер  $i$ -го сектора пропорционален соответствующей величине  $P_s(i)$ . При таком отборе члены популяции с более высокой приспособленностью с большей вероятностью будут чаще выбираться, чем особи с низкой приспособленностью.

После отбора  $n$  выбранных особей подвергаются кроссоверу (иногда называемому рекомбинацией) с заданной вероятностью  $P_c$ .

$n$  строк случайным образом разбиваются на  $n/2$  пары. Для каждой пары с вероятностью  $P_c$  может применяться кроссовер. Соответственно с вероятностью  $1-P_c$  кроссовер не происходит, и неизменные особи переходят на стадию мутации. Если кроссовер происходит, полученные потомки заменяют собой родителей и переходят к мутации.

Одноточечный кроссовер работает следующим образом. Сначала случайным образом выбирается одна из  $l-1$  точек разрыва. Точка разрыва – участок между соседними битами в строке. Обе родительские структуры разрываются на два сегмента по этой точке. Затем соответствующие сегменты различных родителей склеиваются и получаются два генотипа потомков.

Например, предположим, один родитель состоит из 10 нолей, а другой - из 10 единиц. Пусть из 9 возможных точек разрыва выбрана точка 3. Родители и их потомки показаны ниже.

Кроссовер.

Родитель 1    0000000000    000~0000000    >>>    111~0000000  
1110000000    Потомок 1

Родитель 2    1111111111    111~1111111    >>>    000~1111111  
0001111111    Потомок 2

После того, как закончится стадия кроссовера, выполняются операторы мутации. В каждой строке, которая подвергается мутации, каждый бит с вероятностью  $P_m$  изменяется на противоположный. Популяция, полученная после мутации, записывается поверх старой и этим цикл одного поколения завершается. Последующие поколения обрабатываются таким же образом: отбор, кроссовер и мутация.

В настоящее время исследователи ГА предлагают много других операторов отбора, кроссовера и мутации. Вот лишь наиболее распространенные из них. Прежде всего, турнирный отбор. Турнирный отбор реализует  $n$  турниров, чтобы выбрать  $n$  особей. Каждый турнир построен на выборке  $k$  элементов из популяции и выборе

лучшей особи среди них. Наиболее распространен турнирный отбор с  $k=2$ .

Элитные методы отбора (De Jong, 1975) гарантируют, что при отборе обязательно будут выживать лучший или лучшие члены популяции совокупности. Наиболее распространена процедура обязательного сохранения только одной лучшей особи, если она не прошла, как другие, через процесс отбора, кроссовера и мутации. Элитизм может быть внедрен практически в любой стандартный метод отбора.

Двухточечный кроссовер (Cavicchio, 1970; Goldberg, 1989) и равномерный кроссовер (Syswerda, 1989) - вполне достойные альтернативы одноточечному оператору. В двухточечном кроссовере выбираются две точки разрыва, и родительские хромосомы обмениваются сегментом, который находится между двумя этими точками. В равномерном кроссовере каждый бит первого родителя наследуется первым потомком с заданной вероятностью; в противном случае этот бит передается второму потомку.

### **3.6. Применение генетического алгоритма для оптимизации процесса формообразования**

Целью задачи поиска является нахождение оптимальных параметров режимов обработки (резания) станочного оборудования.

Чаще всего этими параметрами являются: скорость, глубина резания и подача ( $V$ ,  $S$  и  $t$ ). Параметры технологического режима варьируются в диапазоне: частота вращения - в диапазоне 400 ... 1600 мин<sup>-1</sup>, подача - в диапазоне 0.02 ... 0.1 мм/об, глубина резания - в диапазоне 0.1 ... 0.5 мм. Диапазоны допустимых значений параметров представляют собой пространство поиска.

Необходимо учитывать шаг изменения соответствующих параметров. Таким образом, пространство параметров поиска будет дискретным и представлять множество потенциальных решений. В рамках данной задачи пространство поиска принимает вид: частота вращения - в диапазоне 400 ... 1600 мин<sup>-1</sup> с шагом 10 мин<sup>-1</sup>, подача

- в диапазоне 0.02 ... 0.1 мм/об с шагом 0.01 мм/об, глубина резания
- в диапазоне 0.1 ... 0.5 мм с шагом 0.1мм.

Целевой функцией является:  $F(V, S, t) = \min F(Ra, Wz, \Delta)$ .  $Ra$  - шероховатость,  $Wz$  - волнистость,  $\Delta$  - некруглость (параметры качества детали).

