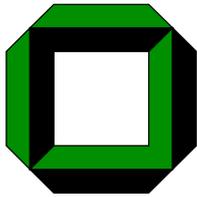


---



Universität Karlsruhe (TH)

Forschungsuniversität – gegründet 1825



# Qualitätsmanagement von Kontexten Verlässlichkeit mit Fuzzy Logik

Diplomarbeit am Institut für Telematik  
Telecooperation Office (TecO)  
Universität Karlsruhe (TH)

Martin Berchtold  
Betreuer: Prof. Dr. Juling  
Betreuer: Prof. Dr. Hartenstein  
Betreuender Mitarbeiter: Dipl. Infor. Tobias Zimmer

8. Januar 2007

---

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Karlsruhe, 31.Dezember 2006

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>4</b>
1.1	Motivation . . . . .	4
1.2	Ziel dieser Arbeit . . . . .	5
1.3	Aufbau und Inhalt . . . . .	5
<b>2</b>	<b>Qualitätsbeurteilung in ubiquitären Systemen und theoretische Annahmen</b>	<b>6</b>
2.1	Vorhandene Qualitätsmetriken in ubiquitären Systemen . . . . .	6
2.1.1	Argumentationen für Qualität von Kontexten . . . . .	6
2.1.2	Konstanten als Maß eines kontextuellen Algorithmus . . . . .	7
2.1.3	Unschärfe Logik um Kontexte zu generieren . . . . .	8
2.2	Theoretische Annahmen über dynamisches Verhalten von Qualitätssystemen . . . . .	8
2.2.1	Qualitätsmaß bei kontextuellen Zustandsübergängen . . . . .	9
2.2.2	Separierung von kontextuellen Informationen durch ein Qualitätsmaß . . . . .	10
2.2.3	Aggregative und fusionierende Schlussmöglichkeiten . . . . .	10
<b>3</b>	<b>Algorithmische Berechnung eines Qualitätsmaßes</b>	<b>12</b>
3.1	Methoden zur Klassifikation - Problematik bei nicht eindeutigen Abbildungen . . . . .	13
3.1.1	Scharfe Mengen und Klassen . . . . .	13
3.1.2	Unschärfe Mengen . . . . .	14
3.1.3	Unschärfe Logik . . . . .	15
3.1.4	Unschärfe Relationen . . . . .	16
3.2	Algorithmische Möglichkeiten zur Datenanalyse - Cluster zur Gruppierung . . . . .	18
3.2.1	Das k-Means Clustering . . . . .	18
3.2.2	Das c-Means Clustering . . . . .	19
3.2.3	Der Gustafson-Kessel Algorithmus . . . . .	20
3.2.4	Der Gath-Geva Algorithmus . . . . .	21
3.2.5	Das Mountain Clustering . . . . .	21
3.2.6	Das subtraktive Clustering . . . . .	22
3.3	Analyse der verschiedenen unscharfen Regelsysteme . . . . .	24
3.3.1	Das Mamdani- und Larsen-Regelsystem . . . . .	26
3.3.2	Das Tsukamoto-Regelsystem . . . . .	27
3.3.3	Das Takagi-Sugeno-Kang (TSK) Regelsystem . . . . .	28
3.4	Automatisierte Adaption mit neuronalen Fuzzy Netzen . . . . .	29
3.4.1	Künstliche neuronale Netze (KNN) . . . . .	29
3.4.2	Fuzzy-Kohonen-Clustering-Netzwerk (FKCN) . . . . .	30
3.4.3	Adaptive-network-based Fuzzy Inference System (ANFIS) . . . . .	31
<b>4</b>	<b>Qualitätsanalyse mit Fuzzy Logik</b>	<b>33</b>
4.1	Spezifikation der unscharfen Logik $\mathbf{S}_Q$ zur Beurteilung von kontextuellen Systemen $\mathbf{A}_C$ . . . . .	35
4.1.1	Erweiterung des Eingangs-Vektors $\mathbf{v}_C$ um kontextuellen Zustand $\mathbf{A}_C(\mathbf{v}_C)$ . . . . .	36
4.1.2	Definition der funktionalen Zugehörigkeit $F_{ij}$ . . . . .	36
4.1.3	Lineare funktionale Konsequenz $f_i$ . . . . .	37
4.1.4	Gewichtete Summe der Regeln . . . . .	39
4.1.5	Normalisierung oder Klassifikation der Sugeno-System Abbildung $\hat{\mathbf{S}}_{Q_{lin}}$ . . . . .	39
4.2	Adaption der Qualität des Kontext Algorithmus durch adaptive Algorithmen . . . . .	44
4.2.1	Initiales subtraktives Clustern zur Bestimmung der Komplexität $o$ . . . . .	44
4.2.2	Methode der kleinsten Quadrate zur initialen Adaption der linearen Konstanten $\mathcal{A} = (a_{ij})$ . . . . .	47
4.2.3	Aufbau des neuronalen Fuzzy Netzes welches das Sugeno-System $\hat{\mathbf{S}}_{Q_{lin}}$ repräsentiert . . . . .	50
4.2.4	Backpropagation-Training zum Tuning der Zugehörigkeitsfunktionen $F_{ij}$ und der linearen Konstanten $\mathcal{A} = (a_{ij})$ . . . . .	51
4.2.5	Hybrider lernender Algorithmus zum Tuning der Zugehörigkeitsfunktionen $F_{vc}$ und der linearen Konstanten $\mathcal{A} = (a_{ik})$ . . . . .	53
4.3	Das Stift-Artefakt . . . . .	55
4.3.1	Fuzzy Logik zur Kontexterkenkung $\mathbf{S}_C$ . . . . .	56
4.3.2	Neuronale Netze als Kontextsystem $\mathbf{N}_C$ . . . . .	64
4.4	Analyse der Resultate der unscharfen Logik $\mathbf{S}_Q$ . . . . .	67

4.4.1	Beobachtung des Systemverhaltens bei kontextuellen Zustandswechselln . . . . .	68
4.4.2	Zustandsabhängige Analyse des Qualitätsmaßes $S_Q$ . . . . .	69
4.4.3	Zustandsunabhängige Analyse des Qualitätsmaßes $S_Q$ . . . . .	71
4.4.4	Vergleich von unscharfen Regelsystemen für Kontext.- $S_C$ und Qualitätsanalyse $S_Q$ . . . . .	74
4.4.5	Kontextuelle Neuronale Netze $N_C$ und unscharfe Qualitätsanalyse $S_Q$ . . . . .	76
4.4.6	Auswirkung der Trainingsdaten auf das resultierende System zur Qualitätsanalyse $S_Q$ . . . . .	78
<b>5</b>	<b>Kontextererkennung und Qualitätsanalyse zweiter Ebene</b>	<b>80</b>
5.1	Allgemeiner Aufbau von Systemen zweiter Ebene . . . . .	82
5.1.1	Fusionierender Schluss aus Kontexten und Qualitäten . . . . .	82
5.1.2	Aggregativer Schluss aus Kontexten und Qualitäten . . . . .	84
5.2	Das Hand-Artefakt . . . . .	86
5.2.1	Erkennung von kontextuellen Zuständen mit dem Hand-Artefakt . . . . .	86
5.2.2	Spezifikation der Qualitätsanalyse für das Hand-Artefakt . . . . .	87
5.2.3	Analyse der Qualitäten . . . . .	88
5.2.4	Schlüsse aus Stift- und Hand-Artefakt . . . . .	89
5.3	Analyse von Resultaten des Qualitätsmaßes zweiter Ebene . . . . .	91
5.3.1	Fusionierende Schlüsse aus Kontexten und Qualitäten . . . . .	92
5.3.2	Aggregative Schlüsse aus Kontexten und Qualitäten . . . . .	93
<b>6</b>	<b>Zusammenfassung und Ausblick</b>	<b>95</b>
6.1	Zusammenfassung . . . . .	95
6.2	Ausblick . . . . .	95

# 1 Einleitung

Das Thema dieser Arbeit ist die Entwicklung eines Systems, das eine Qualitätsanalyse von bestehenden Systemen zur Erkennung von kontextuellen Zuständen ermöglicht. Der Fokus der Ausarbeitung liegt auf der Analyse des entwickelten Systems, welche Merkmale eine Qualitätsanalyse der Verlässlichkeit aus kontextuellen Zuständen extrahieren kann. Es wird auch ausführlich beschrieben wie das System zur Qualitätsanalyse erstellt wird, welche Algorithmen verwendet wurden und wie die formellen Grundlagen aussehen.

Das System, das die Qualitätsanalyse eines bestehenden kontextuellen Systems übernimmt, beruht auf einer unscharfen Logik, deren Grad der Zugehörigkeit ein Maß der Qualität repräsentiert. Das System der Qualitätsanalyse wurde in MATLAB<sup>®</sup> erstellt, evaluiert und analysiert. Die Qualitätsanalyse wurde für ubiquitäre Systeme eingesetzt, die im Telecooperation Office (TecO) entwickelt wurden. Die qualitative Analyse soll eine Verlässlichkeit für solche bestehenden Systeme liefern, über deren kontextuelle Klassifikationen sonst keine Aussagen mehr gemacht werden könnten. Es wurde eine spezielle Auswahl von adaptiven Algorithmen zur automatisierten Erstellung des Systems zur qualitativen Analyse getroffen. Die automatisierte Erstellung resultiert in einem qualitativen System, das eine beste Analyse mit geringst möglichem Verbrauch an Ressourcen liefert.

Der erste Abschnitt der Einleitung soll die Verwendung eines Systems zum Qualitätsmanagement einer Verlässlichkeit zunächst Motivieren. Das Ziel dieser Arbeit wird in einem weiteren Abschnitt festgelegt. Der letzte Abschnitt zeigt in groben Punkten wie diese Ausarbeitung aufgebaut ist und welche inhaltlichen Bestandteile vorzufinden sind.

## 1.1 Motivation

Der Begriff des Kontextes ist im Gebiet der ubiquitären Computertechnik schon lange bekannt und etabliert. Es wird mit Kontext der Zustand des Benutzers, seiner Umgebung und die Beziehung zueinander bezeichnet. Die Definition von Abowed und Dey [AD99] lautet wie folgt: 'Wir definieren Kontext als eine Information die genutzt werden kann, um die Situation einer Einheit zu beschreiben.' Die Systeme, die einen Kontext analysieren, machen Aussagen über einen kontextuellen Zustand, was einer Klassifikation entspricht. Ein kontextueller Zustand ist hierbei der Zustand einer Person oder eines Artefakts bezüglich seines Kontextes. Die Klassifikation ist fehlerbehaftet, da eine algorithmische Abbildung meist nicht eindeutig ist.

Wird ein Mensch aufgefordert eine Aussage über seinen kontextuellen Zustand und den seiner Umgebung zu machen, so kann er einschätzen, zu welchem Grad diese Information richtig ist. Der Mensch kann seine Urteilskraft und die Grenzen seiner Sensorik einschätzen und relativieren. Das explizite Wissen über sich selbst ist hierbei die Grundlage, um die Verlässlichkeit seiner Aussage zu beurteilen. Er kann andere Menschen durch ein Maß der Qualität seiner Aussage, über die Verlässlichkeit deren unterrichten. Die andere Person kann dann urteilen, ob diese Information die sie bekommt, verwertbar ist, oder wie sie sie gewichten muss.

Den maschinellen Systemen stehen keine Informationen zur Verfügung, wie gut die Einschätzung eines kontextuellen Zustandes ist und in welchen Situationen ihre Sensorik fehlerhafte Schlüsse bewirkt. Es wird eine Wissensbasis benötigt, die eine Einschätzung des Systemes von sich selbst und dessen Klassifikationen ermöglicht. Die Verlässlichkeit, die ein System somit durch Wissen über den gemachten Fehler einschätzen kann, kann es anderen maschinellen Systemen über ein Maß der Qualität zukommen lassen. Diese anderen Systeme können dann ihre Entscheidungsgrundlagen ebenso relativieren.

Ein großes Problem in ubiquitären Umgebungen ist die hohe Dimensionalität der kontextuellen Informationen. Es werden Umgebungen mit mehreren hundert eigenständigen Systemen [Wei91] prognostiziert, die Informationen bezüglich kontextueller Zustände liefern. Diese Informationsflut ist ohne Quantifizierung der Informationen nicht zu bewältigen. Es werden verschiedene Möglichkeiten der Quantifizierung vorgeschlagen und zum Teil auch verwendet, wie die Zeit der Erstellung, die Herkunft der Information und wie oft diese seit dem Ursprung schon verarbeitet wurde. In dieser Arbeit wird ein System beschrieben, das Identifikatoren von kontextuellen Zuständen mit deren Verlässlichkeit, einem Maß der Qualität, verbindet.

## 1.2 Ziel dieser Arbeit

Das Ziel dieser Arbeit ist die Definition eines qualitativen Systems, das für vorhandene kontextuelle Systeme die Verlässlichkeit der Klassifikation liefert. Die Verlässlichkeit ist hierbei keine statische Angabe, die die statistische Rate der Erkennung beschreibt, sondern ein dynamisches Qualitätsmaß, das für jeden erkannten kontextuellen Zustand den Fehler seiner Klassifikation repräsentiert.

Es wird kein Wissen über das kontextuelle System benötigt, nur die Werte der Eingabe und den dazu erkannten kontextuellen Zustand. Um das Wissen für die Beurteilung ohne Kenntnis über den Aufbau des kontextuellen Systems dennoch zu erhalten, muss eine automatisierte Erstellung der Architektur verwendet werden. Die Algorithmen, die dazu verwendet werden, müssen bezüglich ihrer Zuverlässigkeit und des Resultates analysiert werden. Auch muss das qualitative System analysiert werden, ob es eine theoretischen Anforderung an dieses genügt.

Das Ergebnis dieser Arbeit wird sein, den theoretischen und praktischen Nachweis zu liefern, dass eine automatisierte Erstellung eines qualitativen Systems möglich ist und auf jeden Algorithmus zur kontextuellen Klassifikation angewandt werden kann. Es wird auch eine weitere Verarbeitung von Maßen der Qualität vorgestellt und nachgewiesen, ob diese eine Reduktion der Dimensionalität in ubiquitären Umgebungen ermöglicht.

## 1.3 Aufbau und Inhalt

Das zweite Kapitel in dieser Arbeit beschreibt einige vorhandene Veröffentlichungen, die als thematisch verwandt angesehen werden können. Die realen Systeme, die eine qualitative Verlässlichkeit dynamisch für jeden kontextuellen Zustand liefern, wurden nicht gefunden, was diese Arbeit wegweisend macht. In der zweiten Hälfte dieses zweiten Kapitels, werden theoretische Annahmen gemacht, wie eine Verlässlichkeit sich in ihrem qualitativen Maß bei verschiedenen Ereignissen verhalten müsste.

Das dritte Kapitel widmet sich der Analyse von Grundlagen und Algorithmen, die zur Konstruktion und zum Aufbau eines Systems zur Qualitätsanalyse verwendet werden können. Es wird immer mit einer simplen Methodik begonnen, dann sukzessive zu komplexeren Strukturen übergegangen und schließlich mit der verwendeten Methode geschlossen. Die Argumentation soll die verwendete Methode als die beste der verfügbaren ausweisen.

Die zentrale Aussage dieser Arbeit wird in Kapitel vier formuliert. Es wird zunächst die Architektur des allgemeinen Systems zur Qualitätsanalyse definiert und formalisiert. Ein weiterer Abschnitt beschreibt die Algorithmen zur automatisierten Spezifikation des qualitativen Systems. Die realen kontextuellen Systeme und in Verbindung damit die qualitativen Systeme, die zur Evaluierung der Verlässlichkeit dienen, werden in einem weiteren Abschnitt beschrieben. Die ausführliche statistische Analyse der Maße der Qualität, die die Verlässlichkeit des zuvor beschriebenen kontextuellen Systems repräsentieren, wird im letzten Abschnitt geliefert.

Ein abstrahierendes Modell für fusionierende und aggregierende Schlüsse aus kontextuellen Informationen wird im Kapitel fünf formalisiert. Es wird ein weiteres reales kontextuelles und damit auch ein qualitatives System beschrieben, das zur Evaluierung der Schlüsse benötigt wird. Der letzte Abschnitt in dem Kapitel, beschreibt die statistischen Analysen der Schlüsse.

Das letzte Kapitel bildet den Abschluss dieser Arbeit und enthält die Zusammenfassung des Beschriebenen und Analysierten. Der Ausblick gibt Anregungen für Vertiefungen des Themas der 'Qualität von Kontext'.

## 2 Qualitätsbeurteilung in ubiquitären Systemen und theoretische Annahmen

Es bedarf zuerst einer Beschreibung des Begriffes der Qualität, bevor die Qualität eines Kontextes, die in dieser Arbeit definiert und verwendet wird, hergeleitet werden kann. Im ersten Abschnitt werden Veröffentlichungen der ubiquitären Computertechnik diskutiert, die mit dem Begriff der Qualität des Kontextes explizit oder implizit umgehen. Im zweiten Teil dieses Kapitels wird theoretisch erörtert, wie sich ein hier verwendetes Qualitätsmaß bei bestimmten Ereignissen verhalten sollte.

### 2.1 Vorhandene Qualitätsmetriken in ubiquitären Systemen

Es existieren nur vereinzelt Systeme, die eine qualitative Metrik nutzen, um die Verlässlichkeit der kontextuellen Erkennung zu beschreiben. Es werden zum Teil statistische Aussagen über die Rate der Erkennung eines gewissen Systems gemacht, die jedoch keine Informationen über ein aktuelles Datum liefern. Eine mittlerweile häufig veröffentlichte Meinung besteht jedoch, dass solche Informationen bezüglich einer Qualität von Kontexten benötigt werden. Welche Informationen von Relevanz sind wird oft diskutiert, denn die Informationsflut der kontextuellen Informationen muss bewältigt werden. Eine Aussage, wie die verschiedenen Informationen erlangt werden sollen, wurde bisher kaum gemacht. Eine Beurteilung der kontextuellen Klassifikation bezüglich ihres Fehlers ist besonders schwierig. Dieser Fehler kann die verschiedensten Ursachen haben, die ohne ein genaueres Wissen über das System zunächst nicht bestimmbar sind.

In den folgenden Abschnitten werden nun zuerst die Arbeiten beschrieben, in denen verschiedene Metriken der Qualität definiert und gefordert werden. Der zweite Abschnitt in diesem Kapitel beschreibt zwei Veröffentlichungen, die eine statistische Analyse zur statischen Beurteilung von Systemen verwenden. Der letzte Abschnitt widmet sich der Verwendung von unscharfer Logik, die in ubiquitären Systemen Einsatz findet.

#### 2.1.1 Argumentationen für Qualität von Kontexten

In verschiedenen Arbeiten, die in der ubiquitären und informationstechnischen Gemeinschaft veröffentlicht wurden, wurde der Begriff der Qualität von Kontexten (eng.: Quality of Context (QoC)) geprägt. Es existieren verschiedene Auffassungen, was unter diesem Begriff zu verstehen ist.

Eine umfassende Definition und eine Abgrenzung gegenüber anderen Begriffen der Qualität findet sich in der Arbeit von Buchholz, Küpper und Schiffers [BKS03]. In diesem Artikel wird zuerst auf die von Mark Weiser [Wei91] formulierten Paradigmen der ubiquitären Computertechnik eingegangen. Anhand dieser Paradigmen wird argumentiert, dass eine solche ubiquitäre Umgebung eine Vielzahl an Sensoren benötigt, die die nötigen Informationen bezüglich der Umwelt zur Verfügung stellen. Diese Sensoren besitzen eine gewisse Qualität der Erkennung, womit auch die gefolgerte kontextuelle Information eine gewisse Qualität der Aussage besitzt. Ein kontextuell bewusster Service (eng.: Context-Aware Service (CAS)) der auf die Benutzerwünsche best möglichst eingeht, wurde in vielen Arbeiten behandelt. Das damit entstehende Problem einer hohen Anzahl von Informationen, wird nun vorgeschlagen, wäre nur zu lösen, wenn eine gewisse Organisation der Services stattfindet. Diese Organisation der Services wäre durch zusätzliche Parameter der QoC möglich.

In den weiteren Abschnitten dieser Arbeit [BKS03] wird der QoC Begriff näher spezifiziert. Es werden die QoC-Parameter spezifiziert, wie die Präzision, die ein Maß der Genauigkeit von Sensoren spezifiziert, die Wahrscheinlichkeit der Korrektheit, die die Probabilität von fehlerhaften Zuständen beschreibt, die Vertraulichkeit, die die Korrektheit einer bestimmten Quelle angibt, die Auflösung, die die Dichte der Informationsgewinnung spezifiziert, und die Aktualität, die das Alter einer kontextuellen Information festhält. Eine größere Anzahl von Parametern wird eingeräumt, jedoch werden die spezifizierten als die wichtigsten angesehen. Es wird auch gegenüber anderen Begriffen wie Qualität des Service (eng.: Quality of Service (QoS)) und der Qualität des Gerätes (eng.: Quality of Device (QoD)) abgegrenzt.

Zuletzt werden Argumente für die Notwendigkeit der QoC aufgelistet, die mit Vereinbarung, Rekonstruktion des Verhaltens von CAS, Selektion von angemessenen Kontext-Versorgern, Adaption der Kontext-Verbesserung und -Verbreitung bezeichnet werden.

In einer weiteren Arbeit, die von Razzaque, Dobson und Nixon [RDN05], wird der QoC Begriff leicht abgewandelt mit Qualität von kontextueller Information (eng.: Quality of Context Information (QoCI))

bezeichnet. Es werden verschiedene kontextuelle Modelle analysiert und diskutiert, in wie weit sie Probleme der Abhängigkeit berücksichtigen. Diese Analyse wird verwendet, um eine Methodik der Modellierung der Qualität von Kontext-Informationen abzuleiten.

In einem speziell auf die QoCI ausgerichteten Abschnitt werden die Fehler, die bei der Erstellung, Interpretation und Präsentation von kontextuellen Informationen entstehen, beschrieben. Es wird argumentiert, dass wenn Applikationen von diesen kontextuellen Informationen abhängen, einige Mechanismen vorhanden sein müssen, aufgrund deren diese Informationen beurteilt werden können. Die meisten Modelle berücksichtigen nicht die Fehler, die in diesen Informationen vorhanden sind. Es wird die Hypothese aufgestellt, dass zukünftige Systeme diversifizierte kontextuelle Informationen benötigen.

Die Veröffentlichung von Zimmer [Zim06] stellt ein konkretes System zum Management von Kontexten vor. Eine Verarbeitung von kontextuellen Informationen in ubiquitären Umgebungen hat die ersten Entwicklungsstufen hinter sich. Es wird nun der nächste Schritt vorgeschlagen, der eine Konvergenz in die Richtung der Vision von Mark Weiser [Wei91] ermöglicht. Die Effekte der Skalierung, die in hochdimensionalen ubiquitären Umgebungen auftreten, werden analysiert und durch zusätzliche Informationen reduziert.

Das System zum Management von kontextuellen Informationen nutzt vier Attribute, um die Rate der Erkennung zu verbessern. Zum einen wird die Lokation des Ursprungs von Kontexten bei deren Verarbeitung berücksichtigt. Ein weiteres Attribut ist der Zeitpunkt der Entstehung, der Informationen zur Skalierung liefert. Werden kontextuelle Informationen verarbeitet, so entstehen durch Ableitung von Informationen erneute Abhängigkeiten zwischen diesen. Es wird über eine genetische Relation von Kontexten (eng.: Genetic Relation of Context (GRC)) ein Baum von Ableitungen aufgebaut, der diese Informationen der Abhängigkeit liefern soll. Das letzte Attribut beschreibt die Verlässlichkeit der Erstellung der kontextuellen Information. Dieses Maß der Qualität wird mit einem unscharfen Regelsystem, ohne Wissen über die Funktionsweise des kontextuellen Algorithmus, extrahiert.

In dieser Arbeit wird nun das System der Verlässlichkeit, das in der Arbeit von [Zim06] vorgeschlagen wurde, ausführlich hergeleitet, beschrieben und analysiert. Es wird eine Methodik der automatisierten Erstellung verwendet, die kein Wissen über die Funktionsweise von kontextuellen Algorithmen benötigt. Diese Verlässlichkeit des Kontextes, die in dieser Arbeit mit Qualität referenziert wird, wird auf verschiedene Aspekte analysiert. Es wird auch die Frage beantwortet, ob eine Ableitung aus mehreren Qualitätsmaßen, einen Mehrwert für das System des Managements von kontextuellen Informationen liefert.

### 2.1.2 Konstanten als Maß eines kontextuellen Algorithmus

Es sind keine Systeme der ubiquitären Computertechnik bekannt, die eine zuvor definierte Qualität des Kontextes generieren und verwenden. Die einzigen Angaben die in dieser Richtung für bestehende Systeme gemacht wurden, sind statistisch bestimmte Wahrscheinlichkeiten der Erkennung.

Mit einem Fehler wird die Zuverlässigkeit eines Systems beschrieben, das von Krause, u.a. [KSSF03] entwickelt wurde. Diese Zuverlässigkeit ist eine statistische Analyse, der richtig oder falschen Erkennung von physiologischen und aktiven kontextuellen Zuständen eines menschlichen Körpers. Es wird ein Armband mit Sensoren verwendet, das am Oberarm angebracht wird. Die Sensordaten werden mit einer Kohonen selbst-organisierenden Karte und einem k-Means Clustering verschiedenen kontextuellen Zuständen zugeteilt. Es wurden verschiedene Datensätze zur Evaluierung verwendet, die bezüglich ihrer repräsentierten Zustände und der verwendeten Sensoren einen Fehler der Quantisierung ergeben.

Eine weitere Arbeit, die von Kim und Lee [KL06] verfasst wurde, referenziert direkt den Begriff der Qualität des Kontextes. Die vorgeschlagenen Dimensionen die von Relevanz sind, die Genauigkeit, die Vollständigkeit, die Darstellung der Konsistenz, die Sicherheit des Zugangs und die Aktualität. Im Zusammenhang mit dem Begriff der Genauigkeit wird ein quadratischer Mittelwert über alle betrachteten Daten berechnet. Eine Schätzung des Intervalls der Vertraulichkeit, das in Kombination mit dem quadratischen Mittel eine Formel für die Genauigkeit ergibt, beschreibt eine statistische Analyse, bei der die Messwerte über einen längeren Zeitraum beobachtet werden. Der Wert der Vollständigkeit beschreibt mit welcher Verfügbarkeit von gewissen kontextuellen Informationen gerechnet werden kann. Auch dies ist eine Angabe einer Qualität bezügliche einer bestimmten Quelle, die statistisch berechnet wird.

Die beiden Arbeiten, die vorgestellt wurden, sind exemplarisch für alle Angaben von qualitativen Infor-

mationen. Es wird statistisch eine Aussage über ein System gemacht, was zur Beurteilung dieser Quelle dient. Diese Informationen sind aber von geringer Relevanz, wenn eine kontextuelle Information, die auf einem aktuellen Datum beruht, zu verarbeiten ist. Die Angaben bezüglich einer Qualität müssen dynamisch für jedes Datum und jeden erkannten kontextuellen Zustand berechnet werden, da eine statische Angabe nur generelle Informationen liefert, die keinen oder nur einen geringen Mehrwert liefern.

### 2.1.3 Unschärfe Logik um Kontexte zu generieren

Eine Möglichkeit über einen kontextuellen Algorithmus eine qualitative Aussage zu machen hat man, wenn die algorithmische Verarbeitung unscharf abläuft. In der Theorie der unscharfen Mengen ist die Zugehörigkeit der Grad, zu dem ein Element der Menge zu dieser gehört. Der Grad der Zugehörigkeit wird in Werten von Null bis Eins gemessen. Diese Grade werden innerhalb unscharfer Systeme miteinander verrechnet, wobei auch die Ausgabe eines solchen Systems einen Zugehörigkeitsgrad besitzt. Wird also durch einen unscharfen Algorithmus auf einen kontextuellen Zustand abgebildet, so gibt der Grad der Zugehörigkeit an, wie gut diese Abbildung ist. Die unscharfe Verarbeitung, die auch Fuzzy genannt wird, beinhaltet also durch die Art der Schlüsse die gemacht werden, ein Maß der Qualität der Aussage.

Es existieren verschiedene Anwendungen die eine unscharfe Logik verwenden, doch meist auf einer höheren Ebene der Schlussfolgerung. Das von Park, u.a. [PLYSK04] vorgeschlagene System, benutzt einen hybriden Ansatz um Domänen-Modelle zu modellieren, die verschiedene vordefinierte Aufgaben unabhängig ausführen. Der hybride Ansatz ist eine Kombination aus Bayesian-Netzwerken und unscharfer Logik. Die unscharfe Logik dient zur Beurteilung der Entscheidungsfindung, ob diese subjektiv zu einem gewissen Grad der Ungewissheit ist und Klassifikationen sich an einer Grenze befinden.

In der Arbeit von Eldin und Wagenaar [EW04], wird wieder eine unscharfe Logik auf der Ebene der Schlussfolgerung benutzt. Sie schlagen ein Modell des zustimmenden Versorgers vor. Dabei beeinflussen die Faktoren die Zustimmung des Benutzers, die Modellierung ihres Einflusses und Evaluierung der Rolle in der Entscheidungsfindung in einer unscharfen Schlussfolgerung.

Die soeben zitierten Arbeiten verwenden unscharfe Logik zur Schlussfolgerung, die mittels der unscharfen Entscheidungsfindung eine Qualität des Kontextes implizit verarbeitet. Diese Systeme befinden sich beide auf einer höheren Ebene der Schlussfolgerung, und argumentieren nicht bezüglich einer Unschärfe der Generierung von Kontextinformationen. Eine solche Methodik wird in den meisten Arbeiten verwendet, die für ubiquitäre Systemen eine unscharfe Logik verwenden. Diese Systeme sind auch auf die Verwendung der unscharfen Schlussfolgerung limitiert, was keine freie Wahl eines kontextuellen Algorithmus oder die Verwendung eines vorhandenen zulässt.

## 2.2 Theoretische Annahmen über dynamisches Verhalten von Qualitätssystemen

Im letzten Abschnitt wurde argumentiert, dass zusätzliche Parameter zu kontextuellen Informationen eine Möglichkeit sind, die Informationsflut in ubiquitären Umgebungen zu bewältigen. Es wurde festgelegt, dass in dieser Arbeit ein System der Verlässlichkeit, das zu jedem erkannten kontextuellen Zustand dessen Qualität bereitstellt, definiert wird. Die Algorithmen, die Sensorwerte auf einen kontextuellen Zustand abbilden, machen generell einen Fehler in der Projektion. Dieser Fehler ist entweder schon in den Sensorwerten vorhanden, entsteht durch die Konstruktion der Systeme oder ergründet sich durch nicht eindeutige Abbildungen. Ist der Fehler in den Sensorwerten vorhanden, so liegt das an der Toleranz der Sensorik (z.B. Temperatur ist  $20^{\circ}\text{C} \pm 0,5^{\circ}$ ) oder an gewissen Fehlerzuständen (z.B. Bereiche der Sättigung) in denen die Sensoren keine brauchbaren Werte liefern. Werden bei der Konstruktion gewisse Fehler gemacht, so kann dies daran liegen, dass gewisse Zustände zu Entwicklungszeiten nicht bekannt sind, bei adaptiven Algorithmen die Daten zum Training nicht vollständig oder fehlerhaft sind oder das Modell, aufgrund das System erstellt wird, nicht korrekt ist. Eine Abbildung auf einen Bereich, der einem gewissen kontextuellen Zustand zugeordnet wird, ist meist nicht eindeutig. Es kann auf die Grenze projiziert werden, die Bereiche können nicht gänzlich voneinander trennbar sein oder es können Schnittmengen der Bereiche vorhanden sein, die keinem Zustand eindeutig zuzuordnen sind.

Wird eine Erkennung von kontextuellen Zuständen für jeden erkannten Zustand bezüglich des Fehlers und somit der Qualität beurteilt, so muss bekannt sein, ob die Erkennung richtig war. Es muss ebenso bekannt sein, was erkannt wurde und auf welcher Grundlage diese Entscheidung gemacht wurde. Sind

diese Informationen vorhanden, so wird davon ausgegangen, dass keine spezifischen Informationen über den Algorithmus der Erkennung vorhanden sein müssen. Es ist jedoch die Frage, anhand welcher Kriterien eine Qualitätsbeurteilung der kontextuellen Erkennung zu evaluieren ist. Wurde ein Maß der Qualität für einen Kontext berechnet, so kann zunächst nicht gesagt werden, ob dieses dem Fehler der Erkennung entspricht, da die Mannigfaltigkeit der Fehlerquellen zu dem Zeitpunkt der Detektion nicht bekannt sind.

Es werden nun Hypothesen aufgestellt, wie ein Maß der Qualität sich bei bestimmten Ereignissen verhalten sollte. Gegen diese Hypothesen kann dann später das System zur Qualitätsanalyse verglichen werden. Eine erste Annahme ist, dass wenn eine Projektion von einem Bereich eines kontextuellen Zustandes in den Bereich eines anderen übergeht, zuerst eine Verschlechterung der Qualität und dann eine Verbesserung beobachtet werden kann. Wurde ein kontextueller Zustand richtig erkannt, so sollte das Maß der Qualität höher sein, als wenn ein Zustand falsch erkannt wird. Ein Schluss aus vorhandenen kontextuellen Informationen müsste auch gewisse Charakteristiken aufweisen. Werden aus kontextuellen Informationen auf neue Informationen geschlossen, so wird aus einzelnen vagen Aspekten auf neue Informationen geschlossen, weshalb eine resultierende Qualität nur so gut sein kann, wie die schlechteste der Grundlagen. Weden mehrere kontextuelle Informationen zur gegenseitigen Bestätigung benutzt, so müsste eine resultierende Qualität mindestens so hoch sein, wie der größte Bestandteil der Bestätigung.

### 2.2.1 Qualitätsmaß bei kontextuellen Zustandsübergängen

Ein Algorithmus der Kontexterkenkung ist eine Abbildung von Informationen aus Sensoren, auf einen Raum von kontextuellen Zuständen. Dieser Raum ist in verschiedene Bereiche eingeteilt, die unterschiedlichen kontextuellen Identifikatoren zugeordnet werden. Ein solcher Bereich kann eine scharfe Grenze gegenüber anderen Bereichen besitzen, eine eindeutige Zuordnung zu einem jeweiligen Bereich ist jedoch nur in dessen Zentrum möglich. Die Abbildungen, die nicht in das Zentrum projizieren, haben in ihrer Aussage eine geringere Verlässlichkeit als zentrale Projektionen. Wird durch den kontextuellen Algorithmus nahe an eine Grenze von Bereichen projiziert, so kann durch den Fehler der in der Abbildung liegt, die Entscheidung für einen kontextuellen Zustand nicht eindeutig sein. Eine Entscheidung muss jedoch getroffen werden, weshalb die nicht Eindeutigkeit durch den Wert der Verlässlichkeit repräsentiert wird.

Eine schematische Darstellung von Bereichen für drei verschiedene kontextuelle Zustände, ist in der Abbildung 1 abgebildet. Die Zentren der Bereiche sind in der jeweiligen Farbe dunkel dargestellt, wobei die Aufhellung der Farben in Richtung der Bereichsgrenzen die Unsicherheit in der Abbildung symbolisiert. Die eindeutige Abbildung auf einen Zustand wird verlassen, wenn in einen anderen Zustand übergegangen

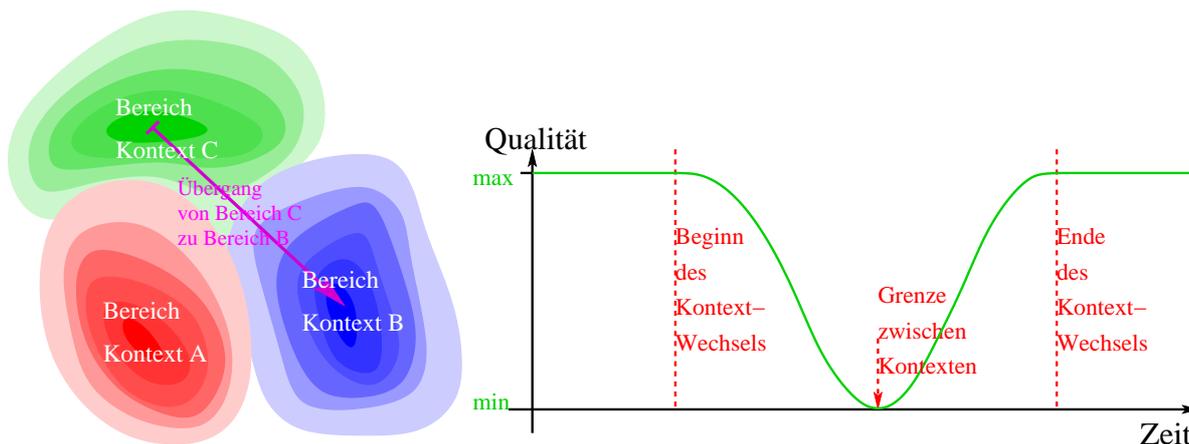


Abbildung 1: Schematischer Raum der kontextuellen Zustände mit Kennzeichnung (rot, grün und blau) der Bereiche der Kontexte und Übergang von einer Bereichsmitte zu einer anderen (links); Verlauf des Maßes der Qualität bei dem Übergang von einem zum anderen Kontext (rechts)

wird. Die klare Zuordnung von Daten zu einem Zustand wird schlechter, je mehr aus dem Zentrum eines Bereiches herausgegangen wird, und wird besser, wenn die Schritte in Richtung des Zentrums eines neuen Bereiches vollzogen werden. Für den lila markierten Weg, der in der Abbildung 1 auf der linken Seite den Übergang von Kontext C zum kontextuellen Zustand B beschreibt, müsste also ein bestimmter Verlauf der Verlässlichkeit festzustellen sein. Die Qualität der Aussage, die eine algorithmische Abbildung für

Kontexte für diesen Übergang von C nach B liefert, ist schematisch in Abbildung 1 auf der rechten Seite einzusehen. Der Beginn des Kontext-Wechsels ist rot markiert und bedeutet das Verlassen des Zentrums von Bereich C. Die geringste Verlässlichkeit der algorithmischen Abbildung sollte an der Grenze der Bereiche festzustellen sein. Eine Verbesserung der Qualität müsste bis zum Ende des Kontext-Wechsels, der wieder rot markiert ist, zu erkennen sein. Ist der Wechsel vollzogen, so werden wieder eindeutig auf das Zentrum zu projizierende Daten von den Sensoren gemessen, was in einer maximalen Qualität resultieren sollte.

Ein solcher Wechsel zwischen kontextuellen Zuständen ist in dieser Hypothese stark idealisiert. Dieser beschriebene geradlinige und kontinuierliche Weg zwischen den Zuständen wird in der Realität kaum zu beobachten sein. Es sind verschiedene Aspekte zu vermuten, die die Geradlinigkeit und die Kontinuität der Übergänge beeinflusst. Ein Aspekt könnte die Rate sein, mit der die Sensoren abgefragt werden, ein anderer die Dynamik der gemessenen Zustände.

### 2.2.2 Separierung von kontextuellen Informationen durch ein Qualitätsmaß

Ein weiterer Aspekt, der durch ein Maß der Qualität als repräsentiert vermutet wird, ist die Separierung von richtigen und falschen kontextuellen Abbildungen. Wenn Informationen von Sensoren auf einen kontextuellen Zustand projiziert werden, so kann, wie schon beschrieben, diese Abbildung nicht eindeutig sein. Die Bereiche im Raum der Zustände könnten nicht gänzlich trennbar sein, es existieren Überlappungen, die keinem Zustand eindeutig zuzuordnen sind, und die Abbildungen, die nahe oder auf eine Bereichsgrenze projizieren, können falsch interpretiert werden. Die Verlässlichkeit eines Algorithmus kann über solche Abbildungen gewisse Aussagen machen, so kann eine hohe Qualität eine eindeutige und eine niedrige eine nicht eindeutige Projektion identifizieren.

Um anhand der Qualität die fehlerhaften von den korrekten Erkennungen zu unterscheiden, muss ein Schwellwert festgelegt werden. Die Qualitäten die eine richtige kontextuelle Zuordnung kennzeichnen, sollten über diesem Schwellwert sein, die unterhalb müssten falsche Abbildungen identifizieren. Das Problem hierbei ist, dass die Beurteilung der kontextuellen Abbildung auch fehlerbehaftet ist, denn manche Abbildungen sind nicht voneinander separierbar. Deshalb können die kontextuellen Zustände nur mit einer gewissen Wahrscheinlichkeit bezüglich ihrer Richtigkeit getrennt werden.

Diese Hypothese kann nur anhand von realen Daten evaluiert werden, da jede Abbildung verschieden ist und die kontextuellen Bereiche nicht weiter theoretisch ergründet werden können. Die Methodik der Konstruktion und der Art eines Systems, das eine Verlässlichkeit berechnet, ist von entscheidender Bedeutung und legt den Fehler der qualitativen Beurteilung fest.

### 2.2.3 Aggregative und fusionierende Schlussmöglichkeiten

In kontextuell sensitiven Umgebungen der ubiquitären Computertechnik, werden nicht nur kontextuelle Informationen aus Sensorik generiert, es werden auch aus vorhandenen Informationen Schlussfolgerungen getroffen. Es können generell zwei Arten der Schlussfolgerung unterschieden werden, die Aggregation und die Fusion. Die Aggregation folgert aus vorhandenen Informationen neue kontextuelle Zustände. Durch die Fusion werden erkannte Zustände gegenseitig bestätigt. Werden Informationen bezüglich kontextuellen Zuständen weiterverarbeitet, so müssen auch deren qualitative Verlässlichkeiten entsprechend abgeleitet werden.

Eine Aggregation eines neuen kontextuellen Zustandes benutzt die Grundlagen des Schlusses zu gewissen Teilen. Die Größe des Gewichtes, das eine Information auf den Schluss hat, bestimmt auch den Anteil den dessen Qualitätsmaß am resultierenden besitzt. Wird eine klassische Logik für den aggregativen Schluss verwendet, so sind die Grundlagen, die Prämisse der Ableitung, gleichwertig für das Resultat, die Konklusion, verantwortlich. Diese Vereinfachung wird auch für die Qualität angenommen, da die eventuellen Gewichtungen des Schlusses nicht bekannt sind.

Die Fusion leitet aus mehreren gleichen Informationen der Prämisse, die Bestätigung dieser Information in der Konklusion ab. Es werden also mehrere gleiche Informationen zu gegenseitigen Bestätigung verwendet. Das Resultat der Fusion von Qualitäten sollte deshalb nicht kleiner sein, als das Maximum der Eingabe. Es wird befürwortet, dass die Fusion von Informationen besser sein sollte als die größte Prämisse, da sonst eine gegenseitige Bestätigung sich in ihrer Qualität nicht verbessert.

## 2.2 THEORETISCHE ANNAHMEN ÜBER DYNAMISCHES VERHALTEN VON QUALITÄTSSYSTEMEN

---

Eine ausführliche Abhandlung zum Thema der Ableitungen findet sich in der Arbeit von Dubois und Parade [DP00], die eine Posibilitätstheorie als Grundlage verwenden. Sie bezeichnen mit dem Begriff der Fusion die Kombination mehrerer Informationen, um eine neue Information abzuleiten. Dieser Begriff ist hier anders verwendet, wie zuvor definiert wurde.

### 3 Algorithmische Berechnung eines Qualitätsmaßes

Die Qualität, die eine Erkennung eines kontextuellen Zustandes besitzt, entspricht dem Fehler einer Klassifikation. Dieser Fehler, den ein kontextueller Algorithmus macht, soll ein System der Qualitätsanalyse repräsentieren. Das System zur Qualitätsanalyse muss spezifiziert werden, was eine Architektur und eine Algorithmik benötigt. Es existieren verschiedene Herangehensweisen, da die Vielzahl der vorhandenen Architekturen und Algorithmen einer Auswahl bedürfen.

Ein Wissen über den Algorithmus der kontextuellen Klassifikation ist nicht gefordert, weshalb eine Möglichkeit der Beurteilung dessen gefunden werden muss, die keinerlei Wissen bedarf. Es können also nur die Informationen genutzt werden, aufgrund deren der kontextuelle Algorithmus seine Entscheidung trifft, und der Identifikator der kontextuellen Klasse, der durch die Klassifikation bestimmt wurde. Zum Zeitpunkt der Erstellung eines qualitativen Systems, müssen sekundäre Informationen über den Algorithmus zur kontextuellen Klassifikation bekannt sein. Ein Datensatz der für jedes Datum eine Zuordnung zu einem kontextuellen Zustand besitzt, kann bezüglich der Richtigkeit der Zuordnung quantifiziert werden. Diese sekundäre Information, ob eine Klassifikation richtig oder falsch war, ermöglicht eine automatisierte Erstellung des qualitativen Systems. Wird ein überwachtes Lernverfahren zur Spezifikation eines qualitativen Systems eingesetzt, so wird der Fehler der in dem Datensatz der sekundären Information vorhanden ist, durch Lernen adaptiert.

Es existieren verschiedene Architekturen, die eine automatisierte Erstellung durch überwachte und nicht überwachte Lernverfahren ermöglichen. Die Architektur sollte jedoch die Repräsentation eines Fehlers schon auf semantischer Ebene berücksichtigen. Eine kontextuelle Klassifikation hat hierbei eine Semantik der Zugehörigkeit zu einer Klasse. Diese Zugehörigkeit findet Verwendung in der Theorie der unscharfen Mengen. Eine unscharfe Menge besitzt Elemente, die nicht nur zu ihr gehören oder nicht, sondern dies zu einem gewissen Grad tun. Die Systeme die auf unscharfen Mengen beruhen, besitzen verschiedene Möglichkeiten der automatisierten Erstellung, was ihren Einsatz zur Qualitätsanalyse ermöglicht.

In diesem Kapitel wird nun zuerst der Einsatz von unscharfen Mengen ergründet, die dann in einer knappen Beschreibung eingeführt werden. In einem weiteren Abschnitt werden Algorithmen beschrieben, mit denen eine automatisierte Identifizierung der Architektur von unscharfen Systemen möglich ist. Es existieren verschiedene unscharfe Systeme, die in einem weiteren Abschnitt aufgeführt werden. Der letzte Abschnitt beschreibt hybride Systeme, die eine Spezifizierung von unscharfen Systemen mit Lernverfahren ermöglichen.

### 3.1 Methoden zur Klassifikation - Problematik bei nicht eindeutigen Abbildungen

Eine Klasse bedeutet für ein Datum eine gewisse Eigenschaft, die es mit anderen Daten teilt. Der Algorithmus, der ein Datum einer Klasse zuteilt, macht in dieser Abbildung einen Fehler. Der Fehler ist für gewöhnlich gering genug, dass auf die richtige Klasse abgebildet wird. Die Eindeutigkeit der Abbildung kann jedoch anhand der zugeordneten Klasse nicht erkannt werden. Auch kann der Fehler, der bei der Klassifikation gemacht wurde, durch den Identifikator der Klasse, selbst wenn es der richtige ist, nicht mehr rekonstruiert werden.

Wird die Abbildung auf eine Klasse durch eine Unschärfe ausgedrückt, so ist nicht nur der Identifikator bekannt, sondern auch der Grad der Zuordnung. Der Grad entspricht somit dem Fehler der nicht eindeutigen Abbildung. Werden also unscharfe Mengen zur kontextuellen Klassifikation eingesetzt, dann kann der Fehler repräsentiert werden.

Eine Spezifikation eines kontextuellen Algorithmus ist jedoch nicht Gegenstand dieser Arbeit, denn diese Systeme existieren bereits. Es gilt ein unabhängiges System zu spezifizieren, das den Fehler eines kontextuellen Systems repräsentiert. Dieses qualitative System liefert also den Grad der Zugehörigkeit zu einer kontextuellen Klasse, die von einem kontextuellen System bestimmt wurde. Der Grad entspricht aber auch in diesem Fall einer Unschärfe, was zur Wahl einer unscharfen Logik führt. Diese unscharfe Logik wird in diesem Kapitel kurz eingeführt, wobei mit der Verbindung von Klassen und scharfen Mengen begonnen und dann zu der Theorie der unscharfen Mengen übergegangen wird. Die unscharfe Logik ist im Bereich der unscharfen Mengen eine Erweiterung der scharfen Logik, deren Beschreibung in unscharfen Relationen in diesem Abschnitt endet. Die Notation die hier verwendet wird, ist für die Theorie der unscharfen Mengen von Hanebeck und Rökler [HR05] spezifiziert und formalisiert worden, wohingegen die allgemeine lineare algebraische Syntax aus den Skripten von Kunle, Aumann und Schober [KuUS00] entnommen wurde.

#### 3.1.1 Scharfe Mengen und Klassen

Werden Daten in Klassen eingeteilt, so muss zu jedem Datum die Zugehörigkeit zu einer Klasse mitgeführt werden. Wird die Einteilung in Klassen durch Mengen vorgenommen, wird die umgekehrte Methode verwendet. Das Datum muss den Identifikator der Klasse nicht mehr mitführen, sondern eine Menge, die die Klasse repräsentiert, enthält das Datum.

Im mathematischen und algebraischen Sinne ist die Deklaration von Klassen als Mengen zu restriktiv. Die Informatik ist da pragmatischer, ein Datum gehört zu einer Klasse, und somit zu einer Menge, oder nicht. Die Syntax der Menge die mathematisch gebraucht wird, ist auch hier verwendet. Wird zum Beispiel die Menge  $K_i$  definiert, die alle Elemente  $x_{i1}, \dots, x_{in}$  der Klasse  $i$  enthält, so ist diese Menge wie folgt definiert:

$$K_i := \{x_{i1}, \dots, x_{in}\}$$

Enthält eine Menge die Elemente einer Klasse die ein bestimmtes Attribut erfüllen, dann müssen die Elemente nicht explizit aufgezählt werden. Wird das Attribut durch ein beliebiges Prädikat  $P$  entschieden, dann kann eine Menge  $K_i$  wie folgt definiert werden:

$$K_i := \{x_j \in X | P(x_j) = true\}$$

Das ursprüngliche Universum  $X$  aus dem Elemente  $x_j$  entnommen werden, kann der Menge der natürlichen  $\mathbb{N}$ , ganzen  $\mathbb{Z}$ , rationalen  $\mathbb{Q}$ , reellen  $\mathbb{R}$  oder komplexen  $\mathbb{C}$  Zahlen entsprechen. Da jede CPU eine endliche Maschine ist, werden diese Mengen nie komplett repräsentiert. Aus diesem Grund sind die Elemente der Mengen, die die Universen repräsentieren, die Daten-Typen, die in Programmiersprachen verwendet werden, wie Integer, Float oder Double.

Ein Datum, das zu einer Klasse gehört, besteht meist nicht nur aus einem Element, sondern ist mehrdimensional. Um Operatoren der linearen Algebra auf ein solches Datum anwenden zu können, werden die Elemente in einem Vektor angeordnet. Mit dieser Forderung erhält man die folgende Menge  $K_i$  als Beispiel:

$$K_i := \{\vec{x}_j \in X_1 \times \dots \times X_m | P(\vec{x}_j) = true\}$$

Die Zugehörigkeit eines  $m$ -dimensionalen Vektors  $\vec{x}_j$  zu den Universen  $X_1, \dots, X_m$  muss auf das Kreuzprodukt dieser  $X_1 \times \dots \times X_m$  erweitert werden.

Die Zuordnung eines Datums zu einer Klasse, und somit zu einer Menge, ist oft nicht eindeutig. Mit scharfen Mengen können graduelle Zugehörigkeiten nicht verwirklicht werden, weshalb die klassische Mengentheorie auf unscharfe Mengen erweitert werden muss.

### 3.1.2 Unscharfe Mengen

Im Gegensatz zu scharfen Mengen ist die Zugehörigkeit zu einer unscharfen Menge nicht eindeutig. Eine unscharfe Menge  $\tilde{A}$  über dem Universum  $U$  ist definiert durch die Funktion  $\mu_{\tilde{A}} : U \rightarrow [0, 1]$ , wobei  $\mu_{\tilde{A}}$  den Grad der Zugehörigkeit zu der unscharfen Menge  $\tilde{A}$  darstellt. Die unscharfe Menge  $\tilde{A}$  wird wie folgt beschrieben:

$$\tilde{A} = \{(x, \mu_{\tilde{A}}(x)) : x \in U\}$$

Wird also von Elementen von unscharfen Mengen gesprochen, so muss immer das Element und der Grad der Zugehörigkeit in einem Tupel angegeben werden. Im Englischen entspricht dem Ausdruck unscharf der Begriff 'fuzzy', der auch in dieser Arbeit gebraucht wird.

Es existiert eine reichhaltige Theorie bezüglich unscharfen Mengen, die eine Erweiterung der scharfen Mengentheorie ist. Es müssen jedoch immer spezielle Definitionen für unscharfe Mengen deklariert werden. So ist eine unscharfe Menge  $\tilde{A}$  gleich einer unscharfen Menge  $\tilde{B}$ , wenn ihre Grade der Zugehörigkeit für alle  $x \in U$  gleich sind  $\mu_{\tilde{A}}(x) = \mu_{\tilde{B}}(x)$ . Im Falle der Gleichheit zweier scharfen Mengen enthalten beide Mengen einfach die gleichen Elemente. Eine unscharfe Menge ist Teilmenge oder gleich einer anderen, wenn die Grade der Zugehörigkeit kleiner oder gleich sind  $\mu_{\tilde{A}}(x) \leq \mu_{\tilde{B}}(x)$  für alle  $x \in U$ . Die Ungleichheit und echte Teilmenge werden analog über die Zugehörigkeit definiert.

Eine Beschränkung einer unscharfen Menge nach unten bezüglich ihres Zugehörigkeitsgrades ist der  $\alpha$ -Schnitt. Dieser Schnitt bedeutet wieder eine scharfe Menge, die nur Elemente enthält die zu einem bestimmten minimalen Grad zu der ursprünglichen unscharfen Menge gehören. In einer Formel wird der  $\alpha$ -Schnitt wie folgt definiert:

$$\tilde{A}_\alpha = \{x \in U : \mu_{\tilde{A}}(x) \geq \alpha\}$$

Wird anstatt dem  $\geq$  ein  $>$  gefordert, so spricht man von einem strengen  $\alpha$ -Schnitt  $\tilde{A}_{\alpha^-}$ .

Die unscharfe Arithmetik beinhaltet die unscharfen Zahlen und die Operatoren zu deren Verrechnung. Sie entsteht durch das Erweiterungsprinzip aus der scharfen Arithmetik.

Eine unscharfe Zahl ist eine unscharfe Größe  $\tilde{A}$  mit  $\mu_{\tilde{A}}(c) = 1$  für genau ein  $c \in \mathbb{R}$  und  $\tilde{A}$  muss konvex sein. Die unscharfe Größe ist eine unscharfe Menge über dem Universum der reellen Zahlen, mit folgendem Grad der Zugehörigkeit:

$$\mu_{\tilde{A}} : \mathbb{R} \rightarrow [0, 1]$$

Das unscharfe Intervall ist wieder eine unscharfe Größe  $\tilde{A}$  die konvex ist, jedoch ist  $\mu_{\tilde{A}}(c) = 1$  für mindestens ein  $c$ . Es existieren auch spezielle unscharfe Zahlen wie z.B. Singletons, die durch folgende funktionale Zugehörigkeit beschrieben werden:

$$\mu_{\tilde{A}}(x) = \begin{cases} c & , \text{ für } x = a \text{ mit } c \in [0, 1] \\ 0 & , \text{ sonst.} \end{cases}$$

Eine parametrische Repräsentation von unscharfen Zahlen und Intervallen bietet die LR-Zerlegung. Diese zerlegt die unscharfe Zahl oder das unscharfe Intervall in einen monoton wachsenden Teil, die L-Funktion, und einen monoton fallenden Teil, die R-Funktion.

Die unscharfen arithmetischen Operatoren werden gemäß dem Erweiterungsprinzip aus den scharfen Operatoren erweitert. Die unscharfen Operatoren sind, für die unscharfen Zahlen  $\tilde{A}$ ,  $\tilde{B}$  und  $\tilde{C}$  anhand ihrer Zugehörigkeiten, wie folgt definiert:

$$\begin{aligned} \mu_{\tilde{A}+\tilde{B}}(x) &= \sup\{\min(\mu_{\tilde{A}}(u), \mu_{\tilde{B}}(v)) \mid u+v = x\} \\ \mu_{\tilde{A}\cdot\tilde{B}}(x) &= \sup\{\min(\mu_{\tilde{A}}(u), \mu_{\tilde{B}}(v)) \mid u \cdot v = x\} \\ \mu_{\tilde{A}-\tilde{B}}(x) &= \sup\{\min(\mu_{\tilde{A}}(u), \mu_{\tilde{B}}(v)) \mid u-v = x\} \\ \mu_{-\tilde{A}}(x) &= \sup\{\mu_{\tilde{A}}(u) \mid -u = x\} \\ \mu_{\tilde{A}/\tilde{B}}(x) &= \sup\{\min(\mu_{\tilde{A}}(u), \mu_{\tilde{B}}(v)) \mid u = v \cdot x\} \end{aligned}$$

Mit der Theorie der unscharfen Mengen ergibt sich also eine Möglichkeit der unscharfen Interpretation von Mengen, Zahlen und arithmetischen Operatoren. Die Probleme von scharfen Mengen und der mit diesen Mengen assoziierten Klassen, die eine nicht immer mögliche eindeutige Zuordnung bedingen, können mit den unscharfen Mengen gelöst werden. Um auch die Aussagekraft der scharfen Logik in eine 'Unschärfe' zu transferieren und dadurch deren Unzulänglichkeiten bei nicht eindeutigen Aussagen zu beseitigen, wird auch eine unscharfe Logik benötigt.

### 3.1.3 Unscharfe Logik

Die unscharfe Logik ist eine Erweiterung der klassischen Aussagenlogik. Die Menge der Wahrheitswerte  $W = \{0, 1\}$  der Aussagenlogik enthält zwei Elemente, 0 entspricht wahr und 1 entspricht falsch. In der unscharfen Logik ist die Menge der Wahrheitswerte ein Intervall  $\widetilde{W} = [0, 1]$  das reelle Zahlen enthält. Auf diese unscharfen Wahrheitswerte lassen sich die klassischen Operatoren der Logik  $\{\neg, \wedge, \vee, \rightarrow, \leftrightarrow\}$  nicht anwenden, weshalb sie für den unscharfen Fall erweitert werden müssen. Diese Erweiterung ist nicht eindeutig, denn es existieren viele unscharfe Operatoren die sich rechtfertigen lassen. Aus diesem Grund werden Axiome festgelegt, denen ein unscharfer Operator genügen muss. Die Operatoren im unscharfen Fall werden Normen genannt, wobei im axiomatisch korrekten Fall ausgegangen von dem UND-Äquivalent der  $T$ -Norm die weiteren Normen bestimmt werden. Das Äquivalent zu dem klassischen ODER ist die  $S$ -Norm, das der Negation die  $N$ -Norm und das der scharfen Implikation die  $I$ -Norm. Der Index der jeweiligen Norm bestimmt um welche Art es sich handelt. Nicht jede Norm erfüllt alle Axiome.

Eine  $T$ -Norm ist eine binäre Operation auf  $[0, 1]$ , also eine Funktion  $T : [0, 1]^2 \rightarrow [0, 1]$  die für  $x, y, z \in [0, 1]$  folgende Axiome erfüllt:

$$\begin{array}{ll} T(x, y) = T(y, x) & \text{Kommutativität (T1)} \\ T(T(x, y), z) = T(x, T(y, z)) & \text{Assoziativität (T2)} \\ x \leq y \Rightarrow T(x, z) \leq T(y, z) & \text{Monotonie (T3)} \\ T(1, x) = x & \text{Einselement (T4)} \end{array}$$

Weitere axiomatische Eigenschaften sind Stetigkeit auf  $(0, 1)^2$  (T5) und Idempotenz  $T(x, x) = x$  (T6), welche aber nicht verpflichtend sind.

Damit ergeben sich die folgenden gängigsten  $T$ -Normen:

$$\begin{array}{ll} \text{Minimumsnorm } T_M : & T_M(x, y) = \min(x, y) & , \text{ erfüllt (T1)-(T6)} \\ \text{Produktnorm } T_P : & T_P(x, y) = x \cdot y & , \text{ erfüllt (T1)-(T5)} \\ \text{Lukasiewicz'sche } T\text{-Norm } T_L : & T_L(x, y) = \max(x + y - 1, 0) & , \text{ erfüllt (T1)-(T5)} \\ \text{Drastisches Produkt } T_D : & T_D(x, y) = \begin{cases} \min(x, y) & , \text{ falls } \max(x, y) = 1 \\ 0 & , \text{ sonst} \end{cases} & , \text{ erfüllt (T1)-(T4)} \end{array}$$

Diese  $T$ -Normen sind total geordnet  $T_D < T_L < T_P < T_M$ , wobei die Minimumsnorm  $T_M$  immer die größte und das drastische Produkt  $T_D$  immer die kleinste  $T$ -Norm ist.

Eine unscharfe Negation ist eine Funktion  $N : [0, 1] \rightarrow [0, 1]$ , die für  $x, y \in [0, 1]$  folgende Axiome erfüllt:

$$\begin{array}{ll} N(0) = 1, N(1) = 0 & \text{Grenzbedingung (N1)} \\ x \leq y \Rightarrow N(x) \geq N(y) & \text{monotones Fallen (N2)} \end{array}$$

Weitere Eigenschaften die nicht gefordert sind, sind Stetigkeit (N3) und die Involution  $N(N(x)) = x$  für alle  $x \in [0, 1]$  (N4).

In Abhängigkeit von der gewählten  $T$ -Norm  $T$  und den Axiomen ergeben sich durch das Erweiterungsprinzip  $N_T = \sup\{y : T(x, y) = 0\}$  folgende Negationen:

$$\begin{array}{ll} \text{Minimumsnorm } T_M : & N_{T_M}(x) = N_\gamma(x) = \begin{cases} 1 & , \text{ falls } x = 0 \\ 0 & , \text{ sonst} \end{cases} & , \text{ erfüllt (N1)-(N2)} \\ \text{Produktnorm } T_P : & N_{T_P}(x) = N_\gamma(x) = \begin{cases} 1 & , \text{ falls } x = 0 \\ 0 & , \text{ sonst} \end{cases} & , \text{ erfüllt (N1)-(N2)} \\ \text{Lukasiewicz'sche } T\text{-Norm } T_L : & N_{T_L}(x) = N(x) = 1 - x & , \text{ erfüllt (N1)-(N4)} \\ \text{Drastisches Produkt } T_D : & N_{T_D}(x) = \begin{cases} 1 & , \text{ falls } x < 1 \\ 0 & , \text{ falls } x = 1 \end{cases} & , \text{ erfüllt (N1)-(N2)} \\ & & \text{nicht Kontradiktionsprinzip} \end{array}$$

Eine  $S$ -Norm ( $T$ -Conorm) ist eine Funktion  $S : [0, 1]^2 \rightarrow [0, 1]$ , die für  $x, y, z \in [0, 1]$  folgende Axiome erfüllt:

$$\begin{aligned} S(x, y) &= S(y, x) && \text{Kommutativität (S1)} \\ S(S(x, y), z) &= S(x, S(y, z)) && \text{Assoziativität (S2)} \\ x \leq y &\Rightarrow S(x, z) \leq S(y, z) && \text{Monotonie (S3)} \\ S(x, 0) &= x && \text{Nullelement (S4)} \end{aligned}$$

Andere Eigenschaften die von  $S$ -Normen erfüllt werden können, ist die Stetigkeit (S5) und die Idempotenz  $S(x, x) = x$  (S6).

Das DeMorgan'sche Tripel  $S(x, y) = N(T(N(x), N(y)))$  bestimmt mit der gewählten  $T$ -Norm und Negation  $N$  die jeweilige  $S$ -Norm. Es wird die Standardnegation  $N(x) = 1 - x$  gewählt, da diese involutiv ist und somit folgende unscharfe Disjunktionen ergibt:

$$\begin{aligned} \text{Maximumsnorm } S_M : & \quad S_M(x, y) = \max(x, y) && , \text{ erfüllt (S1)-(S4),(S6)} \\ \text{Probabilistische Summe } S_P : & \quad S_P(x, y) = x + y - x \cdot y && , \text{ erfüllt (S1)-(S5)} \\ \text{Lukasiewicz'she } S\text{-Norm } S_L : & \quad S_L(x, y) = \min(x + y, 1) && , \text{ erfüllt (S1)-(S5)} \end{aligned}$$

Die unscharfe Implikation  $I$  bedarf mehr Spezifikationen als die bisherigen Normen. Wie schon in der klassischen Aussagenlogik kann die unscharfe Implikation  $I$  über die anderen Operatoren bestimmt werden. In der klassischen Aussagenlogik gilt generell das Kontradiktionsprinzip  $a \wedge \neg a = 0$  und der tertium non datur  $a \vee \neg a = 1$ , weshalb folgende Grundlagen in dieser gleich wären:

$$\begin{aligned} I^S(x, y) &= S(N(x), y) && \text{'nicht-oder'-Implikation} \\ I^Q(x, y) &= S(N(x), T(x, y)) && \text{Q-Implikation} \end{aligned}$$

Eine weitere Möglichkeit der Bestimmung von Implikationen ist die Residuierung, die nicht aus Termen berechnet, sondern aus dem Erweiterungsprinzip bestimmt wird. Die Implikation der Residuierung ergibt sich somit analog zu der  $T$ -Norm aus folgender Menge:

$$I^R(x, y) = \sup\{z : T(x, z) \leq y\} \quad \text{Residuierungs-Implikation}$$

Es existieren noch weitere Implikationsoperatoren, wie der Mamdani-Operator  $I^{MAM}$  oder der Produkt-Operator  $I^P$ , die jedoch weder den Axiomen entsprechen, noch ein Äquivalent zu der scharfen Implikation sind.

Das nun vollständige Normen(Operatoren)-System  $\{T, N, S, I\}$  sorgt für einen funktional vollständigen Umfang von unscharfen Termen, der benötigt wird, um nicht eindeutige Sachverhalte zu formulieren. Diese Sachverhalte sind in dieser Arbeit die Zustände von kontextuellen Systemen, die eine bestimmte Qualität der Erkennung haben.

Es fehlt nun noch eine Möglichkeit aus einer Regel und einem Fakt einen Schluss abzuleiten, was mit den unscharfen Relationen ermöglicht wird.

### 3.1.4 Unscharfe Relationen

Relationen stellen Beziehungen zwischen Objekten her. Eine  $n$ -stellige scharfe Relation ist eine scharfe Teilmenge  $R \subset U_1 \times \dots \times U_n$  des kartesischen Produktes der Universen  $U_1, \dots, U_n$ . Eine  $n$ -stellige unscharfe Relation auf den Universen  $U_1, \dots, U_n$  ist eine unscharfe Menge  $\tilde{R} \in \mathbb{F}(U_1 \times \dots \times U_n)$ , also ein Element der unscharfen Potenzmenge  $\mathbb{F}$ . Die Funktion, die die unscharfe Relation  $\tilde{R}$  charakterisiert, ist wie folgt definiert:

$$\mu_{\tilde{R}} : U_1 \times \dots \times U_n \rightarrow [0, 1]$$

Die Projektion  $\tilde{R} \downarrow U_1 \times \dots \times U_k$  zu der unscharfen Relation  $\tilde{R} \in \mathbb{F}(U_1 \times \dots \times U_n)$  ist für  $k \leq n$  durch folgende funktionale Zugehörigkeit gegeben:

$$\mu_{\tilde{R} \downarrow U_1 \times \dots \times U_k}(u_1, \dots, u_k) = \sup_{(u_{k+1}, \dots, u_n) \in U_{k+1} \times \dots \times U_n} \{\mu_{\tilde{R}}(u_1, \dots, u_n)\}$$

Die Zylinderbildung  $\tilde{R} \uparrow U_1 \times \dots \times U_n \in \mathbb{F}(U_1 \times \dots \times U_n)$  zu der unscharfen Relation  $\tilde{R} \in \mathbb{F}(U_1 \times \dots \times U_k)$  ist durch folgende Gleichheit der funktionalen Zugehörigkeiten gegeben:

$$\mu_{\tilde{R} \uparrow U_1 \times \dots \times U_n}(u_1, \dots, u_n) = \mu_{\tilde{R}}(u_1, \dots, u_k)$$

Die Komposition der Projektion und der Zylinderbildung ist verlustbehaftet, denn die Relation  $\tilde{R} \in \mathbb{F}(U_1 \times \dots \times U_n)$  kann aus der Projektion nicht wieder hergestellt werden. Dieser Sachverhalt wird durch folgende Teilmengenbeziehung beschrieben:

$$\tilde{R} \subseteq (\tilde{R} \downarrow U_1 \times \dots \times U_k) \uparrow U_1 \times \dots \times U_n$$

Der Durchschnitt, die Vereinigung und das Komplement zweier unscharfer Relationen  $\tilde{R}, \tilde{S} \in \mathbb{F}(U_1 \times \dots \times U_n)$  sind, mit den Elementen  $u_1, \dots, u_n$  des kartesischen Produkts der Universen  $U_1 \times \dots \times U_n$ , wie folgt definiert:

$$\begin{aligned} \tilde{R} \cap_T \tilde{S} : \quad & \mu_{\tilde{R} \cap_T \tilde{S}}(u_1, \dots, u_n) = T(\mu_{\tilde{R}}(u_1, \dots, u_n), \mu_{\tilde{S}}(u_1, \dots, u_n)) \\ \tilde{R} \cup_S \tilde{S} : \quad & \mu_{\tilde{R} \cup_S \tilde{S}}(u_1, \dots, u_n) = S(\mu_{\tilde{R}}(u_1, \dots, u_n), \mu_{\tilde{S}}(u_1, \dots, u_n)) \\ \overline{\tilde{R}}^N : \quad & \mu_{\overline{\tilde{R}}^N}(u_1, \dots, u_n) = N(\mu_{\tilde{R}}(u_1, \dots, u_n)) \end{aligned}$$

Es wurde explizit gekennzeichnet welche  $T$ -Norm,  $S$ -Norm und Negation  $N$  verwendet wurde.

Spezielle Relationen sind die Äquivalenz- und Ordnungsrelation, die bestimmten Eigenschaften entsprechen müssen. Für binäre unscharfe Relationen  $\tilde{R} : U \times U \rightarrow [0, 1]$  sind folgende Eigenschaften für alle  $u, v, w \in U$  wie folgt definiert:

$$\begin{array}{ll} \mu_{\tilde{R}}(u, u) = 1 & \text{Reflexivität} \\ \mu_{\tilde{R}}(u, v) = \mu_{\tilde{R}}(v, u) & \text{Symmetrie} \\ u \neq v \Rightarrow \mu_{\tilde{R}}(u, v) = 0 \vee \mu_{\tilde{R}}(v, u) = 0 & \text{Antisymmetrie} \\ \mu_{\tilde{R}}(u, w) \geq T(\mu_{\tilde{R}}(u, v), \mu_{\tilde{R}}(v, w)) & \text{t-Transitivität} \end{array}$$

Eine unscharfe Relation  $\tilde{R}$  heißt unscharfe t-Äquivalenzrelation, genau dann wenn  $\tilde{R}$  reflexiv, symmetrisch und t-transitiv ist.

Die unscharfe Relation  $\tilde{R}$  ist unscharfe t-Halbordnung, wenn  $\tilde{R}$  reflexiv, antisymmetrisch und t-transitiv ist.

Zuletzt wird in diesem Abschnitt die unscharfe Komposition definiert, welche wieder in der unscharfen Variante nicht eindeutig ist. Die Komposition ermöglicht den unscharfen Schluss mittels des Modus Ponens und ist somit die eigentlich wichtigste Relation die hier verwendet wird. Eine häufig verwendete Komposition ist die sup-min-Komposition. Diese ist für die beiden unscharfen Relationen  $\tilde{P} \in \mathbb{F}(U_1 \times \dots \times U_m \times W)$  und  $\tilde{Q} \in \mathbb{F}(W \times V_1 \times \dots \times V_n)$  über die charakteristische Funktion wie folgt definiert:

$$\mu_{\tilde{P} \circ \tilde{Q}}(u_1, \dots, u_m, v_1, \dots, v_n) = \sup_{w \in W} \{ \min \{ \mu_{\tilde{P}}(u_1, \dots, u_m, w), \mu_{\tilde{Q}}(w, v_1, \dots, v_n) \} \}$$

Weitere Möglichkeiten der Komposition von Relationen sind die  $S$ - $T$ -Komposition und die sup- $T$ -Komposition. Die Komposition wird auch oft mit dem eingedeutschten Begriff Inferenz bezeichnet, der hauptsächlich in der Literatur zur Fuzzy Logik auftaucht.

Mit den Ableitungen des unscharfen Schlusses ist die Theorie der unscharfen Mengen noch lange nicht erschöpft. Die Anwendung für unscharfe Mengen, die in dieser Arbeit vorgestellt wird, benötigt jedoch keine tieferen Analysen und Spezifizierungen zum Thema der unscharfen Mengen.

Die Aussagekraft von unscharfen Termen ist von zentraler Bedeutung für die Analyse bestehender kontextueller Systeme bezüglich ihrer Qualität. Die Anzahl von Termen bestimmt die Genauigkeit der Analysen, da Sachverhalte besser erkannt werden, wenn sie möglichst präzise beschrieben werden. Die Systeme, die zur Kontexterkenkung eingesetzt werden, müssen oft mit einer großen Masse von Daten umgehen, weshalb auch die unscharfe Logik der Qualitätsanalyse diesen Anforderungen entsprechen muss. Es wird jedoch aus Gründen der Performanz ein minimaler Satz an unscharfen Termen befürwortet. Um den Datenraum möglichst geschickt zu unterteilen und damit eine Verminderung der Komplexität in den ersten Schritten der Erstellung von unscharfen Termen zu erreichen, werden im folgenden Kapitel verschiedene Verfahren des Clusterings vorgestellt, mit denen dies ermöglicht wird.

### 3.2 Algorithmische Möglichkeiten zur Datenanalyse - Cluster zur Gruppierung

Die Unterteilung einer Datenmenge in Teilmengen, ohne Wissen über die Zugehörigkeit der Daten zu bestimmten Klassen zu besitzen, kann mit einem automatisierten Verfahren des Clusterings gemacht werden. Es werden Daten durch ein Clustering zusammen gruppiert, die eine gewisse Similarität besitzen. Die Similarität kann durch verschiedene Kriterien bestimmt werden. Es existieren Methoden der heuristischen Gruppierung, die die Similarität durch einen mathematischen Abstand bestimmen. Die Daten werden zu Vektoren gruppiert und der Abstand zwischen diesen bestimmt die Ähnlichkeit. Es existieren verschiedene Abstandsmetriken, wie z.B. die euklidische. Die zweite Möglichkeit der Gruppierung ist die der probabilistischen Ähnlichkeit, bei der die Wahrscheinlichkeit für das Maß der Similarität genutzt wird.

Die verschiedenen Algorithmen die zum Clustering eingesetzt werden, ermöglichen es, auch eine automatische Extraktion von Termen aus Daten zu erhalten. Wird ein Cluster identifiziert, so können dessen einzelne Dimensionen über eine Logik verknüpft werden und eine Klasse implizieren. Der logische Term der sich daraus ergibt, kann entscheiden ob ein Vektor zu einer Klasse gehört oder nicht.

Wird eine unscharfe Logik für diese Entscheidung eingesetzt, so erhält man eine unscharfe Zugehörigkeit zu einer Klasse. Eine Klasse kann im Sinne der Informatik vieles bedeuten, was im vorigen Kapitel beschrieben wurde. Es kann also auf eine Klasse impliziert werden, die alle richtig erkannten kontextuellen Zustände beinhaltet. Damit ergibt sich direkt mit der unscharfen Zugehörigkeit zu dieser 'Richtig'-Klasse, die aus einem unscharfen Term gewonnen wurde, die relative Qualität des kontextuellen Zustandes.

Das größte Problem bei der Einteilung von Daten durch Cluster-Algorithmen liegt jedoch in der Trennbarkeit. Häufig sind Daten die verschiedene Klassen implizieren durch keine Similarität voneinander zu unterscheiden. Es macht auch Sinn, dass verschiedene Cluster gemeinsame Teilmengen besitzen, was allerdings nicht durch ein Clustering erkannt werden kann. Trotz diesen Problemen können Cluster Algorithmen zur ersten Trennung von Daten und zur grundlegenden Extraktion von Termen eingesetzt werden. Oft ist die menschliche Einteilung gerade höher dimensionaler Daten kaum möglich, weshalb dann sowieso auf einen Cluster-Algorithmus zurückgegriffen werden muss.

In diesem Abschnitt werden die wichtigsten Algorithmen des Clusterings vorgestellt, wobei mit einem Algorithmus begonnen wird, der die Daten in scharfe Cluster einteilt. Mit diesem scharfen Clustering wird unter anderem die Notwendigkeit eines unscharfen Cluster-Algorithmus ergründet. Die weitere Auflistung der Algorithmen fährt mit einem einfachen Algorithmus zur Einteilung des Datenraumes in Cluster fort und endet bei einem Algorithmus bei dem die Probleme der vorherigen weitestgehend gelöst wurden.

#### 3.2.1 Das k-Means Clustering

Das k-Means Clustering ist eines der am häufigsten verwendeten scharfen Cluster-Verfahren, da es schnell konvergiert und einfach zu implementieren ist. Es muss dem Clustering die Anzahl der Cluster vorgegeben werden, was durch das  $k$  im Namen beschrieben wird. Das Verfahren wird über ein Maß der Ähnlichkeit beschrieben, die Distanz, die eine Zuordnung zu einem nächsten Cluster ermöglicht. Die Wahl der Distanzfunktion ist daher ein grundlegender Bestandteil des k-Means Algorithmus und entscheidet unter Anderem über das Resultat der Gruppierung.

Es können zwei Gruppen der Distanzfunktionen eingeteilt werden, die Gruppe der heuristischen und die der probabilistischen Maße. Es wird hier nun ein heuristisches Maß beschrieben, die euklidischen Distanz, die wie folgt zu berechnen ist:

$$d(\vec{x}_j, \vec{m}_i) = \|\vec{x}_j - \vec{m}_i\|_2 = \sqrt{\sum_{l=1}^p (x_{jl} - m_{il})^2} \quad , \text{ mit } \vec{x}_j \in X = \{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n\} \\ \text{ und } \vec{m}_i \in M = \{\vec{m}_1, \dots, \vec{m}_c\}$$

Der Datensatz der in Cluster eingeteilt werden soll, ist die Menge  $X = \{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n\}$  der  $p$ -dimensionalen Vektoren  $\vec{x}_j$ . Die Menge der Cluster-Zentren ist  $M = \{\vec{m}_1, \dots, \vec{m}_c\}$ , die  $c$  Vektoren  $\vec{m}_i$  enthält.

Die Berechnung der Cluster erfolgt nun iterativ, wobei zuerst eine initiale Menge der Cluster-Zentren gewählt werden muss. Diese initialen Cluster-Zentren bestimmen über Erfolg und Misserfolg des Algorithmus, da falsch initialisierte Vektoren  $\vec{m}_i$  nie zu richtigen Cluster-Zentren konvergieren werden.

Die Iteration des Algorithmus erfolgt über die Berechnung der Distanz  $d(\vec{x}_j, \vec{m}_i)$  eines jeden Vektors  $\vec{x}_j$  zu jedem Cluster-Zentrum  $\vec{m}_i$ , wobei der Vektor  $\vec{x}_j$  zu dem Cluster  $i$  als zugehörig gezählt wird, zu dem er den kleinsten Abstand  $d$  hat. Wurde für jeden Vektor entschieden zu welchem Cluster  $i$  er gehört, so können anhand dieser Teilmengen  $X_i$  der Menge  $X$  die neuen Cluster-Zentren wie folgt berechnet werden:

$$\vec{m}_{i,t+1} = \vec{m}_{i,t} + \alpha(\vec{x}_j - \vec{m}_i) \quad , \text{ für alle } \vec{x}_j \in X_i$$

Der Faktor  $\alpha$  bestimmt die Schrittweite, in der pro Iteration die Vektoren der Mittelwerte in Richtung der Datenvektoren, die zu ihrem repräsentierten Cluster zählen, verschoben werden. Es werden Werte für  $\alpha \in [0, 1]$  von  $< 1$  gewählt, da sonst eine zu sprunghafte Veränderungen der Cluster-Zentren  $i$  pro Veränderung durch Vektoren der Menge  $X_i$  auftreten. Der Algorithmus stoppt, wenn die Cluster-Zentren von der  $t$ 'ten Iteration sich zur  $(t+1)$ 'ten Iteration weniger als  $\varepsilon$  unterscheiden, was die folgende Gleichung ausdrückt:

$$|\vec{m}_{i,t} - \vec{m}_{i,t+1}| < \varepsilon$$

Ein Nachteil des k-Means Clustering ist, dass die Anzahl  $k$  der Cluster vor der Initialisierung des Algorithmus bestimmt werden muss. Dies ist oft ohne visuellen Zugang zu den Daten kaum möglich, da selbst eine richtige Wahl von  $k$  nicht den Erfolg des Verfahrens bedeutet und wenn ein möglicher Erfolg eintritt, dies auch nicht immer mathematisch nachgewiesen werden kann. Wie soll jedoch ein multidimensionaler Datenraum visualisiert werden und macht das für die automatisierte und eigentlich 'nicht überwachte' Datenanalyse überhaupt Sinn? Eine Möglichkeit der Gruppierung und Visualisierung von multidimensionalen Daten ist eine Kohlen selbstorganisierende Karte [Koh81], die zu den topologischen einschichtigen neuronalen Netzen zählt.

Die in dieser Arbeit an Algorithmen der Datenanalyse gestellten Anforderungen, werden durch die scharfe Interpretation der Clustergrenzen durch den k-Means Algorithmus nicht erfüllt. Eine Identifikation eines Clusters durch fließende Grenzen entspricht mehr den Ergebnissen einer Qualitätsanalyse, da nicht nur richtig oder falsch erkannt werden soll, sondern auch sämtliche Zwischenzustände die ein kontextuelles System annehmen kann. Aus diesem Grund werden Algorithmen des unscharfen Clusterings bevorzugt.

### 3.2.2 Das c-Means Clustering

Der c-Means Algorithmus ist eine unscharfe Variante des scharfen k-Means Algorithmus. Dieser unscharfe Cluster-Algorithmus wurde von Dunn [Dun73] formuliert. Die endgültige Verallgemeinerung wurde durch Einführung eines Fuzzifiers  $m$  von Bezdek [Bez73] beschrieben. Die Beschreibung des c-Means Algorithmus in dieser Arbeit stammt aus dem Buch 'Fuzzy Clusteranalyse' von Höppner, Klawonn und Kruse [HKK97].

Der unscharfe c-Means Algorithmus erkennt kugelförmige Datenwolken im p-dimensionalen Raum. Jeder Cluster wird durch einen Vektor identifiziert, der seinen Mittelpunkt spezifiziert. Die Cluster werden als gleich groß angenommen. Damit eine Zugehörigkeit zu einem Cluster durch den Algorithmus berechnet werden kann, muss ein Maß für die Distanz gewählt werden. Der euklidische Abstand beschreibt diese Distanz und wird wie folgt berechnet:

$$d(\vec{x}_j, \vec{m}_i) = \|\vec{x}_j - \vec{m}_i\|_2$$

Die Anzahl der Cluster muss dem Algorithmus vorgegeben werden und ist durch das c im Namen des unscharfen c-Means Algorithmus beschrieben. Bei einer gegebenen Zugehörigkeit  $\mu_i$  zu einem Cluster  $i$  wird der p-dimensionale Vektor  $\vec{m}_i$  des Mittelpunkts durch eine verallgemeinerten Mittelwert berechnet, was durch den englischen Begriff 'Means' im Namen des Algorithmus beschrieben wird. Die Cluster-Zentren werden wie folgt berechnet:

$$\vec{m}_i = \frac{\sum_{j=1}^n \mu_i(\vec{x}_j) \vec{x}_j}{\sum_{j=1}^n \mu_i(\vec{x}_j)}$$

Die Herleitung der Formel für die Mittelwerte  $\vec{m}_i$  findet man in [HKK97]. Eine Möglichkeit zur Definition der Zugehörigkeit  $\mu_i$  des Vektors  $\vec{x}_j$  zum Cluster  $i$ , ist durch folgende Gleichung gegeben:

$$\mu_i(\vec{x}_j) = \frac{1}{d(\vec{x}_j, \vec{m}_i)}$$

Oft wird die Summe der Zugehörigkeiten für alle Vektoren  $x_j$  eines Datensatzes  $X$  zu einem Cluster  $i$  auf Eins normiert, was durch folgende Gleichung beschrieben wird:

$$\sum_{i=1}^c \mu_i(\vec{x}_j) = 1 \text{ für alle } \vec{x}_j \in X$$

Die Zugehörigkeiten zu den Clustern muss nun modifiziert werden, damit sie dieser Forderung der Normierung genügen. Die normalisierten Zugehörigkeiten sind die folgenden:

$$\mu_k(\vec{x}_j) = \frac{1}{\sum_{i=1, i \neq k}^c \left( \frac{d(\vec{x}_j, \vec{m}_k)}{d(\vec{x}_j, \vec{m}_i)} \right)^{\frac{1}{m-1}}}, \text{ mit } m > 1$$

Der Parameter  $m$  ist der erwähnte Fuzzifier und bestimmt eine Gewichtung, wobei ein Wert  $m \approx 1$  eine starke Gewichtung des nächsten Cluster-Zentrums bewirkt.

Der eigentliche Algorithmus berechnet nun iterativ die Cluster-Zentren  $\vec{m}_i$  und die Zugehörigkeiten  $\mu_j$  bis einem Konvergenzkriterium erfüllt wird. Es wird pro Iteration die normalisierte gewichtete Summe jedes Cluster-Zentrums  $\vec{m}_i \in M$  über den gesamten Datensatz  $X = \{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n\}$  berechnet. Vor der ersten Iteration des  $c$ -Means Clusterings muss eine Menge  $M = \{\vec{m}_1, \dots, \vec{m}_c\}$  der initialen Cluster-Zentren festgelegt werden, anhand deren die ersten Zugehörigkeiten  $\mu_j$  berechnet werden können. Ein Konvergenzkriterium, das die Abbruchbedingung des Algorithmus bildet, kann Folgendes sein:

$$|\vec{m}_{i,t} - \vec{m}_{i,t+1}| < \varepsilon$$

Die  $t$ 'te Iteration wird durch den Index  $t$  identifiziert, welche die letzte des Algorithmus war und somit die endgültige Menge der Cluster-Zentren  $M$  und funktionalen Zugehörigkeiten  $\mu$  bestimmt.

Die Probleme des unscharfen  $c$ -Means Clustering ergeben sich bei unterschiedlicher Form, Größe und Dichte der Cluster. Hinsichtlich der Form können nur kugelförmige Clusterformen realisiert werden, da der Algorithmus die Form nicht an die tatsächliche anpasst. Die unterschiedliche Dichte von Clustern hat auf entfernt liegende Datenpunkte nicht die entsprechende Auswirkung, sie werden alleinig nach ihrer Distanz zum Cluster-Zentrum bezüglich ihrer Zugehörigkeit beurteilt. Auch die Größe der Cluster kann oft durch die menschliche Intuition als falsch erachtet werden, obwohl sie durch die durch den Algorithmus geforderten Richtlinien richtig adaptiert werden.

### 3.2.3 Der Gustafson-Kessel Algorithmus

Durch die Veränderung der Metrik, die bei  $c$ -Means benutzt wurde, können bei dem Gustafson-Kessel Algorithmus auch ellipsoide Cluster realisiert werden. Das Gustafson-Kessel Clustering [GK] kann die Clusterform für jeden Cluster separat anpassen. Eine kurze Beschreibung und Analyse des Algorithmus findet sich wieder in [HKK97].

Die Grundlage des Gustafson-Kessel Algorithmus ist eine spezielle Distanz, die wie folgt über eine neue  $A$ -Norm definiert wird:

$$d(\vec{x}_j, \vec{m}_i) := \|\vec{x}_j - \vec{m}_i\|_{A_i} = \sqrt{(\vec{x}_j - \vec{m}_i)^T A_i (\vec{x}_j - \vec{m}_i)}$$

Diese  $A$ -Norm wird über die positiv definite Matrix  $A_i$  charakterisiert, die wie folgt zu berechnen ist:

$$A_i = \varrho \sqrt{\det(S_i)} S_i^{-1}$$

Es wird in [HKK97] ein konstantes Clustervolumen durch  $\det(A_i) = 1$  gefordert. Durch diese Forderung ist nur eine variable Clusterform möglich und keine unterschiedlichen Clustergrößen. Gustafson und Kessel erlauben unterschiedliche Clustergrößen, indem sie für jede Matrix  $A_i$  eine Konstante  $\varrho_i$  einführen mit  $\det(A_i) = \varrho_i$ . Die Matrizen  $S_i$  berechnet man mit folgender Gleichung:

$$S_i = \sum_{j=1}^n \mu_i(\vec{x}_j) \|\vec{x}_j - \vec{m}_i\|_2^2$$

Die Mittelwert-Vektoren  $\vec{m}_i$ , die die Cluster-Zentren bilden, und die Zugehörigkeiten  $\mu_j$  berechnet man mit den gleichen Formeln wie schon zuvor beim  $c$ -Means Clustering, jedoch jetzt mit der veränderten

Metrik der Distanz.

Die Begründung der genannten Formeln ist wieder in [HKK97] zu finden.

Mit der Gustafson-Kessel Methode wurde zwar die Problematik von unterschiedlichen Clusterformen gelöst, jedoch durch die Forderung von konstanten Matrizen  $A_i$ , über den Wert der Determinante von 1 oder  $\rho_i$ , kann die Clustergröße nicht dynamisch an die reale Größe angepasst werden. Das Problem der Datenhäufung im Cluster-Zentrum bleibt auch erhalten, da algorithmisch die Dichte der Cluster noch nicht berücksichtigt wurde.

### 3.2.4 Der Gath-Geva Algorithmus

Der Gath-Geva Algorithmus [GG89] ist eine Erweiterung des Gustafson-Kessel Clustering, die auch auf die Größe und Dichte der Cluster eingeht. Der Gath-Geva Algorithmus interpretiert die Daten als  $p$ -dimensionale normalverteilte Zufallsvariablen, was auch wieder in [HKK97] beschrieben wird.

Gath und Geva verallgemeinern die Ergebnisse der Wahrscheinlichkeitstheorie für den Fall der unscharfen Zugehörigkeit. Sie gehen davon aus, dass die Normalverteilung  $N_i$  mit dem Erwartungswert für das Cluster-Zentrum  $\bar{m}_i$ , der Kovarianzmatrix  $C_i$  und der a-priori Wahrscheinlichkeit  $P_i$  zur Erzeugung der Daten  $\bar{x}_i \in X$  verwendet wurde. Aus dieser Vermutung ergibt sich die a-posteriori Wahrscheinlichkeit  $\mathbb{P}$ , dass das Datum  $\bar{x}_i$  durch die Normalverteilung  $N_i$  erzeugt wurde, durch die Maximum-Likelihood-Methode wie folgt:

$$\mathbb{P}(\bar{x}_j) = \frac{P_i}{(2\pi)^{\frac{p}{2}} \sqrt{\det(\Sigma_i)}} e^{(-\frac{1}{2}(\bar{x}_j - \bar{m}_i)^T \Sigma_i^{-1} (\bar{x}_j - \bar{m}_i))}$$

Die Kovarianzmatrizen  $\Sigma_i$  berechnet man, anhand der Zugehörigkeiten  $\mu_i$  der Datenvektoren  $\bar{x}_j$  des Datensatzes  $X = \{\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n\}$ , wie folgt:

$$\Sigma_i = \frac{\sum_{j=1}^n \mu_i(\bar{x}_j) (\bar{x}_j - \bar{m}_i) (\bar{x}_j - \bar{m}_i)^T}{\sum_{j=1}^n \mu_i(\bar{x}_j)}$$

Die a-priori Wahrscheinlichkeit  $P_i$  wird nach folgender Formel berechnet:

$$P_i = \frac{\sum_{j=1}^n \mu_i(\bar{x}_j)}{\sum_{j=1}^n \sum_{l=1}^c \mu_l(\bar{x}_j)}$$

Die Mittelwerte der Cluster-Zentren  $\bar{m}_i$  und die Zugehörigkeiten  $\mu_i$  werden wieder über die jeweiligen Formeln berechnet, die im Zusammenhang mit dem c-Means Clustering beschrieben wurden. Nun muss noch die Distanzfunktion für den Gath-Geva Algorithmus gewählt werden, die umgekehrt proportional zu der a-posteriori Wahrscheinlichkeit  $\mathbb{P}$  des Vektors  $\bar{x}_j$  ist. Die Distanz des Vektors  $\bar{m}_i$  der Mittelwerte des Cluster-Zentrums  $i$  und des Datenvektors  $\bar{x}_j$  wird anhand folgender Formel berechnet:

$$d^2(\bar{x}_j, \bar{m}_i) := \frac{\sqrt{\det(\Sigma_i)}}{P_i} e^{(-\frac{1}{2}(\bar{x}_j - \bar{m}_i)^T \Sigma_i^{-1} (\bar{x}_j - \bar{m}_i))}$$

Der Algorithmus wird, wie schon zuvor bei dem c-Means Clustering, über eine iterative Berechnung von Cluster-Zentren und Zugehörigkeiten, bis eine Abbruchbedingung erfüllt wird, ausgeführt. Die Abbruchbedingung ist wieder eine Epsilon-Schranke  $\varepsilon$  die, hinsichtlich der Distanz von Mittelwerten einer Iteration  $t$  zu einer Iteration  $t + 1$ , unterschritten wird.

Mit der Methode des Gath-Geva Clusterings können nun Cluster unterschiedlicher Form, Größe und Dichte entsprechend adaptiert werden. Das Problem der bisherig vorgestellten Cluster-Algorithmen, dass eine Anzahl  $c$  der Cluster durch den Entwickler vorgegeben werden muss, ist noch nicht gelöst. Auch sind die Probleme die in Verbindung mit der Initialisierung der Cluster-Zentren entstehen, die schon für das scharfe k-Means Clustering beschrieben wurden, noch nicht gelöst.

### 3.2.5 Das Mountain Clustering

Das Mountain Clustering, das von Yager und Filev [YF94] eingeführt wurde, zerteilt den Datenraum in äquidistante Intervalle um ein Gitter zu erstellen. Die Distanz der Daten zu den einzelnen Knoten, die durch die Einteilung des Gitters im Datenraum entstehen, bietet eine Möglichkeit die Knoten zu identifizieren, die ein großes Potential haben, ein Cluster-Zentrum zu sein. Das Potential eines jeden Knotens

wird über die funktionale Zugehörigkeit der Datenpunkte zu dem Knoten gemessen. Die funktionale Zugehörigkeit, die das Maß der Distanz über alle Datenpunkte evaluiert, ist eine Gauss-Funktion. Durch die Gitter-Einteilung wird eine Vorgabe der Anzahl der Cluster und der initialen Cluster-Zentren nicht mehr benötigt.