Параметры обработки кодируются бинарной строкой  $s$ . Переход из пространства параметров в хеммингово пространство бинарных строк осуществляется кодированием параметров в двоичные целочисленные строки достаточной длины - достаточной для того, чтобы обеспечить желаемую точность. Желаемая точность в этом случае будет условием, которое определяет длину бинарных строк. Для этого пространство параметров должно быть дискретизировано таким образом, чтобы расстояние между узлами дискретизации соответствовало требуемой точности.

Мы изменили классический ГА следующим образом.

1. В качестве алгоритма скрещивания нами был выбран универсальный (или равномерный) кроссовер, наиболее соответствующий специфике данной задачи. В процессе кроссовера выбранная из популяции пара параметров обменивается битами с заданной вероятностью.

2. Мутация происходит таким же образом, как и в классическом ГА, т.е. с определенной вероятностью инвертируются биты параметров обработки.

3. Селекция производится методом «рулетки». При этом вероятность попадания параметров в новую популяцию определяется рейтингом, вычисление которого описано ниже.

Опишем подробно работу модифицированного ГА.

Сначала происходит генерация начальной популяции оптимизируемых параметров, где каждому параметру присваивается случайное значение. Далее циклически запускается механизм «эволюции».

Сначала заданное число раз производится скрещивание выбранных случайным образом представителей популяции. В процес-

се кроссовера первый выбранный член популяции претерпевает процесс мутации.

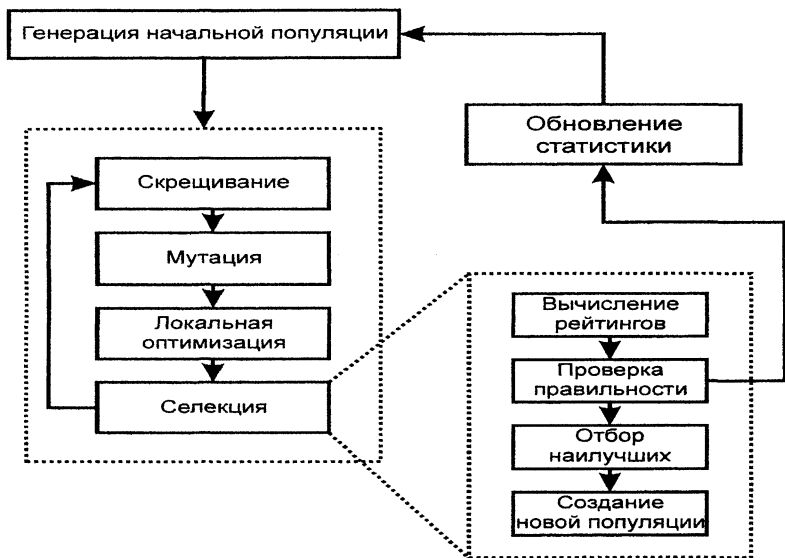
Далее запускается уникальный для нашего метода механизм локальной оптимизации. Его необходимость вызвана тем, что классический ГА дает лишь приближенное решение, а специфика данной задачи требует точного ответа. В процессе локальной оптимизации происходит небольшой перебор в окрестностях параметров. При каждом изменении вычисляется новый рейтинг оптимизируемых параметров. Если полученный рейтинг больше предыдущего, то изменения сохраняются, иначе отменяются. В результате мы получаем улучшенные параметры. В качестве дополнительной оптимизации имеется возможность оптимизировать только параметры с рейтингом выше среднего для данной популяции, что примерно вдвое сокращает временные затраты.

Алгоритм селекции реализует механизм рулетки, при котором каждому члену популяции отводится один сектор, размер которого пропорционален рейтингу и определяет вероятность попадания данного члена в новую популяцию. Сектора реализованы в виде одномерного массива, где в ячейках хранятся номера членов популяции в количестве, определяемом рейтингом. Для параметров с рейтингом выше среднего вероятность увеличивается вдвое, а для наилучших – еще вдвое.

Поскольку ГА является вероятностным алгоритмом, то для объективной оценки необходим сбор статистики. Сбор статистики осуществляется путем замеров времени и числа итераций, необходимых для нахождения ключа.

Схема алгоритма представлена на рис. 3.2.

Изложенные выше методы можно использовать для построения интеллектуальной системы управления технологическим оборудованием машиностроительного производства.