Der Algorithmus des Mountain Clusterings beginnt mit der Einteilung jeder Dimension des Eingaberaumes in äquidistante Knoten des Gitters. Da das Universum der jeweiligen Dimension des Raumes über die Daten, die in Cluster einzuteilen sind, selbst bestimmt wird, muss zuerst das Intervall jeder Dimension bestimmt werden. Das Intervall der Dimension  $j = 1, \dots, s$  wird über alle Datenvektoren  $\vec{x}_k$  mit  $k = 1, \dots, n$  anhand von Minimum und Maximum der Daten, wie folgt berechnet:

$$I_j := [\min_k(x_{kj}), \max_k(x_{kj})]$$

Die Intervalle  $I_j$  werden in jeweils  $r_j$  äquidistante Gitterpunkte  $i_{jl}$  wie folgt eingeteilt:

$$I_1^G [i_{11}, \dots, i_{1r_1}], \dots, I_s^G [i_{s1}, \dots, i_{sr_s}]$$

Die eigentlichen Knoten  $N$  werden nun aus allen möglichen Kombinationen der Gitterpunkte  $i_{jo_j}$  mit  $o_j \in I_j^G$  aller Dimensionen  $s$  zusammengesetzt. Aus welchen Gitterpunkten der einzelne Knotenpunkt  $N_{i=i_{1o_1}, \dots, i_{so_s}}$  besteht, wird durch den Index  $i = (i_{1o_1}, \dots, i_{so_s})$ , mit  $o \in \{1, \dots, r_1\}$  und  $q \in \{1, \dots, r_s\}$ , beschrieben.

Der zweite Schritt des Mountain Algorithmus assoziiert mit jedem Knoten  $N_i$  ein Potential, das wie folgt, anhand der Datenvektoren  $\vec{x}_k$  und der Norm  $\|\cdot\|_p$ , berechnet wird:

$$M^1(N_i) = \sum_{k=1}^n e^{-\alpha \|\vec{x}_k - N_i\|_p}$$

Der Index  $t$  von  $M^t$  beschreibt die  $t$ 'te Iteration des Algorithmus. Die erste Iteration besteht aus der Auswahl des ersten größten Potentials  $M_1^*$  durch das Maximum, wie folgt:

$$M_1^* := \max_i (M^1(N_i))$$

Der mit diesem Potential  $M_1^*$  assoziierte Knoten  $N_1^*$  bildet das erste Cluster-Zentrum. Nachdem ein Cluster-Zentrum  $N_{t-1}^*$  gewählt wurde, müssen die Potentiale  $M^t(N_i)$  aller verbleibenden Knoten  $N_i$  wie folgt erneuert werden:

$$M^t(N_i) := M^{t-1}(N_i) - M_{t-1}^* - 1 \sum_{k=1}^n e^{-\beta \|N_{t-1}^* - N_i\|_2}$$

Jede Iteration  $t$  des Mountain Algorithmus besteht aus Erneuerung der Potentiale  $M^t(N_i)$  für alle noch nicht erwählten Knoten  $N_i$  und Auswahl von neuen Cluster-Zentren  $N_t^*$  durch Berechnung des Maximums.

Das Verfahren stoppt, wenn das Verhältnis vom Potential  $M_1^*$  des ersten Cluster-Zentrums  $N_1^*$  zum Potential  $M_t^*$  des als nächsten zu erwählenden Cluster-Zentrums  $N_t^*$  unter eine Schranke  $\delta$  fällt. Das Verhältnis der Potentiale und der Schranke, ist wie folgt zu berechnen:

$$\frac{M_1^*}{M_t^*} < \delta$$

Es wurden einige Verbesserungen des Mountain Clusterings vorgeschlagen. Eine Methode, die von Lee, Son und Kwon [LSK01] vorgeschlagen wurde, reduziert die Anzahl der Parameter von drei ( $\alpha, \beta, \delta$ ) auf einen ( $\omega$ ). Eine automatisierte Methode der Bestimmung der Parameter ( $\alpha, \beta, \delta$ ) und eine Prozedur der Anpassung der Cluster, wurde von Branco, Lori und Dente [BLD95] beschrieben.

Trotz der vielen Verbesserungen, die zum Algorithmus des Mountain Clusterings entwickelt wurden, bestehen immer noch grundlegende Nachteile des Verfahrens. Für die Qualität der Clustereinteilung ist die Dichte des Gitters von entscheidender Bedeutung. Die Abstände der Gitterpunkte bestimmt jedoch auch die Performanz des Algorithmus.

### 3.2.6 Das subtraktive Clustering

Das Subtractive Clustering wurde von Chiu [Chi96][Chi97] entwickelt und ist eine einfach zu benutzende Methode des Clusterings. Der Algorithmus berechnet die Anzahl der Cluster selbstorganisierend, benötigt

also keine Angabe des Entwicklers. Jeder Datenpunkt, des Datensatzes der in unscharfe Cluster einzuteilen ist, wird als potentielles Cluster-Zentrum abgewogen. Dieser Sachverhalt verbessert das Mountain Clustering, da in diesem eine Auflösung eines Gitters spezifiziert werden muss. Diese Auflösung des Gitters bestimmt auch die Komplexität des Mountain Clusterings, wohingegen beim subtraktiven Clustering die Anzahl der Datenpunkte die Komplexität der Berechnung bestimmt. Das subtraktive Clustering erweitert auch das Kriterium der Akzeptanz oder Verwerfung von potentiellen Cluster-Zentren.

Es wird jeder Vektor des Datensatzes, der in Cluster einzuteilen ist, als potentielles Cluster-Zentrum angenommen. Die Entscheidung, ob sich ein Datenvektor als Cluster-Zentrum eignet, wird anhand einer Berechnung des Potentials getroffen. Das Potential jedes Vektors wird durch eine bestimmte Distanz zu jedem anderen Vektor des Datensatzes berechnet. Die verwendete Distanz eines Vektors zu einem anderen ist zunächst euklidisch, wird dann aber in einer Gauss-Funktion gewichtet, die die Zugehörigkeit zu einem potentiellen Cluster-Zentrum bedeutet. Der Datenvektor der das größte Potential gegenüber allen anderen besitzt wird als erstes Cluster-Zentrum definiert.

Aufgrund dieses ersten Cluster-Zentrums werden die Potentiale aller anderen Vektoren verändert. Von jedem Potential wird das gewichtete Potential des ersten Cluster-Zentrums abgezogen. Die Gewichtung ist wieder eine Gauss-Funktion der Zugehörigkeit zu diesem Cluster-Zentrum.

Das zweite Cluster-Zentrum wird anhand des größten Potentials der verbleibenden Vektoren gewählt. Nun wiederholt sich der Prozess der Veränderung der Potentiale anhand dieses zweiten Cluster-Zentrums.

Die Abwechslung von Auswahl von neuen Cluster-Zentren und Veränderung der Potentiale wird solange wiederholt, bis das zuletzt größte Potential kleiner ist als ein Prozentsatz des ersten.

Das Subtractive Clustering ist laut Chiu [Chi97] effizient und robust im Umgang mit hoch-dimensionalen und verrauschten Daten. Der Benutzer dieser Methode muss nur den in Cluster einzuteilenden Datensatz präsentieren, die Berechnung der Cluster-Zentren und die Anzahl der Cluster werden selbstorganisierend durch den Algorithmus ausgeführt und festgestellt. Da der Algorithmus des subtraktiven Clusterings die Berechnungen der Potentiale über Gauss-Funktionen vornimmt, können diese Funktionen direkt in Zugehörigkeiten zum jeweiligen Cluster umgerechnet werden.

Die eigentliche Methode des subtraktiven Clusterings wird in Kapitel 4.2.1 beschrieben, da sie aufgrund der Vorteile gegenüber den anderen unscharfen Cluster-Algorithmen zum Extrahieren der unscharfen Regeln aus einem Datensatz in dieser Arbeit verwendet wurde.

### 3.3 Analyse der verschiedenen unscharfen Regelsysteme

Werden unscharfe Mengen zur Systembeschreibung benutzt, so wird das Übertragungsverhalten nicht durch ein Gleichungssystem, sondern anhand von Regeln beschrieben. Diese linguistischen IF..THEN..-Regeln werden in einer Regelbasis zusammengefasst. Das Übertragungsverhalten des Systems ist unscharf, wobei die Eingabe und die Ausgabe scharfe Wertebereiche benutzt. Der schematische Aufbau eines unscharfen Regelsystems ist in Abbildung 2 dargestellt, in welcher auch eingesehen werden kann, dass die klassische Rückkopplung eines Reglers hier nicht gefordert ist. Besitzt die Ausgabe eines unscharfen Regelsystems eine aktive Beeinflussung des Systems, so wird die Eingabe wieder um diese verändert. Gerade in dieser Arbeit werden Ausgaben von unscharfen Regelsystemen verwendet, die nur eine bedingte bis keine Beeinflussung der Eingabe bewirken und deshalb nicht als klassisches Regelsystem bezeichnet werden können. Es wird jedoch eine Regelbasis verwendet, die eine Abhängigkeit des Verhaltens zwischen Eingabe und Ausgabe beschreibt, wodurch eine Bezeichnung als regelbasiertes System eigentlich treffender wäre. In dieser Arbeit wird nun mit der Deklaration von Regelsystem ein regelbasiertes System beschrieben, da die Kürze für diesen oftmals wiederholten Ausdruck ausschlaggebend ist.

Das unscharfe Regelsystem muss die reellwertige Eingabe vor der unscharfen Schlussfolgerung erst in unscharfe Mengen umwandeln. Dieser Schritt ist symbolisch durch die Box 'Fuzzifikations-Schnittstelle' in Abbildung 2 dargestellt. Wie die scharfe Eingabe in die unscharfen Werte transferiert werden soll, wird

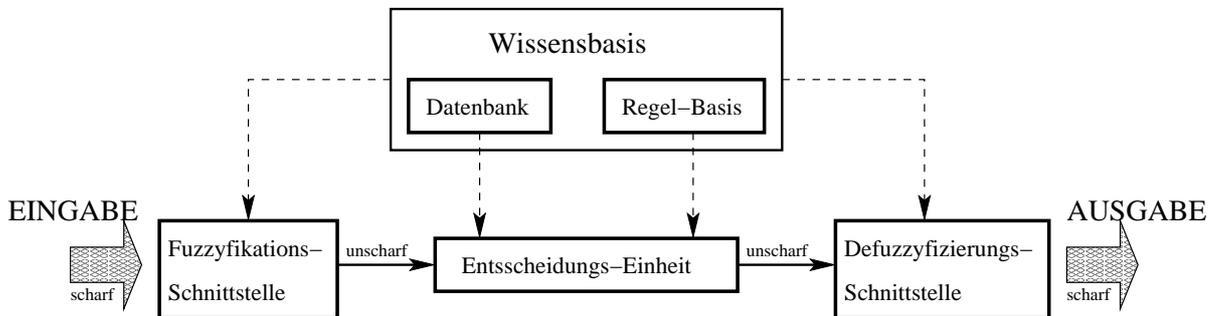


Abbildung 2: Das unscharfe Regel-System

durch die in der Wissensbasis gespeicherten Zugehörigkeitsfunktionen der unscharfen Mengen beschrieben. Das nächste Glied in der Kette von Eingabe zu Ausgabe des Systems ist die 'Entscheidungs-Einheit', welche die regelbasierte Ableitung der unscharfen Werte durchführt. Die Regel besteht aus dem schon erwähnten IF..THEN..-Konstrukt, das in der klassischen scharfen Logik die Implikation beschreibt. Der Vorgang der Ableitung wird, wie auch im scharfen Fall, durch den Modus-Ponens definiert. Der Modus-Ponens ist mit der IF..THEN..-Regel wie folgt darzustellen:

$$\frac{\begin{array}{l} \text{IF } x \text{ is } \tilde{A} \text{ THEN } y \text{ is } \tilde{B} \quad (\text{Regel}) \\ x \text{ is } \tilde{A}' \quad (\text{Fakt}) \end{array}}{y \text{ is } \tilde{B}' \quad (\text{Schluss})}$$

Die Bedingung wird aufgrund der implikativen Natur der Regel als Antezedens und der bedingte Teil als Konsequenz bezeichnet. Mehrere Bestandteile der Antezedens werden über eine geeignete  $T$ -Norm und die einzelnen Teile der Konsequenz über eine geeignete  $S$ -Norm verknüpft. Die funktionale Zugehörigkeit von  $x$  zu der unscharfen Menge  $\tilde{A}$  beschränkt die Zugehörigkeit von  $y$  zur unscharfen Menge  $\tilde{B}$ , was die IF..THEN..-Regel beschreibt. Der 'Fakt' beschreibt die aktuelle Unschärfe die den Schluss über die Regel bestimmt und somit als Resultat eine funktionale Zugehörigkeit der Variablen  $y$  zur unscharfen Menge  $\tilde{A}'$  hat. Wird die linguistisch formulierte Regel durch eine geeignete (Implikations-)Relation  $\tilde{R}$  ersetzt, so wird der Modus Ponens wie folgt formuliert:

$$\frac{\begin{array}{l} (x, y) \text{ is } \tilde{R} \quad (\text{Regel}) \\ x \text{ is } \tilde{A}' \quad (\text{Fakt}) \end{array}}{y \text{ is } \tilde{A}' \circ \tilde{R} \quad (\text{Schluss})}$$

Durch die Methode der Konstruktion dieser (Implikations-)Relation  $\tilde{R}$  werden zu der beschriebenen linguistischen Methode des approximativen Schließens noch zwei weitere Möglichkeiten des Schlusses ermöglicht. Das possibilistische Schließen verwendet eine geeignete Possibilitätsverteilung, um die Relation

$\tilde{R}$  gemäß der minimalen Spezifität zu bestimmen. Wird anstatt der Possibilitätsverteilung eine Evidenzverteilung genutzt, so nennt man die Schussmöglichkeit evidenzgestützt. Das evidenzgestützte Schließen arbeitet nach dem Prinzip der maximalen Kompletierung.

Das Ergebnis des Schlusses ist in den meisten Fällen eine kombinierte funktionale Zugehörigkeit. Um einen reellen Wert zu erhalten, der durch diese funktionale Zugehörigkeit am meisten repräsentiert wird, muss eine 'Defuzzifizierung' (Abbildung 2, rechts) vorgenommen werden. Es existieren verschiedene Möglichkeiten der 'Defuzzifizierung', wie die Maximum-, die Mittelwert-Maximum- und die (Flächen-)Schwerpunkt-Methode.

Die grundsätzliche Einteilung der verschiedenen unscharfen Regelsysteme wird von Jang [Jan93] anhand der Regel- und Schlussfolgerungs-Methode vorgenommen. Wie in Abbildung 3 eingesehen werden

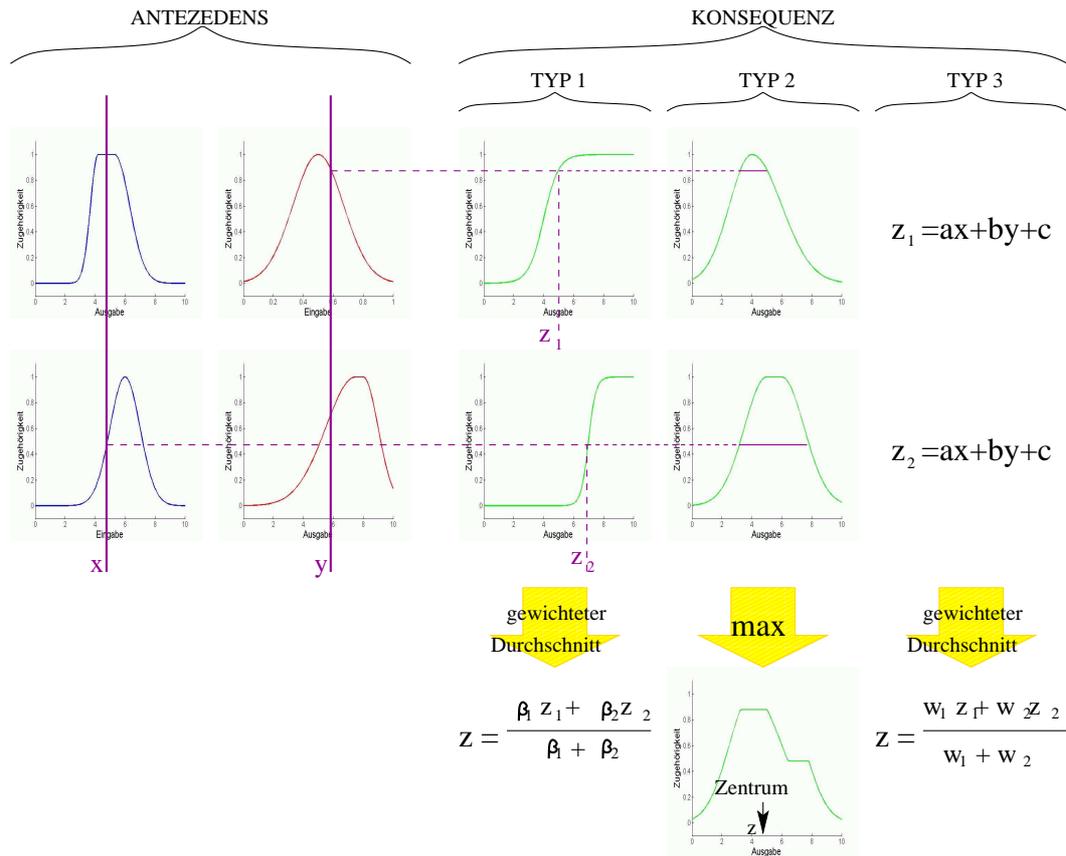


Abbildung 3: Abhängigkeit der Konsequenz von der Antezedens für die verschiedenen Typen der Regelsysteme

kann, unterteilt er die unscharfen Regelsysteme in drei verschiedene Typen. Die meisten unscharfen Regelsysteme können in folgende Typen (Abbildung 3, rechts) eingeteilt werden:

- TYP1: Die Ausgabe dieses Typs wird durch einen gewichteten Mittelwert der einzelnen Regeln anhand der funktionalen Konsequenzen und die jeweiligen über die  $T$ -Norm verknüpften funktionalen Zugehörigkeiten der Antezedenzen berechnet. Die funktionalen Zugehörigkeiten der Konsequenzen müssen streng monoton sein.
- TYP2: Die einzelnen funktionalen Zugehörigkeiten der Konsequenzen werden über die  $T$ -Norm verknüpften Zugehörigkeitsfunktionen der Antezedenzen nach oben beschränkt. Diese beschränkten Zugehörigkeiten werden anhand ihres Maximums miteinander verknüpft, was die resultierende funktionale Zugehörigkeit aller Regeln ergibt. Eine gewählte 'Defuzzifikations'-Methode bildet die Berechnungsvorschrift für den Ausgabewert.
- TYP3: Die spezielle Implikation die im Rahmen der unscharfen Takagi-Sugeno-Kang (TSK) Regelsysteme [TS85][SK88] benutzt wird bedingt in der Konsequenz eine Konstante oder eine lineare Funktion die

von den Eingabewerten abhängig ist. Die endgültige Ausgabe des Regelsystems ist die gewichtete normalisierte Summe der Antezedenzen und Konsequenzen.

In den Folgenden Abschnitten werden nun die gängigsten unscharfen Regelsysteme beschrieben. Die Vorteile und Nachteile der einzelnen Systeme werden abgewogen, unter anderem anhand der automatisierbaren Erstellbarkeit des Systems.

### 3.3.1 Das Mamdani- und Larsen-Regelsystem

Die Mamdani- [MA75] und Larsen-Regelsysteme [LY90] gehören zum Typ 2 der unscharfen Regelsysteme. Die Zugehörigkeitsfunktionen für die Antezedenzen und die Konsequenzen der Regeln sollten stetig sein. Es wird approximatives Schließen verwendet, bei dem die Regeln über linguistische Terme formuliert werden. Eine Regel mit zwei Eingabewerten und einem Ausgabewert ist mittels dem Schluss des Modus Ponens folgendermaßen zu formulieren:

$$\frac{\begin{array}{l} \text{IF } x \text{ is } \tilde{A} \text{ AND } y \text{ is } \tilde{B} \text{ THEN } z \text{ is } \tilde{C} \quad (\text{Regel}) \\ x \text{ is } \tilde{A}' \text{ AND } y \text{ is } \tilde{B}' \quad (\text{Fakt}) \end{array}}{z \text{ is } \tilde{C}' \quad (\text{Schluss})}$$

Die unscharfe Regel wird als unscharfe Relation  $\tilde{R}$  auf den ursprünglichen Universen  $X \times Y \times Z$  interpretiert. Die unscharfen Eingabewerte werden als unscharfe Relation  $\tilde{S}$  auf  $X \times Y$  definiert. Mit diesen Relationen muss der Modus Ponens umformuliert werden, was wie folgt aussieht:

$$\frac{\begin{array}{l} (x, y, z) \text{ is } \tilde{R} \quad (\text{Regel}) \\ (x, y) \text{ is } \tilde{S} \quad (\text{Fakt}) \end{array}}{z \text{ is } \tilde{S} \circ \tilde{R} \quad (\text{Schluss})}$$

Da es gewünscht ist, dass  $\tilde{S}$  und  $\tilde{R}$  gleichzeitig Gültigkeit besitzen, ist es ermöglicht die Verkettungsregel für unscharfe Regeln wie folgt einzusetzen:

$$\tilde{C}' = \tilde{S} \circ \tilde{R}$$

Die unscharfe Relation  $\tilde{R}$  wird über die einfachste Methode des unscharfen kartesischen Produkts als größtmögliche unscharfe Relation bestimmt. Die Methode ist die Max-Min-Inferenzmethode und wird für die gegebene Regel mit den funktionalen Zugehörigkeiten  $\mu$  wie folgt formuliert:

$$\begin{aligned} \mu_{\tilde{C}'}(y) &= \max_{x \in X, y \in Y} \min [\mu_{\tilde{R}}(x, y, z), \mu_{\tilde{A}}(x), \mu_{\tilde{B}}(y)] \\ &= \max_{x \in X, y \in Y} \min [\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(y), \mu_{\tilde{C}}(z), \mu_{\tilde{A}}(x), \mu_{\tilde{B}}(y)] \\ &= \min [\mu_{\tilde{A}}(a), \mu_{\tilde{B}}(b), \mu_{\tilde{C}}(z)], \text{ da } x = a \text{ und } y = b \text{ konst.} \end{aligned}$$

Werden die konstanten Eingabewerte  $x = a$  und  $y = b$  angenommen, so ergibt sich der Aktivierungsgrad  $\beta$  der Formel wie folgt:

$$\beta = \min [\mu_{\tilde{A}}(a), \mu_{\tilde{B}}(b)]$$

Werden mehrere Regeln  $n$  im unscharfen Regelsystem benutzt, so muss die unscharfe Relation  $\tilde{R}$  auf die unscharfen Relationen  $\tilde{R}_i$  mit  $i = 1, \dots, n$  erweitert werden. Auch sind die unscharfen Ausgabewerte  $\tilde{C}'$  auf  $n$  Werte  $\tilde{C}'_i$  zu erweitern, was die Bestimmung des gesamten Ausgabewertes nicht gerade vereinfacht. Der gesamte unscharfe Ausgabewert  $\tilde{C}'$  ist die Vereinigung aller unscharfen Einzelergebnisse, also  $\tilde{C}' = \cup_i \tilde{C}'_i$ . Wird diese Vereinigung auf die funktionalen Zugehörigkeiten der Ausgaben übertragen, so erhält man die Formel  $\mu_{\tilde{C}'}(z) = \max_i [\mu_{\tilde{C}'_i}(z)]$ .

Zur Berechnung eines scharfen Ausgabewertes muss eine 'Defuzzyfikations'-Methode gewählt werden. Die gewählte Methode wird dann auf die funktionale Zugehörigkeit der vereinigten unscharfen Ausgaben  $\tilde{C}'_i$  angewandt. Ist die Flächen-Schwerpunkt-Methode die Wahl, so ergibt sich die folgende scharfe Ausgabe:

$$z^* = \frac{\int_Z z \mu_{\tilde{C}'}(z) dz}{\int_Z \mu_{\tilde{C}'}(z) dz}$$

Wenn  $\mu_{\tilde{C}_r}(z)$  aus einer Menge gewichteter Singletons  $z_i$  besteht, dann vereinfacht sich die Berechnung der scharfen Ausgabe mit den Aktivierungsgraden  $\beta_i$ , wie folgt:

$$z_{singleton}^* = \sum_{z_i \in Z} \frac{\beta_i}{\sum_{z_i \in Z} \beta_i} z_i$$

Die Methode der Mamdani-Regelsysteme kann anhand des Beispiels beschrieben werden, das in Abbildung 3 für den zweiten Typ dargestellt ist. Die unscharfen Werte der Antezedens jeder Regel werden über eine  $T$ - oder  $S$ -Norm verknüpft. Bezüglich den typischen Normalformen die in Logiken existieren, sollte die  $T$ -Norm gewählt werden. In dem vorliegenden Beispiel wurde die Minimum-Norm  $T_M$  gewählt. Das Minimum aller funktionalen Zugehörigkeiten der Antezedens einer Regel bestimmt, anhand der gewählten Implikation, die Anpassung der unscharfen Werte der Konsequenz. Da ein Mamdani-Regelsystem der Gegenstand der Beschreibung ist, wird eine Mamdani-Implikation  $I^{MAM}$  benutzt, die jedoch den Axiomen der unscharfen Implikation nicht entspricht. Die Mamdani-Implikation  $I^{MAM}$  ist das Minimum, was die funktionalen Zugehörigkeiten einer Konsequenz jeder Regel auf das Minimum der Zugehörigkeiten einer beliebigen scharfen Eingabe beschränkt. Bestehen die Konsequenzen aus mehr als einem unscharfen Wert, so werden diese, aus gleichen Gründen wie bei der Verknüpfung der Antezedens, mit einer  $S$ -Norm verknüpft. Die gewählte  $T$ -Norm ist ein Minimum, weshalb für die  $S$ -Norm das Gegenstück, das Maximum genommen wird. Die somit bestimmte Konsequenz jeder Regel bildet die funktionale Zugehörigkeit der Ausgabe, indem über alle Funktionen das Maximum bestimmt wird. Mit der Flächen-Schwerpunkt-Methode über dieser funktionalen Zugehörigkeit der Ausgabe, ergibt sich die scharfe Ausgabe des Mamdani-Regelsystems.

Die Methode die Mamdani [MA75] und Larsen [LY90] beschrieben haben, eignet sich besonders zur manuellen Erstellung der Regelbasis und der funktionalen Zugehörigkeiten. Eine automatisierte Erzeugung eines solchen Regelsystems ist mit Cluster-Algorithmen und künstlichen neuronalen Netzen ebenso möglich. Die 'Defuzzifikations'-Methode benötigt jedoch zusätzlichen Aufwand der Berechnung, der bei anderen Regelsystemen eingespart werden kann.

### 3.3.2 Das Tsukamoto-Regelsystem

Das Tsukamoto-Regelsystem [Tsu90] gehört zu den Regelsystemen des Typs 1 und setzt somit streng monotone Zugehörigkeitsfunktionen für die Konsequenzen voraus, Dreiecks- oder Glockenkurven sind nicht erlaubt. Bei gleichem Approximationsfehler und größerer Anzahl von Eingabevariablen benötigt die Tsukamoto-Methode weniger Regeln als die Mamdani- und Larsen-Methode.

Die Schlussmöglichkeit des Modus Ponens die schon bei der Mamdani- und Larsen-Methode verwendet wurde, wird auch bei den Tsukamoto-Regelsystemen angewandt. Die Regeln werden auch wieder mit linguistischen Konstrukten des approximativen Schließens formuliert. Die unscharfe Ausgabe ist jedoch nicht eine funktionale Zugehörigkeit, sondern eine Menge von Singletons  $z_i$ , ein Singleton als Konsequenz der  $i$ 'ten Regel. Die reelle Ausgabe wird über die folgende normalisierte gewichtete Summe berechnet:

$$z_{singleton}^* = \sum_{z_i \in Z} \frac{\beta_i}{\sum_{z_i \in Z} \beta_i} z_i$$

Die Aktivierungsgrade  $\beta_i$  werden, für die Anzahl  $n$  der Regeln, wie schon bei der Mamdani- und Larsen-Methode anhand der funktionalen Zugehörigkeiten der Antezedenzien, durch folgende Formel berechnet:

$$\beta_i = \min \left[ \mu_{\tilde{A}_{1i}}(x_1), \dots, \mu_{\tilde{A}_{ni}}(x_n) \right]$$

Ein Beispiel für die Tsukamoto-Methode ist in Abbildung 3 für den Typ 1 dargestellt. Die Singletons werden wieder über die gewählte  $T$ -Norm anhand der funktionalen Zugehörigkeiten der Antezedenzien und der Eingabe des Regelsystems an der funktionalen Zugehörigkeit der Konsequenzen abgelesen.

Die Festlegung der funktionalen Zugehörigkeiten der Antezedenzien auf streng monoton fallende Funktionen, kann für gewisse Anwendungen zu restriktiv sein. Das Tsukamoto-Regelsystem ist recht exotisch und wurde hier nur aus Gründen der Vollständigkeit angegeben. Die Vorteile die das Regelsystem gegenüber der Mamdani- und Larsen-Methode besitzt, kann auch das folgende Takagi-Sugeno-Kang (TSK) Regelsystem aufweisen.

### 3.3.3 Das Takagi-Sugeno-Kang (TSK) Regelsystem

Das TSK-Regelsystem entstand in Arbeiten von Takagi und Sugeno [TS85] sowie von Sugeno und Kang [SK88]. Die Methode des TSK-Regelsystems wurde mehr zur Systemidentifikation und unscharfen Modellierung als zur Konstruktion unscharfer Regeln entwickelt. Dies ist auch der Grund weswegen eine automatisierte Entwurfsmethode bei den TSK-Systemen zu bevorzugen ist. Das Übertragungsverhalten der TSK-Methode bedarf der Spezifizierung vieler Variablen, die kaum manuell zu bestimmen sind.

Der größte Unterschied zu den bisher beschriebenen unscharfen Regelsystemen ist die Konsequenz der Implikation. Die Konsequenz jeder Regel ist eine lineare Funktion die von den reellwertigen Eingabegrößen abhängt oder eine Konstante. Der unscharfe Schluss der in der TSK-Methode verwendet wird muss daher funktional sein und beinhaltet zugleich die 'Defuzzifizierung' des Resultates. Die funktionalen Zugehörigkeiten die die Übertragung der reellen Eingabe auf die unscharfen Werte der Verarbeitung übernehmen, bilden wieder die Antezedens jeder implikativen Regel. Die Antezedens jeder Regel bildet nun jeweils ein Gewicht, das die Verknüpfung aller funktionalen Zugehörigkeiten der jeweiligen Antezedens durch eine geeignete  $T$ -Norm ist. Es wird von Takagi, Sugeno und Kang eine Produkt-Norm  $T_P$  vorgeschlagen, da diese in einem stetigen Verlauf des Gewichtes resultiert. Das Resultat des TSK-Regelsystems wird, mittels der Multiplikation der gewichtenden Antezedens mit der jeweiligen linearen oder konstanten Konsequenz und der Summierung dieser aller Regeln berechnet. Um dieses Resultat in den gewünschten Wertebereich zu transferieren, wird eine Normalisierung durch das Teilen dieses durch die Summe aller Gewichte erreicht.

Eine grafische Abstrahierung eines Beispiels findet man in der Abbildung 3 für den Typ 3.

Die TSK-Methode erlaubt speziell eine automatisierbare Erstellung des Regelsystems. Die Performanz der Übertragung von Eingabe zu Ausgabe ist wegen der Vereinigung von unscharfem Schluss und 'Defuzzifikation' sehr gut.

Eine Bestimmung eines Qualitätsmaßes kann meist nur durch adaptive Methoden erreicht werden. Aus diesem Grund ist ein TSK-Regelsystem zu bevorzugen. Da eine Ausführung eines Regelsystems auf einer in der Performanz beschränkten Hardware-System erwünscht ist, wurde das TSK-Regelsystem für die Verwendung in dieser Arbeit gewählt.

### 3.4 Automatisierte Adaption mit neuronalen Fuzzy Netzen

Unschärfe Systeme sind universelle Approximatoren, sie können jede reelle stetige Funktion infinitesimal genau nachbilden, was in dem Buch von Wang [Wan98] nachgewiesen wurde. Eine infinitesimal genaue Approximation benötigt eine infinite Anzahl von unscharfen Regeln. Die Genauigkeit steht jedoch in Konkurrenz mit der Komplexität des unscharfen Regelsystems. Eine reduzierte Anzahl von Regeln ist jedoch in der Approximation der Funktion eventuell nicht genügend genau. Die Bestimmung einer befriedigenden Näherung, an das Optimum der Approximation und an die beste Effizienz der Berechnung des resultierenden Regelsystems, ist die zentrale Aufgabe bei der Erstellung eines unscharfen Regelsystems.

Die Genauigkeit von unscharfen Regelsystemen in der Approximation von Funktionen hängt jedoch nicht nur von der Art und Anzahl der Regeln ab, sondern auch von den funktionalen Zugehörigkeiten der Ein- und Ausgabe. Während es noch manuell möglich ist, das Verhalten von Ausgabe zu Eingabe über Regeln zu formulieren, ist es umso schwieriger sinnvolle Zugehörigkeitsfunktionen zu bestimmen. Ein Umweg der Anpassung der Funktionen über einen Experten, auf Basis von Daten oder Systemverhalten, ist generell ineffizienter als eine automatisierte Anpassung. Wird also eine automatisierte Erstellung von unscharfen Regelsystemen angestrebt, so kann mit einer geschickten Wahl von adaptiven Algorithmen das Problem der Approximation mit einer begrenzten Anzahl von Regeln gelöst und trotzdem eine gute Effizienz des resultierenden Regelsystems erreicht werden.

Ein unscharfes Regelsystem ist eine Abbildung von der Eingabe auf die Ausgabe des Systems die deterministisch und statisch ist. Es ist zunächst nicht vorgesehen, dass sich ein unscharfes Regelsystem dynamisch verhält und gewisse automatische Anpassungen durchführen kann. Ein künstliches neuronales Netz hingegen hat genau diese Eigenschaft der automatisierten Anpassung. Diese neuronalen Netze lernen ihre Abbildung aufgrund von Eingabedaten und dynamischem Ausgabeverhalten. Um nun die Vorteile der neuronalen und unscharfen Methodiken zu vereinen wurden hybride Systeme entwickelt. Ein hybrides Neuro-Fuzzy-System hat die adaptiven Eigenschaften von neuronalen Netzen und die regelbasierte Mächtigkeit von unscharfen Mengen.

Den Aufbau und die Verwendung von unscharfen Regelsystemen wurde in vorangegangenen Kapiteln hergeleitet und beschrieben, nicht jedoch das Prinzip der neuronalen Netze. Eine kleine Einführung in die Theorie der neuronalen Netze soll nun den Übergang zur Konstruktion von hybriden Systemen erleichtern.

#### 3.4.1 Künstliche neuronale Netze (KNN)

Ein künstliches neuronales Netz besteht aus Knoten den Neuronen und Verbindungen zwischen diesen. Jedes Neuron ist mit einer Funktion assoziiert, die das Übertragungsverhalten festlegt. Die Eingabe des Neurons wird durch die Verbindungen, den Kanten des Netzes, mit Gewichten multipliziert. Die Neuronen sind in Schichten aufgebaut mit einer Eingabe- und einer Ausgabe-Schicht.

Wie die einzelnen Schichten von Neuronen verbunden sind wird durch die Netztopologie bestimmt. Sind die Neuronen nur in eine Richtung verbunden, so hat das neuronale Netz eine vorwärts gerichtete Netztopologie (Abbildung 4a). Bestehen auch Verbindungen von Neuronen innerhalb einer Schicht, dann ist das Netz eines mit 'Wettbewerb' zwischen den Neuronen (Abbildung 4b). Wenn die Verbindungen nicht nur Informationen in der Richtung von Eingabe zu Ausgabe weiter leiten, so spricht man von einem dynamischen neuronalen Netz (Abbildung 4c). Ein rekurrentes neuronales Netz (Abbildung 4d) hat Verbindungen von Neuronen der Eingabe- mit der Ausgabe-Schicht.

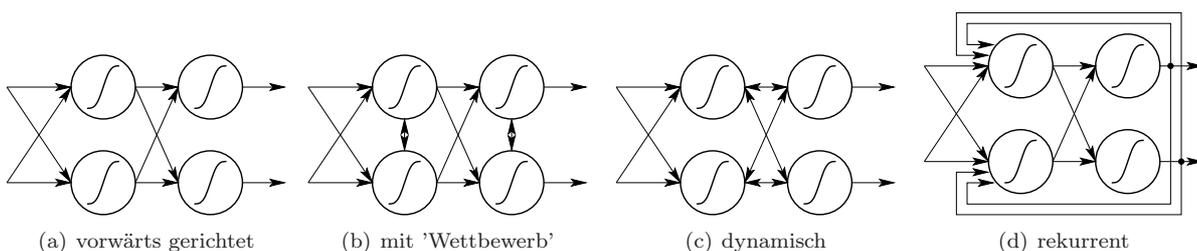


Abbildung 4: Verschiedene Arten des Aufbaus von neuronalen Netzwerken

Ein neuronales Netz erhält die Information, wie die Abbildung zwischen Eingabe und Ausgabe spezifiziert werden muss, durch Lernen. Es existieren zunächst zwei verschiedene Arten wie die Daten, die gelernt werden sollen, dem Netz zur Verfügung gestellt werden. Sind die Daten an die das Netz adaptiert werden soll vollständig vorhanden, so kann die Veränderung des Netzes in einem Schritt durchgeführt werden. Diese Methode wird Offline-Lernen genannt. Wird das Netz an jedes Datum separat adaptiert, spricht man von Online-Lernen.

Die folgenden groben Einteilungen der Lernverfahren beinhalten immer die Offline- und die Online-Lernmethode. Das überwachte Lernen benötigt zu jedem Datum der Eingabe ein weiteres Datum der gewünschten Ausgabe. Die neuronalen Netze, die keine Informationen über die gewünschte Ausgabe erhalten, werden mit einem unüberwachten Lernverfahren trainiert. Diese Netze lernen das Muster das die Daten beinhalten selbst-organisierend. Bei bestärkendem Lernen wird dem Netz eine Verbesserung oder Verschlechterung der Ausgabe für die jeweiligen Daten gemeldet.

Durch das Lernen werden bei den meisten neuronalen Netzen nur die Gewichte der Verbindungen verändert. Das Übertragungsverhalten des einzelnen Neurons bleibt normalerweise gleich. Die Neuronen beinhalten im Allgemeinen eine Schwellwertfunktion, die die Summe aller Eingaben des Neurons als Argument erhält. Wenn die Schwellwertfunktion in ihrem Wert eine bestimmte Grenze (Schwellwert) überschreitet, dann ändert sich die Ausgabe des Neurons von einem niedrigen zu einem hohen Pegel. In der Fachsprache wird dies als 'Feuern' eines Neurons bezeichnet.

Diese Einführung in die Theorie der künstlichen neuronalen Netze ist mehr als knapp und soll hier nur den Übergang zu dem Thema der neuronalen Fuzzy Netze erleichtern. Eine ausführliche Beschreibung von neuronalen Netzen mit mathematischen Analysen und der Einführung in den Stuttgarter-Neuronale-Netze-Simulator ist in dem Buch von Zell [Zel94] gegeben.

### 3.4.2 Fuzzy-Kohonen-Clustering-Netzwerk (FKCN)

Das Fuzzy-Kohonen-Clustering-Netzwerk (FKCN) ist die unscharfe Variante der Kohonen [Koh81] selbst-organisierenden Karten (SOM). Die SOM sind einschichtige neuronale Netze deren Neuronen in einem festen Gitter angeordnet sind. Die Position eines Neurons im Gitter bestimmt seine Nachbarn. Der Aufbau des Gitters definiert die Topologie, weshalb die SOM auch topologische Karten genannt werden. Mit jedem Neuron der SOM ist ein Vektor assoziiert, der Codebook-Vektor (CV) genannt wird. Dieser CV hat die gleiche Dimension wie die Vektoren der Eingabe. Diese CV werden zum Beginn des Trainings-Algorithmus initialisiert, wobei die anfänglichen Werte ausschlaggebend für die Topologie des finalen Netzes sind. Jeder neue Eingabevektor wird mit allen CV verglichen. Der CV, der die kleinste Distanz zur Eingabe besitzt, wird als 'Gewinner' bezeichnet. Das Maß zur Berechnung der Distanz ist zu wählen, ein mögliches ist das euklidische. Der CV der 'gewonnen' hat, bildet das Zentrum der Veränderung des Netzes durch den Eingabevektor. Dieser Gewinner wird am stärksten durch den Eingabevektor beeinflusst. Die Nachbarschaft zum jeweiligen 'Gewinner'-Neuron bestimmt die Gewichtung der Anpassung der CVs, je größer die Distanz desto kleiner die Veränderung.

Der Offline-Algorithmus zum Training der FKCN beginnt, wie der Online-Algorithmus der SOM auch, mit der Initialisierung der CV. Enthalten die CV schlechte Anfangswerte, so wird die Topologie der Karte sich nicht richtig entfalten. Sind die CV nicht zufällig genug initialisiert und gleichmäßig über die Karte verteilt, so können die Eingabevektoren nur bestimmte Bereiche der Karte beeinflussen. Die Topologie der gesamten Karte kann dann nicht richtig entwickelt werden oder die Konvergenz zu einer guten Topologie ist zu langsam.

Mit jedem Neuron des FKCN ist nicht nur ein CV assoziiert, sondern auch ein Wert  $\mu$ , der die Zugehörigkeit der Eingabe zum jeweiligen Neuron beschreibt. Die Zugehörigkeit  $\mu_{ij}$  des Eingabevektors  $\vec{x}_i$  zu dem CV  $\vec{w}_j$  für die  $k$ -te Iteration des FKCN Trainings-Algorithmus, ist wie folgt zu berechnen:

$$\mu_{ij}^{(k)} = \left( \sum_{r=1}^M \left( \frac{\|\vec{x}_i - \vec{w}_j^{(k)}\|}{\|\vec{x}_i - \vec{w}_r^{(k)}\|} \right)^{\left(\frac{2}{q-1}\right)} \right)^{-1}$$

Der Exponent  $q = q_0 - k\Delta q$  ist der monoton fallende Kontrastfaktor, mit  $q \in [0, 1]$ . Durch die Lernrate  $\eta$  beeinflusst die Zugehörigkeit  $\mu$  die Veränderung  $\Delta$  der CV. Es wird also aus der Zugehörigkeit  $\mu_{ij}^{(k)}$  die

Lernrate  $\eta_{ij}^{(k)}$  zur  $k$ -ten Iteration wie folgt berechnet:

$$\eta_{ij}^{(k)} = \left(\mu_{ij}^{(k)}\right)^q$$

Mit der Lernrate  $\eta$  kann nun die Veränderung der CV wie folgt bestimmt werden:

$$\Delta \bar{w}_j^{(k)} = \frac{\sum_{i=1}^N \eta^{(k)} (\bar{x}_i - \bar{w}_j^{k-1})}{\sum_{i=1}^N \eta^{(k)}}$$

Der  $N$ -wertige Datensatz der Vektoren  $\bar{x}_i$ , der für das Offline-Training verwendet wird, wird so lange für die Anpassung der CV eingesetzt, bis folgende Bedingung zum Abbruch der Iteration erfüllt ist:

$$\sum_{i=1}^N \sum_{j=1}^M \|\mu_{ij}^{(k)} - \mu_{ij}^{(k-1)}\| < \varepsilon$$

Das FKCN ist eigentlich eher den unscharfen Cluster-Verfahren zuzuordnen, obwohl die Architektur die eines einschichtigen neuronalen Netzwerks ist. Die Daten die dem Netzwerk zum Training dienen, formen die Topologie durch Anpassung der Codebook-Vektoren. Eine Interpretation der Karte und der Daten, die auf diese abgebildet werden, muss nach dem Training noch durch den Entwickler geschehen. Aus diesem Grund eignen sich topologische Karten sehr gut zur Visualisierung und Gruppierung von Daten über die kein Wissen besteht.

Im Fall einer Qualitätsanalyse von kontextuellen Zuständen ist aber für das Training eines neuronalen Netzes explizites Wissen vorhanden. Für einen Datensatz und den dafür erkannten kontextuellen Zustand kann angegeben werden, ob der Zustand richtig oder falsch ist. Diese absolute Klassifizierung bezüglich der Qualität ist ein starkes Argument für ein überwachtes Lernverfahren. Für die topologischen Karten, die auf Kohonen [Koh81] zurück gehen, existieren keine überwachten Lernverfahren, da diese nicht dafür konstruiert wurden.

Eine Architektur von unscharfen neuronalen Netzwerken, die mit überwachten Lernverfahren trainiert werden können, wird im nächsten Abschnitt vorgestellt.

### 3.4.3 Adaptive-network-based Fuzzy Inference System (ANFIS)

Das Adaptive-network-based Fuzzy Inference System (ANFIS) das von Jang [Jan93] entwickelt wurde, ist eine Methodik zur Konstruktion von neuronalen Fuzzy Netzen. Dieses Verfahren wird vornehmlich für TSK-Regelsysteme eingesetzt, kann aber auf die meisten unscharfen Regelsysteme angewandt werden. Das neuronale Netz, das die Repräsentation des unscharfen Regelsystems übernimmt, ist ein vorwärts gerichtetes Netz, das stückweise differenzierbar ist. Das ANFIS System ist eine Architektur für neuronale Fuzzy Netze, die zunächst keine Anforderungen an das Lernverfahren hat. Es wird jedoch das typische Lernen mit Gradientenabstieg, das für vorwärts gerichtete mehrschichtige Neuronale Netze üblich ist, eingesetzt.

Wird von einem TSK-Regelsystem [TS85][SK88] als Grundlage der Konstruktion ausgegangen, so kann das neuronale Fuzzy Netz auf zwei verschiedene Arten [JS95] aufgebaut werden. Die erste Art ergibt sich aus der folgenden allgemeinen Formel für ein TSK-System:

$$y^* = \frac{\sum_i w_i f_i}{\sum_i w_i}$$

Werden die Gewichte  $w_i$  einzeln durch die Summe aller Gewichte normalisiert, so ergibt sich die zweite Art ein ANFIS aufzubauen wie folgt:

$$y^* = \frac{\sum_i w_i f_i}{\sum_i w_i} = \sum_i \frac{w_i}{\sum_i w_i} f_i = \sum_i \bar{w}_i f_i$$

Die erste Methode ergibt ein Netzwerk mit 4-Schichten von Neuronen, die zweite Methode eines mit 5-Schichte.

Für ein Beispiel mit einer zweidimensionalen Eingabe  $(x, y)$  mit jeweils zwei Zugehörigkeitsfunktionen und den zwei linguistischen Regeln

Regel 1: IF  $x$  is  $X_1$  AND  $y$  is  $Y_1$  THEN  $f_1 = a_1x + b_1y + c_1$   
 Regel 2: IF  $x$  is  $X_2$  AND  $y$  is  $Y_2$  THEN  $f_2 = a_2x + b_2y + c_2$

ergeben sich die zwei Arten des ANFIS wie in Abbildung 5 dargestellt, die erste Methode links und die zweite rechts. Die ANFIS Repräsentation der TSK-Regelsysteme ist unter der Bedingung von Gauss'schen

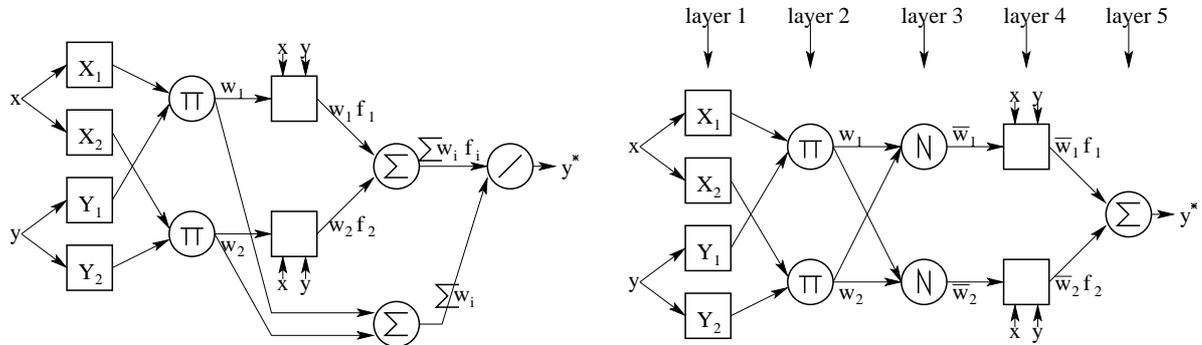


Abbildung 5: ANFIS für ein TSK-Regelsystem der ersten Art (links) und der zweiten Art (rechts)

Glockenkurven als funktionale Zugehörigkeiten und dem Produkt als AND-Operator  $T_P$  funktional identisch.

Eine andere Möglichkeit der Konstruktion eines ANFIS ist mit einem Typ-1-Regelsystem, dem Tsukamoto-Regelsystem [Tsu90] das zuvor vorgestellt wurde. Wird wieder das Beispiel mit zwei Eingängen und jeweils zwei funktionale Zugehörigkeiten genommen, so ergibt sich nach Jang [Jan93], mit den zwei Regeln und

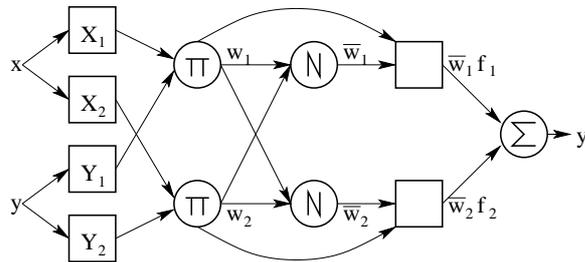


Abbildung 6: ANFIS für ein Tsukamoto-Regelsystem des Typ 1

den streng monotonen Zugehörigkeitsfunktionen der Ausgabe, das ANFIS dessen Schema in Abbildung 6 dargestellt ist.

Wird ein Mamdani-Regelsystem [MA75] als Ausgangspunkt für die Konstruktion eines ANFIS benutzt, so müssen gewisse Einschränkungen beachtet werden. Die 'Defuzzifikation' der Typ-2-Regelsysteme ist auf eine diskrete Version der Flächen-Schwerpunkt-Methode festgelegt, die eine approximative Näherung des Zentrums ergibt. Diese Festlegung verkompliziert die Konstruktion des ANFIS erheblich [Jan93], weshalb die Typ-2-Regelsysteme nicht als Grundlage benutzt werden.

Das ANFIS wird mit der Methode des Gradientenabstiegs trainiert, wobei es zwei verschiedene Möglichkeiten gibt. Die eine Möglichkeit ist ein reines Backpropagation-Training, bei dem der Fehler zurück durch das Netz gereicht wird, bis zu der Stelle an der das Training angreifen soll. Die zweite Möglichkeit ist ein hybrides Training, das eine Kombination von Backpropagation und der Methode der kleinsten Quadrate ist. Diese hybride Methode ist nur auf ein ANFIS der Typ-3-Regelsysteme anwendbar.

Die Konstruktionsmethode, die ein ANFIS bietet, resultiert in neuronalen Fuzzy Netzen, die schnell konvergierend trainiert werden können. Mit einer Möglichkeit der Regelextraktion aus den Trainingsdaten, bildet die ANFIS-Methode eine gute Möglichkeit der automatisierten Erstellung von qualitätsanalytischen unscharfen Regelsystemen.

---

## 4 Qualitätsanalyse mit Fuzzy Logik

Wird eine algorithmische Erkennung von Kontexten aufgrund sensorischer Daten ausgeführt, die nicht mit unscharfen Mengen operiert, so können nur scharfe Zustände erkannt werden. Diese scharfen Zustände werden mit Klassenidentifikatoren assoziiert, welche kommuniziert und abgeleitet zu weiteren Zuständen führen können. Diese weitere Verarbeitung durch Kommunikation und Schlussfolgerung hat als Grundlage nur den Klassenidentifikator zur Verfügung. Ein Algorithmus, der mit Daten von Sensoren arbeitet, hat aber meist einen Fehler in der Erkennung eines Zustandes, der entweder schon in den Daten vorhanden ist oder durch eine nicht eindeutige Abbildung des kontextuellen Algorithmus verursacht wird. Könnte ein solcher Fehler gemessen und dem jeweiligen kontextuellen Zustand zugeordnet werden, so könnten Algorithmen der weiteren Verarbeitung diese zusätzliche Information als ergänzende Entscheidungsgrundlage nutzen.

Die Theorie der unscharfen Mengen kennt beliebig viele Zwischenstufen der Zugehörigkeit zu Mengen. Werden diese Mengen als kontextuelle Zustände angesehen, so kann eine wesentlich genauere Aussage über die Zugehörigkeit zu einem kontextuellen Klassenidentifikator gemacht werden. Ein kontextueller Zustand wird dann nicht mehr nur erkannt oder nicht erkannt, sondern wird zu einem gewissen Grad zugeordnet. Dieser Grad kann interpretiert Auskunft über den Fehler geben, der in der Erkennung des kontextuellen Zustandes gemacht wurde.

Es existieren viele Systeme die kontextuelle Zustände erkennen, diese auch kommunizieren und zu weiteren Zuständen ableiten. Auch Systeme die mit unscharfen Mengen operieren, finden in der Erkennung von Kontexten eine Anwendung. In dieser Arbeit wird nun eine Methodik vorgestellt die versucht diese beiden Ansätze zu kombinieren. Zum einen soll ein bestehendes System, das einen kontextuellen Zustand wahrnehmen kann, nicht verändert werden, da dieses meist für die gedachte Anwendung optimiert wurde. Zum anderen ist eine Beschränkung des Entwicklers auf den ausschließlichen Umgang mit unscharfen Mengen zu restriktiv.

Diese Methodik versucht nun eine Kategorisierung bezüglich der Qualität der Kontexte, die von anderen Systemen erkannt wurden, durch Analyse der Eingabedaten und des zugehörigen kontextuellen Zustandes zu erreichen. Da diese Methodik also keinerlei Wissen über die inneren Abläufe des Algorithmus haben kann, musste eine vom Entwickler gesteuerte Adaption an die kontextuellen Systeme gefunden werden. Da unscharfe Logiken in ihren graduellen Zugehörigkeiten eine Art der Qualität beinhalten, erfüllen sie den Anspruch der Repräsentation einer Qualität des Kontextes. Durch verschiedene adaptive Systeme können bestimmte unscharfe Logiken automatisiert an die bestehenden kontextuellen Systeme angepasst werden. Mögliche Algorithmen die dies für unscharfe Logiken oder auch unabhängig von Logiken als Klassifikatoren leisten, wurden in Kapitel 3 vorgestellt.

Mit den nun hergeleiteten Randbedingungen und Ansprüchen kann das System, das die Qualitätsanalyse übernimmt und Gegenstand dieser Arbeit ist, definiert werden. Der grundsätzliche Aufbau und die Verbindung mit bestehenden Systemen zur Kontexterkenennung ist in Abbildung 7 dargestellt. Das qualitative

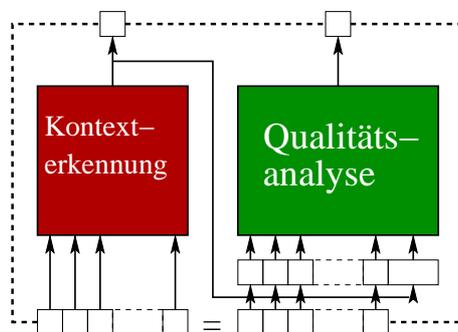


Abbildung 7: Kontexterkenennung und Qualitätsanalyse gekoppelt im Gesamtsystem

System ist ein Sugeno-Regelsystem das als neuronales Netz interpretierbar ist. Die Spezifikation dieses unscharfen Regelsystems, das die Abbildung von sensorischen Eingabedaten und kontextuellen Zustand auf eine qualitative Klasse übernimmt, ist im ersten Abschnitt in diesem Kapitel beschrieben.

Der zweite Abschnitt beschäftigt sich mit der Anpassung des Sugeno-Regelsystems aus dem ersten Abschnitt an das jeweilig gegebene System der Kontexterkenkung mittels adaptiven Algorithmen. Einige dieser adaptiven Algorithmen sind selbst-organisierend, andere benötigen eine Menge, die die gewünschte absolute Qualität pro Eingabedatum enthält, als Referenz. Diese Menge der gewünschten Qualität des endgültigen Sugeno-Regelsystems, das man nach der Anpassung durch die adaptiven Algorithmen erhalten sollte, ist absolut, enthält also nur die Werte für die schlechteste und die beste Qualität.

In der bisherigen Arbeit wurden sämtliche Annahmen theoretisch analysiert und mathematisch ergründet. Im dritten Abschnitt wird nun ein bestehendes System beschrieben, das verschiedene Algorithmen zur Erkennung von kontextuellen Zuständen verwendet, um eine praktische Grundlage zur Evaluierung des Systems der Qualitätsanalyse zu erhalten. Die kontextuellen Algorithmen sind eine unscharfe Logik (Kapitel 4.3.1) und ein neuronales Netz (Kapitel 4.3.2), welche oberflächlich in ihrem Aufbau beschrieben werden.

Anhand dieses Systemes, das Sift-Artefakt genannt wird, können nun im letzten Abschnitt in diesem Kapitel die Ausgaben der Sugeno-Regelsysteme analysiert werden, die jeweils an die zwei kontextuellen Algorithmen angepasst wurden. Es wird erörtert, ob die theoretischen Annahmen praktisch nachgewiesen wurden (Kapitel 4.4.1), welche kontextuellen Zustände mit welchen sensorischen Daten zu guten oder schlechten Qualitätsmaßen führen (Kapitel 4.4.2 und Kapitel 4.4.3) und wie kontextuelle Zwischenergebnisse mit der Qualitätsanalyse korrelieren (Kapitel 4.4.4 und Kapitel 4.4.5). Es werden auch die Auswirkungen der Trainingsdaten (Kapitel 4.4.6) auf das resultierende System der Qualitätsanalyse statistisch analysiert.

### 4.1 Spezifikation der unscharfen Logik $S_Q$ zur Beurteilung von kontextuellen Systemen $A_C$

Zur qualitativen Analyse der kontextuellen Werte, die durch einen beliebigen Algorithmus zur Bestimmung von kontextuellen Zuständen berechnet werden, wird ein unscharfes Regelsystem eingesetzt. Dieses unscharfe Regelsystem ist vom Typ der TSK-Regelsysteme, die von Takagi, Sugeno und Kang [TS85][SK88] entwickelt wurden. In dem Rahmen des in MATLAB verfügbaren ANFIS-Moduls ist dieses TSK-Modell als Sugeno-(Regel)System referenziert, welche Bezeichnung auch in dieser Arbeit verwendet wird. Die Analyse der berechneten kontextuellen Zustände erfolgt auf Basis der Eingabewerte, die auch zur Bestimmung der kontextuellen Zustände benutzt werden, und des Identifikators, der den kontextuellen Zustand identifiziert. Die Berechnungen eines solchen Sugeno-Systems sind Werte einer linearen Menge, deren Grenzen nicht konkret zu bestimmen sind, da die Abbildung des Regelsystems bei gewissen untypischen Eingabewerten den gewünschten Wertebereich verlässt. Aus diesem Grund wird zu dem Sugeno-Regelsystem eine normalisierende Funktion definiert, die die Werte, die einen gewünschten Wertebereich verlassen, wieder auf diesen zurück projiziert.

Das Sugeno-System zur Qualitätsanalyse, wird mit der Bezeichnung  $\hat{S}_{Q_{lin}}$  identifiziert und ist in der Abbildung 8 auf der rechten Seite in der Mitte dargestellt. Dort ist auch zu sehen, wie das Sugeno-System  $\hat{S}_{Q_{lin}}$  mit einem beliebigen algorithmischen System  $A_C$  zur kontextuellen Klassifikation verbunden ist.

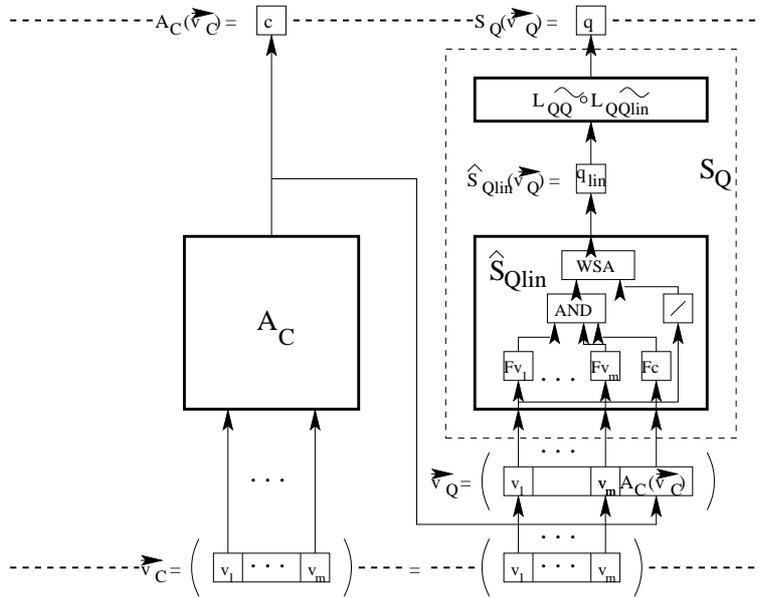


Abbildung 8: Zusammenhänge von kontextuellem Algorithmus  $A_C$  und Spezifikationen der qualitativen Analyse  $S_Q$  innerhalb der Kontext-Schicht des kontextsensitiven Modells

Die Bezeichnung  $\hat{S}_{Q_{lin}}$  des Sugeno-Systems bezieht sich auf die Menge der Resultate  $Q_{lin}$ , die eine Erweiterung der gewünschten Menge  $Q = [0, 1]$  ist, und die Verwendung eines Sugeno-Systems, das eine Instanz des gesamten Systems  $S_Q$  der Qualitätsanalyse ist. Es wird eine Beziehung von Teilmengen spezifiziert, die eine untere Schranke  $Q = [0, 1]$  der Mindestgröße der Menge  $Q_{lin}$  und eine obere Schranke  $\mathbb{R}$  der Maximalgröße verwendet, was die Teilmengenbeziehung  $Q \subset Q_{lin} \subset \mathbb{R}$  ergibt.

Das System  $A_C$  zur Bestimmung der kontextuellen Zustände ist nicht bekannt, weshalb das syntaktische Symbol  $A$  einen allgemeinen Algorithmus symbolisiert. Die Menge  $C$  spezifiziert den resultierenden Wertebereich, auf den durch die Abbildung  $A_C$  des kontextuellen Algorithmus abgebildet wird. In der Menge  $C$  sind die Identifikatoren der Kontextklassen enthalten, die die jeweiligen kontextuellen Zustände identifizieren. Eine allgemeine Belegung der Menge  $C$  und die Spezifikation der Abbildung  $A$  sind wie folgt definiert:

$$A_C : \begin{cases} \mathcal{V}_1 \times \dots \times \mathcal{V}_M & \longrightarrow C, \text{ mit } C = \{c_1, \dots, c_N\} \text{ und } N \geq 2 \\ (v_1, v_2, \dots, v_m) & \longmapsto A_C(v_1, v_2, \dots, v_m) \end{cases} \quad (\text{Definition 4.1.0.1})$$

Die Anzahl der Elemente der Menge  $C$  wird größer oder gleich Zwei angenommen, da sonst eine kontextuelle Klassifikation trivial wäre.

Die auf das Intervall  $[0, 1]$  normalisierende Funktion wird mit  $\mathbf{L}_{QQ_{lin}}$  bezeichnet, da sie die Werte der Menge  $Q_{lin}$  auf die Menge  $Q = [0, 1]$  der qualitativen Werte abbildet. Die Abbildung  $\mathbf{L}_{QQ_{lin}}$  wird in zwei Teile aufgespalten, eine Funktion  $\mathbf{L}_{\tilde{Q}Q_{lin}}$  und eine Funktion  $\mathbf{L}_{Q\tilde{Q}}$ . Das syntaktische Symbol  $\mathbf{L}$  steht hier jeweils für eine funktionale Interpretation von unscharfer Logik und Prädikatenlogik. Wird anstatt einer Normalisierung auf das kontinuierliche Intervall  $[0, 1]$  eine Abbildung auf eine Menge  $Q_K$  von Qualitätsklassen bevorzugt, so wird die Funktion  $\mathbf{L}_{Q_KQ_{lin}}$  verwendet. Die Spezifikationen der Funktionen  $\mathbf{L}_{QQ_{lin}}$  und  $\mathbf{L}_{Q_KQ_{lin}}$  und eine ausführliche Erläuterung sind in Kapitel 4.1.5 gegeben.

Die Zusammenfassung des Sugeno-Systems  $\widehat{\mathbf{S}}_{Q_{lin}}$  und der normalisierenden Funktion  $\mathbf{L}_{QQ_{lin}}$  ist durch folgende Abbildung  $\mathbf{S}_Q$  definiert:

$$\mathbf{S}_Q : \begin{cases} \mathcal{V}_1 \times \dots \times \mathcal{V}_M \times C & \longrightarrow Q, \text{ mit } Q = [q_1, \dots, q_P] \\ (v_1, v_2, \dots, v_m, c) & \longmapsto \mathbf{S}_Q(v_1, v_2, \dots, v_m, c) \end{cases} \quad (\text{Definition 4.1.0.2})$$

Diese Kombination ist in Abbildung 8 einzusehen, wobei die einzelnen Bestandteile durch eine gestrichelte Linie zusammengefasst und mit  $\mathbf{S}_Q$  gekennzeichnet sind.

Die Eingaben der Systeme zur kontextuellen Klassifikation  $\mathbf{A}_C$  und der qualitativen Analyse  $\mathbf{S}_Q$  sind jeweils Vektoren. Der Vektor  $\vec{\mathbf{v}}_C \in \mathcal{V}_1 \times \dots \times \mathcal{V}_M \subset \mathbb{R}^M$  (Definition 4.1.0.3) ist die Grundlage für die

$$\vec{\mathbf{v}}_C := (v_1, v_2, \dots, v_m) \quad (\text{Definition 4.1.0.3})$$

Entscheidung von  $\mathbf{A}_C$  für einen kontextuellen Zustand  $c \in C$ . Dieser Vektor  $\vec{\mathbf{v}}_C$  wird um den Identifikator  $c$  erweitert, was im nächsten Abschnitt beschrieben wird, und bildet die Eingabe  $\vec{\mathbf{v}}_Q$  des qualitativen Systems  $\mathbf{S}_Q$ .

In den folgenden Abschnitten wird nun zunächst das Sugeno-Regelsystem zur Qualitätsanalyse definiert. Im letzten Abschnitt in diesem Kapitel wird die Normalisierung auf das Intervall  $[0, 1]$  spezifiziert und eine klassifizierende Alternative beschrieben.

#### 4.1.1 Erweiterung des Eingangs-Vektors $\mathbf{v}_C$ um kontextuellen Zustand $\mathbf{A}_C(\mathbf{v}_C)$

Der Eingangsvektor der Sugeno-Funktion  $\widehat{\mathbf{S}}_{Q_{lin}}$  unterscheidet sich um eine Stelle vom Eingangsvektor der kontextuellen Klassifizierung  $\mathbf{A}_C$ . Der Vektor  $\vec{\mathbf{v}}_Q$ , aufgrund dessen die Entscheidung der Qualitätsklasse getroffen wird, setzt sich aus dem Vektor  $\vec{\mathbf{v}}_C$ , der die Kontextklasse bedingt, und der Klasse selbst zusammen. Die Kontextklasse wird durch die eindeutige Abbildung von kontextuellem Eingangsvektor und Spezifikation der kontextuellen klassifizierenden Funktion  $\mathbf{A}_C$  identifiziert. Somit ergibt sich der  $(M+1)$ -dimensionale Vector  $\vec{\mathbf{v}}_Q \in \mathcal{V}_1 \times \dots \times \mathcal{V}_M \times C \subset \mathbb{R}^{M+1}$ , der in Definition 4.1.1.4 und in Abbildung 8 dargestellt ist.

$$\begin{aligned} \vec{\mathbf{v}}_Q &:= (v_1, v_2, \dots, v_m, \mathbf{A}_C(v_1, v_2, \dots, v_m)) \\ &= (\vec{\mathbf{v}}_C, \mathbf{A}_C(\vec{\mathbf{v}}_C)) \end{aligned} \quad (\text{Definition 4.1.1.4})$$

Dieser Vektor  $\vec{\mathbf{v}}_Q$  ist der eigentliche Eingabevektor des Sugeno-Regelsystems  $\widehat{\mathbf{S}}_{Q_{lin}}$ . Er bedingt nicht nur die Ausgabe der Qualitätsanalyse, sondern bestimmt mittels den adaptiven Algorithmen, die in Kapitel 4.2 vorgestellt werden, auch den inneren Aufbau der Abbildung  $\vec{\mathbf{v}}_Q$ .

Die Universen der Eingabe  $\mathcal{V}_1, \dots, \mathcal{V}_M$  enthalten die die möglichen Eingabewerte  $v_1, \dots, v_m$ , die eine Herkunft aus Sensoren besitzen. Die Sensorwerte werden für gewöhnlich zuerst durch eine Vorverarbeitung zusammengefasst, zwei Beispiele hierfür findet man in Kapitel 4.3.

#### 4.1.2 Definition der funktionalen Zugehörigkeit $F_{ij}$

Die Qualitätsanalyse wird über ein Sugeno-System  $\widehat{\mathbf{S}}_{Q_{lin}}$  durchgeführt, welches eine besondere Variante eines unscharfen Regelsystems ist. Jede unscharfe Menge  $\tilde{A} = \{(x, \mu_{\tilde{A}}(x)) : x \in \mathcal{U}\}$  wird durch ein Tupel definiert, das sich aus Elementen  $x$  eines Universums  $\mathcal{U}$  und einem Grad der Zugehörigkeit  $\mu_{\tilde{A}}(x)$  zusammen setzt. Der Grad der Zugehörigkeit  $\mu_{\tilde{A}}(x)$  zur unscharfen Menge  $\tilde{A}$  ist eine Funktion ( $\mu_{\tilde{A}} : \mathcal{U} \rightarrow [0, 1]$ )

die auf die scharfe Menge  $[0, 1]$  abbildet. Das 'Unschärfen' eines Elementes  $x$  aus  $\mathcal{U}$  geschieht also durch Ergänzung des Elementes  $x$  um dessen Grad der Zugehörigkeit  $\mu_{\tilde{A}}(x)$  zur jeweiligen unscharfen Menge  $\tilde{A}$ , was zu dem identifizierenden Tupel  $(x, \mu_{\tilde{A}}(x))$  führt.

Das Prinzip der Sugeno-Regelsysteme beruht auf dieser funktionalen Zugehörigkeit. Im Gegensatz zu gewöhnlichen unscharfen Logiken ist aber die Implikation eine grundsätzlich andere, denn die Konsequenz (Sukzedenz) ist hier nicht eine bestimmt verknüpfte Anzahl von unscharfen Mengen, sondern eine Konstante oder eine lineare Funktion. Um Verwechslungen vorzubeugen wird deshalb bei der Definition des Sugeno-Regelsystems  $\hat{S}_{Q_{lin}}$  und anderen Sugeno-Systemen, die in dieser Arbeit verwendet werden, nicht das übliche Funktionssymbol der Zugehörigkeit  $\mu_{\tilde{A}}$  verwendet, sondern ein unabhängiges Funktionssymbol  $F$ .

Der Grad der Zugehörigkeit zu dem Sugeno-Regelsystem zur Qualitätsanalyse  $\hat{S}_{Q_{lin}}$  wird durch eine Gauss-Funktion (Definition 4.1.2.6) gemessen.

Die Gauss-Funktion wird erhalten, durch die Streichung des normalisierenden Faktors einer Dichtefunktion  $P(x)$  (Gleichung 4.1.2.1) der Normalverteilung - die klassische Bedingung, des Flächeninhaltes der Kurve von  $\int_{-\infty}^{\infty} P(x)dx = 1$ , die in der Stochastik gebraucht wird, trifft bei der hier gebrauchten Gauss-Funktion nicht zu - die durch den Mittelwert  $\mu$ , der Wert für  $x$  in dem sie den maximalen Funktionswert annimmt, und die Varianz  $\sigma^2$ , die Streuung der Funktion, definiert ist.

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{(2\sigma^2)}} \quad (\text{Gleichung 4.1.2.1})$$

Die Gauss-Funktion, die hier verwendet wird, ist in ihrem maximalen Wert beschränkt und in ihrem Mittelwert  $\mu$  auf Eins normiert ( $F(\mu) = 1$ ).

Jedes Element des Eingangsvektors  $\vec{v}_Q$  des Sugeno-Systems  $\hat{S}_{Q_{lin}}$  bedingt mehrere Zugehörigkeitsfunktionen. Die Anzahl  $o$  der Zugehörigkeitsfunktionen pro Element  $v_i$  des Vectors  $\vec{v}_Q$  und die jeweiligen Varianzen  $\sigma_{ij}^2$  und Mittelwerte  $\mu_{ij}$ , werden durch ein subtraktives Clustering bestimmt. Der Ablauf dieses subtraktiven Clusterings und die Parametrisierung des Verfahrens werden in Kapitel 4.2.1 erläutert.

Die jeweiligen Gauss-Funktionen haben somit eine eindeutige Assoziierung mit dem jeweiligen Element  $v_i$  des Eingangsvektors  $\vec{v}_Q$ , weshalb ihr Index  $i$  diesen identifiziert (Definition 4.1.2.6).

$$F_{ij}(v_i) := e^{-\frac{(v_i - \mu_{ij})^2}{(2\sigma_{ij}^2)}} \quad (\text{Definition 4.1.2.6})$$

Alle Zugehörigkeitsfunktionen des Systems  $\hat{S}_{Q_{lin}}$  sind in einer Matrix (Definition 4.1.2.7), die das System

$$F := \begin{pmatrix} F_{11} & \cdots & F_{1o} \\ \vdots & \ddots & \vdots \\ F_{m1} & \cdots & F_{mo} \\ F_{c1} & \cdots & F_{co} \end{pmatrix} \quad (\text{Definition 4.1.2.7})$$

zu einer Hälfte spezifiziert, zusammengefasst. Jeder Zeilenvektor der Matrix  $F$  ist mit einem anderen Element des Eingangsvektors  $\vec{v}_Q$  assoziiert, welches nochmals explizit in Definition 4.1.2.8 und Definition 4.1.2.9 dargestellt ist.

$$\vec{F}_i(v_i) := (F_{i1}(v_i), \dots, F_{io}(v_i)) \quad (\text{Definition 4.1.2.8})$$

$$\vec{F}_c(\mathbf{A}_C(v_1, v_2, \dots, v_m)) := (F_{c1}(\mathbf{A}_C(\vec{v}_C)), \dots, F_{co}(\mathbf{A}_C(\vec{v}_C))) \quad (\text{Definition 4.1.2.9})$$

Die Zugehörigkeitsfunktionen bilden die Gewichte der gewichten normalisierten Summe die das Resultat der Sugeno-Qualitätsanalyse  $\hat{S}_{Q_{lin}}$  ergibt, was detailliert in Kapitel 4.1.4 beschrieben wird.

### 4.1.3 Lineare funktionale Konsequenz $f_i$

In der klassischen Logik ist die scharfe Implikation eine Funktion ( $\rightarrow: \{0, 1\}^2 \rightarrow \{0, 1\}$ ) die aus einer Menge von Wahrheitswerten wieder auf eine Menge von Wahrheitswerten abbildet. Im unscharfen Fall muss nicht nur die Implikation sondern auch das ganze Operatorensystem  $\{\wedge, \vee, \neg\}$  erweitert werden. So entspricht die  $T$ -Norm dem klassischen  $\wedge$ , die  $S$ -Norm dem  $\vee$  und die unscharfe Negation  $N$  der scharfen Negation  $\neg$ . Das Operatorensystem im unscharfen Fall ist jedoch nicht eindeutig. Es werden ausgehend von einer bestimmten  $T$ -Norm eine zugehörige Negation  $N$  und  $S$ -Norm definiert. In dem hier

verwendeten Sugeno-Regelsystem wird die Produktnorm  $T_P(x, y) = xy$  verwendet, die mit der unscharfen Standardnegation  $N(x) = 1 - x$  die probabilistische Summe  $S_P(x, y) = x + y - xy$  bedingt. Über die verschiedenen scharfen Definitionen wird dann auch im unscharfen Fall die Implikation hergeleitet. Wird die scharfe Implikation über 'nicht-oder' als Ausgangspunkt der unscharfen Implikation gewählt, so ergibt sich mit dem unscharfen Operatorensystem  $\{T_P, S_P, N\}$  die Implikation  $I_P^S(x, y) = S_P(N(x), y) = 1 - x + xy$ .

Das Sugeno-Regelsystem ist hier eine Ausnahme. Es verwendet eine spezielle Implikation, die nicht von den unscharfen Wahrheitswerten abhängig ist und deren Konsequenz keine unscharfe Menge ist. Die Konsequenz der Implikation ist eine Konstante oder eine lineare Funktion, die von den ursprünglichen Elementen der Universen abhängig ist. Die lineare funktionale Konsequenz, die hier verwendet wird ist also zunächst nicht von den Zugehörigkeitsfunktionen  $F_{ij}(v_i)$  abhängig, sondern von dem eigentlichen Eingangswert  $v_i$ . Somit wird die spezielle Sugeno-Implikation  $I_S$ , die in dem Sugeno-System zur Qualitätsanalyse  $\widehat{\mathbf{S}}_{Q_{lin}}$  verwendet wird, über die Definition 4.1.3.10 spezifiziert. In dieser speziellen Sugeno-Implikation

$$\text{if } F_{1j}(v_1) \text{ and } \dots \text{ and } F_{mj}(v_m) \text{ and } F_{cj}(\mathbf{A}_C(\vec{\mathbf{v}}_C)) \text{ then } y = f_j(\vec{\mathbf{v}}_Q), \text{ mit } i=1, \dots, o$$

$$I_S(T_P(F_{1j}(v_1), \dots, F_{mj}(v_m), F_{cj}(\mathbf{A}_C(\vec{\mathbf{v}}_C))), f_j(\vec{\mathbf{v}}_Q)) \quad (\text{Definition 4.1.3.10})$$

wird die Antezendenz  $T_P(F_{1j}(v_1), \dots, F_{mj}(v_m), F_{cj}(\mathbf{A}_C(\vec{\mathbf{v}}_C)))$  der unscharfen Regel zum Gewicht, das mit der linearen Konsequenz  $f_j(\vec{\mathbf{v}}_Q)$  multipliziert wird. Dieser Sachverhalt führt über alle Regeln zur Ausgabe des Sugeno-Systems  $\widehat{\mathbf{S}}_{Q_{lin}}$  deren Berechnung in Kapitel 4.1.4 erläutert wird. Die linearen Funktionen  $f_j$  (Definition 4.1.3.11) sind eindeutig mit jeder Implikation, also jeder Regel des Sugeno-Regelsystems, verbunden. Die Funktionen  $f_j$  der Konsequenzen bilden aus den Universen der Eingänge auf eine

$$f_j(\vec{\mathbf{v}}_Q) := a_{1j}v_1 + a_{2j}v_2 + \dots + a_{mj}v_m + a_{(m+1)j}\mathbf{A}_C(\vec{\mathbf{v}}_C) + a_{(m+2)j} \quad (\text{Definition 4.1.3.11})$$

Teilmenge der reellen Zahlen ab, was in Definition 4.1.3.12 spezifiziert ist. Die einzelnen Konstanten  $a_{ij}$  der linearen Konsequenz  $f_j$  - mit  $j \in \{1, \dots, o\}$  und  $o$  Anzahl der Regeln des Systems  $\widehat{\mathbf{S}}_{Q_{lin}}$  - werden initial durch eine Methode der kleinsten Quadrate (Kapitel 4.2.2) bestimmt. Das Training und der Aufbau des neuronalen Fuzzy Netzes das die Verfeinerung der initialen linearen Gewichte  $a_{ij}$  von  $f_j$  ermöglicht,

$$f_j : \begin{cases} \mathcal{V}_1 \times \dots \times \mathcal{V}_M \times C & \longrightarrow Q_{lin}, \text{ mit } [0, 1] \subset Q_{lin} \subset \mathbb{R} \\ (v_1, v_2, \dots, v_m, c) & \longmapsto f_j(v_1, v_2, \dots, v_m, c) \end{cases} \quad (\text{Definition 4.1.3.12})$$

wird in Kapitel 4.2 hergeleitet. Die Konstanten  $a_{ij}$ , die wieder eindeutig mit jedem Element  $v_i$  des Eingangsvektors  $\vec{\mathbf{v}}_Q$  und der implizierten Konsequenz  $f_j$  korreliert sind, werden in einer Matrix  $A$  (Definition 4.1.3.13) aufgelistet, da eine Vektor-Matrix-Multiplikation die jeweiligen linearen Funktionen  $f_j$  der

$$A := \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1o} \\ a_{21} & a_{22} & \dots & a_{2o} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mo} \\ a_{(m+1)1} & a_{(m+1)2} & \dots & a_{(m+1)o} \\ a_{(m+2)1} & a_{(m+2)2} & \dots & a_{(m+2)o} \end{pmatrix} \quad (\text{Definition 4.1.3.13})$$

Konsequenzen erzeugt. Das Ergebnis dieser Multiplikation (Definition 4.1.3.14) ist ein Vector  $\vec{\mathbf{f}}(\vec{\mathbf{v}}_Q)$ , welcher abhängig vom jeweiligen Eingangsvektor  $\vec{\mathbf{v}}_Q$  ist, der die einzelnen Funktionswerte  $f_j(\vec{\mathbf{v}}_Q)$  enthält. Da der letzte Zeilenvektor der Matrix  $A$  die jeweiligen Konstanten enthält, die kein multiplikatives Äquivalent im qualitativen Eingangsvektor  $\vec{\mathbf{v}}_Q$  haben, muss der Vektor um das neutrale Element der Multiplikation erweitert werden (Definition 4.1.3.14), um die Vektor-Matrix-Multiplikation überhaupt

$$\begin{aligned} \vec{\mathbf{f}}(\vec{\mathbf{v}}_Q) &:= (\vec{\mathbf{v}}_Q, 1)A \\ &= (f_1(\vec{\mathbf{v}}_Q), \dots, f_o(\vec{\mathbf{v}}_Q)) \end{aligned} \quad (\text{Definition 4.1.3.14})$$

erst zu ermöglichen. Die somit berechneten Konsequenz-Funktionswerte  $f_j(\vec{\mathbf{v}}_Q)$  werden in einer gewichteten normalisierten Summe mit den Gewichten multipliziert und normalisiert. Diese Summe bildet das Resultat des qualitativen Sugeno-Regelsystems  $\widehat{\mathbf{S}}_{Q_{lin}}$ .

#### 4.1.4 Gewichtete Summe der Regeln

Die Regeln einer unscharfen Logik implizieren eine unscharfe Größe, die im Wertebereich  $[0, 1]$  liegt. Der unscharfe Schluss ist, wie zuvor die unscharfen Operatoren aus den scharfen, aus dem scharfen Schluss, dem Modus Ponens, erweitert. Im Modus Ponens ist mit einer Regel und einem Fakt ein Schluss gegeben:

$$\frac{\text{IF } a \text{ THEN } b \quad (\text{Regel})}{\frac{a \quad (\text{Fakt})}{b} \quad (\text{Schluss})}$$

Im unscharfen Fall gibt es wieder verschiedene Möglichkeiten des Schließens, wie das approximative, das possibilistische und das evidenzgestützte Schließen. Im intuitiv verständlichsten Fall des approximativen Schließens wird der unscharfe Modus Ponens linguistisch wie folgt formuliert:

$$\frac{\text{IF } x \text{ is } \tilde{A} \text{ THEN } y \text{ is } \tilde{B} \quad (\text{Regel})}{\frac{x \text{ is } \tilde{A}' \quad (\text{Fakt})}{y \text{ is } \tilde{B}'} \quad (\text{Schluss})}$$

Dieser unscharfe Schluss, der durch eine unscharfe Menge  $\tilde{B}'$  beschrieben wird, wird dann meist wieder zurück in eine scharfe Aussage transferiert, was durch ein Defuzzifizierungsverfahren erreicht wird. Ein Beispiel einer Defuzzifizierungsmethode ist die Maximum Methode:

$$y_{sup}^* = \arg \sup_y \mu_{\tilde{B}'}(y)$$

Die scharfe Aussage ist nun wieder ein Element  $y^*$  des ursprünglichen Universums, ohne die Dualität mit dem Zugehörigkeitsgrad in einem unscharfen Tupel zu besitzen. Der Wert  $y^*$  wird durch das gewählte Defuzzifizierungsverfahren, über den Grad der Zugehörigkeit zur unscharfen Menge  $\tilde{B}'$ , berechnet.

In dem verwendeten Sugeno-Regelsystem  $\hat{S}_{Q_{lin}}$  wird der Zwischenschritt des unscharfen Schlusses ausgelassen und gleich eine Defuzzifizierung vorgenommen. Aus diesem Grund bildet auch die Funktion  $\hat{S}_{Q_{lin}}$  auf eine scharfe Menge  $Q_{lin}$  ab.

Zuerst werden über die einzelnen Zugehörigkeitsfunktionen  $F_{ij}$  Gewichte berechnet. Hierbei definiert wieder die Regel die Bildung des jeweiligen Gewichts  $w_j(\vec{\nu}_Q)$  (Definition 4.1.4.15), das vom Eingangsvektor  $\vec{\nu}_Q$  abhängig ist. Das Regelsystem  $\hat{S}_{Q_{lin}}$  beinhaltet für jede Zugehörigkeitsfunktion eines Eingangs eine Regel. Eine Regel verknüpft die Werte der Zugehörigkeitsfunktionen  $F_{1j}(v_1), \dots, F_{mj}(v_m), F_{cj}(\mathbf{A}_C(\vec{\nu}_C))$  über eine Produktnorm  $T_P$ , was schon zuvor in Definition 4.1.3.10 beschrieben wurde.

$$w_j(\vec{\nu}_Q) := \left( \prod_{i=1}^m F_{ij}(v_i) \right) F_{cj}(\mathbf{A}_C(\vec{\nu}_C)) \quad (\text{Definition 4.1.4.15})$$

Die Anzahl  $o$  der Zugehörigkeitsfunktionen  $F_{ij}(x)$  pro Eingang ist für jedes Eingangselement  $v_i$  gleich und entspricht der Anzahl  $o$  der Regeln und somit auch den linearen Funktionen  $f_j$  der Konsequenzen. Das Resultat des Sugeno-Systems  $\hat{S}_{Q_{lin}}$  wird über eine gewichtete Summe (Definition 4.1.4.16) berechnet, wobei das Gewicht  $w_r(\vec{\nu}_Q)$  mit der linearen Konsequenz  $f_j(\vec{\nu}_Q)$  der Regel  $j$  multipliziert und die Summe über alle  $j$  gebildet wird. Diese gewichtete Summe wird noch normalisiert, indem durch die Summe aller Gewichte  $w_j(\vec{\nu}_Q)$  geteilt wird.

$$\hat{S}_{Q_{lin}}(\vec{\nu}_Q) := \frac{\sum_{j=1}^o w_j(\vec{\nu}_Q) f_j(\vec{\nu}_Q)}{\sum_{j=1}^o w_j(\vec{\nu}_Q)} \quad (\text{Definition 4.1.4.16})$$

Die Menge  $Q_{lin}$ , auf die durch  $\hat{S}_{Q_{lin}}$  abgebildet wird, ist keine Menge, die Werte innerhalb festzulegender Grenzen besitzt. Sie ist größer als das gewünschte Intervall  $[0, 1]$  eines interpretierbaren Qualitätsmaßes ( $[0, 1] \subset Q_{lin}$ ). Um das Resultat nun noch auf dieses Intervall  $[0, 1]$  zu beschränken, muss eine Normalisierung der Werte von  $\hat{S}_{Q_{lin}}(\vec{\nu}_Q)$  vorgenommen werden.

#### 4.1.5 Normalisierung oder Klassifikation der Sugeno-System Abbildung $\hat{S}_{Q_{lin}}$

Das Ergebnis des Sugeno-Regelsystem  $\hat{S}_{Q_{lin}}(\vec{\nu}_Q)$  kann an der oberen Schnittstelle (Abbildung 8, obere gestrichelte Linie) des gesamten Systems der Kontexterkenkung und Qualitätsanalyse, ohne weitere

Klassifizierung an das nächste System weiter gegeben werden, denn die Informationen einer 'Qualität' sind schon in Elementen der Menge  $Q_{lin}$  enthalten. Ein Problem tritt jedoch auf wenn die Projektion von  $\widehat{\mathbf{S}}_{Q_{lin}}$  außerhalb der Menge  $[0, 1]$ , der unteren Beschränkung von  $[0, 1] \subset Q_{lin}$ , liegt. Was ist die qualitative Bedeutung eines Wertes über Eins oder unter Null? Die Interpretation dieser Werte muss noch durch eine Funktion übernommen werden, die Werte aus  $Q_{lin}$ , die nicht in  $[0, 1]$  liegen wieder auf  $[0, 1]$  abbildet. Die formalen Grundlagen, die nun und auch in anderen Kapiteln verwendet werden, entstammen dem Skriptum 'Formale Systeme' von Deussen [Deu02]. Die Elemente der Theorie der unscharfen Mengen entstammen wieder dem Skriptum von Hanebeck [HR05].

Die Abbildung des qualitativen Systems  $\mathbf{S}_Q$ , setzt sich aus dem Sugeno-System  $\widehat{\mathbf{S}}_{Q_{lin}}$ , das auf die Menge  $Q_{lin}$  abbildet, der Funktion  $\mathbf{L}_{\widetilde{Q}Q_{lin}}$ , die aus der Menge  $Q_{lin}$  auf die unscharfe Menge  $\widetilde{Q}$  abbildet, und der Funktion  $\mathbf{L}_{Q\widetilde{Q}}$ , welche wieder aus der unscharfen Menge  $\widetilde{Q}$  auf die scharfe Menge  $Q = [0, 1]$  des resultierenden Intervalls projiziert, wie folgt zusammen:

$$\mathbf{S}_Q : \begin{cases} \mathcal{V}_1 \times \dots \times \mathcal{V}_M \times C & \longrightarrow & Q, \text{ mit } Q = [0, 1] \\ (v_1, v_2, \dots, v_m, c) & \longmapsto & \mathbf{L}_{Q\widetilde{Q}} \circ \mathbf{L}_{\widetilde{Q}Q_{lin}} \circ \widehat{\mathbf{S}}_{Q_{lin}}(v_1, v_2, \dots, v_m, c) \end{cases} \quad (\text{Definition 4.1.5.17})$$

Die Zusammensetzung der einzelnen Funktionen erfolgt über eine Konkatenation der einzelnen Abbildungen, wobei die Hintereinanderausführung der Funktion  $\mathbf{L}_{\widetilde{Q}Q_{lin}}$  und  $\mathbf{L}_{Q\widetilde{Q}}$  die normalisierende Abbildung  $\mathbf{L}_{Q_{lin}}$  ergibt. Die Eingabe ist der qualitative Vektor  $\vec{v}_Q$  der, wie zuvor definiert wurde, die Erweiterung des Eingabevektors  $\vec{v}_C$  um den kontextuellen Identifikator  $c$  ist.

Die Ausgabe des Sugeno-Systems  $\widehat{\mathbf{S}}_{Q_{lin}}$  ist ein Element  $q_{lin}$  der scharfen Menge  $Q_{lin}$ , da der spezielle Schluss, der bei dem Sugeno-System und auch anderen unscharfen Regelsystemen verwendet wird, eine Defuzzifikation beinhaltet. Das Sugeno-System wird jedoch durch die im nächsten Abschnitt beschriebenen adaptiven Algorithmen auf eine gewünschte Ausgabe einer Menge mit zwei Elementen, im hier verwendeten Fall die Menge  $\{0, 1\}$ , hinsichtlich seiner Architektur erstellt und trainiert. Diese gewünschte Abbildung auf die Menge  $\{0, 1\}$ , Null für eine falsche Klassifikation des kontextuellen Algorithmus  $\mathbf{A}_C$  und Eins für eine richtige, ist fehlerbehaftet. Der Fehler des qualitativen Sugeno-Systems  $\widehat{\mathbf{S}}_{Q_{lin}}$  entspricht dem Fehler des kontextuellen Systems  $\mathbf{A}_C$ . Aus diesem Grund wird die Ausgabe des Sugeno-Systems wieder unscharf interpretiert, denn der Zugehörigkeitsgrad zu einer unscharfen Menge  $\widetilde{Q}$  entspricht dem Fehler des Systems. Die Abbildung  $\mathbf{L}_{\widetilde{Q}Q_{lin}}$  auf die unscharfe Menge  $\widetilde{Q}$ , ist wie folgt definiert:

$$\mathbf{L}_{\widetilde{Q}Q_{lin}} : \begin{cases} Q_{lin} & \longrightarrow & \widetilde{Q}, \text{ mit } \widetilde{Q} = \{\widetilde{0}, \widetilde{1}\} \\ q_{lin} & \longmapsto & \mathbf{L}_{\widetilde{Q}Q_{lin}}(q_{lin}) = ((q_{lin}, \mu_{\widetilde{0}}(q_{lin})), (q_{lin}, \mu_{\widetilde{1}}(q_{lin}))) \end{cases} \quad (\text{Definition 4.1.5.18})$$

Die unscharfe Menge  $\widetilde{Q}$ , die wieder den Fehler des Sugeno-Systems  $\widehat{\mathbf{S}}_{Q_{lin}}$  bei der eigentlichen Abbildung auf die Menge  $\{0, 1\}$  repräsentiert, besteht deshalb aus zwei unscharfen Zahlen  $\widetilde{0} = \{(x, \mu_{\widetilde{0}}(x)) : x \in Q_{lin}\}$  und  $\widetilde{1} = \{(x, \mu_{\widetilde{1}}(x)) : x \in Q_{lin}\}$  mit  $\widetilde{Q} = \{\widetilde{0}, \widetilde{1}\}$ . Die beiden Zugehörigkeitsfunktionen  $\mu_{\widetilde{0}}(x)$  (Definition 4.1.5.19) und  $\mu_{\widetilde{1}}(x)$  (Definition 4.1.5.20) werden durch eine LR-Funktion (Definition 4.1.5.21) bestimmt.

$$\mu_{\widetilde{0}}(x) := \begin{cases} L(-x) & \text{für } x < 0 \\ 1 & \text{für } x = 0 \\ R(x) & \text{für } x > 0. \end{cases} \quad (\text{Definition 4.1.5.19})$$

$$\mu_{\widetilde{1}}(x) := \begin{cases} L(1-x) & \text{für } x < 1 \\ 1 & \text{für } x = 1 \\ R(x-1) & \text{für } x > 1. \end{cases} \quad (\text{Definition 4.1.5.20})$$

Die Funktion  $\mathbf{L}_{\widetilde{Q}Q_{lin}}(x) = ((x, \mu_{\widetilde{0}}(x)), (x, \mu_{\widetilde{1}}(x)))$ , die auf die unscharfe Zahlenmenge  $\widetilde{Q}$  abbildet, hat als Rückgabewert zwei Tupel, die die unscharfen Zahlen  $\widetilde{0}$  und  $\widetilde{1}$  bedeuten. Graphisch sind die beiden Dreieckszahlen, deren LR-Funktion in Definition 4.1.5.21 spezifiziert ist, in Abbildung 9 durch ihre Zugehörigkeitsfunktionen dargestellt.

$$L(x) = R(x) = \begin{cases} (1-x) & \text{für } 0 < x < 1 \\ 0 & \text{sonst.} \end{cases} \quad (\text{Definition 4.1.5.21})$$

Um nun eine Projektion der unscharfen Zahlen  $\widetilde{0}$  und  $\widetilde{1}$  auf Elemente der Menge  $Q$  zu erhalten, wurde die Funktion  $\mathbf{L}_{Q\widetilde{Q}}$  (Definition 4.1.5.22) definiert. Sie bildet die Elemente der unscharfen Zahlenmenge  $\widetilde{Q}$

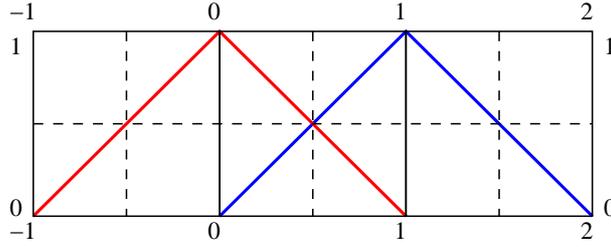


Abbildung 9: Zugehörigkeitsfunktionen  $\mu_{\tilde{0}}$  und  $\mu_{\tilde{1}}$  für unsharp Zahlen  $\tilde{0}$  und  $\tilde{1}$

auf die Menge  $Q = [0, 1]$ , wie folgt ab:

$$\mathbf{L}_{Q\tilde{Q}} : \begin{cases} \tilde{Q} & \longrightarrow Q, \text{ mit } Q = [0, 1] \\ ((q_{lin}, \mu_{\tilde{0}}(q_{lin})), (q_{lin}, \mu_{\tilde{1}}(q_{lin}))) & \longmapsto \mathbf{L}_{Q\tilde{Q}}(((q_{lin}, \mu_{\tilde{0}}(q_{lin})), (q_{lin}, \mu_{\tilde{1}}(q_{lin})))) = q \end{cases} \quad (\text{Definition 4.1.5.22})$$

Die Zuordnung von unsharp Zahl zu einem repräsentierenden Element erfolgt über Interpretation eines Zugehörigkeitsgrades. Um aus zwei unsharp Zahlen und einem Element der Zahlen einen neuen Grad der Zugehörigkeit abzuleiten, ist ein zu interpretierender Term  $t = S_M(N_\gamma(\mu_{\tilde{0}}(x)), \mu_{\tilde{1}}(x))$  formuliert worden. Es wird die Interpretation  $\mathcal{I}$  des Terms  $t$  auf Belegungen  $q_{lin}$  der Menge  $Q_{lin}$ , wie folgt definiert:

$$\mathcal{I}(Q_{lin}, q_{lin})(S_M(N_\gamma(\mu_{\tilde{0}}(q_{lin})), \mu_{\tilde{1}}(q_{lin}))) \in [0, 1] \quad (\text{Definition 4.1.5.23})$$

Diese Interpretation  $\mathcal{I}$  des Terms  $t$  ist kongruent mit der folgenden Definition der Funktion  $\mathbf{L}_{Q\tilde{Q}}$ :

$$\begin{aligned} \mathbf{L}_{Q\tilde{Q}}((q_{lin}, \mu_{\tilde{0}}(q_{lin})), (q_{lin}, \mu_{\tilde{1}}(q_{lin}))) &:= \max(1 - \mu_{\tilde{0}}(q_{lin}), \mu_{\tilde{1}}(q_{lin})) \quad (\text{Definition 4.1.5.24}) \\ &\equiv \mathcal{I}(Q_{lin}, q_{lin})(S_M(N_\gamma(\mu_{\tilde{0}}(x)), \mu_{\tilde{1}}(x))) \end{aligned}$$

Die Beziehung der Kongruenz, die zwischen der Funktion  $\mathbf{L}_{Q\tilde{Q}}$  und der Interpretation  $\mathcal{I}(Q_{lin}, q_{lin})$  besteht, bildet den semantischen Übergang von der unsharp Logik zu einer Abbildung, die in ein qualitatives System  $S_Q$  eingegliedert werden kann.