**Рис.3.2. Схема генетического алгоритма**

Интеллектуальные системы управления строятся как самообучающиеся, самонастраивающиеся системы с гибкими процедурами принятия решений об управлении, как системы, основанные на знаниях и формирующие новые знания в процессе управления и функционирования. В задачи интеллектуальной системы управления входят: экспертная оценка ситуации управления; синтез управляющих воздействий; анализ среды функционирования и т.д. Класс интеллектуальных систем управления соответствует следующим пяти принципам [20]:

- наличие взаимодействия управляющей системы с реальным внешним миром с использованием информационных каналов связи; интеллектуальные системы управления получают из него знания и влияют на него. Выполнение этого принципа позволяет организовать канал связи для извлечения знаний и организации целесообразного поведения;

- принципиальная открытость системы с целью повышения интеллектуальности и совершенствования собственного поведения (открытость системы обеспечивается наличием самонастройки, самоорганизации и самообучения). Система знаний интеллектуальной системы управления состоит из двух частей: поступающие знания и проверенные знания. Этот принцип позволяет организовать пополнение и приобретение знаний;

- наличие механизмов прогноза изменений среды функционирования и собственного поведения системы в динамически меняющемся внешнем мире. В соответствии с этим принципом интеллектуальная система управления не полностью интеллектуальна, если она не обладает возможностью прогноза изменений внешнего мира и собственного поведения;

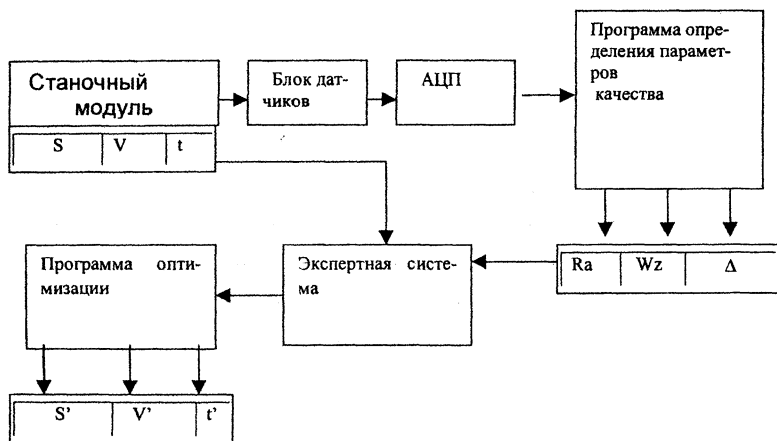
- наличие у системы структуры построения, соответствующей принципу IPDI (Increase of Precision with Decrease of Intelligence): точность управления тем выше, чем меньше интеллектуальность системы. Это путь построения сложных интеллектуальных систем управления в случае, когда неточность знаний о модели объекта управления или о его поведении может быть скомпенсирована за счет повышения интеллектуальности создаваемой системы;

- сохранение функционирования при разрыве связей или потере управляющих воздействий от вышестоящих уровней иерархии.

Рассмотренные свойства актуальны для изучаемого класса объектов.

Интеллектуальная система управления базируется на информационной системе анализа функционирования технологического оборудования по параметрам качества формообразования с применением искусственных нейронных сетей, являющихся базой знаний системы и оптимизации управляющих воздействий (параметров режимов резания ( $v$ ,  $S$  и  $f$ )) с использованием генетического алгоритма.





**Рис.3.3.Схема интеллектуальной системы управления станочным модулем**

Система состоит из следующих функциональных блоков:

- информационной системы анализа функционирования технологического оборудования по параметрам качества формообразования с применением искусственных нейронных сетей. Она реализована в виде программы определения параметров качества;
- системы оптимизации управляющих воздействий на основе генетического алгоритма. На рис.3.3 она представлена программой оптимизации;
- экспертной системы.

Генетические алгоритмы не гарантируют обнаружение глобального решения за кратчайшее время. Генетические алгоритмы не гарантируют и того, что глобальное решение будет найдено, но они хороши для поиска «достаточно хорошего» решения задачи «достаточно быстро».