Die Zuordnung von Elementen aus  $Q_{lin}$  zu Elementen der Menge  $Q = [0, 1]$  ist jedoch semantisch noch nicht korrekt, denn Werte von  $q_{lin} < -0,5$  oder  $q_{lin} > 1,5$  werden durch die Funktion  $\mathbf{L}_{Q\tilde{Q}}$  mehr Eins oder Null zugeordnet. Der maximale Fehler, der durch die Normalisierung noch semantisch korrekt repräsentiert werden kann, ist also für falsche und richtige Klassifikationen ein Wert von 0,5. Es müssen also die Bereiche  $q_{lin} < -0,5$  und  $q_{lin} > 1,5$  durch einen speziellen Term  $e$  abgeschnitten werden, der einen besonderen, nicht durch die Funktion  $\mathbf{L}_{Q\tilde{Q}}$  semantisch korrekt interpretierbaren Fehlerzustand identifiziert. Der Term  $e = N_\gamma(S_M(\mu_{\tilde{0}}(x), \mu_{\tilde{1}}(x)))$  entspricht dem Komplement der unsharp Menge  $\tilde{Q}$ , was die folgende semantische Interpretation ergibt:

$$\mathcal{I}(Q_{lin}, q_{lin})(N_\gamma(S_M(\mu_{\tilde{0}}(x), \mu_{\tilde{1}}(x)))) \in [0, 1] \quad (\text{Definition 4.1.5.25})$$

Der Term  $e$  spezifiziert eine funktional Zugehörigkeit  $\mu_{\tilde{E}}$  zu einer neuen unsharp Menge  $\tilde{E} = \{(x, \mu_{\tilde{E}}) : x \in Q_{lin}\}$ , die die unsharp Repräsentation eines Fehlerzustands  $\epsilon$  ist. Dieser Fehlerzustand  $\epsilon$  ist einziges Element der Menge  $E = \{\epsilon\}$  der Fehlerzustände. Die funktionalen Zugehörigkeiten zu den unsharp Zahlen  $\tilde{0}$  und  $\tilde{1}$ , der unsharp Menge  $\tilde{Q}$ , und zu der unsharp Fehlermenge  $\tilde{E}$ , sind in der Abbildung 10 dargestellt. Wird nun die Funktion  $\mathbf{L}_{Q\tilde{Q}}$  um die Menge  $E$  der Fehlerzustände erweitert, so kann die resultierende Funktion  $\mathbf{L}_{(Q \cup E)\tilde{Q}}$  nicht mehr einfach über eine kongruente Beziehung von unsharp Logik und Funktion definiert werden. Es muss eine Entscheidung getroffen werden, ob die Interpretation eines neuen Terms  $s = S_M(\mu_{\tilde{0}}(x), \mu_{\tilde{1}}(x))$ , der die funktionale Zugehörigkeit zu der unsharp Menge  $\tilde{Q}$  definiert, oder die des Terms  $e$  größer ist. Hierfür wird ein Prädikat  $P_{\leq}$  benötigt, die ein Wahrheitswert für die Interpretationen  $\mathcal{I}$  liefert. Das Prädikat  $P_{\leq}(\mathcal{I}(Q_{lin}, q_{lin})(e), \mathcal{I}(Q_{lin}, q_{lin})(s))$  entscheidet also, ob die Interpretation  $\mathcal{I}$  des Terms  $e$  kleiner oder gleich der Interpretation  $\mathcal{I}$  des Terms  $s$  ist. Wenn diese Fall eintritt, so ist das Prädikat wahr, ansonsten ist es falsch. Die Funktion  $\mathbf{L}_{(Q \cup E)\tilde{Q}}$  ist nun, anhand der

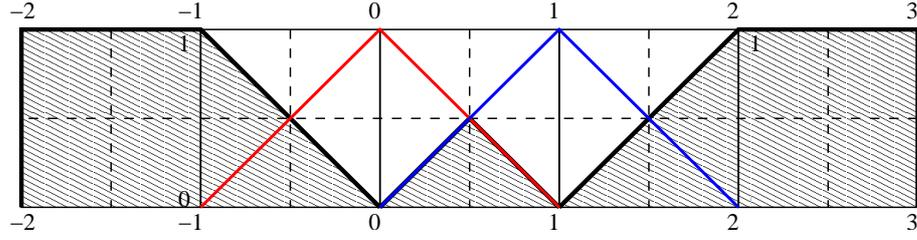


Abbildung 10: Zugehörigkeitsfunktionen  $\mu_{\tilde{0}}$ ,  $\mu_{\tilde{1}}$  und  $\mu_{\tilde{E}}$  für unscharfe Zahlen  $\tilde{0}$ ,  $\tilde{1}$  und unscharfe Menge  $\tilde{E}$

Interpretation  $\mathcal{I}$  und des Prädikates  $P_{\leq}$ , wie folgt definiert:

$$\mathbf{L}_{(Q \cup E)\tilde{Q}}((q_{lin}, \mu_{\tilde{0}}(q_{lin}), (q_{lin}, \mu_{\tilde{1}}(q_{lin}))) := \begin{cases} \max(1 - \mu_{\tilde{0}}(q_{lin}), \mu_{\tilde{1}}(q_{lin})) & \text{wenn } P_{\leq}(\mathcal{I}(Q_{lin}, q_{lin})(e), \\ & \mathcal{I}(Q_{lin}, q_{lin})(s)) = true \\ \epsilon & \text{sonst.} \end{cases} \quad (\text{Definition 4.1.5.26})$$

Es muss nun noch in der gesamten Abbildung des qualitativen Systems  $\mathbf{S}_Q$  die Funktion  $\mathbf{L}_{Q\tilde{Q}}$  mit der Funktion  $\mathbf{L}_{(Q \cup E)\tilde{Q}}$  ersetzt werden, was die folgende mathematische Abbildung ergibt:

$$\mathbf{S}_Q : \begin{cases} \mathcal{V}_1 \times \dots \times \mathcal{V}_M \times C & \longrightarrow & Q \cup E, \text{ mit } Q = [0, 1] \\ (v_1, v_2, \dots, v_m, c) & \longmapsto & \mathbf{L}_{(Q \cup E)\tilde{Q}} \circ \mathbf{L}_{\tilde{Q}Q_{lin}} \circ \hat{\mathbf{S}}_{Q_{lin}}(v_1, v_2, \dots, v_m, c) \end{cases} \quad (\text{Definition 4.1.5.27})$$

Diese Abbildung  $\mathbf{S}_Q$  eines Systems zur Qualitätsanalyse wird in den folgenden Abschnitten verwendet. Wird durch das qualitative System  $\mathbf{S}_Q$  auf den Fehlerzustand  $\epsilon$  projiziert, so wird dieses Datum aus den Analysen entfernt, da diese Kategorisierung keine analytischen Informationen enthält, sondern lediglich nicht durch das System beurteilbar ist.

### Klassifizierung der Qualität

Eine Alternative zur Normalisierung der Ausgabe des Sugeno-Systems  $\hat{\mathbf{S}}_{Q_{lin}}$  ist eine Klassifizierung. Bei der Klassifizierung wird nicht auf eine Menge  $Q = [0, 1]$  abgebildet, sondern auf eine Menge  $Q_K = \{q_1, \dots, q_P\}$  von Identifikatoren von Qualitätsklassen. Diese Methode wird für die Verwendung der Qualitätsanalyse auf limitierten Ressourcen und verteilten Wissensbasen, da ein kontextueller Identifikator weniger Speicherplatz und Aufwand an Kommunikation benötigt, als Elemente der Menge  $Q$ .

Um eine Klassifikation zu erreichen, wird die Funktion  $\mathbf{L}_{(Q \cup E)\tilde{Q}}$  durch die Funktion  $\mathbf{L}_{(Q_K \cup E)\tilde{Q}}$  ersetzt. Die Konkatination der Funktion  $\mathbf{L}_{(Q_K \cup E)\tilde{Q}}$  und der Funktion  $\mathbf{L}_{Q\tilde{Q}}$  ergibt die klassifizierende Funktion  $\mathbf{L}_{Q_K Q_{lin}}$ . Die Abbildung des neuen qualitativen System  $\mathbf{S}_{Q_K}$ , das einen Identifikator einer Klasse als Ausgabe liefert, ist wie folgt definiert:

$$\mathbf{S}_{Q_K} : \begin{cases} \mathcal{V}_1 \times \dots \times \mathcal{V}_M \times C & \longrightarrow & Q_K, \text{ mit } Q_K = \{q_1, \dots, q_P\} \\ (v_1, v_2, \dots, v_m, c) & \longmapsto & \mathbf{L}_{(Q_K \cup E)\tilde{Q}} \circ \mathbf{L}_{\tilde{Q}Q_{lin}} \circ \hat{\mathbf{S}}_{Q_{lin}}(v_1, v_2, \dots, v_m, c) \end{cases} \quad (\text{Definition 4.1.5.28})$$

Um die Tupel der Funktion  $\mathbf{L}_{Q\tilde{Q}}$ , die auch schon bei der Normalisierung verwendet wurde, noch Klassen zuzuordnen wird ein Prädikat definiert, das eine scharfe Entscheidung über Zugehörigkeit zu einer Klasse treffen kann. Die Eingabemenge auf der das Prädikat operiert, ist wieder die Interpretation des Terms  $t$ . Das Prädikat  $P_{\in}(t, [q_i])$  liefert für eine Belegung  $q_{lin}$  den Wahrheitswert  $true$ , wenn  $\mathcal{I}(Q_{lin}, q_{lin})(t) \in [q_i]$ . In diesem Fall ist  $P_{\in}(\mathcal{I}(Q_{lin}, q_{lin})(t), [q_i]) = true$  wenn  $[q_i] < \mathcal{I}(Q_{lin}, q_{lin})(t) \leq [q_i]$ . Damit die Projektion auf die Menge  $Q = \{q_1, \dots, q_P\}$  der Klassenidentifikatoren  $q_i$  eindeutig ist, muss die Menge  $Q$  eine Totalordnung besitzen. Somit werden die Klassengrenzen -  $[q_i]$  untere und  $[q_i]$  obere Klassengrenze der Klasse  $[q_i]$  - nach folgender Ordnung bestimmt:

$$[q_1] < [q_1] = [q_2] < [q_2] = \dots = [q_P] < [q_P]$$

Ist für eine Belegung  $q_{lin}$  die Interpretation  $\mathcal{I}(Q_{lin}, q_{lin})(t)$  nicht Element einer Klasse  $[q_i]$ , für alle  $i \in \{1, 2, \dots, P\}$ , so wird sie der dem Fehlerzustand  $\epsilon$  zugeordnet.

Die Funktion  $\mathbf{L}_{\{Q \cup E\} \tilde{Q}}$  ist mit den Prädikaten  $P_{\in}$  und  $P_{\leq}$ , welches auch wieder auf dem Termen  $s$  und  $e$  operiert, wie folgt definiert:

$$\mathbf{L}_{\{Q \cup E\} \tilde{Q}}((q_{lin}, \mu_0(q_{lin}), (q_{lin}, \mu_1(q_{lin}))) = \begin{cases} q_1 & \text{wenn } P_{\in}(\mathcal{I}(Q_{lin}, q_{lin})(t), [q_1]) \\ & \text{und } P_{\leq}(\mathcal{I}(Q_{lin}, q_{lin})(e), \mathcal{I}(Q_{lin}, q_{lin})(s)) \\ q_2 & \text{wenn } P_{\in}(\mathcal{I}(Q_{lin}, q_{lin})(t), [q_2]) \\ & \text{und } P_{\leq}(\mathcal{I}(Q_{lin}, q_{lin})(e), \mathcal{I}(Q_{lin}, q_{lin})(s)) \\ \vdots & \vdots \\ q_P & \text{wenn } P_{\in}(\mathcal{I}(Q_{lin}, q_{lin})(t), [q_1]) \\ & \text{und } P_{\leq}(\mathcal{I}(Q_{lin}, q_{lin})(e), \mathcal{I}(Q_{lin}, q_{lin})(s)) \\ \epsilon & \text{sonst.} \end{cases}$$

(Definition 4.1.5.29)

Das qualitative System  $S_{Q_K}$  wird in dieser Arbeit nicht verwendet, da hier nur Offline-Analysen gemacht werden, die nicht in ihren Ressourcen limitiert sind und deren Qualitäten auch nicht in einem Netzwerk für verteilte Wissensbasen kommuniziert werden. Ein Beispiel für eine qualitative Klassifikation, für die Klassen  $[0]$ ,  $[0, 1]$ ,  $[0, 2]$ ,  $[0, 3]$ ,  $[0, 4]$ ,  $[0, 5]$ ,  $[0, 6]$ ,  $[0, 7]$ ,  $[0, 8]$ ,  $[0, 9]$  und  $[1]$ , ist in der Abbildung 11 mit dem Fehlerzustand  $\epsilon$  dargestellt.

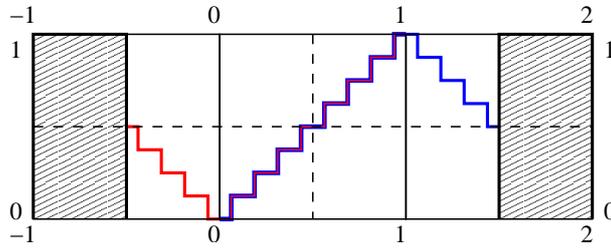


Abbildung 11: Klassifizierte Interpretation der Zugehörigkeitsgrade von Termen  $t$  und  $e$  resultierend in Klassenwerten aus  $Q$  und  $E$

## 4.2 Adaption der Qualität des Kontext Algorithmus durch adaptive Algorithmen

In diesem Kapitel werden die Algorithmen zur automatisierten Spezifikation des Sugeno-Regelsystems  $\hat{\mathbf{S}}_{Q_{lin}}$  beschrieben. Da das Sugeno-System  $\hat{\mathbf{S}}_{Q_{lin}}$  eine kaum durch den Entwickler definierbare Abbildung  $\hat{\mathbf{S}}_{Q_{lin}} : \mathcal{V}_1 \times \dots \times \mathcal{V}_M \times C \rightarrow Q_{lin}$  ist, weil das Design des zu beurteilenden kontextuellen Algorithmus  $\mathbf{A}_C$  nicht bekannt ist, wird nun in diesem Kapitel eine Methode vorgestellt, die zur automatisierten Spezifikation des Systems zur Qualitätsanalyse  $\hat{\mathbf{S}}_{Q_{lin}}$  verwendet wurde.

Das Werkzeug, das zur Spezifizierung des Sugeno-Systems  $\hat{\mathbf{S}}_{Q_{lin}}$  verwendet wird, ist das von MATLAB bereitgestellte ANFIS-Modul. Hierbei steht ANFIS für 'Adaptive-Network-based Fuzzy Inference System' [Jan93], das eine Methodik der Fusion von neuronalen Netzwerken und einer Sugeno Fuzzy-Logik bedeutet. Es werden also Parameter die in einer Sugeno Fuzzy-Logik verwendet werden durch Training eines korrespondierenden neuronalen Netzes adaptiert.

Der erste Schritt zum Erstellen eines neuronalen Netzes ist die Bestimmung des Aufbaus. Die Anzahl der Neuronen bestimmt die Komplexität der resultierenden Logik. Ein selbst-organisierendes Clusterverfahren bestimmt die Architektur des neuronalen Netzes und der initialen Parameter der nichtlinearen Zugehörigkeitsfunktionen  $F_{ij}$ . Im Rahmen des ANFIS-Moduls wird das subtraktive Clustering verwendet.

Die linearen funktionalen Konsequenzen  $f_i$  des Sugeno-Systems  $\hat{\mathbf{S}}_{Q_{lin}}$  können, anhand der zuvor bestimmten initialen Zugehörigkeitsfunktionen, in einem überwachten Verfahren der kleinsten Quadrate bestimmt werden. Die Menge, die zu jedem Datum der Eingabe eine gewünschte Ausgabe enthält, muss vom Entwickler bereitgestellt werden, was die überwachten Verfahren ausmacht. Diese Referenz-Menge ist absolut, sie setzt sich aus den Grenzen des Intervalls  $[0, 1]$  zusammen. Die Null steht hierbei für die falsche und die Eins für die richtige Klassifikation des zu beurteilenden kontextuellen Algorithmus  $\mathbf{A}_C$ . Der Identifikator der kontextuellen Klasse ist in jedem Datum der Eingabe enthalten, was die Besonderheit der in dieser Arbeit vorgestellten Methodik bedeutet. Dieses letzte Quadrate-Verfahren ist auch eine weitere Grundlage zum Aufbau des neuronalen Fuzzy Netzes.

Zum eigentlichen Training des neuronalen Fuzzy Netzes stehen zwei Verfahren im ANFIS-Modul zur Verfügung, zum einen ein reines Backpropagation-Training und zum anderen eine hybride Methode. Das Backpropagation-Training ist ein überwachtetes Gradientenabstiegsverfahren, das ausgehend von dem Abstand der gewünschten und der realen Ausgabe des Netzes ein Minimum des Fehlers zu erreichen versucht. Die gewünschte Ausgabe ist wieder in der zuvor beschriebenen Referenz-Menge enthalten. Die hybride Methode fusioniert das Backpropagation-Training und die letzte Quadrate Methode.

In folgenden Abschnitten werden nun die angesprochenen Verfahren anhand des zuvor beschriebenen qualitativen Sugeno-Regelsystems  $\hat{\mathbf{S}}_{Q_{lin}}$  beschrieben. Die reale Anwendung der Adaptionen des qualitativen Systems  $\hat{\mathbf{S}}_{Q_{lin}}$  für bestehende kontextuelle Systeme  $\mathbf{A}_C$  wird in Kapitel 4.3 und Kapitel 5.2 beschrieben.

### 4.2.1 Initiales subtraktives Clustern zur Bestimmung der Komplexität $o$

Zu Beginn der Adaption des Sugeno-Systems  $\hat{\mathbf{S}}_{Q_{lin}}$  an die Eingangsdaten und die gewünschte Menge der Ausgangsdaten muss die Komplexität des multidimensionalen Eingaberaumes bestimmt werden. Dies wird durch Teilung des Datensatzes, an den es zu adaptieren gilt, in eine gewisse Anzahl von Clustern erreicht. Die Anzahl der Cluster ist für den Algorithmus, der die Cluster bilden soll nicht bekannt. Aus diesem Grund muss ein Verfahren benutzt werden, das diese Aufgabe selbst-organisierend lösen kann. Hier wird ein subtraktives Clustering [Chi96] verwendet, das nicht nur die Anzahl der Cluster berechnet, sondern auch die initialen Gauss-Funktionen bestimmt.

Es wird ein Datensatz für das Training des unscharfen Sugeno-Systems verwendet, der alle kontextuellen Zustände beinhaltet. Sind gewisse Zustände in diesem Datensatz nicht enthalten, so kann das Resultat der Adaption für diese Zustände nicht vorhergesagt werden. Im Idealfall wird ein solcher Vektor bei der Normalisierung des Resultates des Sugeno-Systems  $\hat{\mathbf{S}}_{Q_{lin}}$  auf die Menge der Fehlerzustände  $E$  abgebildet, was aber keinesfalls garantiert werden kann. Mit dieser Forderung an die Vollständigkeit des zu adaptierenden Datensatzes wird die Matrix  $V_T$  der Trainingsdaten folgendermaßen festgelegt:

$$V_T = (\bar{\mathbf{v}}_{Q,t}^T, \bar{\mathbf{v}}_{Q,t+1}^T, \dots, \bar{\mathbf{v}}_{Q,t+n}^T)$$

$$\begin{aligned}
 &= \begin{pmatrix} v_{1t} & v_{1(t+1)} & \cdots & v_{1(t+n)} \\ v_{2t} & v_{2(t+1)} & \cdots & v_{2(t+n)} \\ \vdots & \vdots & \ddots & \vdots \\ v_{mt} & v_{m(t+1)} & \cdots & v_{m(t+n)} \\ A_C(\vec{\mathbf{v}}_{C,t}) & A_C(\vec{\mathbf{v}}_{C,t+1}) & \cdots & A_C(\vec{\mathbf{v}}_{C,t+n}) \end{pmatrix} \\
 &= \begin{pmatrix} v_{1t} & v_{1(t+1)} & \cdots & v_{1(t+n)} \\ v_{2t} & v_{2(t+1)} & \cdots & v_{2(t+n)} \\ \vdots & \vdots & \ddots & \vdots \\ v_{mt} & v_{m(t+1)} & \cdots & v_{m(t+n)} \\ v_{(m+1)t} & v_{(m+1)(t+1)} & \cdots & v_{(m+1)(t+n)} \end{pmatrix} \quad (\text{Definition 4.2.1.1})
 \end{aligned}$$

Die Matrix  $V_T$  setzt sich also aus den transponierten Vektoren  $\vec{\mathbf{v}}_{Q,k}^T$  zusammen, die in Definition 4.1.1.4 spezifiziert wurden und die Eingaben des Sugeno-Systems  $\widehat{\mathbf{S}}_{Q_{lin}}$  der Qualitätsanalyse bilden. Der Index  $k$  ist Element der Indexmenge  $T = \{t, t+1, \dots, t+n\}$ , die die äquidistanten Zeitpunkte der Erzeugung der Vektoren  $\vec{\mathbf{v}}_{Q,k}^T$  bestimmt. Da für diese Matrix  $V_T$ , an die es das Sugeno-System  $\widehat{\mathbf{S}}_{Q_{lin}}$  zu adaptieren gilt, eine explizite Unterscheidung von Werten der Eingabe  $v_{ik}$  und kontextuellem Identifikator  $A_C(\vec{\mathbf{v}}_{C,k})$  nicht mehr notwendig ist, wird der Identifikator von nun an mit einem Wert der Eingabe  $v_{(m+1)k} := A_C(\vec{\mathbf{v}}_{C,k})$  identifiziert. Um die Identifikatoren der kontextuellen Klassen zu erhalten, die aufgrund des zuvor definierten Vektors  $\vec{\mathbf{v}}_{C,k} = (v_{1k}, \dots, v_{mk})$  (Definition 4.1.0.3) berechnet werden, muss ein Algorithmus  $\mathbf{A}_C$  der kontextuellen Klassifikation vorhanden sein. Mögliche Spezifikationen für einen solchen Algorithmus  $\mathbf{A}_C$  werden in Kapitel 4.3 beschrieben.

Anhand dieses Datensatzes  $V_T$ , der auch die Grundlage der weiteren adaptiven Algorithmen in diesem Kapitel 4.2 ist, kann nun das subtraktive Clustering erklärt werden.

### Methoden des subtraktiven Clusterings

Der Algorithmus des subtraktiven Clusterings [Chi96] schätzt die Zentren der Cluster des Datensatzes  $V_T$  und deren Verteilungen. Hierzu wird der Datensatz  $V_T$  anhand der Zugehörigkeit zu den Grenzen Null und Eins des Intervalls der Qualität  $Q = [0, 1]$  aufgeteilt, was zu zwei Matrizen  $V_{T0}$  und  $V_{T1}$  führt. Dies ist der einzige Bezug zu der Referenz-Menge der gewünschten Ausgabe des ansonsten selbst-organisierenden Algorithmus des subtraktiven Clusterings. Um die getrennte Berechnung der Cluster für die zwei klassenspezifischen Matrizen  $V_{T0}$  und  $V_{T1}$  nicht parallel beschreiben zu müssen, wird der Algorithmus des subtraktiven Clusterings anhand der gesamten Matrix  $V_T$  beschrieben.

Der Algorithmus des subtraktiven Clusterings arbeitet nur auf normierten Vektoren, da Parameter für Radien benutzt werden, die sonst für jede Dimension des Vektors separat definiert werden müssten. Die Normierung bezüglich des einheitlichen Hyperkubus erfolgt mittels einer noch zu spezifizierenden Norm  $\|\cdot\|$ . Eine mögliche Norm für die praktische Anwendung, die in Kapitel 4.3 beschrieben wird, ist die Maximums-Norm  $\|\cdot\|_{max}$ . Es wird die Norm der Zeilenvektoren  $\vec{\mathbf{v}}_{T,i} := (v_{it}, v_{i(t+1)}, \dots, v_{i(t+n)})$  berechnet und damit die jeweilige Dimension der ursprünglichen Vektoren  $\vec{\mathbf{v}}_{Q,k}^T$ , auf das Intervall  $[0, 1]$  normalisiert. Die Transformation der Matrix  $V_T$ , des in Cluster einzuteilenden Datensatzes, in die normierte Matrix  $\mathcal{V}_T$ , ist wie folgt definiert:

$$\begin{aligned}
 \mathcal{V}_T &= \begin{pmatrix} \frac{1}{\|\vec{\mathbf{v}}_{T,1}\|} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\|\vec{\mathbf{v}}_{T,m+1}\|} \end{pmatrix} \mathbf{V}_T \\
 &= \begin{pmatrix} \frac{v_{1t}}{\|\vec{\mathbf{v}}_{T,1}\|} & \cdots & \frac{v_{1(t+n)}}{\|\vec{\mathbf{v}}_{T,1}\|} \\ \vdots & \ddots & \vdots \\ \frac{v_{(m+1)t}}{\|\vec{\mathbf{v}}_{T,m+1}\|} & \cdots & \frac{v_{(m+1)(t+n)}}{\|\vec{\mathbf{v}}_{T,m+1}\|} \end{pmatrix} \\
 &= (\mathbf{v}_{Q,t}^T, \mathbf{v}_{Q,t+1}^T, \dots, \mathbf{v}_{Q,t+n}^T) \quad (\text{Definition 4.2.1.2})
 \end{aligned}$$

Die normierte Matrix  $\mathcal{V}_T$  setzt sich also aus den transponierten normierten Vektoren  $\mathbf{v}_{Q,i}^T$  zusammen, die nun in das subtraktive Clustering einfließen.

Es wird jeder Vektor  $\mathbf{v}_{Q,i}^T$  mit  $i \in T$  zunächst als potentielles Zentrum eines Clusters angenommen.

Für alle weiteren Vektoren  $\mathbf{v}_{Q,j}^T$  mit  $j \neq i$  wird ein Maß der Ähnlichkeit, ein Potential  $P_i(\mathcal{V}_T)$  (Definition 4.2.1.3) zum potentiellen Zentrum  $\mathbf{v}_{Q,i}^T$  des Clusters berechnet. Die einzelnen Summanden des Potentials

$$P_i(\mathcal{V}_T) = \sum_{j=t, j \neq i}^{t+n} e^{-\alpha \|\mathbf{v}_{Q,i}^T - \mathbf{v}_{Q,j}^T\|_2^2} \quad , \quad \text{mit} \quad \alpha = \frac{4}{r_a^2} \quad (\text{Definition 4.2.1.3})$$

$P_i$  sind Gauss-Glocken, die anhand der Varianz  $\sigma^2$ , die im Faktor  $\alpha = \frac{1}{2\sigma^2}$  steckt, und des jeweilig betrachteten Mittelwert-Vektors  $\mathbf{v}_{Q,i}^T$  eine starke Beeinflussung des gesamten Potentials  $P_i$  bewirken, wenn sich der betrachtete Vektor  $\mathbf{v}_{Q,j}^T$  nahe am potentiellen Cluster-Zentrum  $\mathbf{v}_{Q,i}^T$  befindet. Die Varianz wird durch einen Parameter  $r_a$  definiert, der einen Radius der Nachbarschaft bestimmt. Somit ist der Wert  $\alpha$  durch den Radius  $r_a$  bestimmt, der mathematisch eher ein Durchmesser ist und deshalb noch halbiert werden muss ( $\frac{r_a}{2}$ ), was die folgende Gleichung ergibt:

$$\frac{1}{\alpha} = \left(\frac{r_a}{2}\right)^2 = \frac{r_a^2}{4}, \quad \text{somit folgt} \quad \alpha = \frac{4}{r_a^2}$$

Vektoren  $\mathbf{v}_{Q,j}^T$ , die außerhalb dieses Radius  $r_a$  liegen haben nur einen geringen Einfluss auf das Potential  $P_i(\mathcal{V}_T)$ .

Nun wird der Vektor  $\mathbf{v}_{Q,i}^T$  ausgewählt, der das höchste Potential  $P_i(\mathcal{V}_T)$  besitzt. Dieser bildet das erste feste Cluster-Zentrum  $\mathbf{v}_{Q,1}^{*T}$  mit dem Zugehörigen Potential  $P_1^*(\mathcal{V}_T)$ . Mit diesem festgelegten ersten Cluster-Zentrum werden die Potentiale der anderen Vektoren  $\mathbf{v}_{Q,j}^T$  mit  $j \in T$  nun erneuert. Von jedem vorherigen Potential  $P_j(\mathcal{V}_T)$  wird das Potential  $P_1^*(\mathcal{V}_T)$ , multipliziert mit mit einer neuen Exponentialfunktion  $e^{-\beta \|\mathbf{v}_{Q,i}^T - \mathbf{v}_{Q,1}^{*T}\|_2^2}$ , abgezogen und dem Potential  $P_j(\mathcal{V}_T)$  als neuer Wert zugewiesen ( $\Leftarrow$ ), was in Definition 4.2.1.4 beschrieben wird. Der Wert  $\beta = \frac{4}{r_b^2}$  wird durch einen neuen Radius  $r_b$  definiert, welcher

$$P_j(\mathcal{V}_T) \Leftarrow P_j(\mathcal{V}_T) - P_1^*(\mathcal{V}_T) e^{-\beta \|\mathbf{v}_{Q,j}^T - \mathbf{v}_{Q,1}^{*T}\|_2^2} \quad , \quad \text{mit} \quad \beta = \frac{4}{r_b^2} \quad (\text{Definition 4.2.1.4})$$

einen Radius der Nachbarschaft beschreibt, der ein Mindestmaß des Abstandes zum ersten Cluster-Zentrum definiert. Vektoren  $\mathbf{v}_{Q,j}^T$  die sich nahe am ersten Cluster-Zentrum  $\mathbf{v}_{Q,1}^{*T}$  befinden, werden eine große Reduktion des Potentials  $P_j(\mathcal{V}_T)$  durch den Faktor  $e^{-\beta \|\mathbf{v}_{Q,i}^T - \mathbf{v}_{Q,1}^{*T}\|_2^2}$  erhalten und deshalb sehr wahrscheinlich nicht zu einem neuen Cluster-Zentrum erwählt werden. Um zu dicht gepackte Cluster zu vermeiden wird ein Verhältnis von  $r_b$  zu  $r_a$  von  $r_b = 1,25 \cdot r_a$  [Chi96] gewählt. Wurden alle Potentiale reduziert, so wird wiederum der Vektor  $\mathbf{v}_{Q,j}^T$  mit dem höchsten Potential  $P_j(\mathcal{V}_T)$  als zweites Cluster-Zentrum  $P_2^*(\mathcal{V}_T)$  erwählt.

Dieser Algorithmus wird in der Abwechslung von Erneuerung der Potentiale und Auswahl neuer Cluster-Zentren solange wiederholt, bis eine Terminations-Bedingung erfüllt ist. Die allgemeine Formel zur Erneuerung der Potentiale  $P_l(\mathcal{V}_T)$  in Abhängigkeit des zuvor gewählten Cluster-Zentrums  $\mathbf{v}_{Q,k}^{*T}$  wird in

$$P_l(\mathcal{V}_T) \Leftarrow P_l(\mathcal{V}_T) - P_k^*(\mathcal{V}_T) e^{-\beta \|\mathbf{v}_{Q,i}^T - \mathbf{v}_{Q,k}^{*T}\|_2^2} \quad , \quad \text{mit} \quad \beta = \frac{4}{r_b^2} \quad (\text{Definition 4.2.1.5})$$

Definition 4.2.1.5 beschrieben. Das Stopp-Kriterium des Algorithmus des subtraktiven Clusterings wird für das Potential  $P_o^*(\mathcal{V}_T)$  des letzten Cluster-Zentrums  $\mathbf{v}_{Q,o}^{*T}$  auf typische Weise  $P_o^*(\mathcal{V}_T) < 0,15 \cdot P_1^*(\mathcal{V}_T)$  [Chi96] gesetzt. Alle erneuerten Potentiale  $P_l(\mathcal{V}_T)$ , die aufgrund dieses letzten Cluster-Zentrums  $\mathbf{v}_{Q,o}^{*T}$  berechnet wurden, erfüllen nicht diese Bedingung.

### Umwandlung initialer Cluster in Zugehörigkeitsfunktionen

Mit der so erhaltenen Menge  $Z = \{\mathbf{v}_{Q,1}^*, \dots, \mathbf{v}_{Q,o}^*\}$  der Cluster-Zentren können nun die zugehörigen Gauss-Funktionen berechnet werden. Die Menge  $Z$  setzt sich nun wieder aus allen Cluster-Zentren zusammen, die zuvor, aufgrund der Funktionsweise des Algorithmus des subtraktiven Clusterings, für die Mengen  $V_{T0}$  und  $V_{T1}$  getrennt berechnet wurden. Da die Vektoren  $\mathbf{v}_{Q,k}^*$  mit  $k \in O = \{1, \dots, o\}$  aufgrund der algorithmischen Spezifikation aus normalisierten Vektoren  $\mathbf{v}_{Q,l}$  mit  $l \in T$  bezüglich des Einheits-Hyperkubus berechnet wurden, sind die Gauss-Glocken, die nun bestimmt werden, auch normiert.

Somit ergibt sich die Zugehörigkeitsfunktion für eines der Cluster-Zentren  $\mathbf{v}_{Q,i}^{*T}$  mit einer Gauss-Funktion wie folgt:

$$F_{\mathbf{v}_{Q,k}^*}(\mathbf{x}) = e^{-\alpha \|\mathbf{x}^T - \mathbf{v}_{Q,k}^{*T}\|_2^2}, \quad \text{mit} \quad \dim(\mathbf{x}) = \dim(\mathbf{v}_{Q,k}^*) \quad (\text{Definition 4.2.1.6})$$

Wird die Menge  $Z$  der Cluster-Zentren wieder als Matrix formuliert, so kann die Gauss-Funktion (Definition 4.2.1.7) für jedes Element und jedes Cluster-Zentrum separat definiert werden. Hierbei ergibt sich aus dem jeweiligen Element  $v_{ij}^*$ , des jeweiligen Vektors  $\vec{\nabla}_{Q,j}^*$ , der Mittelwert  $v_{ij}^*$ , der normierten Gauss-Funktion  $F_{ij}^*(x_i)$ .

$$F_{ij}^*(x_i) = e^{-\frac{(x_i - v_{ij}^*)^2}{(2\sigma_{ij}^{2*})}}, \text{ mit } \sigma_{ij}^{2*} = \frac{1}{2\alpha} \quad (\text{Definition 4.2.1.7})$$

Da diese Normalverteilungen nur auf normierten Eingabevektoren  $\mathbf{v}_{Q,j}$  mit  $j \in T$  definiert sind, das eigentliche Sugeno-System  $\hat{S}_{Q_{lin}}(\vec{\nabla}_Q)$  zur Qualitätsanalyse aber auf den ursprünglichen Universen  $\vec{\nabla}_Q \in \mathcal{V}_1 \times \dots \times \mathcal{V}_M \times C$  definiert ist, muss noch eine Aufhebung der Normierung formuliert werden. Die Elemente der Cluster-Zentren  $v_{ij}^*$  werden zu Mittelwerten  $\mu_{ij}$  der eigentlichen Gauss-Funktionen  $F_{ij}(x_i)$  umformuliert, indem die normierten Mittelwerte  $v_{ij}^*$  mit der Norm dieser Dimension multipliziert werden, was die Gleichung  $\mu_{ij} = \|\vec{\nabla}_{T,i}\| v_{ij}^*$  ergibt. Ähnliches muss mit der normierten Varianz  $\sigma_{ij}^{2*}$  geschehen, was die Gleichung  $\sigma_{ij}^2 = \|\vec{\nabla}_{T,i}\| \sqrt{\sigma_{ij}^{2*}}$  ergibt. Die somit erhaltenen Gauss-Funktionen des Sugeno-Regelsystems  $\hat{S}_{Q_{lin}}$  sind zuerst definiert worden in Definition 4.1.2.6, welche wie folgt wiederholt werden:

$$F_{ij}(v_i) = e^{-\frac{(v_i - \mu_{ij})^2}{(2\sigma_{ij}^2)}}$$

Die berechneten Cluster definieren nun nicht nur die initialen Zugehörigkeitsfunktionen  $F_{ij}(x_i)$ , sondern auch die Anzahl  $o$  der implikativen Regeln des Sugeno-Regelsystems  $\hat{S}_{Q_{lin}}$  und damit auch die Komplexität des neuronalen Fuzzy Netzes (Abbildung 12), das zur Verfeinerung der Parameter ( $\sigma_{ij}^2, \mu_{ij}, a_{ij}$ ) verwendet wird.

#### 4.2.2 Methode der kleinsten Quadrate zur initialen Adaption der linearen Konstanten

$$\mathcal{A} = (a_{ij})$$

Die Methode der kleinsten Quadrate ist ein Verfahren, das in der Stochastik zur Schätzung in der Regressionsanalyse [DS81] benutzt wird. Diese Methode versucht eine Kurve in einen Satz von Daten möglichst genau passend einzufügen. Die Kurve wird anhand des Abstandes, also dem Fehler der Kurve, zu den einzelnen Daten eingepasst, was auch recht gut durch den Begriff der kleinsten Fehlerquadrate beschrieben wird. Dieser Abstand (Fehler) wird über ein lineares Gleichungssystem formuliert, dessen Lösung das Resultat der Methode ist. Der Datensatz, an den es die Kurve zu adaptieren gilt, resultiert, anhand seiner Größe, in einem überbestimmten Gleichungssystem. Diese überbestimmte lineare Gleichung erfordert eine numerische Lösung.

In dem Fall des qualitativen Sugeno-Systems  $\hat{S}_{Q_{lin}}$  sollen die linearen Konsequenzen  $f_i$ , in Abhängigkeit von den nichtlinearen Antezedenzien  $w_j$ , an die Daten  $V_T$  angepasst werden, die wie folgt spezifiziert sind:

$$V_T = \begin{pmatrix} v_{1t} & v_{1(t+1)} & \cdots & v_{1(t+n)} \\ v_{2t} & v_{2(t+1)} & \cdots & v_{2(t+n)} \\ \vdots & \vdots & \ddots & \vdots \\ v_{mt} & v_{m(t+1)} & \cdots & v_{m(t+n)} \\ v_{(m+1)t} & v_{(m+1)(t+1)} & \cdots & v_{(m+1)(t+n)} \end{pmatrix}$$

Der Fehler  $\varepsilon$  (Definition 4.2.2.11) des Sugeno-Systems ergibt sich aus dem Abstand von der vom Entwickler bereitzustellenden gewünschten Ausgabe  $u_k$  und der tatsächlichen Ausgabe des Systems. Die Werte  $u_k$  der gewünschten Ausgaben werden in einem Vektor  $\mathbf{u}$  zusammengefasst, dessen jeweiliges Element  $u_k$  sich eindeutig mit einem Spaltenvektor  $\vec{\nabla}_{Q,k}$  der Matrix  $V_T$  der Daten assoziiert. Die einzelnen Fehler  $\varepsilon_k$  pro Eingabevektor  $\vec{\nabla}_{Q,k}$  werden für die ganze Matrix  $V_T$  aufsummiert und ergeben den gesamten Fehler  $E$  (Definition 4.2.2.12) des Sugeno-Systems für den Trainingsdatensatz. Diesen gesamten Fehler  $E$  gilt es anhand eines linearen Gleichungssystems zu minimieren, indem die linearen Gewichte  $a_{ij}$  der linearen Konsequenzen  $f_i$  für dieses Minimum berechnet werden. Um die Formulierung des linearen Gleichungssystems zu ermöglichen, muss zunächst die normalisierte gewichtete Summe des Sugeno-Systems in Operatoren der linearen Algebra umformuliert werden. Die linearen Gewichte  $a_{ij}$  müssen außerdem

noch vom Rest der Formel isoliert werden, was die Gleichung 4.2.2.1 ergibt. Da sich die Werte der nicht-linearen Antezedenzien  $w_j$  durch die Methode der kleinsten Quadrate nicht ändern, und somit auch die normalisierten  $\omega_j$  nicht, werden sie in dem Vektor  $\vec{\mathbf{b}}$  als Konstanten  $b_i$  zusammengefasst.

$$\begin{aligned}
 \widehat{\mathbf{S}}_{Q_{lin}}(\vec{\mathbf{v}}_Q) &= \frac{\sum_{j=1}^O w_j(\vec{\mathbf{v}}_Q) f_j(\vec{\mathbf{v}}_Q)}{\sum_{j=1}^O w_j(\vec{\mathbf{v}}_Q)} \quad , \text{ mit } \omega_j(\vec{\mathbf{v}}_Q) := \frac{w_j(\vec{\mathbf{v}}_Q)}{\sum_{j=1}^O w_j(\vec{\mathbf{v}}_Q)} \text{ folgt} \\
 &= \sum_{j=1}^O \omega_j(\vec{\mathbf{v}}_Q) f_j(\vec{\mathbf{v}}_Q) \\
 &= \sum_{j=1}^O (\omega_j(\vec{\mathbf{v}}_Q) a_{1j} v_1 + \omega_j(\vec{\mathbf{v}}_Q) a_{2j} v_2 + \dots + \omega_j(\vec{\mathbf{v}}_Q) a_{(m+1)j} v_{m+1} + \omega_j(\vec{\mathbf{v}}_Q) a_{(m+2)j}) \\
 &= \langle (\omega_1(\vec{\mathbf{v}}_Q) v_1, \dots, \omega_1(\vec{\mathbf{v}}_Q), \omega_2(\vec{\mathbf{v}}_Q) v_1, \dots, \omega_2(\vec{\mathbf{v}}_Q), \dots, \omega_o(\vec{\mathbf{v}}_Q) v_1, \dots, \omega_o(\vec{\mathbf{v}}_Q)) , \\
 &\quad (a_{11}, \dots, a_{1o}, a_{21}, \dots, a_{2o}, \dots, a_{(m+1)1}, \dots, a_{(m+1)o}, a_{(m+2)1}, \dots, a_{(m+2)o}) \rangle \\
 &\quad , \text{ mit } \vec{\mathbf{b}} := (\omega_1(\vec{\mathbf{v}}_Q) v_1, \dots, \omega_o(\vec{\mathbf{v}}_Q)) \text{ und } \vec{\mathbf{a}} := (a_{11}, \dots, a_{(m+2)o}) \text{ folgt} \\
 &= \vec{\mathbf{b}} \vec{\mathbf{a}}^T \quad \text{(Gleichung 4.2.2.1)}
 \end{aligned}$$

Auch alle linearen Gewichte  $a_{ij}$  werden, gemäß der Ordnung der Zeilenvektoren der Matrix  $A$  (Definition 4.1.3.13), in einem Vektor  $\vec{\mathbf{a}}$  vereint. Diese Formulierung der normalisierten gewichteten Summe in einer Gleichung 4.2.2.1 der linearen Algebra, liefert die Grundlage die Ausgabe des qualitativen Sugeno-Systems  $\widehat{\mathbf{S}}_{Q_{lin}}$  für einen beliebigen Vektor  $\vec{\mathbf{v}}_Q$  zu berechnen. Die Berechnung des gesamten Fehlers  $E$  wird jedoch über den gesamten Datensatz  $V_T$  ausgeführt, weshalb eine Matrix  $B_T$  (Definition 4.2.2.9) definiert wurde. Sie beinhaltet die Vektoren  $\vec{\mathbf{b}}_k$  als Zeilenvektoren, die die Konstanten  $b_{kj}$  der gewichteten Elemente des Datenvektors  $\vec{\mathbf{v}}_{Q,k}$  enthalten.

$$\begin{aligned}
 B_T &:= \begin{pmatrix} \omega_1(\vec{\mathbf{v}}_{Q,t}) v_{1t} & \dots & \omega_1(\vec{\mathbf{v}}_{Q,t}) & \dots & \omega_o(\vec{\mathbf{v}}_{Q,t}) v_{1t} & \dots & \omega_o(\vec{\mathbf{v}}_{Q,t}) \\ \omega_1(\vec{\mathbf{v}}_{Q,t+1}) v_{1(t+1)} & \dots & \omega_1(\vec{\mathbf{v}}_{Q,t+1}) & \dots & \omega_o(\vec{\mathbf{v}}_{Q,t+1}) v_{1(t+1)} & \dots & \omega_o(\vec{\mathbf{v}}_{Q,t+1}) \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ \omega_1(\vec{\mathbf{v}}_{Q,t+n}) v_{1(t+n)} & \dots & \omega_1(\vec{\mathbf{v}}_{Q,t+n}) & \dots & \omega_o(\vec{\mathbf{v}}_{Q,t+n}) v_{1(t+n)} & \dots & \omega_o(\vec{\mathbf{v}}_{Q,t+n}) \end{pmatrix} \\
 &= \begin{pmatrix} b_{t,1} & \dots & b_{t,(m+2)} & \dots & b_{t,o} & \dots & b_{t,((m+2)o)} \\ b_{(t+1),1} & \dots & b_{(t+1),(m+2)} & \dots & b_{(t+1),o} & \dots & b_{(t+1),((m+2)o)} \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ b_{(t+n),1} & \dots & b_{(t+n),(m+2)} & \dots & b_{(t+n),o} & \dots & b_{(t+n),((m+2)o)} \end{pmatrix} \\
 &= \begin{pmatrix} b_{11} & \dots & b_{1p} \\ b_{21} & \dots & b_{2p} \\ \vdots & \ddots & \vdots \\ b_{(n+1)1} & \dots & b_{(n+1)p} \end{pmatrix} \\
 &= \begin{pmatrix} \vec{\mathbf{b}}_1 \\ \vec{\mathbf{b}}_2 \\ \vdots \\ \vec{\mathbf{b}}_{n+1} \end{pmatrix} \quad \text{(Definition 4.2.2.9)}
 \end{aligned}$$

Mit dieser Matrix  $B_T$  kann der Vektor der realen Ausgaben des qualitativen Sugeno-Systems  $\widehat{\mathbf{S}}_{Q_{lin}}$  für den gesamten Datensatz  $V_T$  wie folgt berechnet werden.

$$\begin{pmatrix} \widehat{\mathbf{S}}_{Q_{lin}}(\vec{\mathbf{v}}_{Q,t}) \\ \widehat{\mathbf{S}}_{Q_{lin}}(\vec{\mathbf{v}}_{Q,t+1}) \\ \vdots \\ \widehat{\mathbf{S}}_{Q_{lin}}(\vec{\mathbf{v}}_{Q,t+n}) \end{pmatrix} = B_T \vec{\mathbf{a}}^T \quad \text{(Definition 4.2.2.10)}$$

Es wurden nun alle Voraussetzungen hergeleitet um das lineare Gleichungssystem der Methode der kleinsten Quadrate aufzustellen.

### Methode der kleinsten Quadrate

Die Grundlage der Methode sind die Fehler  $\varepsilon$ , die es zu minimieren gilt. Der Fehler den das Sugeno-System  $\widehat{\mathbf{S}}_{Q_{lin}}$  bei der Eingabe  $\vec{\mathbf{v}}_{Q,k}^T$  bezüglich der gewünschten Ausgabe  $u_k$  macht, berechnet man wie folgt:

$$\varepsilon(\vec{\mathbf{v}}_{Q,k}) := u_k - \widehat{\mathbf{S}}_{Q_{lin}}(\vec{\mathbf{v}}_{Q,k}) \quad (\text{Definition 4.2.2.11})$$

Werden die Einzelfehler für alle Vektoren  $\vec{\mathbf{v}}_{Q,k}$  der Datenmenge  $V_T$  aufsummiert, so erhält man den gesamten Fehler  $E$  wie folgt:

$$E(V_T) := \sum_{k=t}^{t+n} \varepsilon^2(\vec{\mathbf{v}}_{Q,k}) = \sum_{k=t}^{t+n} \left( u_k - \widehat{\mathbf{S}}_{Q_{lin}}(\vec{\mathbf{v}}_{Q,k}) \right)^2 = \sum_{k=t}^{t+n} \left( u_k - \sum_{j=1}^{p=(m+2)o} b_{kj} a_j \right)^2 \quad (\text{Definition 4.2.2.12})$$

Es wird nun versucht diesen Fehler  $E$  zu minimieren, wodurch die Konvergenz  $\widehat{\mathbf{S}}_{Q_{lin}}(\vec{\mathbf{v}}_{Q,k}^T) \rightarrow u_k$  gegen die gewollte Ausgabe  $u_k$  erreicht wird. Die erste Ableitung einer Funktion  $E$  ergibt für ihr Minimum einen Wert von null, wodurch dieses Eindeutig identifiziert wird. Der Gradient der Fehlerfunktion  $E$  für das jeweilige lineare Gewicht  $a_j$  berechnet man anhand folgender Gleichung:

$$\frac{\partial E}{\partial a_i} = \sum_{k=t}^{t+n} 2 \left( u_k - \sum_{j=1}^{p=(m+2)o} b_{kj} a_j \right) (-b_{ki}) = 0 \quad (\text{Gleichung 4.2.2.2})$$

Das lineare Gewicht  $a_{ij}$  besaß bisher zwei Indizes, die zum einen die Dimension  $i$  des Eingabevektors und zum anderen die jeweilig Regel  $j$  identifiziert, welche jedoch jetzt auf einen Index reduziert wurden, der das jeweilige Element des Vektors  $\vec{\mathbf{a}}^T$  identifiziert. Durch diesen Gradienten ergibt sich die Gleichung 4.2.2.3, die die Gradienten für alle  $a_j \in \vec{\mathbf{a}}$  simultan berechnen lässt. Die Herleitung dieser Gleichung 4.2.2.3 findet man auch in der Diplomarbeit von Manfred Männle [Män95] in anderer Notation.

$$\begin{aligned} 2(B_T \vec{\mathbf{a}}^T - \mathbf{u})^T B_T &= 0 \\ B_T^T (B \vec{\mathbf{a}}^T - \mathbf{u}) &= 0 \\ B_T^T B_T \vec{\mathbf{a}}^T &= B_T^T \mathbf{u} \\ \vec{\mathbf{a}}^T &= ((B_T^T B_T)^{-1} B_T^T) \mathbf{u} \end{aligned} \quad (\text{Gleichung 4.2.2.3})$$

Mit der Gleichung 4.2.2.3 erhält man nun das überbestimmte Gleichungssystem, dessen Lösung gerade die Lösung der Methode der kleinsten Quadrate ist. Eine numerisch stabile Lösung liefert die Singulärwertzerlegung (eng.: Singular Value Decomposition (SVD)), die auch im Rahmen des von MATLAB<sup>®</sup> bereitgestellten ANFIS-Modul benutzt wird.

### Die Singulärwertzerlegung des linearen Gleichungssystems

Die Singulärwertzerlegung (SVD) ermöglicht es die Matrix  $B_T$  in drei Matrizen mit bestimmten Eigenschaften zu zerlegen. Die Matrix  $B_T = U\mathcal{D}\mathcal{V}^T$  wird so zerlegt, dass die Multiplikation der Matrix  $U$  mit ihrer Transponierten  $U^T U = \mathcal{E}$  die Einheitsmatrix ergibt, die Matrix  $\mathcal{D}$  eine Diagonalmatrix ist und die Matrix  $\mathcal{V}$  orthogonal ist. Durch diese Zerlegung ergibt sich folgende Lösung des ursprünglichen Gleichungssystems (Gleichung 4.2.2.3):

$$\begin{aligned} \vec{\mathbf{a}}^T &= (\mathcal{V}\mathcal{D}^T U^T U \mathcal{D}\mathcal{V}^T)^{-1} \mathcal{V}\mathcal{D}^T U^T \mathbf{u} && (B_T = U\mathcal{D}\mathcal{V}^T) \\ &= (\mathcal{V}\mathcal{D}^T \mathcal{D}\mathcal{V}^T)^{-1} \mathcal{V}\mathcal{D}^T U^T \mathbf{u} && (U^T U = \mathcal{E}) \\ &= ((\mathcal{D}^T \mathcal{D}) \mathcal{V}\mathcal{V}^T)^{-1} \mathcal{V}\mathcal{D}^T U^T \mathbf{u} && (\mathcal{D}^T \mathcal{D} \text{ diagonal}) \\ &= (\mathcal{D}^T \mathcal{D})^{-1} \mathcal{V}\mathcal{D}^T U^T \mathbf{u} && (\mathcal{V} \text{ orthogonal} \Rightarrow \mathcal{V}\mathcal{V}^T = \mathcal{E}) \\ &= \mathcal{V}(\mathcal{D}^T \mathcal{D})^{-1} \mathcal{D}^T U^T \mathbf{u} \\ &= \mathcal{V}\mathcal{D}^{-1} (\mathcal{D}^T)^{-1} \mathcal{D}^T U^T \mathbf{u} \\ &= \mathcal{V}\mathcal{D}^{-1} U^T \mathbf{u} \end{aligned} \quad (\text{Gleichung 4.2.2.4})$$

Die Methode der Singulärwertzerlegung und ein QR-Algorithmus zu deren Ausführung, wird in dem Buch 'Numerische Mathematik' [DH93] beschrieben.

Die nun somit berechneten Parameter  $a_{ij}$  der linearen Konsequenzen  $f_j$  bilden einen weiteren Bestandteil des neuronalen Fuzzy Netzes, das im nächsten Abschnitt beschrieben wird. Die weitere Anpassung

der Parameter  $a_{ij}$  wird über ein Backpropagation-Training des neuronalen Fuzzy Netzes oder mit weiteren Iterationen der Methode der kleinsten Quadrate durch veränderte Parameter  $b_{kl}$  in einer hybriden Trainings-Methode erreicht.

#### 4.2.3 Aufbau des neuronalen Fuzzy Netzes welches das Sugeno-System $\widehat{S}_{Qlin}$ repräsentiert

Das Feintuning der Parameter  $(\sigma_{ij}^2, \mu_{ij}, a_{ij})$  wird mittels eines reinen Backpropagation-Trainings oder einer hybriden Methode erreicht. Die Hybridität ergibt sich aus einer Kombination von Backpropagation Training der nichtlinearen Funktionsparameter  $(\sigma_{ij}^2, \mu_{ij})$  und einer von den veränderten Zugehörigkeitsfunktionen abhängigen Optimierung der linearen Funktionsparametern  $(a_{ij})$  durch die Methode der kleinsten Quadrate.

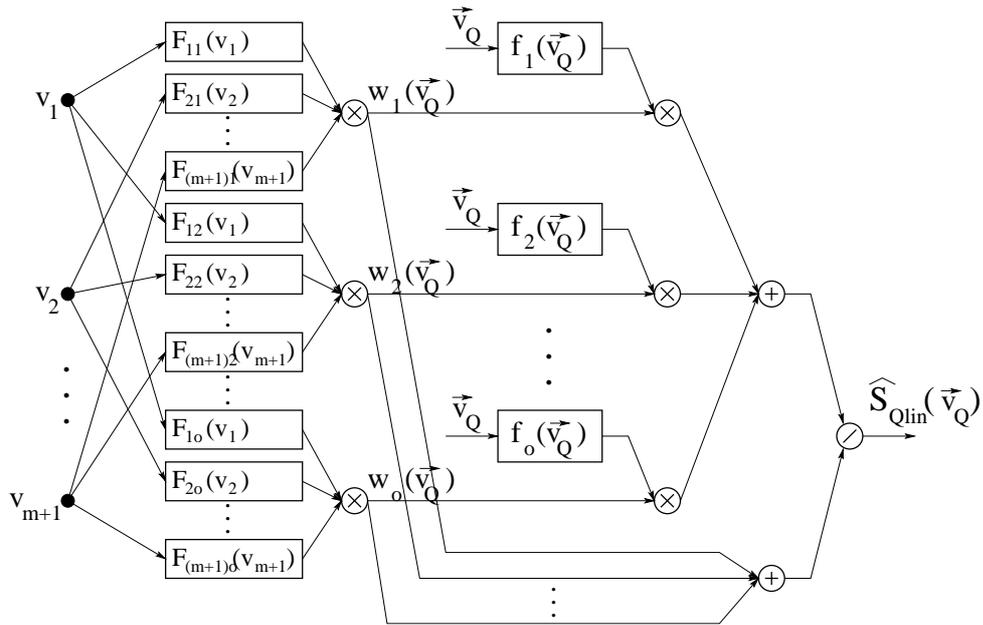


Abbildung 12: Neuronales Fuzzy Netz, das das allgemeine qualitative Sugeno-Regelsystem repräsentiert; diejenigen Neuronen die rechteckig symbolisiert sind, können durch ein Verfahren des Trainings verändert werden, wobei die runden Neuronen nicht veränderliche Operatoren symbolisieren

Zunächst wird ein neuronales Fuzzy Netz (Abbildung 12, [OL02]) aufgebaut, das das durch subtraktives Clustering und der Methode der kleinsten Quadrate erhaltene bisherige Sugeno-Regelsystem  $\widehat{S}_{Qlin}$  repräsentiert. Der Vektor  $\vec{v}_{Q,k}^T$  wird in seine Elemente aufgeteilt, an die jeweiligen Neuronen der ersten Schicht übergeben. Diese Neuronen haben keine funktionale Komponente, sondern verteilen die Elemente des Vektors  $\vec{v}_{Q,k}^T$  nur auf die zuständigen Zugehörigkeitsfunktionen der zweiten Schicht des neuronalen Fuzzy Netzes. Diese Schicht von Neuronen ist die einzige, die bei der hybriden Methode durch das Backpropagation Training beeinflusst wird.

Die Funktionen  $F_{ij}$  der Neuronen der zweiten Schicht sind die Gauss-Funktionen, die zuvor mittels des subtraktiven Clusterings berechnet wurden. Die Menge  $o$  der Cluster bestimmt den Aufbau des neuronalen Fuzzy Netzes, anhand der durch sie bestimmten Anzahl von implizierenden Regeln. Dies ist die Schicht deren Parameter  $(\sigma_{ij}^2, \mu_{ij})$  durch das Backpropagation Training, gegenüber der gewollten Ausgabe  $\mathbf{u}$ , angepasst werden soll. Um diesen Fehler zu minimieren wird über das Gradientenabstiegs-Verfahren ein Minimum in der mehrdimensionalen Fläche der Fehler gesucht.

Die Neuronen der dritten Schicht symbolisieren eine einfache Operation, die Multiplikation der Funktionswerte  $F_{ij}(\vec{v}_{Q,k}^T)$  der Zugehörigkeitsfunktionen, die die Gewichte  $w_j(\vec{v}_{Q,k})$  ergeben. Diese Ergebnisse werden verdoppelt und einmal direkt an die fünfte Schicht und einmal über die Neuronen der vierten an die fünfte Schicht von Neuronen weitergegeben.

In die vierte Schicht von Neuronen werden neben den Ausgaben der dritten Schicht noch die Funktionswerte  $f_j(\vec{\nabla}_{Q,k})$  der linearen Konsequenzen mit eingespeist. Da die linearen Konsequenzen der Regeln für die hybride Methode nur indirekt vom Backpropagation-Training beeinflusst werden, stehen sie in der graphischen Darstellung (Abbildung 12) leicht außerhalb des eigentlichen neuronalen Fuzzy Netzes. Die Neuronen führen wieder, wie die der Schicht drei, eine Multiplikation der Eingangswerte aus, was die Formel  $w_j(\vec{\nabla}_{Q,k})f_j(\vec{\nabla}_{Q,k})$  als einen Bestandteil der gewichteten normalisierten Summe  $\widehat{\mathbf{S}}_{Q_{lin}}(\vec{\nabla}_{Q,k})$  ergibt.

Vor der Schicht des Neurons das die Ausgabe des neuronalen Fuzzy Netzes berechnet, befindet sich eine Schicht die zwei Neuronen beinhaltet. Diese berechnen jeweils Summen, das eine bildet die gewichtete Summe  $\sum_{j=1}^o w_j(\vec{\nabla}_{Q,k})f_j(\vec{\nabla}_{Q,k})$  und das andere die Summe des normalisierenden Wertes  $\sum_{j=1}^o w_j(\vec{\nabla}_{Q,k})$ .

Zur Berechnung der Ausgabe, der neuronalen Abstraktion des Sugeno-Regelsystems  $\widehat{\mathbf{S}}_{Q_{lin}}$ , muss in der letzten Schicht das Neuron der Ausgabe eine Division ausführen. Dieses Neuron bildet den letzten operationellen Bestandteil der normalisierten gewichteten Summe  $\sum_{j=1}^o w_j(\vec{\nabla}_{Q,k})f_j(\vec{\nabla}_{Q,k}) / \sum_{j=1}^o w_j(\vec{\nabla}_{Q,k})$ , die das Resultat des Sugeno-Regelsystems ist.

Da der Aufbau des neuronalen Fuzzy Netzes nun beschrieben ist, kann der reine Backpropagation-Algorithmus, der zum Tuning der funktionalen Zugehörigkeiten und linearen Konsequenzen benutzt wird, definiert werden. Die Veränderung der Parameter der Zugehörigkeitsfunktionen bedingt auch bei der hybriden Methode eine Änderung der linearen Konsequenzen mittels der Methode der kleinsten Quadrate. Dieser Zusammenhang wird nochmals im Detail beschrieben.

#### 4.2.4 Backpropagation-Training zum Tuning der Zugehörigkeitsfunktionen $F_{ij}$ und der linearen Konstanten $\mathcal{A} = (a_{ij})$

Die Optimierung der veränderbaren Parameter  $(\sigma_{ij}^2, \mu_{ij}, a_{ij})$  kann anhand des neuronalen Fuzzy Netzes (Abbildung 12) über ein Backpropagation-Training des Netzes durchgeführt werden. Das Backpropagation-Training überträgt den Fehler, der sich zwischen Eingabe  $V_T$  und gewollter Ausgabe  $\mathbf{u}$  ergibt, zurück durch das neuronale Netz zu der Schicht der Neuronen, die aufgrund dessen verändert werden sollen. In dem hier vorliegenden Fall sind nur zwei Schichten des Netzes adaptierbar, da die anderen Operatoren sind. Somit sind die Parameter  $(\sigma_{ij}^2, \mu_{ij})$  der nichtlinearen Zugehörigkeitsfunktionen  $F_{ij}$  und die Gewichte  $(a_{ij})$  der linearen Konsequenz  $f_j$  zu verändern.

Die Adaption wird durch einen Gradientenabstieg [Män00] berechnet, indem in einer Fehlerfunktion  $E$  (Definition 4.2.4.16) versucht wird ein möglichst globales Minimum zu finden. Die Fehlerfunktion  $E(V_T)$  wird über den zu lernenden Datensatz  $V_T$  berechnet, wobei der Einzelfehler  $\varepsilon(\vec{\nabla}_{Q,k})$  (Definition 4.2.4.17)

$$E(V_T) := \frac{1}{2} \sum_{k=t}^{t+n} \varepsilon^2(\vec{\nabla}_{Q,k}) \quad (\text{Definition 4.2.4.16})$$

aus der Distanz der Ausgabe  $\widehat{\mathbf{S}}_{Q_{lin}}(\vec{\nabla}_{Q,k})$  bezüglich des Eingabevektors  $\vec{\nabla}_{Q,k}$  und der gewünschten Ausgabe  $u_k$  für diesen Vektor berechnet wird. Das initiale Sugeno-System  $\widehat{\mathbf{S}}_{Q_{lin}}$ , das den Referenzwert zu Beginn des Trainings ergibt, wurde zuvor durch das subtraktive Clustering und die Methode der kleinsten

$$\varepsilon(\vec{\nabla}_{Q,k}) = u_k - \widehat{\mathbf{S}}_{Q_{lin}}(\vec{\nabla}_{Q,k}) \quad (\text{Definition 4.2.4.17})$$

Quadrate bestimmt. Um nun die Gradienten für die einzelnen Parameter  $(\sigma_{ij}^2, \mu_{ij}, a_{ij})$  zu bestimmen, muss eine Formel definiert werden, die zum einen die Position des Parameters im neuronalen Netz (Abbildung 12) wiedergibt und zum anderen den Fehler  $\varepsilon(\vec{\nabla}_{Q,k})$  des derzeitigen Vektors  $\vec{\nabla}_{Q,k}$  enthält.

#### Ableitung der Parameter der linearen Konsequenzen

Aus den Forderungen an die Herleitung der Gradienten ergibt sich für die Parameter  $a_{ij}$  der ersten Funktionen  $f_j$ , die in der Rückrichtung von Ausgang zu Eingang des neuronalen Netzes auftreten, die Gleichung der Gradienten [Män95][Män00] wie folgt:

$$\frac{\partial E}{\partial a_{ij}} = (-1) \sum_{k=t}^{t+n} \varepsilon(\vec{\nabla}_{Q,k}) \frac{\sum_{l=1}^o w_l(\vec{\nabla}_{Q,k}^T) \frac{\partial f_l(\vec{\nabla}_{Q,k}^T)}{\partial a_{ij}}}{\sum_{l=1}^o w_l(\vec{\nabla}_{Q,k}^T)}$$

$$= \sum_{k=t}^{t+n} \frac{-\varepsilon(\vec{\nabla}_{Q,k}) w_j(\vec{\nabla}_{Q,k}^T) v_{ik}}{\sum_{l=1}^o w_l(\vec{\nabla}_{Q,k}^T)}, \text{ da } \frac{\partial f_l(\vec{\nabla}_{Q,k}^T)}{\partial a_{ij}} = v_{ik} \quad (\text{Gleichung 4.2.4.5})$$

Der Index  $j$  beschreibt in dieser Gleichung die jeweilige Regel, deren Konsequenz die lineare Funktion  $f_j$  ist.

#### Ableitung der Parameter der nichtlinearen Antezedenzien

Die beiden Gradienten  $\frac{\partial E}{\partial \mu_{ij}}$  und  $\frac{\partial E}{\partial \sigma_{ij}}$  der nichtlinearen funktionalen Zugehörigkeit  $F_{ij}$  berechnet man durch die gleichen Positionen im neuronalen Fuzzy Netz, weshalb ihre Zusammensetzungen ähnlich sind. Beide Gradienten sind vom derzeitigen Parameter des anderen abhängig, da sie Bestandteil der gleichen Funktion  $F_{ij}$  sind. Der Gradient [Män95][Män00] für den Mittelwert  $\mu_{ij}$  der Gauss-Funktion ergibt sich somit wie folgt:

$$\frac{\partial E}{\partial \mu_{ij}} = \sum_{k=t}^{t+n} \left( -\varepsilon(\vec{\nabla}_{Q,k}) \sum_{l=1}^o f_l(\vec{\nabla}_{Q,k}^T) \frac{\partial w_l(\vec{\nabla}_{Q,k}^T)}{\partial \mu_{ij}} \right) \quad (\text{Gleichung 4.2.4.6})$$

Die Ableitung des gesamten Gradienten  $\frac{\partial E}{\partial \mu_{ij}}$  für den Mittelwert  $\mu_{ij}$  ergibt sich schrittweise über die Ableitungen [Män95] der einzelnen Bestandteile. Der erste Schritt ist die Definition der normalisierten Gewichte  $\omega_j$ , die wie folgt aufgebaut sind:

$$\omega_j(\vec{\nabla}_{Q,k}^T) := \frac{w_j(\vec{\nabla}_{Q,k}^T)}{\sum_{l=1}^o w_l(\vec{\nabla}_{Q,k}^T)} \quad (\text{Definition 4.2.4.20})$$

Für diese normalisierten Gewichte kann nun der erste partielle Gradient  $\frac{\partial \omega_l(\vec{\nabla}_{Q,k}^T)}{\partial \mu_{ij}}$  hergeleitet werden, was für den Index  $l \neq j$  zu folgender Gleichung führt:

$$\frac{\partial \omega_l(\vec{\nabla}_{Q,k}^T)}{\partial \mu_{ij}} = \frac{-w_l(\vec{\nabla}_{Q,k}^T)}{\left( \sum_{s=1}^o w_s(\vec{\nabla}_{Q,k}^T) \right)^2} \frac{\partial w_j(\vec{\nabla}_{Q,k}^T)}{\partial \mu_{ij}} \quad (\text{Gleichung 4.2.4.7})$$

Ist der Index  $l = j$  so ist die Ableitung nach  $\mu_{ij}$  wie folgt beschrieben:

$$\begin{aligned} \frac{\partial \omega_j(\vec{\nabla}_{Q,k}^T)}{\partial \mu_{ij}} &= \frac{\left( \sum_{s=1}^o w_s(\vec{\nabla}_{Q,k}^T) \right) \frac{\partial w_j(\vec{\nabla}_{Q,k}^T)}{\partial \mu_{ij}} - w_j(\vec{\nabla}_{Q,k}^T) \frac{\partial w_j(\vec{\nabla}_{Q,k}^T)}{\partial \mu_{ij}}}{\left( \sum_{s=1}^o w_s(\vec{\nabla}_{Q,k}^T) \right)^2} \\ &= \frac{\left( \sum_{s=1}^o w_s(\vec{\nabla}_{Q,k}^T) \right) - w_j(\vec{\nabla}_{Q,k}^T)}{\left( \sum_{s=1}^o w_s(\vec{\nabla}_{Q,k}^T) \right)^2} \frac{\partial w_j(\vec{\nabla}_{Q,k}^T)}{\partial \mu_{ij}} \end{aligned} \quad (\text{Gleichung 4.2.4.8})$$

Der nächste Schritt ergibt sich anhand der normalisierten Gewichte  $\omega$ , die die eigentlichen Gewichte  $w_j$  enthalten. Deren Gradient  $\frac{\partial w_j(\vec{\nabla}_{Q,k}^T)}{\partial \mu_{ij}}$  leitet sich folgendermaßen ab:

$$\frac{\partial w_j(\vec{\nabla}_{Q,k}^T)}{\partial \mu_{ij}} = \prod_{\substack{l=1 \\ l \neq j}}^o F_{lj}(v_{lk}) \frac{\partial F_{ij}(v_{ik})}{\partial \mu_{ij}} \quad (\text{Gleichung 4.2.4.9})$$

Zuletzt in der Hierarchie des neuronalen Fuzzy Netzes steht die Gauss'sche Zugehörigkeitsfunktion  $F_{ij}$ , deren Ableitung den letzten Gradienten in Gleichung 4.2.4.10 ergibt.

$$\begin{aligned} \frac{\partial F_{ij}(v_{ik})}{\partial \mu_{ij}} &= e^{-\frac{(v_{ik}-\mu_{ij})^2}{(2\sigma_{ij}^2)}} \frac{(v_{ik}-\mu_{ij})}{\sigma_{ij}^2} \\ &= F_{ij}(v_{ik}) \frac{(v_{ik}-\mu_{ij})}{\sigma_{ij}^2} \end{aligned} \quad (\text{Gleichung 4.2.4.10})$$

Der Gradient  $\frac{\partial E}{\partial \sigma_{ij}}$  der Varianz  $\sigma_{ij}$  ist durch die Gleichung 4.2.4.11 beschrieben, der sich bis auf den partiellen Gradienten der Gauss'schen Zugehörigkeitsfunktion  $F_{ij}$  mit der Ableitung des Gradienten  $\frac{\partial E}{\partial \mu_{ij}}$  für den Mittelwert  $\mu_{ij}$  deckt. Aufgrund der Ähnlichkeit der partiellen Gradienten werden diese an

$$\frac{\partial E}{\partial \sigma_{ij}} = \sum_{k=t}^{t+n} \left( -\varepsilon(\vec{\nabla}_{Q,k}) \sum_{l=1}^o f_l(\vec{\nabla}_{Q,k}^T) \frac{\partial \omega_l(\vec{\nabla}_{Q,k}^T)}{\partial \sigma_{ij}} \right) \quad (\text{Gleichung 4.2.4.11})$$

dieser Stelle nicht beschrieben, sondern nur der letzte Gradient  $\frac{\partial F_{ij}(v_{ik})}{\partial \sigma_{ij}}$  über die Gleichung 4.2.4.12 abgeleitet.

$$\begin{aligned} \frac{\partial F_{ij}(v_{ik})}{\partial \sigma_{ij}} &= e^{-\frac{(v_{ik}-\mu_{ij})^2}{(2\sigma_{ij}^2)}} \frac{(v_{ik}-\mu_{ij})^2}{\sigma_{ij}^3} \\ &= F_{ij}(v_{ik}) \frac{(v_{ik}-\mu_{ij})^2}{\sigma_{ij}^3} \end{aligned} \quad (\text{Gleichung 4.2.4.12})$$

### Erneuerung der Parameter nach jeder Epoche

Mit dem Gradientenabstieg wird iterativ innerhalb der Fehlerfläche, die sich durch die Funktion  $E(V_T)$  und den zu bestimmenden variablen Parametern  $(\sigma_{ij}^2, \mu_{ij}, a_{ij})$  ergibt, ein Minimum gesucht. Das Verfahren stoppt, wenn ein Minimum in der Fläche gefunden wurde. Die Iteration erfolgt mittels mehreren Trainings-Epochen, bei denen jeweils die Werte der Gradienten für den Datensatz  $V_T$  in Abhängigkeit von der gewünschten Ausgabe  $\mathbf{u}$  berechnet wird. Nach jeder Epoche werden die Parameter  $(\sigma_{ij}^2, \mu_{ij}, a_{ij})$  mit den berechneten Gradienten erneuert, wobei eine Schrittweite  $\lambda$  den Bewegungsradius in der Fehlerfläche festlegt. Die Erneuerungen [Chi96] pro Epoche sind wie folgt definiert:

$$\mu_{ij} \leftarrow \mu_{ij} - \lambda \frac{\partial E}{\partial \mu_{ij}}, \quad \sigma_{ij} \leftarrow \sigma_{ij} - \lambda \frac{\partial E}{\partial \sigma_{ij}} \quad \text{und} \quad a_{ij} \leftarrow a_{ij} - \lambda \frac{\partial E}{\partial a_{ij}} \quad (\text{Definition 4.2.4.27})$$

Die Identifikation eines Minimums ist recht problematisch, da nicht zwischen lokalen und globalen Minima unterschieden werden kann. Auch können Plateaus nicht klar bestimmt werden. Hierbei ist ein anfänglicher Wert der Parameter wichtig, der schon nahe an einem Minimum liegt, was durch die anfängliche Methode der kleinsten Quadrate und das subtraktive Clustering wahrscheinlich ist.

Auch die Wahl einer guten Schrittweite ist für den Erfolg des Verfahrens essentiell. Wird eine zu große Schrittweite gewählt, so können Minima übersprungen werden oder es wird innerhalb einer Schlucht hin und her gesprungen ohne ihr Minimum zu erreichen. Wird die Schrittweite zu klein gewählt, so werden sehr viele Iterationen gebraucht bis ein Minimum erreicht wird. Oft wird die Schrittweite pro Epoche verändert, wobei mit einer hohen angefangen und diese pro Iteration verkleinert wird. Auch ein Momentum-Term kann die Konvergenz des Verfahrens verbessern.

### 4.2.5 Hybrider lernender Algorithmus zum Tuning der Zugehörigkeitsfunktionen $F_{vc}$ und der linearen Konstanten $\mathcal{A} = (a_{ik})$

Bei dem hybriden Ansatz zum Tuning der Parameter  $(\sigma_{ij}^2, \mu_{ij})$  der nichtlinearen funktionalen Zugehörigkeit  $F_{ij}$  und der Gewichte  $a_{ij}$  der linearen Konsequenzen  $f_j$ , wird eine Mischung aus Backpropagation-Training und Methode der kleinsten Quadrate benutzt.

In der Rückwärtsrichtung von Ausgang zum Eingang des neuronalen Fuzzy Netzes (Abbildung 12) werden mittels des in Kapitel 4.2.4 beschriebenen Backpropagation-Trainings neue Werte für die Parameter  $(\sigma_{ij}^2, \mu_{ij})$  berechnet. Hierzu werden die schon definierten Gradienten  $\frac{\partial E}{\partial \mu_{ij}}$  (Gleichung 4.2.4.6) und  $\frac{\partial E}{\partial \sigma_{ij}}$  (Gleichung 4.2.4.11) benötigt, die pro Epoche des Trainings erneuerte funktionale Zugehörigkeiten  $F_{ij}$  liefern.

Da die funktionalen Zugehörigkeiten  $F_{ij}$  in der Rückwärtsrichtung neu berechnet wurden, können nun in der Vorwärtsrichtung neue lineare Konsequenzen  $f_j$  berechnet werden. Die Methode der kleinsten Quadrate bildet das Minimum der linearen Gleichung über  $\vec{a}_j$ , die wiederum abhängig ist von den funktionalen Zugehörigkeiten die über die Gewichte in der Matrix  $W_i$  repräsentiert sind.

Eine Epoche, die sich über die gesamte Matrix  $V_T$  bildet, setzt sich also aus einer Rückwärtsbewegung des

Backpropagation-Training und einer Vorwärtsbewegung der Methode der kleinsten Quadrate zusammen. Diese hybride Methode ist in der Epoche zu beenden, in der eine Verschlechterung gegenüber dem Fehler einer Check-Menge erreicht wird.

### 4.3 Das Stift-Artefakt

Die theoretischen Annahmen, die in Kapitel 2.2.1 gemacht wurden, erfordern eine Verifikation anhand eines realen Systems, das die dynamischen Daten aus Sensoren erhält. Das Sugeno-System  $\hat{\mathbf{S}}_{Q_{lin}}$  ergibt sich anhand einer Datenmenge mittels adaptiver Algorithmen (Kapitel 4.2). Diese Datenmenge muss durch ein reales System geliefert werden, da reale dynamische Verhaltensweisen von Sensoren kaum simuliert werden können und eine Simulation auch keine analytische Grundlage liefern würde. Es muss auch ein kontextueller Algorithmus  $\mathbf{A}_C$  vorhanden sein, der hinsichtlich seiner Qualität pro Datum zu beurteilen ist. Dieser Algorithmus  $\mathbf{A}_C$ , der die Erkennung der kontextuellen Zustände leistet, kann in seiner Analyse nicht vorhergesagt werden, da Wissen über seine Funktionsweise nicht gefordert ist.

Mit einem somit geforderten realen System können dann dynamische Verhaltensweisen einer trainierten Sugeno-Logik  $\hat{\mathbf{S}}_{Q_{lin}}$  für Qualitäten analysiert werden, und dies nicht nur für Wechsel von kontextuellen Zuständen. Es kann auch ein Verhalten der Logik innerhalb kontextueller Zustände beobachtet werden und dieses mit anderen kontextuellen Zuständen verglichen werden.

Ein System das verschiedene Algorithmen der kontextuellen Analyse von sensorischen Daten bereitstellt, ist das Stift-Artefakt. Dieses besteht aus einer Hardware-Komponente, einen im TecO entwickelten Particle Computer, und einer variablen Software-Komponente, die wiederum aus einer Einheit zur Kommunikation und einer Algorithmik zur Kontexterkenkung zusammengesetzt ist. Der verwendete Particle Computer ist in Abbildung 13 in einzelnen Teilen (links) und montiert auf einem Stift (rechts) dargestellt. Die Partikel-Hardware besteht aus einem Partikel 2/29, der ein drahtloses Kommunikationsmodul

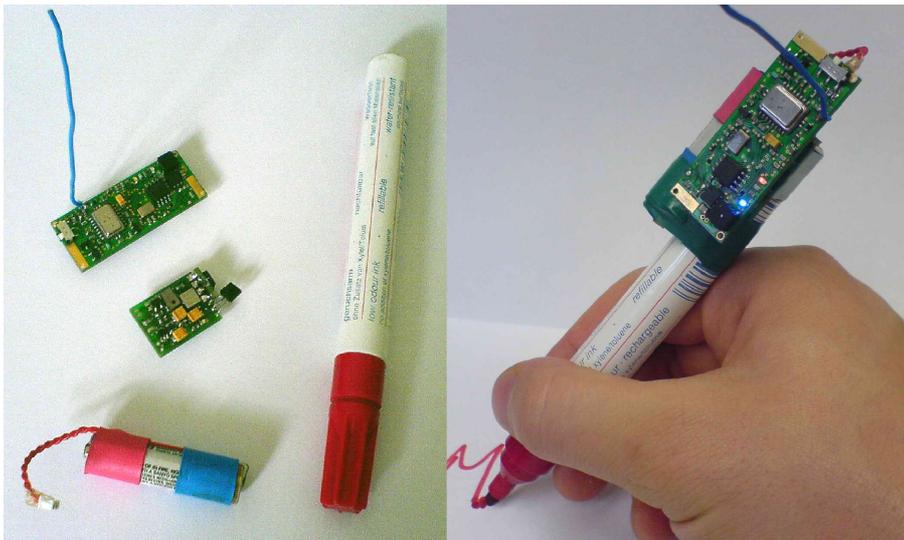


Abbildung 13: Bestandteile des Stift-Artefakts in Einzelteilen (links) und montiert auf dem Stift (rechts)

TR1001, einen Mikrocontroller PIC18LF6720 und eine Schnittstelle für andere Module enthält, und einem Sensor-Board 1/92, das verschiedene Sensoren vereint. Auf dem Sensor-Board befinden sich auch die verwendeten Beschleunigungs-Sensoren ADXL 210, die Werte für Beschleunigungen in x-, y- und z-Richtung liefern. Um eine ausreichende Stromversorgung für einige Stunden zu gewährleisten wurde eine AAA-Batterie gewählt. Die spezifischen technischen Daten zu dieser Hardware, befinden sich auf der Webseite [Par06] der Particle Computer.

Die Software die auf dem Mikrocontroller ausgeführt wird, wurde von Pablo Estévez im Rahmen seiner Master-These [Est06] entwickelt. Diese Software besteht entweder aus einem neuronalen Netz  $\mathbf{N}_C^{NeuroPen}$  oder einem unscharfen Sugeno-Regelsystem  $\mathbf{S}_C^{FuzzyPen}$  jeweils mit Klassifikation, welche beide zur Erkennung von drei verschiedenen kontextuellen Zuständen eingesetzt werden. Beide Algorithmen werden mit Modulen von Matlab erstellt und in C-Code übersetzt auf dem Stift-Artefakt ausgeführt. Das kontextuelle Sugeno-Regelsystem  $\mathbf{S}_C^{FuzzyPen}$  wird in Kapitel 4.3.1 beschrieben und das neuronale Netz  $\mathbf{N}_C^{NeuroPen}$  zur Kontexterkenkung wird in Kapitel 4.3.2 beschrieben. Die ausführliche Anleitung zur Erstellung der beiden Systeme findet sich in den Arbeiten von Estévez [Est05][Est06].

### 4.3.1 Fuzzy Logik zur Kontextererkennung $S_C$

Das unscharfe System zur Kontextererkennung  $S_C^{FuzzyPen}$  für das Stift-Artefakt unterscheidet sich kaum vom unscharfen System zur Qualitätsanalyse  $S_Q^{FuzzyPen}$ , da in beiden Systemen die Zuordnung von vorverarbeitetem Sensor-Datum zu Kontextklasse oder Qualitätsmaß über eine unscharfe Sugeno-Logik erfolgt. Der Vektor der vorverarbeiteten Eingabe  $\vec{v}_C^{FuzzyPen} := (v_1, v_2, v_3)$  des kontextuellen Sugeno-Systems  $\hat{S}_{C_{lin}}^{FuzzyPen}$  ist wieder um eine Dimension kleiner, als der Vektor der qualitativen Klassifikation  $\vec{v}_Q^{FuzzyPen} := (v_1, v_2, v_3, v_4)$ , denn dieser beinhaltet noch den Identifikator des erkannten kontextuellen Zustandes  $v_4 := S_C^{FuzzyPen}(\vec{v}_C^{FuzzyPen})$ .

Das System zur Kontextererkennung ist im Umfang der Regeln, die die unscharfe Logik bestimmen, wesentlich komplexer, da zur Reduktion des Daten-Raumes eine Aufteilung anhand eines Gitters eingesetzt wird und nicht ein subtraktives Clustering. Diese Gitter-Partitionierung und die Auswirkung auf das resultierende kontextuelle Sugeno-System  $\hat{S}_{C_{lin}}^{FuzzyPen}$  werden im Folgenden beschrieben.

Das gesamte System zur kontextuellen Klassifikation und qualitativen Analyse ist mit dessen Schnittstellen in Abbildung 14 dargestellt. Die weiteren Systeme, die in der Hierarchie der Verarbeitung von

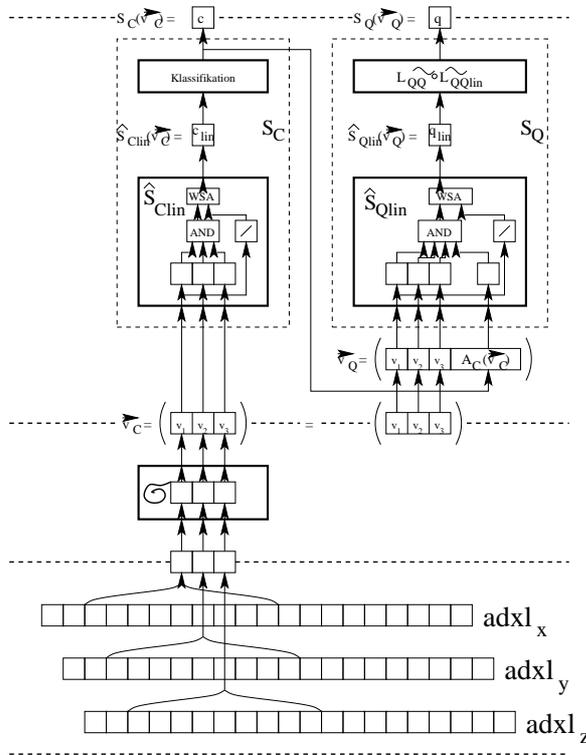


Abbildung 14: Zusammenhänge von kontextueller Fuzzy Logik  $S_C$  und Spezifikationen der qualitativen Analyse  $S_Q$  für das Stift-Artefakt

Sensordatum hin zu Klassenidentifikatoren auftreten, sind in der gleichen Abbildung dargestellt. Zuerst werden die Daten der Sensoren in einen kontinuierlichen Speicher eingespeist, der bei ihrem Auslesen wieder frei wird. Das Auslesen erfolgt durch den ersten Schritt der Verarbeitung, die als Resultat einen Vektor an der Schnittstelle des Systems übergibt. Die einzelnen Komponenten dieser Hierarchie werden nun in folgenden Abschnitten für dieses Stift-Artefakt, mit dessen kontextuellem Sugeno-System beschrieben. Wie dieses Sugeno-System zur Kontextererkennung automatisiert erstellt wird, ist auch in diesem Kapitel erklärt.

#### Auslesen der Sensoren für die Beschleunigung

Die Grundlage der Kontextererkennung, und aufgrund dessen auch der Qualitätsanalyse, bilden die Daten der Sensoren, die über den Mikrocontroller zu äquidistanten Zeitpunkten gemessen werden. In dem betrachteten Fall des Stift-Artefaktes messen die verwendeten Sensoren (adxl) die Beschleunigung in x-, y- und z-Richtung. Sie bilden einen kontinuierlichen Strom von Messwerten, der bei Aktivierung des Arte-

faktes zum Zeitpunkt 0 beginnt und bei der Deaktivierung zum Zeitpunkt  $\mathcal{T}$  endet. Somit ergibt sich für jede Richtung eine Menge von Messwerten, die wie folgt definiert ist:

$$X := \{x_0, x_1, \dots, x_t, x_{t+1}, \dots, x_{t+n}, \dots, x_{\mathcal{T}}\} \quad (\text{Definition 4.3.1.1})$$

$$Y := \{y_0, y_1, \dots, y_t, y_{t+1}, \dots, y_{t+n}, \dots, y_{\mathcal{T}}\} \quad (\text{Definition 4.3.1.2})$$

$$Z := \{z_0, z_1, \dots, z_t, z_{t+1}, \dots, z_{t+n}, \dots, z_{\mathcal{T}}\} \quad (\text{Definition 4.3.1.3})$$

Die Mengen  $X$ ,  $Y$  und  $Z$ , werden über den gesamten Zeitraum der Aktivität des Artefaktes gebildet. Zur Verarbeitung dieser Messwerte müssen diese Mengen nun in kleinere Mengen unterteilt werden, die in einem kleineren Zeitraum  $T$  gemessen wurden. Diese Teilmengen können dann an der unteren Schnittstelle (Abbildung 14, zweit unterste gestrichelte Linie) der Vorverarbeitung an diese weiter gegeben werden. Die Teilmengen sind über den Zeitraum  $T := \{t, t+1, \dots, t+n\}$ , der eine Breite eines Fensters der Betrachtung beschreibt, zu bilden und aus Gründen der Weiterverarbeitung als Vektoren  $\vec{\mathbf{x}}_T^{Pen}$ ,  $\vec{\mathbf{y}}_T^{Pen}$  und  $\vec{\mathbf{z}}_T^{Pen}$  zu bezeichnen. Die Breite dieses Fensters und sogleich die Größe des Zeitraumes ist  $n+1$ . Die Anzahl der Messwerte  $n+1$  die für die Vorverarbeitung und danach für den jeweiligen kontextuellen Klassifikator benötigt werden, hängt von dem erwünschten Resultat der Klassifikation und dem Abstand  $\tau$  zwischen den einzelnen Messwerten ab. Ist ein hochdynamischer kontextueller Zustandswechsel zu erwarten, so muss  $n+1$  klein sein. Sind die Übergänge von einem zum anderen kontextuellen Zustand eher langsam, so kann  $n$  größer gewählt werden.

Der zeitliche Abstand  $\tau := j - i$  zwischen zwei Messwerten zum Zeitpunkt  $i$  und dem darauf folgenden Zeitpunkt  $j$ , ist für alle sukzessiven Messwerte gleich, was die zuvor geforderte Äquidistanz ausdrückt. Dieser Abstand  $\tau$  ist ebenso für die Größe  $n+1$  von Bedeutung wie die Dynamik der kontextuellen Zustandswechsel, da diese beiden korrelieren.

#### Vorverarbeitung der Sensordaten durch Berechnung der Standardabweichung

Die Messdaten sind Werte der Beschleunigung, die auf kontextuelle Zustände eines Stiftes abgebildet werden sollen. Die Zustände, die zu erkennen sind, sind 'liegend', 'schreibend' und 'bewegt'. Die Aufgabe der Vorverarbeitung der Messdaten für den Algorithmus zur Klassifikation ist nun eine Reduktion der Dimensionalität von  $n+1$  auf 1, und dies für jede Richtung des Raumes (x, y und z), deren Beschleunigung gemessen wird. Es können auch mehrere parallele Schritte der Vorverarbeitung benötigt werden, je nach dem welche Informationen der Algorithmus der kontextuellen Klassifikation benötigt. In dem zu beschreibenden Fall einer unscharfen Sugeno-Logik wird nur eine Methodik der Reduktion benutzt, da zum einen eine höhere Dimension des Eingabevektors  $\vec{\mathbf{v}}_C^{FuzzyPen}$ , des kontextuellen Sugeno-Systems  $\hat{\mathbf{S}}_{C_{lin}}^{FuzzyPen}$ , eine wesentlich größere Komplexität dieses bedingt und zum anderen dieses System mit einer geschickten Reduktion eine gute Erkennung leistet. Eine zu hohe Dimensionalität des Eingabevektors, und einer somit erhöhten Information dessen, kann für viele Algorithmen der Klassifikation sogar ein schlechteres Resultat bedingen.