## ЛИТЕРАТУРА

1. Автоматизация систем управления / В.Я. Ротач, В.Ф. Кузищин, А.С. Ключев и др.; Под ред. В.Я. Ротача. - М.: Энергоатомиздат, 1984. - 272 с.
2. Автоматизированное проектирование систем автоматического управления/ Я.Я. Алексанкин, А.Э. Бржозовский, В.А. Жданов и др.; Под ред. В.В. Солодовникова. - М.: Машиностроение, 1990. - 332 с.
3. А. с. СССР № 831532, кл.В 23 Q 15/00, 1979. Способ оптимизации процесса резания / А.И. Обабков.
4. Алексеев О.Г. Комплексное применение методов дискретной оптимизации. - М.: Наука, 1987. - 248 с.
5. Анализ и оптимальный синтез на ЭВМ систем управления / Под ред. А.А.Воронова и И.А.Опурка. - М.: Наука, 1984. - 344 с.
6. Андреев В.Н., Каплан Р.Я., Хрусталеv Б.С. Опыт оптимизации режимов механической обработки резанием. - Л.: ЛДНТП, 1982. - 20 с.
7. Аугустайтис В.В., Кубилене М.А., Кульветис Г.П. Применение аналитических вычислений на ЭВМ для расчета нелинейных колебаний станков // Динамика станочных систем ГАП: Материалы Всесоюз. конф. Тольятти, 1988. - С. 100-101.
8. Афанасьев В.Н., Колмановский В.Б., Носов В.Р. Математическая теория конструирования систем управления: Учеб. пособие для вузов. - М.: Высшая школа, 1989. - 447 с.
9. Безъязычный В.Ф. Методика определения оптимальных режимов резания при точении с учетом эксплуатационных свойств изделий // Производительность обработки и надежность деталей машин: Межвуз. сб. науч. тр. - Ярославль: Рыбинский авиац. ин-т, 1981. - С.11-14.
10. Беллман Р., Заде Л. Вопросы принятия решений в расплывчатых условиях// Вопросы анализа и процедуры принятия решений. М.: Мир, 1976. С. 172-215.

11. Бобров В.Ф. Основы теории резания металлов. - М.: Машиностроение, 1975. – 341 с.
12. Бобров В.Ф., Спиридонов Э.С. Оптимизация режима резания при точении // Станки и инструмент. – 1980. - №10. - С.22-23.
13. Борисов А.Н., Крумберг О.А., Федоров И.П. Принятие решений на основе нечетких моделей. Примеры использования. - Рига:Зинатне, 1990. - 184 с.
14. Бржозовский Б.М., Бровкова М.Б. К вопросу оценивания качества и определения оптимальной настройки технологического объекта в реальном времени // Исследования станков и инструментов для обработки сложных и точных поверхностей: Межвуз. науч. сб. – Саратов: СГТУ, 2003. - С. 131-137.
15. Бржозовский Б.М., Бровкова М.Б., Мартынов В.В. Определение параметров качества изделий при механообработке в реальном времени на основе нейросетевых алгоритмов // Нелинейная динамика и прикладная синергетика: Материалы Междунар. конф. – Комсомольск-на-Амуре, 2003. - С.62-65.
16. Бушуев В.В. Станочное оборудование автоматизированного производства. М.: Изд-во МГТУ «Станкин», 1993. – 584 с.
17. Вейц В.Л., Коловской М.З., Кочура А.Е. Динамика управляемых машинных агрегатов. - М.: Наука, 1984. – 352 с.
18. Вильсон А.Л., Иорданян Р.В. Производственный метод оценки качества станочных систем // Динамика станочных систем ГАП: Материалы Всесоюз. конф. - Тольятти, 1988. – С. 231-232.
19. Гантман С.А., Передрей Ю.М., Усанкин Н.Г. Управление качеством изделий в приборостроении статистическими методами: Учеб. пособие. – Пенза: Пенз. политехн. ин-т, 1980. – 117 с.
20. Городецкий Ю.И. Анализ и синтез динамического качества фрезерных станков: Материалы V Междунар. науч.-техн. конф. по динамике технологических систем. - Ростов-н/Д, 1997. - С.6-8.
21. Ерофеев А.А. Теория автоматического управления: Учебник для вузов. – СПб.: Политехника, 2002. – 302 с.