Aus genannten Gründen und wegen der Unterscheidbarkeit der kontextuellen Zustände des Stift-Artefaktes anhand der Abweichungen der Messdaten voneinander, wurde die Standardabweichung, als Reduktion der Dimensionalität für jede Richtung der gemessenen Beschleunigung, gewählt. Für die x-Richtung berechnet sich diese Standardabweichung, so wie analog für die Richtungen y und z, wie folgt:

$$\sigma(\vec{\mathbf{x}}_T^{Pen}) = \sqrt{\frac{1}{n+1} \sum_{i=t}^{t+n} (x_i - \mu(\vec{\mathbf{x}}_T^{Pen}))^2} \quad (\text{Gleichung 4.3.1.1})$$

Die in der Gleichung der Standardabweichung  $\sigma$  verwendete Funktion des Mittelwertes  $\mu$  - welches Funktionssymbol nicht mit dem der funktionalen Zugehörigkeit zu unscharfen Mengen zu verwechseln ist - ist in folgender Gleichung beschrieben:

$$\mu(\vec{\mathbf{x}}_T^{Pen}) = \frac{1}{n+1} \sum_{i=t}^{t+n} x_i \quad (\text{Gleichung 4.3.1.2})$$

Der 3-dimensionale Vektor  $\vec{\mathbf{v}}_C^{FuzzyPen}$  (Definition 4.3.1.6), der an der Schnittstelle des gesamt Systems zur Kontexterkenkung und Qualitätsanalyse (Abbildung 14, zweit oberste gestrichelte Linie) des Stift-Artefaktes übergeben wird, ist Element der Universen  $\mathcal{V}_1 \times \mathcal{V}_2 \times \mathcal{V}_3$  der Standardabweichungen.

$$\vec{\mathbf{v}}_C^{FuzzyPen} \in \mathcal{V}_1 \times \mathcal{V}_2 \times \mathcal{V}_3 \subset \mathbb{R} \quad (\text{Definition 4.3.1.6})$$

Die einzelnen Elemente sind wie folgt, über die Standardabweichungen der zum Zeitraum  $T$  gemachten Messwerte der Beschleunigungen in x-, y- und z-Richtung, zu berechnen:

$$\begin{aligned} \vec{\mathbf{v}}_C^{FuzzyPen} &:= (\sigma(x_t, \dots, x_{t+n}), \sigma(y_t, \dots, y_{t+n}), \sigma(z_t, \dots, z_{t+n})) \\ &= (\sigma(\vec{\mathbf{x}}_T^{Pen}), \sigma(\vec{\mathbf{y}}_T^{Pen}), \sigma(\vec{\mathbf{z}}_T^{Pen})) \\ &= (v_1, v_2, v_3) \end{aligned} \quad (\text{Definition 4.3.1.7})$$

Im folgenden Abschnitt werden nun das kontextuelle Sugeno-System  $\widehat{\mathbf{S}}_{C_{lin}}^{FuzzyPen}$  und der Unterschied zum qualitativen Sugeno-System  $\widehat{\mathbf{S}}_{Q_{lin}}^{FuzzyPen}$  beschrieben, die beide in einer Variante der Software des Stift-Artefaktes kombiniert sind. Die beiden Systeme werden pro Vektor  $\vec{\mathbf{v}}_C^{FuzzyPen}$  sequentiell ausgeführt, wobei der kontextuelle Klassifikator  $\mathbf{S}_C$  das vierte Element des erweiterten qualitativen Eingabevektors  $\vec{\mathbf{v}}_Q^{FuzzyPen}$  liefert und somit zuerst ausgeführt wird.

### Das kontextuelle Sugeno-System

Die Komplexität und Dimensionalität der Eingabe bestimmen nicht nur die Anzahl der Regeln der Sugeno-Systeme, sondern auch die gewählte Methode des Clusterings, die im Rahmen der adaptiven Algorithmen verwendet wird. Wie im Allgemeinen ein qualitatives Sugeno-System  $\widehat{\mathbf{S}}_{Q_{lin}}$  durch adaptive Algorithmen zu erstellen ist, wurde in Kapitel 4.2 beschrieben. In diesem Fall wurde eine Methode des subtraktiven Clusterings gezeigt, die für die Erstellung der qualitativen Sugeno-Systeme verwendet werden soll. Im Rahmen der kontextuellen Sugeno-Systeme wird eine andere Methode zur Erstellung des initialen Sugeno-Systems zur Kontexterkenkung verwendet, die Gitter-Partitionierung. Diese Methode teilt das Universum der Eingabe gleichmäßig in eine gewählte Anzahl von homogenen Zugehörigkeitsfunktionen auf. Der Mittelwert der ersten Zugehörigkeitsfunktion einer Dimension der Eingabe wird auf die untere Grenze des entsprechenden Universums gesetzt. Die letzte Zugehörigkeitsfunktion erhält als Mittelwert die obere Grenze dieses Universums. Alle weiteren Funktionen der Zugehörigkeit werden in dem universalen Intervall der Dimension äquidistant verteilt. Die Spreizung der Funktionen wird so bestimmt, dass sich die benachbarten Funktionen in ihrem Funktionswert von  $\frac{1}{2}$  schneiden. Die Art der funktionalen Zugehörigkeit kann ebenso gewählt werden wie die Anzahl der Funktionen pro Dimension des Eingabevektors. Auch die durch die Spezifizierung der Sugeno-Logik gegebenen Konsequenzen der Implikation, ob linear oder konstant, können durch den Entwickler spezifiziert werden. Diese Freiheitsgrade sind bei der Methode des subtraktiven Clusterings nicht gegeben, da Gaussfunktionen und lineare Konsequenz gefordert werden. Die Anzahl der funktionalen Zugehörigkeiten wird durch den Algorithmus des subtraktiven Clusterings selbst-organisierend berechnet.

In der Arbeit von Estévez [Est05] werden lineare Trapez-Zugehörigkeitsfunktionen und lineare Konsequenzen vorgeschlagen, was auch in eigenen Versuchen zu guten Resultaten der Klassifikation geführt hat. Die Vorteile der Effizienz, die lineare Trapez-Funktionen gegenüber Gauss-Funktionen liefern, werden durch den wesentlich komplexeren Regelsatz eigentlich wieder aufgehoben, denn in der regelbasierten Verknüpfung der funktionalen Zugehörigkeiten werden alle Kombinationen der verschiedenen Dimensionen verwirklicht. Die Anzahl  $o$  der Cluster, die durch das subtraktive Clustering bestimmt wurde hat direkt die Anzahl  $o$  der Regeln bestimmt. Die Anzahl  $o$  der Regeln, die durch die Methode der Gitter-Partitionierung bedingt werden, berechnet man anhand der Dimensionalität  $m$  des Eingabevektors und der Anzahl  $o_i$  der funktionalen Zugehörigkeiten pro Dimension  $i$  wie folgt:

$$o = \prod_{i=1}^m o_i \quad (\text{Gleichung 4.3.1.3})$$

Da dieses kontextuelle Sugeno-Regelsystem als Softwarekomponente Bestandteil des Mikrocontroller basierten Stift-Artefaktes ist, muss eine erhöhte Komplexität des Regelsatzes gegenüber einer effizienteren Berechnung der Zugehörigkeit abgewogen werden. Auch sind unterschiedliche Resultate der Klassifikation der Sugeno-Logik zu beobachten, was ebenso einer Entscheidung des Entwicklers bedarf. Diese Abwägungen sind jedoch nicht Bestandteil dieser Arbeit, da diese die Qualitätsanalyse und nicht die Kontexterkenkung ergründet. Aus diesem Grund wurde den Vorschlägen von Estévez [Est05], zur Konstruktion des kontextuellen Sugeno-Systems  $\widehat{\mathbf{S}}_{C_{lin}}^{FuzzyPen}$ , Folge geleistet.

Das kontextuelle Sugeno-Regelsystem  $\widehat{\mathbf{S}}_{C_{lin}}^{FuzzyPen}$  besitzt 3 lineare Zugehörigkeitsfunktionen pro Dimension des Eingabevektors, was eine Anzahl der Regeln von  $o = 3 \cdot 3 \cdot 3 = 27$  ergibt. Die Anzahl  $o$  der linearen Konsequenzen entspricht der Anzahl  $o$  der Regeln, wie auch schon zuvor bei dem qualitativen

Sugeno-Regelsystem. Die funktionalen Zugehörigkeiten sind Trapez-Fuktionen, die wie folgt zu berechnen sind:

$$F_{ij}(v_i) := \begin{cases} L(\frac{\underline{m}-x}{\alpha}) & \text{für } x < \underline{m} \\ 1 & \text{für } \underline{m} \leq x \leq \overline{m} \\ R(\frac{\overline{m}-x}{\alpha}) & \text{für } x > \overline{m}. \end{cases} \quad (\text{Definition 4.3.1.9})$$

Die Trapez-Funktionen  $F_{ij}$  sind symmetrisch und werden anhand der LR-Funktion berechnet, die schon bei der Normalisierung des linearen Resultates des qualitativen Sugeno-Systemes durch die Definition 4.1.5.21 beschrieben wurde und hier wiederholt wird:

$$L(x) = R(x) = \begin{cases} (1 - x) & \text{für } 0 < x < 1 \\ 0 & \text{sonst.} \end{cases}$$

Die Gitter-Partitionierung bedingt statische Parameter der funktionalen Zugehörigkeiten nach deren Erstellung. Die Parameter der linearen Konsequenzen sind somit die einzigen Funktionen, die durch die überwachten lernenden Verfahren verändert werden. Die überwachten Verfahren benötigen wieder einen Vektor  $\vec{u}$  von gewünschten Ausgabedaten als Referenz. Die gewünschten Ausgaben sind durch die Identifikatoren der Klassen vom Entwickler zur Verfügung zu stellen. Im Falle des Stift-Artefaktes ist die Klasse 'liegend' durch 0, 'schreibend' durch 127 und 'bewegt' durch 255 zu identifizieren. Um bei großen Datensätzen die Bezeichnung eines jeden Datum  $\vec{v}_{C, u_i}^{FuzzyPen}$  der Eingabe mit dem entsprechenden Identifikator  $u_i$  zu erleichtern, wird empfohlen Datensätze für jeden kontextuellen Zustand separat zu erstellen. Daraus ergeben sich die folgenden Matrizen  $V_0$ ,  $V_{127}$  und  $V_{255}$ :

$$V_0 = \begin{pmatrix} v_{1u_1} & v_{1u_2} & \cdots & v_{1u_k} \\ v_{2u_1} & v_{2u_2} & \cdots & v_{2u_k} \\ \vdots & \vdots & \ddots & \vdots \\ v_{mu_1} & v_{mu_2} & \cdots & v_{mu_k} \end{pmatrix} \quad (\text{Definition 4.3.1.10})$$

$$V_{127} = \begin{pmatrix} v_{1u_{k+1}} & v_{1u_{k+2}} & \cdots & v_{1u_{k+l}} \\ v_{2u_{k+1}} & v_{2u_{k+2}} & \cdots & v_{2u_{k+l}} \\ \vdots & \vdots & \ddots & \vdots \\ v_{mu_{k+1}} & v_{mu_{k+2}} & \cdots & v_{mu_{k+l}} \end{pmatrix} \quad (\text{Definition 4.3.1.11})$$

$$V_{255} = \begin{pmatrix} v_{1u_{k+l+1}} & v_{1u_{k+l+2}} & \cdots & v_{1u_n} \\ v_{2u_{k+l+1}} & v_{2u_{k+l+2}} & \cdots & v_{2u_n} \\ \vdots & \vdots & \ddots & \vdots \\ v_{mu_{k+l+1}} & v_{mu_{k+l+2}} & \cdots & v_{mu_n} \end{pmatrix} \quad (\text{Definition 4.3.1.12})$$

Der Vektor der gewünschten Ausgabe des kontextuellen Sugeno-Regelsystems ist wie folgt aus den einzelnen Identifikatoren  $C = \{0, 127, 255\}$  aufgebaut:

$$\begin{aligned} \vec{u} &= (u_1, u_2, \dots, u_k, \quad u_{k+1}, \dots, u_{k+l}, \quad u_{k+l+1}, \dots, u_n) \\ &= (\underbrace{0, 0, \dots, 0}_{k \text{ - mal}}, \quad \underbrace{127, \dots, 127}_{l \text{ - mal}}, \quad \underbrace{255, \dots, 255}_{(n - (k + l)) \text{ - mal}}) \end{aligned} \quad (\text{Definition 4.3.1.13})$$

Diese Unterscheidung bezüglich den einzelnen kontextuellen Zuständen ist bei der Sugeno-Logik zur Qualitätsanalyse nicht notwendig, da deren gewünschte Ausgabe sich direkt aus dem Unterschied von gewünschter  $u_i$  und realer Ausgabe  $\mathbf{S}_C^{FuzzyPen}(\vec{v}_{C, u_i}^{FuzzyPen})$  des kontextuellen Sugeno-Systems  $\mathbf{S}_C^{FuzzyPen}$  ergibt.

Das neuronale Fuzzy Netz entspricht wieder der Interpretation des initialen kontextuellen Sugeno-Regelsystems, jedoch Aufgrund der erhöhten Anzahl der Regeln dessen, gegenüber einem qualitativen Sugeno-System, in einer komplexeren Netzstruktur. Dieses neuronale Fuzzy Netz ist in einer abstrahierten Form des realen Netzes in Abbildung 15 auf der rechten Seite einzusehen. Dieses neuronale Fuzzy Netz wird mit der hybriden Methode anhand der Vereinigung der Matrizen  $V_0$ ,  $V_{127}$  und  $V_{255}$  trainiert. Der Fehler  $E(V_0, V_{127}, V_{255})$ , den es zu verbessern gilt, berechnet sich, wie schon in Definition 4.2.4.16 beschrieben, durch die Einzelfehler (Definition 4.2.4.17) der realen Ausgabe pro Eingabevektor zu gewünschter Ausgabe. Nach jeder Epoche des Trainings kann dieser Fehler über die Trainingsdaten verbessert werden, was aber nur in einem geringen Umfang geschieht und für eine Anzahl von 10 Epochen in Abbildung 16 dargestellt ist. Das trainierte kontextuelle Sugeno-System ist wieder im Schluss und der Defuzzifizierung

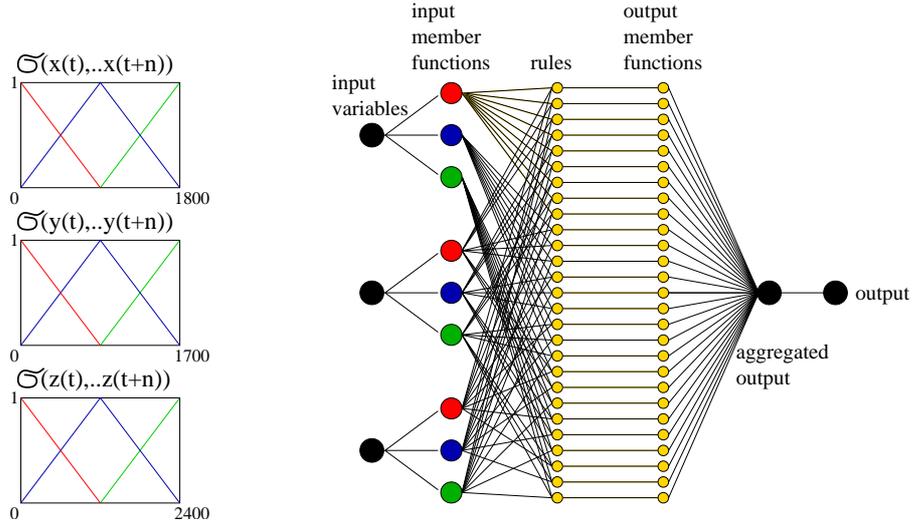


Abbildung 15: Funktionale Zugehörigkeiten  $F_{ij}$  der Eingabe  $v_i$  für  $i = 1, 2, 3$  (links); neuronales Fuzzy Netz zum Training des kontextuellen Sugeno-Systems  $\hat{\mathbf{S}}_{C_{lin}}^{FuzzyPen}$  (rechts)

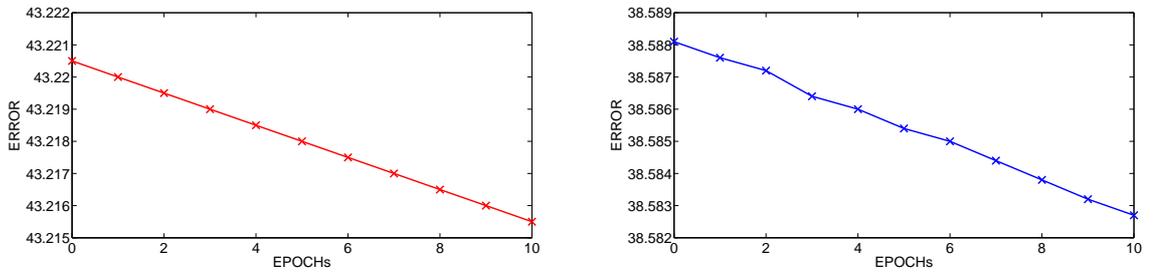


Abbildung 16: Fehler  $E(V_0, V_{127}, V_{255})$  (links) und Fehler  $E(V_{CHECK_{0,127,255}})$  (rechts) einer Check-Datenmenge für jeweils 10 Epochen des Trainings

gleich dem qualitativen Sugeno-System und benutzt eine normalisierte gewichtete Summe. Diese gewichtete normalisierte Summe ist für das kontextuelle Sugeno-System  $\hat{\mathbf{S}}_{C_{lin}}^{FuzzyPen}$  wie folgt zu berechnen:

$$\hat{\mathbf{S}}_{C_{lin}}^{FuzzyPen}(\vec{\nabla}_C^{FuzzyPen}) := \frac{\sum_{j=1}^{27} w_j(\vec{\nabla}_C^{FuzzyPen}) f_j(\vec{\nabla}_C^{FuzzyPen})}{\sum_{j=1}^{27} w_j(\vec{\nabla}_C^{FuzzyPen})} \quad (\text{Definition 4.3.1.14})$$

Diese Abbildung des kontextuellen Sugeno-Systems (Definition 4.3.1.15) resultiert in Werten der Menge  $C_{lin}$ . Diese Menge  $C_{lin}$  beinhaltet die Identifikatoren der kontextuellen Klassen und ist ähnlich wie  $[0, 1] \subset Q_{lin}$  durch die Teilmenge  $\{0, 127, 255\} \subset C_{lin}$  nach unten beschränkt. Die obere Beschränkung ist die

$$\hat{\mathbf{S}}_{C_{lin}}^{FuzzyPen} : \begin{cases} \mathcal{V}_1 \times \dots \times \mathcal{V}_M & \longrightarrow Q_{lin}, \text{ mit } \{0, 127, 255\} \subset C_{lin} \subset \mathbb{R} \\ (v_1, v_2, \dots, v_m) & \longmapsto \hat{\mathbf{S}}_{C_{lin}}(v_1, v_2, \dots, v_m) \end{cases} \quad (\text{Definition 4.3.1.15})$$

Menge der reellen Zahlen mit  $C_{lin} \subset \mathbb{R}$ . Die Werte der Menge  $C_{lin}$  werden nun noch einem eindeutigen Identifikator der jeweiligen Klasse zugeordnet, was in Abbildung 14 auf der linken Seite mit der Box 'Klassifikation' symbolisiert wird.

### Klassifikation des Resultates des kontextuellen Sugeno-Systems

Die Klassifikation der Elemente aus der Menge  $C_{lin}$  ist wieder mit der Klassifikation der Elemente aus der Menge  $Q_{lin}$  zu vergleichen, die in Kapitel 4.1.5 beschrieben wurde. Es wird wieder der Zwischenschritt über eine unscharfe Menge von unscharfen Zahlen gegangen, jedoch ist die grösste Zugehörigkeit zu einer Zahl direkt Indiz für die Klasse. Die unscharfe Menge der unscharfen Zahlen ist  $\tilde{C} := \{\tilde{0}, \tilde{127}, \tilde{255}\}$ , wobei die unscharfe Zahl  $\tilde{0}$  durch  $\tilde{0} := \{(x, \mu_{\tilde{0}}(x)) : x \in C_{lin}\}$ ,  $\tilde{127}$  durch  $\tilde{127} := \{(x, \mu_{\tilde{127}}(x)) : x \in C_{lin}\}$  und

$\widetilde{255}$  durch  $\widetilde{255} := \{(x, \mu_{\widetilde{255}}(x)) : x \in C_{lin}\}$  definiert wird. Die funktionalen Zugehörigkeiten werden wieder durch LR-Darstellungen (Definition 4.3.1.16-Definition 4.3.1.18) definiert.

$$\mu_{\widetilde{0}}(x) := \begin{cases} L(\frac{-x}{127}) & \text{für } x < 0 \\ 1 & \text{für } x = 0 \\ R(\frac{x}{127}) & \text{für } x > 0. \end{cases} \quad (\text{Definition 4.3.1.16})$$

$$\mu_{\widetilde{127}}(x) := \begin{cases} L(1 - \frac{x}{127}) & \text{für } x < 127 \\ 1 & \text{für } x = 127 \\ R(\frac{x}{127} - 1) & \text{für } x > 127. \end{cases} \quad (\text{Definition 4.3.1.17})$$

$$\mu_{\widetilde{255}}(x) := \begin{cases} L(\frac{255-x}{127}) & \text{für } x < 255 \\ 1 & \text{für } x = 255 \\ R(\frac{x-255}{127}) & \text{für } x > 255. \end{cases} \quad (\text{Definition 4.3.1.18})$$

Die linearen LR-Funktionen sind schon in Definition 4.1.5.21 beschrieben worden. Die grafische Darstellung der funktionalen Zugehörigkeiten der unscharfen Zahlenmenge  $\widetilde{C}$  ist in Abbildung 17 einzusehen. Eine Entscheidung zu welcher unscharfen Zahl ein Element der Menge  $C_{lin}$  die grösste oder gleiche Zuge-

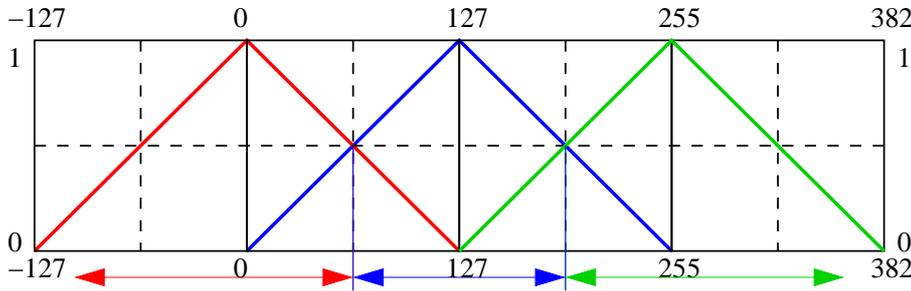


Abbildung 17: Funktionale Zugehörigkeiten für die unscharfen Zahlen  $\widetilde{0}$  (rot),  $\widetilde{127}$  (blau) und  $\widetilde{255}$  (grün)

hörigkeit besitzt, wird über das schon zuvor in Kapitel 4.1.5 benutzte Prädikat  $P_{\leq}$  getroffen. Da auch ein totaler Unterschied oder syntaktische Gleichheit von Größen der Menge  $C_{lin}$  entschieden werden muss, müssen noch zwei weitere Prädikate  $P_{\underline{=}}$  und  $P_{<}$  eingeführt werden. Das Prädikat  $P_{<}$  entscheidet ob der Wert  $x$  echt kleiner ist als der Wert  $y$ . Die syntaktische Gleichheit der Werte  $x$  und  $y$  wird mit dem Prädikat  $P_{\underline{=}}(x, y)$  entschieden.

Es wird kein Umweg über die Interpretation einer unscharfen Logik gegangen, denn es muss kein beschreibender Term in unscharfen Logik-Operatoren formuliert werden. Die funktionalen Zugehörigkeiten werden deshalb direkt als normale Funktionen in die Prädikate eingesetzt. Aus dieser Vereinfachung ergibt sich die klassifizierende Funktion  $\mathbf{K}_C^{Pen}$  wie folgt:

$$\mathbf{K}_C^{Pen}(x) := \begin{cases} [0] & \text{für } P_{<}(\mu_{\widetilde{127}}(x), \mu_{\widetilde{0}}(x)) \\ [127] & \text{für } P_{\leq}(\mu_{\widetilde{0}}(x), \mu_{\widetilde{127}}(x)) \\ & \text{und } P_{<}(\mu_{\widetilde{255}}(x), \mu_{\widetilde{127}}(x)) \\ [255] & \text{für } P_{\leq}(\mu_{\widetilde{127}}(x), \mu_{\widetilde{255}}(x)) \\ \emptyset & \text{für } P_{\underline{=}}(\mu_{\widetilde{0}}(x), 0) \\ & \text{und } P_{\underline{=}}(\mu_{\widetilde{127}}(x), 0) \\ & \text{und } P_{\underline{=}}(\mu_{\widetilde{255}}(x), 0) \end{cases} \quad (\text{Definition 4.3.1.19})$$

Die Funktion  $\mathbf{K}_C^{Pen}$  bildet aus der Menge  $C_{lin}$  in die Menge der kontextuellen Klassen  $C = \{[0], [127], [255]\}$  ab, was in Definition 4.3.1.20 beschrieben wird.

$$\mathbf{K}_C^{Pen} : \begin{cases} C_{lin} & \longrightarrow & C \\ c & \longmapsto & \mathbf{K}_C^{Pen}(c) \end{cases}, \text{ mit } C = \{[0], [127], [255]\} \quad (\text{Definition 4.3.1.20})$$

Aus der Konkatination der zuvor beschriebenen Abbildung des kontextuellen Sugeno-Systems  $\widehat{\mathbf{S}}_{C_{lin}}^{FuzzyPen}$  und der soeben definierten klassifizierenden Funktion  $\mathbf{K}_C^{Pen}$  ergibt sich die Abbildung des gesamten Systems zur kontextuellen Klassifizierung  $\mathbf{S}_C^{FuzzyPen}$ , welche wie folgt aus den einzelnen Bestandteilen zusammengesetzt ist:

$$\mathbf{S}_C^{FuzzyPen} : \begin{cases} \mathcal{V}_1 \times \mathcal{V}_2 \times \mathcal{V}_3 & \longrightarrow & C, \text{ mit } C = \{[0], [127], [255]\} \\ (v_1, v_2, v_3) & \longmapsto & \mathbf{K}_C^{Pen}(\widehat{\mathbf{S}}_{C_{lin}}^{FuzzyPen}(v_1, v_2, v_3)) \end{cases} \quad (\text{Definition 4.3.1.21})$$

Diese Abbildung  $\mathbf{S}_C^{FuzzyPen}$  bildet die Standardabweichungen der Sensordaten der Beschleunigung in x-, y- und z-Richtung auf die Menge der kontextuellen Klassen  $C$  ab. Der Identifikator der erkannten Kontextklasse wird an der Ausgangs-Schnittstelle (Abbildung 14, obere gestichelte Linie) des Stift-Artefaktes mit dem Identifikator der Qualität an weitere Artefakte oder Systeme weiter gegeben. Dieser Identifikator wird auch von dem System zur qualitativen Klassifikation als Ergänzung des Eingabevektors desselben benötigt, was in dem nächsten Abschnitt beschrieben wird.

### Das qualitative Sugeno-System

Das im letzten Abschnitt erarbeitete kontextuelle Sugeno-System führt nicht nur zur Erweiterung der Eingabe des qualitativen Sugeno-Systems, sondern bildet auch die gewünschte Ausgabe  $u$ , die für die überwachten adaptiven Algorithmen gebraucht wird. Der Eingabevektor  $\vec{\nabla}_Q^{FuzzyPen} \in \mathcal{V}_1 \times \mathcal{V}_2 \times \mathcal{V}_3 \times C \subset \mathbb{R}$  ist, wie schon in Kapitel 4.1.1 für den allgemeinen Fall beschrieben, um die Ausgabe des kontextuellen Systems  $\mathbf{S}_C^{FuzzyPen}$  erweitert, was in Definition 4.3.1.22 beschrieben ist.

$$\begin{aligned}
 \vec{\nabla}_Q^{FuzzyPen} &= (\sigma(x_t, \dots, x_{t+n}), \sigma(y_t, \dots, y_{t+n}), \sigma(z_t, \dots, z_{t+n}), \\
 &\quad \mathbf{S}_C(\sigma(x_t, \dots, x_{t+n}), \sigma(y_t, \dots, y_{t+n}), \sigma(z_t, \dots, z_{t+n})) \\
 &= (\sigma(\vec{\mathbf{x}}_T^{Pen}), \sigma(\vec{\mathbf{y}}_T^{Pen}), \sigma(\vec{\mathbf{z}}_T^{Pen}), \\
 &\quad \mathbf{S}_C(\sigma(\vec{\mathbf{x}}_T^{Pen}), \sigma(\vec{\mathbf{y}}_T^{Pen}), \sigma(\vec{\mathbf{z}}_T^{Pen}))) \\
 &= (v_1, v_2, v_3, \mathbf{S}_C^{FuzzyPen}(v_1, v_2, v_3)) \\
 &= (\vec{\nabla}_C^{FuzzyPen}, \mathbf{S}_C^{FuzzyPen}(\vec{\nabla}_C^{FuzzyPen})) \quad (\text{Definition 4.3.1.22})
 \end{aligned}$$

Den gewünschten Ausgabewert  $u_i$ , der für den Eingabevektor  $\vec{\nabla}_{Q,u_i}^{FuzzyPen}$  den Referenzwert bildet, berechnet man durch das kontextuelle Sugeno-Systems mit der Funktion  $e$ , wie folgt:

$$e(\vec{\nabla}_{Q,u_i}^{FuzzyPen}, u_i) = \begin{cases} 1 & \text{für } P_{\pm}(v_{4,i}, u_i) \\ 0 & \text{sonst} \end{cases} \quad (\text{Definition 4.3.1.23})$$

Die Menge der Trainings-Daten (Definition 4.3.1.24) ergibt sich aus den Daten (Definition 4.3.1.10-Definition 4.3.1.12) die schon zuvor für das Training des kontextuellen Sugeno-Systems verwendet wurden.

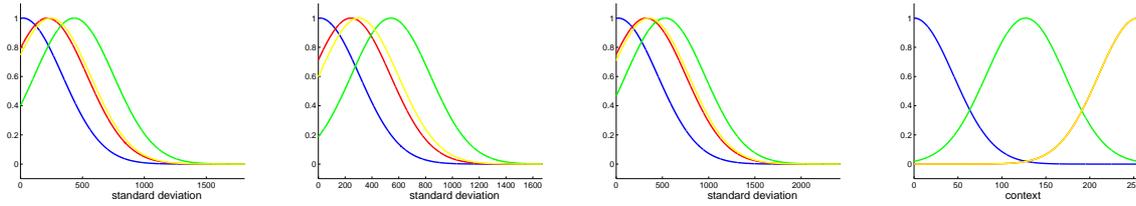
$$V = \begin{pmatrix} v_{1u_1} & v_{1u_2} & \cdots & v_{1u_k} & v_{1u_{k+1}} & v_{1u_{k+2}} & \cdots & v_{1u_{k+l}} & v_{1u_{k+l+1}} & v_{1u_{k+l+2}} & \cdots & v_{1u_n} \\ v_{2u_1} & v_{2u_2} & \cdots & v_{2u_k} & v_{2u_{k+1}} & v_{2u_{k+2}} & \cdots & v_{2u_{k+l}} & v_{2u_{k+l+1}} & v_{2u_{k+l+2}} & \cdots & v_{2u_n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ v_{mu_1} & v_{mu_2} & \cdots & v_{mu_k} & v_{mu_{k+1}} & v_{mu_{k+2}} & \cdots & v_{mu_{k+l}} & v_{mu_{k+l+1}} & v_{mu_{k+l+2}} & \cdots & v_{mu_n} \end{pmatrix} \quad (\text{Definition 4.3.1.24})$$

Die getrennte Betrachtung der Daten für jeden kontextuellen Zustand ist hier nicht notwendig. Die gewünschte Ausgabe wird direkt mit der Funktion  $e$  (Definition 4.3.1.23) berechnet, was für alle  $u_i$  den den Vektor  $\mathbf{u}$  ergibt:

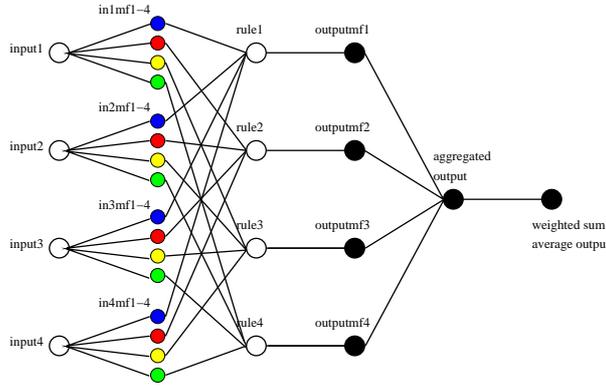
$$\mathbf{u} = (e(\vec{\nabla}_{Q,u_1}^{FuzzyPen}, u_1), e(\vec{\nabla}_{Q,u_2}^{FuzzyPen}, u_2), \dots, e(\vec{\nabla}_{Q,u_n}^{FuzzyPen}, u_n)) \quad (\text{Definition 4.3.1.25})$$

Mit den somit definierten Daten für Eingabe und gewünschter Ausgabe des Trainings kann die automatisierte Erstellung des qualitativen Sugeno-Systems begonnen werden. Der Ablauf ist in Kapitel 4.2 beschrieben und wird im Falle der Qualitätsanalyse der Algorithmen des Stift-Artefaktes dementsprechend ausgeführt. Das neuronale Fuzzy Netz der qualitativen Sugeno-Logik, welches durch das subtraktive Clustering und der Methode der kleinsten Quadrate spezifiziert wird, ist in Abbildung 18(b) in einer vereinfachten Form einzusehen. Die Gauss-Funktionen die die funktionalen Zugehörigkeiten bilden, sind in der initialen Form, die durch das subtraktive Clustering erhalten wird, in Abbildung 18(a) dargestellt. Das qualitative Sugeno-System verbessert sich in seiner Beurteilung der Trainings-Daten bei jeder Epoche des hybriden Trainings. Um einer zu starken Anpassung durch das Training an die Daten vorzubeugen, wird immer mit einer von den Trainings-Daten unabhängigen Datenmenge verglichen. Wie sich der Fehler des qualitativen Sugeno-Systems pro Epoche verändert und wie sich dagegen der Fehler für die unabhängige Datenmenge verhält, ist in Abbildung 19 dargestellt. Aus diesen beiden Diagrammen kann eine zu starke Spezifizierung des Systems, der Punkt ab dem der Fehler für die Check-Daten wieder länger stetig steigt, ab der Epoche 8075 festgestellt werden. Dies ist auch die Epoche, bei der das Training des qualitativen Sugeno-Systems beendet wird.

Das aus dem Training resultierende Sugeno-System  $\hat{\mathbf{S}}_{Q_{lin}}^{FuzzyPen}$  muss noch in seiner Ausgabe normalisiert werden, was für den allgemeinen Fall in Kapitel 4.1.5 beschrieben wurde. Die Normalisierung wird



(a) Zugehörigkeitsfunktionen für Eingänge der Fuzzy Logik zur Qualitätsanalyse - Varianzen für (1)  $adxl_x$ , (2)  $adxl_y$ , (3)  $adxl_z$  und (4) Kontextklasse



(b) Neuronales Netz zum Training der Fuzzy Logik zur Qualitätsanalyse

Abbildung 18: Spezifikation der Eingansfunktionen nach dem subtraktiven Clustering und des zugehörigen Neuronalen Netzes zum Training der Sugeno Fuzzy Logik

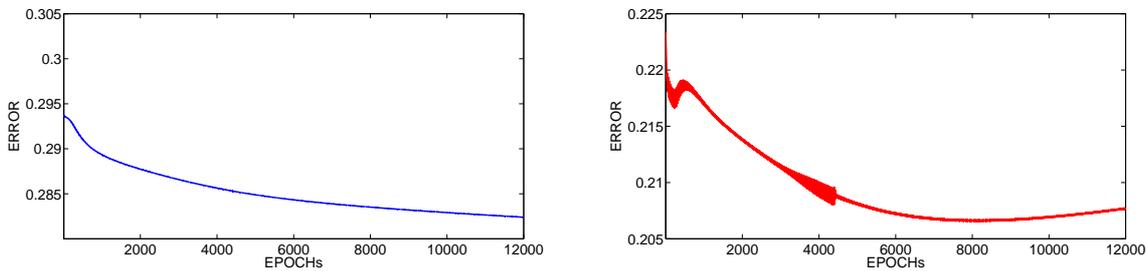


Abbildung 19: Fehler  $E(V_{TRAIN})$  (links) und Fehler  $E(V_{CHECK})$  (rechts) einer Check-Datenmenge für jeweils 1200 Epochen des Trainings und der Kontrolle

durch die Funktion  $\mathbf{L}_{\tilde{Q}_{Q_{lin}}}$  durchgeführt, die auf die Menge  $Q = [0, 1]$  abbildet. Die Abbildung des qualitativen Systems ist also die hintereinander Ausführung des Sugeno-Systems und der normalisierenden Funktion, die wie folgt beschrieben ist:

$$\mathbf{S}_Q^{FuzzyPen} : \begin{cases} \mathcal{V}_1 \times \mathcal{V}_2 \times \mathcal{V}_M \times C & \longrightarrow Q, \text{ mit } Q = \{q_1, \dots, q_P\} \\ (v_1, v_2, v_3, c) & \longmapsto \mathbf{L}_{\tilde{Q}_{Q_{lin}}}(\hat{\mathbf{S}}_{Q_{lin}}^{FuzzyPen}(v_1, v_2, v_3, c)) \end{cases} \quad (\text{Definition 4.3.1.26})$$

Die verwendete Abbildung des qualitativen Sugeno-Systems ist für das Stift-Artefakt und das kontextuelle Sugeno-System wie folgt definiert:

$$\hat{\mathbf{S}}_{Q_{lin}}^{FuzzyPen} : \begin{cases} \mathcal{V}_1 \times \mathcal{V}_2 \times \mathcal{V}_3 \times C & \longrightarrow Q_{lin}, \text{ mit } [0, 1] \subset Q_{lin} \subset \mathcal{R} \\ (v_1, v_2, v_3, c) & \longmapsto \hat{\mathbf{S}}_{Q_{lin}}(v_1, v_2, v_3, c) \end{cases} \quad (\text{Definition 4.3.1.27})$$

Das das nun spezifizierte System  $\mathbf{S}_Q^{FuzzyPen}$  zur Qualitätsanalyse, wird im folgenden Kapitel 4.4 statistisch analysiert. Ein weiteres System zur Erkennung von kontextuellen Zuständen wird im nächsten Abschnitt beschrieben. Dieses System resultiert in einer anderen Architektur des qualitativen Sugeno-Systems, die auch erläutert wird.

### 4.3.2 Neuronale Netze als Kontextsystem $N_C$

Es wurde ein Sugeno-System vorgestellt, das für ein Stift-Artefakt kontextuelle Zustände erkennt. Da die qualitative Bewertung dieser Zustandserkennung ebenfalls mit einem Sugeno-System gemacht wird, das zu evaluieren ist, wird in diesem Abschnitt eine nicht-unscharfe Alternative für ein kontextuelles System vorgestellt. Das System, das auch für die statistische Analyse verwendet wird, ist ein vorwärts gerichtetes neuronales Netz. Dieses neuronale Netz, das nun spezifiziert wird, wurde auch von Estévez [Est05] vorgeschlagen. Die Sensoren, die verwendet werden, und die kontextuellen Zustände, auf die abgebildet wird, sind die gleichen wie bei dem kontextuellen Sugeno-System.

Das neuronale Netz  $\hat{N}_{C_{lin}}^{NeuroPen}$  bildet auch wieder auf eine Menge  $C_{lin}$  ab, deren Werte durch eine Funktion  $K_C^{Pen}$  (Definition 4.3.1.20) den Klassen der kontextuellen Zustände zugeordnet wird. Diese beiden Abbildungen des neuronalen Netzes und der klassifizierenden Funktion, werden in der Abbildung des gesamten kontextuellen Neuro-Systems  $N_C^{NeuroPen}$  wie folgt zusammengefasst:

$$N_C^{NeuroPen} : \begin{cases} \mathcal{V}_1 \times \mathcal{V}_2 \times \mathcal{V}_3 & \longrightarrow C, \text{ mit } C = \{[0], [127], [255]\} \\ (v_1, v_2, v_3) & \longmapsto K_C^{Pen}(\hat{N}_{C_{lin}}^{NeuroPen}(v_1, v_2, v_3)) \end{cases} \quad (\text{Definition 4.3.2.28})$$

Beide Abbildungen und ihre Zusammenfassung, sind in der Abbildung 20 auf der linken Seite dargestellt.

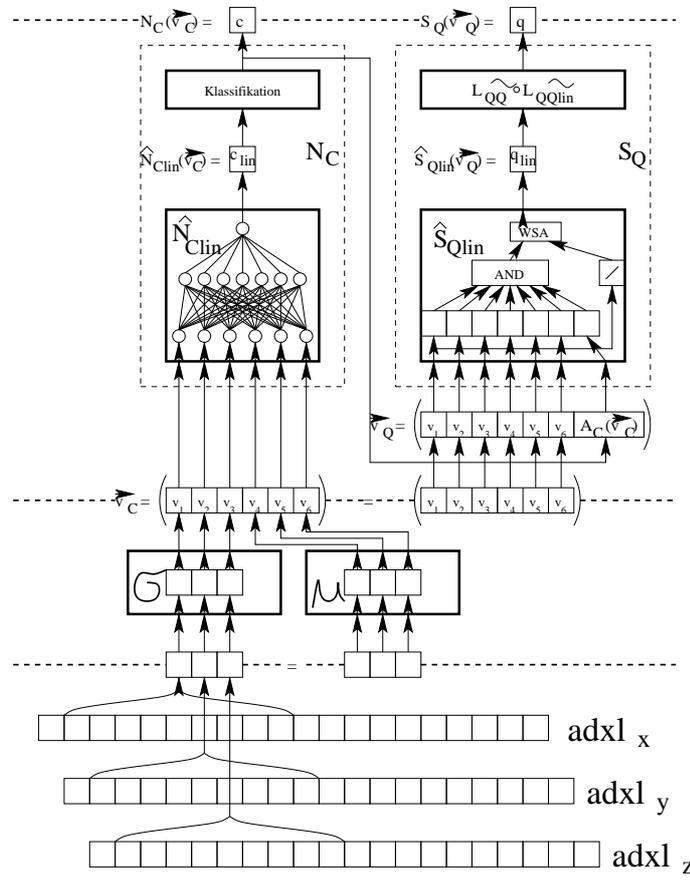


Abbildung 20: Zusammenhänge von kontextuellem neuronalen Netz  $N_C$  und Spezifikationen der qualitativen Analyse  $S_Q$  für das Stift-Artefakt

Es ist auch die Verbindung von kontextuellem und qualitativen System, miteinander und mit den Schichten der Vorverarbeitung und der Sensordaten, in dieser Abbildung einzusehen.

Die Sensordaten, die verwendet werden, sind die gleichen, wie bei den schon beschriebenen Systemen des Stift-Artefaktes. Die Eingabe der Vorverarbeitung sind also die Vektoren  $\vec{x}_T^{Pen}$ ,  $\vec{y}_T^{Pen}$  und  $\vec{z}_T^{Pen}$ , die schon definiert sind.

### Vorverarbeitung der Sensordaten

Die Vorverarbeitung für das kontextuelle neuronale Netz besteht wieder aus der Berechnung der Standardabweichung und nun noch aus den Mittelwerten, der Vektoren  $\vec{\mathbf{x}}_T^{Pen}$ ,  $\vec{\mathbf{y}}_T^{Pen}$  und  $\vec{\mathbf{z}}_T^{Pen}$ . Durch die Bearbeitung ergibt sich nun der folgende Vektor, der die Eingabe des neuronalen Netzes  $\hat{\mathbf{N}}_{C_{lin}}^{NeuroPen}$  und somit auch des qualitativen Systems  $\mathbf{S}_Q$  ist:

$$\begin{aligned} \vec{\mathbf{v}}_C^{NeuroPen} &:= (\mu(x_t, \dots, x_{t+n}), \mu(y_t, \dots, y_{t+n}), \mu(z_t, \dots, z_{t+n}), \sigma(x_t, \dots, x_{t+n}), \sigma(y_t, \dots, y_{t+n}), \sigma(z_t, \dots, z_{t+n})) \\ &= (\mu(\vec{\mathbf{x}}_T^{Pen}), \mu(\vec{\mathbf{y}}_T^{Pen}), \mu(\vec{\mathbf{z}}_T^{Pen}), \sigma(\vec{\mathbf{x}}_T^{Pen}), \sigma(\vec{\mathbf{y}}_T^{Pen}), \sigma(\vec{\mathbf{z}}_T^{Pen})) \\ &= (v_1, v_2, v_3, v_4, v_5, v_6) \end{aligned} \quad (\text{Definition 4.3.2.29})$$

Dieser Vektor wird wieder an der Schnittstelle (Abbildung 20, mittlere gestrichelte Linie) des kontextuellen und qualitativen Systems, an diese Systeme weitergegeben.

### Kontextuelles neuronales Netz

Das neuronale Netz, das für die Erkennung der kontextuellen Zustände eingesetzt wird, ist ein vorwärts gerichtetes Netz mit sechs Neuronen der Eingabeschicht. Jedes Neuron der Eingabeschicht ist mit einem Element des Eingabevektors  $\vec{\mathbf{v}}_C^{NeuroPen}$  assoziiert. Die einzige versteckte Schicht besteht aus sieben Neuronen. Die Schicht der Ausgabe besteht aus einem Neuron, dessen Werte klassifiziert die Ausgabe des kontextuellen Systems bilden.

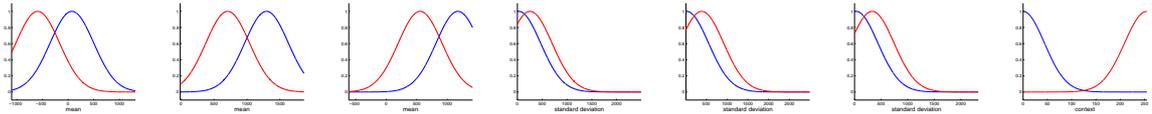
Das Netz wird mit der Methode von absteigenden Gradienten trainiert, die Backpropagation genannt wird. Diese Trainingsmethode wurde in Kapitel 4.2.4 im Zusammenhang mit einem ANFIS des qualitativen Systems erklärt. Die Trainingsdaten sind auch ähnlich zu erhalten, wie die des kontextuellen Sugeno-Systems, jedoch sind nun die 6-dimensionalen Vektoren  $\vec{\mathbf{v}}_C^{NeuroPen}$  zu verwenden. Die gewünschte Ausgabe, gegen die durch das Training konvergiert werden soll, ist dieselbe wie bei dem kontextuellen Sugeno-System. Das Werkzeug, das zum Training und zur Erstellung des neuronalen Netzes verwendet wurde, ist das 'NNTool' von MATLAB<sup>®</sup>.

Es wird hier nicht näher auf das neuronale Netz eingegangen, da neuronale Netze nicht Bestandteil dieser Arbeit sind. Es wurde nun nur kurz erwähnt um was für ein neuronales Netz es sich handelt, das die Grundlage für die Qualitätsanalyse ist. Die genauen Kenntnisse sind auch für einen Entwickler nicht notwendig, der ein System für ein neuronales Netz zur Qualitätsanalyse erstellen möchte, da die automatisierte Erstellung keinerlei Wissen benötigt.