22. Заде Л.А. Понятие лингвистической переменной и его применение к принятию приближенных решений. - М.:Мир, 1976. -165 с.
23. Заковоротный В.Л. Динамическая диагностика и управление процессами обработки резанием: Материалы V Междунар. науч.-техн. конф. по динамике технологических систем. - Ростов-н/Д, 1997. -С.8-11.
24. Изерман Р. Цифровые системы управления: Пер. с англ. - М.: Мир, 1984. – 541 с.
25. Искусственный интеллект: В 3 кн. Кн.2. Модели и методы: Справочник/ Под ред. Д.А. Поспелова. - М.: Радио и связь, 1990.- 304 с.
26. Итоги науки и техники: физические и математические модели нейронных сетей. М.: ВИНТИ, 1990. Т. 1. 115 с.
27. Клушин М.И. Резание металлов. Элементы теории пластического деформирования срезаемого слоя. - М.: Машгиз, 1958. - 455 с.
28. Комарцова Л.Г., Максимов А.В. Нейрокомпьютеры: Учеб. пособие для вузов. – М.: Изд-во МГТУ им. Н.Э.Баумана, 2002. – 320 с.
29. Кофман А.Введение в теорию нечетких множеств. - М.: Радио и связь, 1982. - 432 с.
30. Кутин А.А. Оптимизация технологических параметров токарной обработки при нестационарном резании // Вестник машиностроения. – 1981. - № 8. - С.48-52.
31. Макаров А.Д. Оптимизация процессов резания. - М.: Машиностроение, 1976. - 278 с.
32. Малышев Н.Г., Берштейн Л.С., Боженюк А.В.Нечеткие модели для экспертных систем в САПР. - М.: Энергоатомиздат, 1991. - 220 с.
33. Маслеников И.А., Соколов Ю.А. Структурно-параметрическая оптимизация токарной операции с использованием целевой функции // СТИН. - 1997. - №1. - С.23-26.

34. Математическая теория оптимальных процессов / Л.С. Понтрягин, В.Г. Болтянский, Р.В. Гамкрелидзе, Е.Ф. Мищенко. – М.: Наука, 1976. – 392 с.
35. Медведев Д.Д. Автоматизированное управление процессом обработки резанием. – М.: Машиностроение, 1980. – 143 с.
36. Мелихов А.Н., Бернштейн Л.С., Коровин С.Я. Ситуационные советующие системы с нечеткой логикой. – М.: Наука, 1990. – 272 с.
37. Монахова Е. «Нейрохирурги» с Ордынки// PC Week/RE. 1995. №9.
38. Нечеткие множества в моделях управления и искусственного интеллекта / Под ред. Д.А. Поспелова. – М.: Наука, 1986. – 312 с.
39. Нечеткие множества и теория возможностей. Последние достижения / Под ред. Р.Ягера. – М.: Радио и связь, 1986. – 408 с.
40. Нильсон Н. Принципы искусственного интеллекта. – М.: Радио и связь, 1985. – 376 с.
41. Орловский С.А. Проблемы принятия решений при нечеткой исходной информации. – М.: Наука, 1981. – 206 с.
42. Осуга С. Обработка знаний: Пер. с япон. – М.: Мир, 1989. – 293 с.
43. Подураев В.Н. Технология физико-химических методов обработки. – М.: Машиностроение, 1985. – 264 с.
44. Подураев В.Н., Барзов А.А., Горелов В.А. Технологическая диагностика резания методом акустической эмиссии. – М.: Машиностроение, 1988. – 56 с.
45. Поцелуев А.В. Статистический анализ и синтез сложных динамических систем. – М.: Машиностроение, 1984. – 208 с.
46. Представление и использование знаний: Пер. с япон. / Под ред. Х. Уэно, М. Исидзука. – М.: Мир, 1989. – 220 с.
47. Прикладные нечеткие системы / Под ред. Т. Тэтано, К. Асаи, М. Сугэно. – М.: Мир, 1993. – 368 с.
48. Реклейтис Г., Рейвиндран А., Рэгсдел К. Оптимизация в технике: В 2 кн. Кн.1 / Пер. с англ. – М.: Мир, 1986. – 350 с.
49. Реклейтис Г., Рейвиндран А., Рэгсдел К. Оптимизация в технике: В 2 кн. Кн.2 / Пер. с англ. – М.: Мир, 1986. – 320 с.

50. Рыжкин А.А., Шучев К.Г. Физические аспекты оптимизации режимов изнашивания инструментальных режущих материалов// Материалы V Междунар. науч.-техн. конф. по динамике технологических систем. - Ростов-н/Д, 1997. - С.18-20.
51. Рыжов Э.В., Аверченков В.И. Оптимизация технологических процессов механической обработки. - Киев: Наукова думка, 1989. - 192 с.
52. Сбоев В.Н. Экономические аспекты применения систем автоматического выбора оптимальной скорости резания // Прогрессивные технологические процессы в машиностроении: Сб. научн. тр. - Томск: Изд-во ТПУ, 1997. - С.63-67.
53. Сердюк А.И. Влияние режимов резания на эффективность работы ГПС // СТИН. - 1997. - №5. - С.5-8.
54. Старков В.К. Дислокационные представления о резании металлов. М.: Машиностроение, 1979. - 160 с.
55. Старков В.К. Технологические методы повышения надежности обработки на станках с ЧПУ. - М.: Машиностроение, 1984. - 120 с.
56. Техничко-экономические аспекты повышения эффективности операций механообработки / В.А.Гречишников, С.В.Лукина, С.В.Григорьев, И.М.Толкачева: Материалы V Междунар. науч.-техн. конф. по динамике технологических систем. - Ростов-н/Д, 1997. - С.48-49.
57. Уоссерманн Ф. Нейрокомпьютерная техника: теория и практика. - М.: Изд-во МГТУ им. Н.Э.Баумана, 1995. - 184 с.
58. Цыпкин Я.З. Адаптация и обучение в автоматических системах. - М.: Наука, 1968. - 280 с.
59. Maitre F. Relations entre energies de coupe et endommagement des outils. - Wear. - 1980. - Vol.62. - №1. - P.139-160.
60. Michelbach J. Verschleißgleichung - Notwendigkeit für fertigungstechnisch und fertigungswirtschaftlich optimales Zerspanen. - Zwf, 1978. - Vol.73. - №10. - P. 504-509.
61. <http://www.inftech.webservis.ru/it/database/oo/ar2.html>
62. [lil.newmail.ru](http://lil.newmail.ru)
63. [www.homes.arealcity.com](http://www.homes.arealcity.com)
64. [www.idisys.iae.nsk.su/fuzzy\\_book/content.html](http://www.idisys.iae.nsk.su/fuzzy_book/content.html).

## ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ.....	3
1. ВВЕДЕНИЕ В ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ.....	5
1.1. Краткая история развития искусственного интеллекта.....	5
1.2. Основные направления искусственного интеллекта.....	11
1.3. Нечеткая логика.....	25
1.4. Нечеткие множества.....	28
1.5. Нечеткая и лингвистическая переменные.....	37
1.6. Нечеткие высказывания и нечеткие модели систем.....	39
1.7. Логико-лингвистическое описание систем, нечеткие модели.....	42
1.8. Основные аспекты интеллектуализации автоматизированных систем.....	43
2. ИСПОЛЬЗОВАНИЕ СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ ОПРЕДЕЛЕНИЯ ПАРАМЕТРОВ КАЧЕСТВА ДЕТАЛИ В МЕХАНООБРАБОТКЕ.....	47
2.1. Оценка качества и определения оптимальной настройки технологического объекта в реальном времени.....	47
2.2. Распознавание качества формообразования как объект применения нейросетевых методов.....	55
2.3. Выбор нейросетевого алгоритма в контексте решаемой задачи...	57
2.4. Структура сети.....	61
2.5. Обучение нейросети.....	66
2.6. Применение ИНС встречного распространения в контексте задачи определения параметров качества.....	70
3. ПРИМЕНЕНИЕ СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ ОПТИМИЗАЦИИ В МЕХАНООБРАБОТКЕ.....	81
3.1. Оптимизация в технике: общие вопросы.....	81
3.2. Оптимизация в металлообработке.....	90
3.3. Использование генетических алгоритмов для оптимизации.....	100
3.4. Символьная модель простого генетического алгоритма.....	103
3.5. Работа простого генетического алгоритма.....	105
3.6. Применение генетического алгоритма для оптимизации процессов формообразования.....	107
ЛИТЕРАТУРА.....	113

Учебное издание

**Бровкова Марина Борисовна**

**СИСТЕМЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА  
В МАШИНОСТРОЕНИИ**

Учебное пособие

Редактор О.А.Панина

Лицензия ИД № 06268 от 14.11.01

Подписано в печать 08.06.04

Формат 60×84 1/16

Бум. тип.

Усл. печ.л. 7,0 (7,5)

Уч.-изд.л 7,2

Тираж 100 экз.

Заказ 272

С 138

Саратовский государственный технический университет

410054 г. Саратов, ул. Политехническая, 77

Копипринтер СГТУ, 410054 г. Саратов, ул. Политехническая, 77