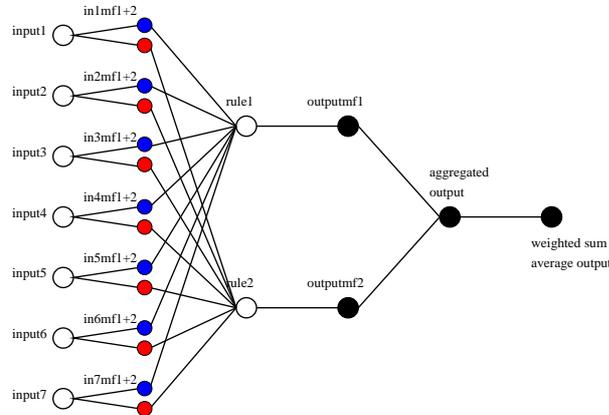
### System zur Qualitätsanalyse des neuronalen Netzes

Die Erstellung des ANFIS erfolgt, wie bei allen qualitativen Sugeno-Systemen, über ein subtraktives Clustering, das die Anzahl der Cluster und somit die Anzahl und Art der Regeln des unscharfen Systems bestimmt. Die Daten auf denen die adaptiven Algorithmen, also auch das subtraktive Clustering, arbeiten, sind die gleichen, die schon zum Training des neuronalen Netzes  $\vec{\mathbf{v}}_C^{NeuroPen}$  eingesetzt wurden. Da das neuronale Netz eine sehr gute Adaption an die Trainingsdaten verwirklicht, und dadurch auch der Fehler des Netzes sehr gering ist, stehen kaum Daten zur Verfügung die falsch klassifiziert werden. Die falsche Klassifikation des kontextuellen Systems, das in diesem Fall ein neuronales Netz mit Klassifikation des Resultates ist, bildet ein Trainingsdatum, dass das System zur Qualitätsanalyse auf eine minimale Qualität abbilden sollte. Stehen also kaum Datenpaare zur Verfügung, die beim Training des qualitativen Systems eine Beeinflussung gegen eine minimale Qualität bewirken, so wird das System generell eine hohe Qualitätsbeurteilung liefern. Dies ist die erwünschte Repräsentation der kontextuellen Klassifikation, denn wenn ein kontextuelles System sehr gut in der Erkennung von Zuständen ist, so sollte dies durch die Qualitätsanalyse berücksichtigt werden.

Eine Abstraktion des ANFIS, das die neuronale Repräsentation des Sugeno-Systems ist, kann in Abbildung 21(b) eingesehen werden. Die funktionalen Zugehörigkeiten für die sieben Eingänge des ANFIS, die durch das subtraktive Clustering bestimmt und mit dem hybriden Training verfeinert wurden, sind in Abbildung 21(a) dargestellt. Es ist zu erkennen, dass nur zwei Zugehörigkeitsfunktionen pro Eingang vorhanden sind, was eine Anzahl von zwei Clustern und zwei Regeln bedeutet. Diese Anzahl ist halb so groß wie bei dem qualitativen Sugeno-System das zur Analyse des kontextuellen Sugeno-Systems verwendet wurde. Der Grund hierfür liegt an der geringen Anzahl von falsch klassifizierten Datenpaaren, die durch das subtraktive Clustering von den richtigen Klassifikationen getrennt werden müssen. Das resultierende ANFIS ist deshalb weniger komplex als das vorhergehende, obwohl es um drei Eingänge erweitert ist. Die Anzahl der Epochen des Trainings ist schwer zu bestimmen, denn so gut wie die Klassifikation



(a) Zugehörigkeitsfunktionen für Eingänge der Fuzzy Logik zur Qualitätsanalyse - Mittelwerte für (1)  $adxl_x$ , (2)  $adxl_y$ , (3)  $adxl_z$ , Standardabweichung für (4)  $adxl_x$ , (5)  $adxl_y$ , (6)  $adxl_z$  und (4) Kontextklasse



(b) Neuronales Netz zum Training der Fuzzy Logik zur Qualitätsanalyse

Abbildung 21: Spezifikation der Eingansfunktionen nach dem subtraktiven Clustering und des zugehörigen Neuronalen Netzes zum Training der Sugeno Fuzzy Logik

des kontextuellen Systems  $\mathbf{N}_C$  auf den Trainingsdaten ist, so schlecht ist sie auf den Daten des Checks. Dieser Sachverhalt kann in den Grafiken der statistischen Analyse, die in Kapitel 4.4.5 beschrieben ist, nachvollzogen werden. Wird beim Training des ANFIS der Fehler für die Trainingsdaten kontinuierlich besser, kann für die Daten des Checks eine Divergenz festgestellt werden. Es kann also kaum eine Aussage gemacht werden, wann eine Epoche mit einem guten Resultat für beide Datensätze erreicht wurde. Angesichts der reduzierten Komplexität der Regeln gegenüber dem vorherigen ANFIS, wurde ein hybrides Training des Netzes über 4000 Epochen verwendet.

Das qualitative System  $\mathbf{S}_Q^{NeuroPen}$  entspricht ansonsten dem qualitativen System  $\mathbf{S}_Q^{FuzzyPen}$ , das zur Qualitätsanalyse des kontextuellen Sugeno-Systems eingesetzt wurde. Im nächsten Abschnitt werden die Qualitäten, die durch dieses System berechnet werden, einer statistischen Analyse unterzogen. Es wird auch tiefer ergründet, was die Unterschiede gegenüber der Qualitätsanalyse für ein unscharfes Regelsystem sind.

## 4.4 Analyse der Resultate der unscharfen Logik $S_Q$

Es existieren viele Methoden um Daten zu analysieren. In dieser Arbeit wurden die Methoden, wie Clustering, lineare Regression und neuronale Netze hauptsächlich zur Identifikation eines unscharfen Regelsystems eingesetzt. Die verschiedenen Algorithmen der Clusteranalyse und der linearen Regression sind jedoch auch zur Datenanalyse zu verwenden. Ein neuronales Netz kann auch zur Aufbereitung der Informationen, die in den Daten liegen, dienen. In diesem Zusammenhang wurden die selbst-organisierenden Karten angesprochen.

Um die Resultate eines qualitativen Systems  $S_Q$  zu untersuchen, werden nun Methoden der Statistik eingesetzt, da präzise Informationen über die Daten vorhanden sind. Es werden drei Teilbereiche der Statistik unterscheiden, die deskriptive, die induktive und die explorative Statistik. Die deskriptive Statistik wird auch beschreibende Statistik genannt und umfasst Methoden um Tabellen, grafische Darstellungen und Kennzahlen aus einer Menge von Daten zu erstellen. Mit dem Begriff der induktiven Statistik werden Methoden der Statistik zusammengefasst, die aus Stichproben die globalen Eigenschaften der Daten ableiten. Die Grundlage für diesen statischen Bereich bildet die Wahrscheinlichkeitstheorie. Die explorativen Methoden sind eine Mischform der deskriptiven und induktiven Statistik, mit denen Zusammenhänge in den Daten gesucht werden. Diese Zusammenhänge bilden Hypothesen, die in induktiven Testverfahren und empirischen Versuchsplanungen als statistisch abgesichert bestätigt werden.

Es werden nun explorative Methoden eingesetzt, die Methoden der anderen statistischen Teilgebiete kombinieren. In dieser Richtung werden verschiedene Herangehensweisen an die Daten der bearbeiteten Werte von Sensoren und den zugehörigen kontextuellen Zuständen beschrieben, deren absolute und qualitative Verlässlichkeit bekannt sind. Die Analysen werden auf das im letzten Abschnitt beschriebene Stift-Artefakt angewandt.

Zuerst wird die Ausgabe des qualitativen Sugeno-Systems  $\widehat{S}_{Q_{lin}}$  bei Übergängen von einem zum anderen kontextuellen Zustand untersucht. Die theoretischen Annahmen was bei einem Zustandswechsel geschehen müsste, wurden im Kapitel 2.2.1 getroffen.

In einem weiteren Abschnitt (Kapitel 4.4.2) werden die qualitätsanalytischen Ausgaben für jeden kontextuellen Zustand getrennt betrachtet. Es wird versucht die Hypothese zu untermauern, dass einige kontextuelle Zustände des Stift-Artefaktes besser in ihrer Qualität zu beurteilen sind, als andere. Es wird auch untersucht, in wie weit falsch klassifizierte Daten bei welchen kontextuellen Zuständen gut durch die Qualitätsanalyse erkannt werden.

Werden die Ergebnisse der Qualitätsanalyse im Gesamten betrachtet (Kapitel 4.4.3), müssen weitere Gütekriterien zur statistischen Analyse benutzt werden. Es werden Wahrscheinlichkeiten für tatsächlich falsch oder richtig klassifizierte kontextuelle Zustände verglichen und somit das Qualitätsmaß auf Nutzen untersucht. Mit diesen Wahrscheinlichkeiten wird eine Methodik vorgestellt, wie ein Schwellwert für das Qualitätsmaß, um die Daten zu filtern, festgelegt werden kann.

Eine Möglichkeit der Kontexterkenennung auf dem Stift-Artefakt die vorgestellt wurde, ist eine unscharfes Regelsystem. Die Ausgabe dieses kontextuellen Sugeno-Systems  $\widehat{S}_{C_{lin}}$  wurde wieder unscharf interpretiert und mit der korrespondierenden Ausgabe des qualitativen Sugeno-Systems  $\widehat{S}_{Q_{lin}}$  verglichen, was in Kapitel 4.4.4 beschrieben wird.

Wird ein neuronales Netz zur Kontexterkenennung eingesetzt wird, bieten sich auch Möglichkeiten des Vergleichs mit der qualitativen Analyse. Es wird wieder statistisch ein Schwellwert für das Qualitätsmaß bestimmt und die Wahrscheinlichkeiten für tatsächliche Richtigkeit und Falschheit verglichen. Diese Vergleiche werden in Kapitel 4.4.5 analysiert.

Im letzten Abschnitt (Kapitel 4.4.6) werden die Auswirkungen von verschiedenen Trainingsdaten auf das qualitative Regelsystem  $S_Q$  analysiert. Es wird die Frage erörtert, was ein nicht vollständiger Satz an Trainingsdaten für die resultierende Qualitätsanalyse  $S_Q$ , die auf kompletten Daten operiert, an Resultaten bewirkt.

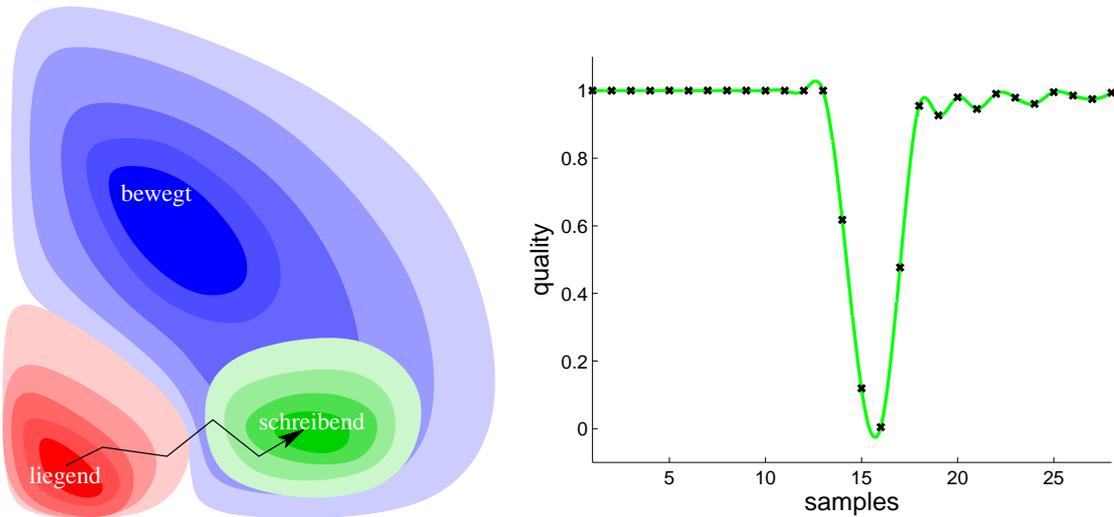


Abbildung 22: Werte der Qualität (schwarze x) für einen Wechsel des kontextuellen Zustandes 'liegend' zu dem Zustand 'schreibend' (rechts), interpoliert mit einem Spline (grüne Linie); Abstraktion eines möglichen Verlaufs für diesen Wechsel in einem 2-dimensionalen Zustandsraum (links)

#### 4.4.1 Beobachtung des Systemverhaltens bei kontextuellen Zustandswechseln

In Kapitel 2.2.1 wurden Vermutungen getroffen, wie sich ein Qualitätsmaß beim Übergang von einem kontextuellen Zustand zum nächsten verhalten könnte. Es wurde angenommen, dass wenn der kontextuelle Algorithmus sich in der Projektion auf die Kontextklasse aus dem Mittelpunkt der eindeutigen Klassenzuordnung heraus bewegt, eine Verschlechterung durch die Qualitätsanalyse festzustellen sein müsste. Um diese Annahme in der Praxis nachzuweisen, wurden wieder Daten von kontextuellen Zuständen des Stift-Artefaktes benutzt. Hierbei treten jedoch Probleme auf, denn dieser Nachweis ist mit den hoch dynamischen Werten der Beschleunigungs-Sensoren und den daraus gewonnenen Standardabweichungen sehr schwer zu treffen. Die Frequenz, in der die Sensorwerte ausgelesen werden, ist hier von entscheidender Bedeutung, da sie die Geschwindigkeit bestimmen, in der der Übergang von einem zum anderen kontextuellen Zustand vonstatten geht. Es ist jedoch der Fall, dass Bewegungen sich sehr schnell verändern können und deshalb selbst eine hohe Abtastung nicht ausreichend ist. Die Breite der Fenster in denen die Standardabweichung berechnet wird, kann auch variiert werden, um dadurch eine höhere Rate der aufkommenden Werte zu erreichen. Dies ist aber auch nur bedingt praktikabel, da eine zu kleine Breite zu Abweichungen führt, die nicht mehr aussagekräftig sind.

Die Bewegungen sind nicht linear in ihren Übergängen, denn es werden immer andere Bewegungen gemacht, die mal mehr oder weniger zu einem Zustand zu zählen sind, andere Zustände tangieren oder in sie übergehen. Um diesen Sachverhalt einzuschränken, werden nun nur Übergänge von der Bewegungslosigkeit, dem kontextuellen Zustand 'liegend', zu einem der beiden Zustände der Bewegung, zu analytischen Zwecken benutzt.

Ein guter Übergang vom kontextuellen Zustand 'liegend' in den Zustand 'schreibend' ist in Abbildung 22 auf der rechten Seite abgebildet. Es ist schön zu sehen, wie in vier Schritten die Qualität zuerst schlechter und dann besser wird. Die Werte der Qualität wurden mit einer Spline-Interpolation (grüne Linie) zu einem Graph verbunden, der einen eindeutigen Umschwung zwischen den Kontexten zeigt. Der Graph ist in der rechten Hälfte nach dem Umschwung stärker variierend, da der kontextuelle Zustand 'schreibend' nicht immer so gut in seiner Qualität der Erkennung ist, wie der Zustand 'liegend'. Dennoch sind die Werte der Qualität für den Zustand 'schreibend' deutlich von dem Übergang zwischen den Zuständen zu unterscheiden. Um die Abstufungen der Werte zu erläutern, wurde ein möglicher Verlauf der Standardabweichungen (schwarze Linie) in einer 2-dimensionalen Graphik (Abbildung 22, links) visualisiert. Die Einbettung des kontextuellen Zustandes 'schreibend' in den Zustand 'liegend', wird im nächsten Abschnitt der zustandsabhängigen Analyse ergründet.

Eine weitere Art der Zustandsübergänge ist in der Abbildung 23 auf der rechten Seite zu betrachten, bei der im Übergang kurzzeitig ein anderer Zustand eingetreten ist. Ausgehend vom Zustand 'liegend', wird in der beginnenden Bewegung zuerst das Qualitätsmaß schlechter, um dann wieder ein wenig besser

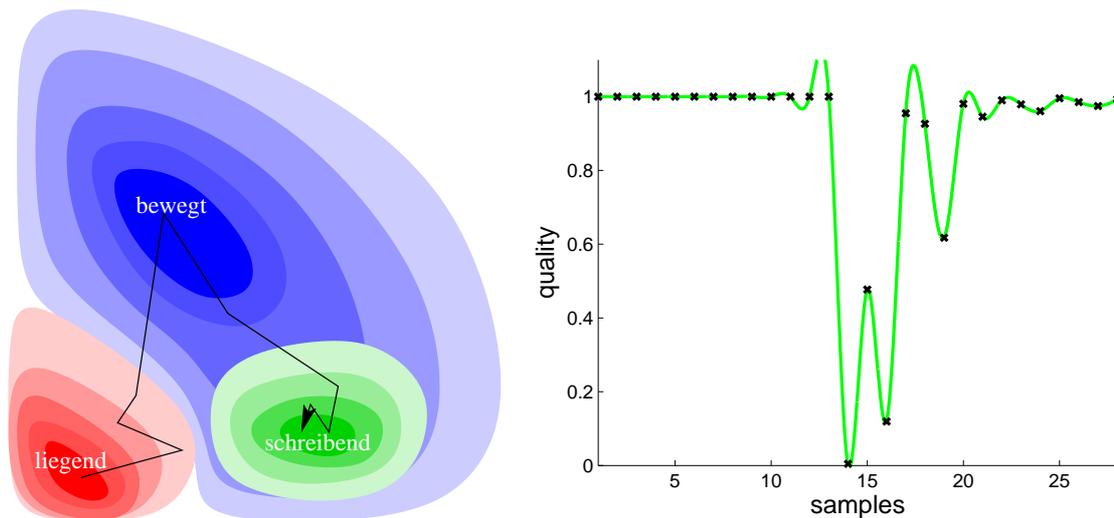


Abbildung 23: Werte der Qualität (schwarze x) für einen Wechsel des kontextuellen Zustandes 'liegend', über den Zustand 'bewegt', hin zu dem Zustand 'schreibend' (rechts), interpoliert mit einem Spline (grüne Linie); Abstraktion eines möglichen Verlaufs für diesen Wechsel in einem 2-dimensionalen Zustandsraum (links)

zu werden. Dies ist mit einer kurzen Phase der geringen Beschleunigung zu erklären, die immer in einer Bewegung auftreten kann. In den weiteren Werten wird dann eine Verbesserung der Kontexterkenkung festgestellt, die in einem gut erkannten kontextuellen Zustand 'bewegt' entstanden sind. Im weiteren Verlauf werden die Qualitäten wieder umso niedriger, je mehr sich die Bewegung dem Zustand 'schreibend' nähert. Die Bewegung endet dann im Zustand 'schreibend', welcher der angestrebte kontextuelle Zustand ist. Ein möglicher Verlauf der Bewegung im kontextuellen Zustandsraum ist, wieder in einer zweiten Grafik in der Abbildung 23 auf der linken Seite dargestellt.

Diese soeben analysierten Werte der Qualität für einen Wechsel des kontextuellen Zustandes sind, wie begründet, sehr selten in einer solchen Eindeutigkeit zu beobachten. Werden andere Artefakte in Verbindung mit einer Qualitätsanalyse verwendet, die trägere Sensoren und somit langsamere Übergänge von einem zum anderen Zustand besitzen, so wird dieser Effekt stetiger auftreten.

Die weiteren Analysen beziehen sich auf andere Aspekte des Qualitätsmaßes, die einen größeren Mehrwert bringen und auch besser mit dem Stift-Artefakt zu beobachten sind. Im nächsten Abschnitt wird eine statistische Analyse der Qualitätsmaße für die einzelnen kontextuellen Zustände beschrieben.

#### 4.4.2 Zustandsabhängige Analyse des Qualitätsmaßes $S_Q$

Das qualitative System, das zu analysieren ist, beurteilt ein kontextuelles unscharfes Regelsystem des Stift-Artefaktes, wie es in Kapitel 4.3.1 beschrieben wurde. Das qualitative Sugeno-System  $\hat{S}_{Q_{lin}}$  wurde mit einem Datensatz trainiert, der eine gute Auswahl von allen kontextuellen Zuständen enthält. Die Analyse die in diesem Abschnitt beschrieben wird, referenziert auf einen von den Trainingsdaten unabhängigen Datensatz des Testes. Dieser Datensatz ist wesentlich kleiner als jener, der zum Training des ANFIS eingesetzt wurde, da die Graphen dadurch übersichtlicher werden.

Es wird nun dieser Datensatz des Testes in die einzelnen Daten des jeweiligen kontextuellen Zustandes unterteilt. Das Wissen über den Datensatz ist von entscheidender Relevanz für die statistische Analyse. In dem Fall der Testdaten ist immer bekannt, welcher kontextuelle Zustand zu welchem Datum geführt hat, weshalb auch bekannt ist, ob der Algorithmus zur kontextuellen Klassifikation den richtigen Zustand erkannt hat oder nicht. Hierbei ist jedoch zu beachten, dass eine falsche Klassifikation dennoch auf dem gegebenen Datum eine gute Klassifikation sein kann. Es ist schwer einen Datensatz zu erhalten, bei dem nur eindeutige Bewegungsmuster mit dem Stift-Artefakt gemacht wurden. In einer Bewegung die auf den Zustand 'schreibend' zurück zu führen ist, werden sicherlich Muster auftreten, die zu dem Zustand 'bewegt' gezählt werden können. Der kontextuelle Zustand 'liegend' ist jedoch recht eindeutig, da er sich zu den anderen über die Bewegungslosigkeit deutlich unterscheidet. Diese Hypothese kann sehr gut in den zustandsabhängigen Darstellungen der nicht klassifizierten, aber auf das Intervall  $[0, 1]$  normalisier-

ten, Ausgabe des qualitativen Sugeno-Systems  $\widehat{S}_{Q_{lin}}$  (Abbildung 24) eingesehen werden. Der Zustand

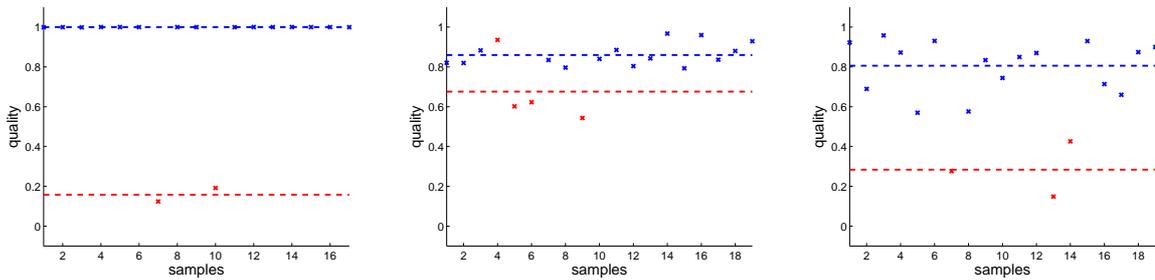


Abbildung 24: Nicht klassifizierte Ausgabe des qualitativen Sugeno-Regelsystems  $\widehat{S}_{Q_{lin}}$  für die kontextuellen Zustände 'liegend' (links), 'bewegt' (mittig) und 'schreibend' (rechts); kontextuell falsche sind rot und richtig klassifizierte Werte sind blau markiert, mit den jeweiligen Mittelwerten als gestrichelte Linien

'liegend' zeigt eine überwiegend richtige Beurteilung der Klassifikation des Kontextsystems. Das qualitative System spiegelt diesen Sachverhalt mit einer generell hohen Qualität wieder, wobei die zwei falsch klassifizierten Datenpaare auch als solche durch die Qualitätsanalyse erkannt wurden. Die Qualität der falsch erkannten kontextuellen Zustände ist für den Zustand 'liegend' sehr niedrig, was auch zu erwarten war. Die statistischen Mittel über die Qualität der richtig (Abbildung 24 links, blaue gestrichelte Linie) und der falsch (Abbildung 24 links, rote gestrichelte Linie) klassifizierten Zustände haben einen großen Abstand zueinander. Eine solche eindeutige Beurteilung des Fehlers des kontextuellen Systems durch das qualitative System ist für die anderen Zustände nicht zu erwarten.

Die Qualität für den kontextuellen Zustand 'bewegt' (Abbildung 24, mittig) ist schon wesentlich varrierender als beim Zustand 'liegend', wobei jedoch die Qualität noch generell sehr hoch ist. Es wird kein Wert der Qualität festgestellt, der zu Null zu zählen wäre und somit falsch sein müsste. Es besteht jedoch eine Schwelle, unter der der Wert der Qualität eine falsche Klassifikation vermuten lässt. Für diesen Datensatz des Testes lassen sich 75% dieser durch eine solche Grenze eines Schwellwertes heraus filtern, was natürlich aufgrund der Menge von nur 19 Datenpaaren nicht repräsentativ ist. Die Diskussion bezüglich eines Schwellwertes, um die statistische Wahrscheinlichkeit zu erhöhen, dass nur richtig erkannte kontextuelle Zustände sich in einer Resultatmenge befinden, muss in einer zustandsübergreifenden Analyse geführt werden, da die einzelnen Zustände jeweils andere Schwellen benötigen. Diese Hypothese wird bei der Betrachtung der Abstände der Mittelwerte (Abbildung 24 mittig, rote und blaue gestrichelte Linie) bestätigt, die wesentlich näher beieinander liegen als noch bei dem Zustand 'liegend'.

Ein Datum fällt bei der Betrachtung der Werte der Qualität besonders auf, da es eine hohe Qualität besitzt und dennoch dem falschen Zustand 'schreibend' zugeordnet wurde. Dieses vierte Datum der Testmenge muss aber, wie schon angedeutet wurde, nicht falsch klassifiziert worden sein, da sich die Zustände 'bewegt' und 'schreibend' den Raum der Muster in den Daten teilen. Es ist anzunehmen, dass der kontextuelle Zustand 'bewegt' eine Obermenge der Muster des Zustandes 'schreibend' ist. Das kontextuelle System muss diese Obermenge von Bewegungsmustern best möglich von der Untermenge der Zustände 'schreibend' isolieren, um eine gute Klassifikation zu erhalten. In einer längeren Bewegung werden aber die Grenzen dieser Teilmengen fließend überschritten, da die eindeutige Bewegung, wie schon angesprochen, nicht existiert.

Der Zustand 'schreibend' zeigt die größte Varianz in den Werten der Qualität (Abbildung 24, rechts), was durch seine Einbettung in Muster des Zustandes 'bewegt' zu vermuten war. Die falsch klassifizierten Daten sind gut von den richtig klassifizierten durch das Qualitätsmaß zu unterscheiden, was wieder durch die Mittelwerte für richtig und falsch eingesehen werden kann. Für den gegebenen Datensatz ist die Zuordnung zu der in Kapitel 4.1.5 eingeführten unscharfen Null  $\bar{0}$  oder unscharfen Eins  $\bar{1}$  eindeutig, was natürlich nicht immer so sein muss.

Generell ist die Qualität des kontextuellen Zustandes 'schreibend' niedriger als die der andern Zustände. Dies deutet darauf hin, dass dieser Zustand am unsichersten erkannt wird, was die zuvor aufgestellte Hypothese stützt.

Wie festgestellt wurde, haben verschiedene kontextuelle Zustände unterschiedliche Auswirkungen auf das Qualitätsmaß. Der Zustand 'liegend' ist fast absolut zu beurteilen bezüglich seiner Qualität, da richtig

klassifizierte Zustände eine Qualität um Eins haben und falsch klassifizierte deutlich zu Null zu zuordnen sind. Bei den beiden anderen kontextuellen Zuständen ist das Qualitätsmaß wesentlich variierender, was somit auch die größere Unsicherheit beim erkennen dieser Zustände widerspiegelt.

Aus dieser zustandsabhängigen Analyse ist aber nicht heraus zu lesen, ob eine Beurteilung bezüglich der Qualität eines kontextuellen Klassifikators einen Mehrwert bedeutet. Die Frage ob eine Filterung bezüglich der Qualität eine Reduktion des statistischen Fehlers des kontextuellen Systems bewirken kann, wird im nächsten Abschnitt beantwortet.

#### 4.4.3 Zustandsunabhängige Analyse des Qualitätsmaßes $S_Q$

Der Datensatz der im letzten Abschnitt zu analytischen Zwecken benutzt wurde, ist für die Aussagen die in diesem Abschnitt getroffen werden nicht umfangreich genug. Die Zuordnung des realen kontextuellen Zustandes des Stift-Artefaktes ist nicht mehr zur Trennung der Daten erforderlich, sondern bildet eine Hilfe bei der Visualisierung und den statistischen Analysen. Eine gute Menge an Daten des Testes für den Zustands unabhängigen Fall, ist bei einem regulären Gebrauch des Artefaktes, mit einer zufälligen Abfolge der Zustände, zu erhalten.

Die wieder nicht klassifizierte, aber auf das Intervall  $[0, 1]$  normalisierte (Kapitel 4.1.5), Ausgabe des qualitativen Sugeno-Regelsystems  $\hat{S}_{Q_{lin}}$  ist in Abbildung 25 auf der linken Seite dargestellt. Für die zufällige Abfolge der Zustände wurden die Mittelwerte der Qualität für richtig (blau) und falsch (rot) klassifizierte kontextuelle Zustände mit gestrichelten Linien eingezeichnet. In der Grafik der Qualitäten kann auch wieder die variierende Varianz der Werte der Qualitätsanalyse von unterschiedlichen kontextuellen Zuständen festgestellt werden. Um die statistische Richtigkeit der durch das Qualitätsmaß

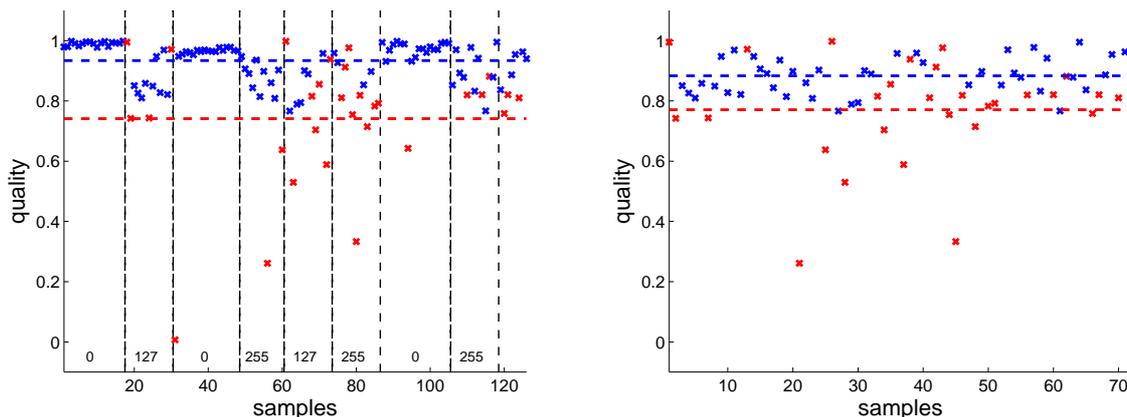


Abbildung 25: Werte der Qualität und Mittelwerte für richtige (blau) und falsche (rot) Klassifikation für eine zufällige Abfolge von kontextuellen Zuständen von 'liegend', 'bewegt' und 'schreibend' (links), die durch die Identifikatoren der Klasse (0,127,255) gekennzeichnet sind; Werte der Qualität und Mittel ohne den Zustand 'liegend' (rechts)

beurteilten Kontexte festzustellen, muss bekannt sein, ob die kontextuelle Erkennung richtig oder falsch war. Mit diesem Wissen kann eine Schranke ermittelt werden, über der die Qualitätsmaße zu einem möglichst hohen Prozentsatz tatsächlich einen richtig erkannten Kontext positiv beurteilen. Unterhalb dieser Schranke sollten hingegen nicht viele Werte vorkommen, die eigentlich eine reale Richtigkeit besitzen. Diese zwei Kriterien sind gegeneinander abzuwägen und bedürfen einer Gewichtung durch den Entwickler der Anwendung, die das Qualitätsmaß nutzt. Es kann in einem Fall wichtiger sein, möglichst alle falsch klassifizierten Kontexte heraus zu filtern und dadurch einen stark reduzierten Datensatz zu erhalten. In einer anderen Anwendung ist es sinnvoller die Schranke nicht zu hoch zu setzen, da das Aufkommen von Informationen bezüglich der kontextuellen Zustände ohnehin schon gering ist. Es kann auch der Fall sein, dass eine zweite Instanz der Verarbeitung auch mit niedrigen Maßen der Qualität gute Informationen über den Kontext erhält.

Es soll nun hier eine statistische Möglichkeit evaluiert werden, wie eine solche Schranke gesetzt werden könnte. Die Methodik, die nun genannt wird, ist jedoch auch zur statistischen Analyse im Allgemeinen zu verwenden.

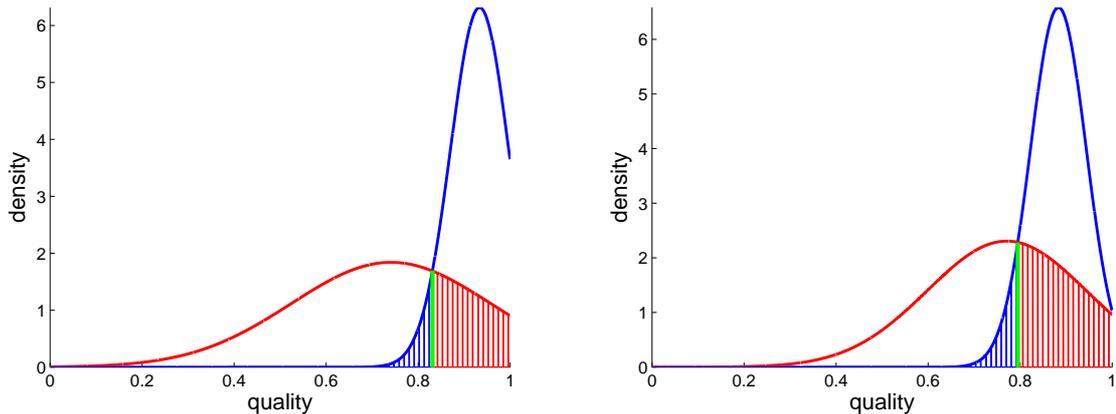


Abbildung 26: Dichtefunktionen der Normalverteilungen für kontextuell richtig (blau) und falsch (rot) klassifizierte Qualitätsmaße, mit Kennzeichnung der Schnittmengen (rot/blau schraffiert) und eines möglichen Schwellwertes (grün); Dichtefunktionen für alle kontextuellen Zustände des Stift-Artefaktes (links) und für kontextuelle Zustände 'bewegt' und 'schreibend' (rechts)

Die Qualitäten sind durch die zusätzliche Information der tatsächlichen Richtigkeit oder Falschheit ausgestattet in zwei Mengen einzuteilen, die der richtigen und der falschen Qualitätsmaße. Diese beiden Mengen besitzen jeweils eine bestimmte Dichte (Definition 4.4.3.1) der Verteilung, die als normalverteilt

$$\varphi_{\mu,\sigma^2}(x) := \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (\text{Definition 4.4.3.1})$$

angenommen wird. Mit den Mittelwerten, die schon in der Abbildung 25 eingezeichnet sind, und der Berechnung einer Varianz  $\sigma^2$  für jede der beiden Mengen, können Dichtefunktion der Normalverteilungen bestimmt werden. Je nach Größe der beiden Mengen sind die Werte der Varianz und des Mittels nur bedingt aussagekräftig, da die Repräsentanten mehr oder weniger typisch für den gesamten Raum der Qualitätsmaße sind. Wird die Methode der maximalen Wahrscheinlichkeit (eng.: maximum likelihood estimation) zur Bestimmung der Varianz und des Mittels jeder Menge eingesetzt, so wird anhand der repräsentierenden Datenwerte die Verteilung aller Daten geschätzt. Sind ausreichend viele repräsentierende Datenwerte in den beiden Mengen vorhanden, so konvergiert die maximale Wahrscheinlichkeit gegen die Mittelwerte und die Varianzen die nicht geschätzt wurden. Aus diesem Grund wurde die Schätzung der maximalen Wahrscheinlichkeit eingesetzt, die für die beiden Mengen die Dichtefunktionen in Abbildung 26 auf der linken Seite ergibt.

Durch diese beiden Funktionen der Dichte können die Wahrscheinlichkeiten für vier verschiedene Fälle unterschieden werden, wenn der Schwellwert (Abbildung 26, grüne Linien) durch den Schnitt der beiden Funktionen festgelegt wird. Die Wahrscheinlichkeit (Gleichung 4.4.3.1), durch Qualitätsmaße über diesem Schwellwert wirklich richtige kontextuelle Klassifikationen zu erhalten, ist die Differenz des Integrals der

$$\Phi_{\mu,\sigma^2}(\bar{s}) := \int_{-\infty}^{\bar{s}} \varphi_{\mu,\sigma^2}(x) dx \quad \text{oder} \quad \Phi_{\mu,\sigma^2}(\underline{s}) := \int_{\underline{s}}^{\infty} \varphi_{\mu,\sigma^2}(x) dx, \quad s \in [0, 1] \quad (\text{Definition 4.4.3.2})$$

Dichte (Definition 4.4.3.2, links) von  $-\infty$  bis zum Schwellwert für die Verteilung der falsch klassifizierten und des Integrals (Definition 4.4.3.2, rechts) von Schwellwert bis  $\infty$  der richtig klassifizierten Dichtefunktion. Diese Wahrscheinlichkeit gibt an, dass die durch das Qualitätsmaß  $q$  als richtig erkannten kontextuellen Zustände  $c$  auch wirklich richtig sind. Wird also diese Wahrscheinlichkeit der richtigen Positiven für die in Abbildung 26 auf der linken Seite dargestellten Funktionen der Dichte berechnet, so ergibt sich mit folgender Gleichung für die Schwelle  $s = 0,81$  folgender Wert:

$$\begin{aligned} \mathbb{P}(c = true | q > s) &= \Phi_{\mu_{true}, \sigma_{true}^2}(\underline{s}) - \Phi_{\mu_{false}, \sigma_{false}^2}(\underline{s}) \\ &= 0,9479 - 0,3391 = 0,6088 \end{aligned} \quad (\text{Gleichung 4.4.3.1})$$

Die Kontexte  $c$  die durch das Qualitätsmaß  $q$  unterhalb des Schwellwertes  $s = 0,81$  als falsch erkannt

wurden und dies auch wirklich sind, werden als richtige Negative bezeichnet und werden, analog zu den richtigen Positiven, durch eine Differenz der abgeschnittenen Dichten wie folgt berechnet:

$$\begin{aligned} \mathbb{P}(c = false|q < s) &= \Phi_{\mu_{false}, \sigma_{false}^2}(\bar{s}) - \Phi_{\mu_{true}, \sigma_{true}^2}(\bar{s}) \\ &= 0,6609 - 0,0521 = 0,6088 \end{aligned} \quad (\text{Gleichung 4.4.3.2})$$

Es fehlen noch die beiden Fälle der durch das Qualitätsmaß  $q > s = 0,81$  als richtig eingeschätzten Kontexte, die jedoch real kontextuell falsch klassifiziert wurden, und der falsch eingeschätzten Kontexte, die jedoch eine reelle Richtigkeit besitzen. Die Wahrscheinlichkeit der falschen Positiven ist durch die blaue schraffierte Fläche in Abbildung 26 auf der linken Seite eingezeichnet und wird durch folgende Gleichung in ihrem Wert berechnet:

$$\begin{aligned} \mathbb{P}(c = true|q < s) &= \Phi_{\mu_{false}, \sigma_{false}^2}(s) \\ &= 0,0521 \end{aligned} \quad (\text{Gleichung 4.4.3.3})$$

Die Wahrscheinlichkeit eines falschen Negativen ist mit der roten schraffierten Fläche gekennzeichnet und wird in ihrem Wert analog wie folgt bestimmt:

$$\begin{aligned} \mathbb{P}(c = false|q > s) &= \Phi_{\mu_{false}, \sigma_{false}^2}(s) \\ &= 0,3391 \end{aligned} \quad (\text{Gleichung 4.4.3.4})$$

Es wurde mit den bisherigen statistischen Analysen gezeigt, dass durch die Qualitätsanalyse und eine geschickte Wahl des Schwellwertes die richtig von den falsch klassifizierten Kontexten mit hohen Wahrscheinlichkeiten zu trennen sind. Der Schwellwert wurde durch den Schnittpunkt der beiden Dichtefunktionen festgelegt, wobei eine Berechnung durch die Schätzung der maximalen Wahrscheinlichkeit für hinreichend viele Qualitätsmaße, die nun nicht bezüglich ihrer tatsächlichen Richtigkeit und Falschheit voneinander getrennt werden, in dem Mittelwert der Normalverteilung gegen diesen Schwellwert konvergiert.

Da im letzten Abschnitt festgestellt wurde, dass der kontextuelle Zustand 'liegend' generell richtig in der Qualität eingestuft wird, soll nun analysiert werden, in wie weit die soeben betrachteten Wahrscheinlichkeiten sich ändern, wenn dieser Zustand aus der statistischen Analyse genommen wird. Die Qualitätsmaße für die Zustände 'bewegt' und 'schreibend' sind ohne den Zustand 'liegend' in der Abbildung 25 auf der rechten Seite, wieder mit den Mittelwerten für tatsächlich richtig und falsch klassifizierte Qualitätsmaße, abgebildet. Die Dichtefunktionen der Normalverteilungen werden wieder mit der Schätzung der maximalen Wahrscheinlichkeit bestimmt und sind in der Abbildung 26 auf der rechten Seite einzusehen. Wie ersichtlich ist, sind die Schnittflächen der Dichten größer als bei den vorherigen, was durch die wegfallende Trennung der eindeutigen Qualitätsmaße des Zustandes 'liegend' zu vermuten war. In dieser Betrachtung sind auch die Wahrscheinlichkeiten der statistischen Analyse nicht so gut wie zuvor. Die Wahrscheinlichkeit der richtigen Positiven ist wieder gleich der Wahrscheinlichkeit der richtigen Negativen, mit  $\mathbb{P}(c = true|q > s) = \mathbb{P}(c = false|q < s) = 0,4828$ , was nur noch eine 50%ige Chance der richtigen Erkennung von falschen und richtigen kontextuellen Klassifikationen durch das Qualitätsmaß bedeutet. Die Wahrscheinlichkeiten der falschen Positiven  $\mathbb{P}(c = richtig|q < s) = 0,0729$  und der falschen Negativen  $\mathbb{P}(c = falsch|q > s) = 0,4443$  sind dann natürlich im Gegenzug größer als zuvor.

Um die Relevanz eines großen Satzes an Daten mit möglichst guter Streuung für die statistische Analyse der Häufigkeiten zu verdeutlichen, werden nun noch die Qualitätsmaße für einen bestens getrennten Datensatz durch die Qualitätsanalyse analysiert. Die Werte der Qualität sind für diesen kleineren Satz an Daten in der Abbildung 27 einzusehen. Es kann erkannt werden, dass die richtig von den falsch erkannten kontextuellen Zuständen durch das Qualitätsmaß und die Wahl eines guten Schwellwertes eindeutig voneinander zu trennen sind. Die Dichtefunktionen die mit der maximalen Wahrscheinlichkeit für diese Qualitäten bestimmt wurden, sind in der gleichen Abbildung 27 auf der rechten Seite dargestellt. Der Schwellwert  $s = 0,81$  ist wieder durch den Schnittpunkt der beiden Dichtefunktionen zu bestimmen, wobei der Wert der gleiche ist wie schon bei den beiden anderen Datensätzen. Die Wahrscheinlichkeiten der richtigen Positiven und der falschen Negativen ist mit  $\mathbb{P}(c = true|q > s) = \mathbb{P}(c = false|q < s) = 0,8112$  wie zu erwarten sehr groß. Es kann also die Menge der richtig erkannten Kontexte von den falsch erkannten durch das Qualitätsmaß mit einem hohen Prozentsatz von 81% voneinander getrennt werden.

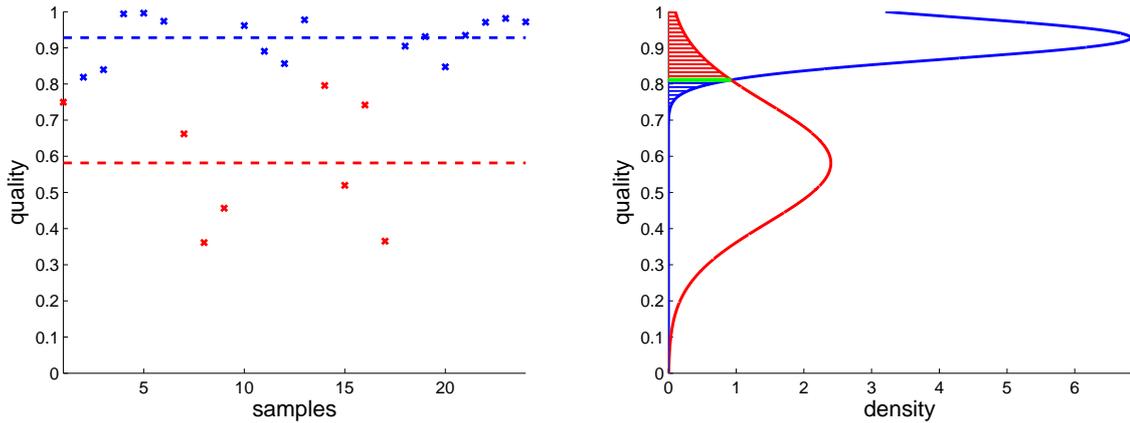


Abbildung 27: Werte der Qualität und Mittelwerte für richtige (blau) und falsche (rot) Klassifikation für eine zufällige Abfolge von kontextuellen Zuständen von 'liegend', 'bewegt' und 'schreibend' eines reduzierten Satzes an Daten (links); Dichtefunktionen (rot/blau) und Schwellwert (grün) der Qualitätsmaße (rechts)

Um die Frage zu beantworten, wie weit das qualitative Sugeno-System den Fehler des kontextuellen Systems repräsentiert, werden die Ausgaben durch eine geschickte Transformation im nächsten Abschnitt direkt miteinander verglichen.

#### 4.4.4 Vergleich von unscharfen Regelsystemen für Kontext.- $S_C$ und Qualitätsanalyse $S_Q$

In den vorangegangenen Abschnitten der Analyse des qualitativen Sugeno-Systems  $\widehat{S}_{Q_{lin}}$  wurden Aspekte des Qualitätsmaßes eingegangen, die Übergänge zwischen den einzelnen kontextuellen Zuständen, Kontext abhängige und unabhängige Resultate betrachteten. Diese Aspekte sind für die Anwendungen, für die das Qualitätsmaß eine Information liefern kann, von größerer Relevanz, als für die Information, die im Maß eigentlich vorhanden ist. Es werden nicht nur Zustände mit der Qualitätsanalyse erkannt, die mit bestimmten Sensordaten eine bessere oder schlechtere Qualität bedeuten, sondern auch der Fehler der Klassifikation des kontextuellen Systems repräsentiert.

Das bisherig analysierte System zur kontextuellen Klassifikation ist ein unscharfes Sugeno-Regelsystem  $\widehat{S}_{C_{lin}}$ , dessen Ausgabe bestimmten Identifikatoren von Klassen zugeordnet wird. Wird jedoch nicht nur der Identifikator der Klasse als ergänzende Information für die statistische Analyse verwendet, sondern eine spezielle Interpretation der Elemente der Menge  $C_{lin}$ , so besteht eine direkte Vergleichbarkeit zwischen kontextuellem und qualitativem System.

Die Interpretation der kontextuellen Ausgabe erfolgt über eine Menge von unscharfen Zahlen  $\widetilde{C} = \{\widetilde{0}, \widetilde{127}, \widetilde{255}\}$  (Abbildung 28), die sich jedoch von der Menge, die in Kapitel 4.3.1 beschrieben wurde,

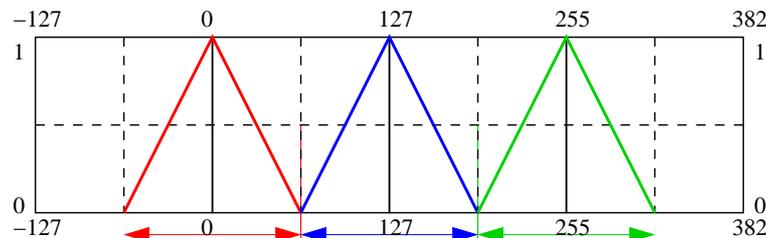


Abbildung 28: Funktionale Zugehörigkeiten für die neuen unscharfen Zahlen  $\widetilde{0}$  (rot),  $\widetilde{127}$  (blau) und  $\widetilde{255}$  (grün)

durch eine geringere Breite der Dreieckszahlen unterscheidet. In diesem Abschnitt ist der Grad der Zugehörigkeit zu einer unscharfen Zahl  $\widetilde{0}$ ,  $\widetilde{127}$  oder  $\widetilde{255}$  ausschlaggebend und nicht, zu welcher das Element aus

$C_{lin}$  die größte Zugehörigkeit besitzt. Der Grad der Zugehörigkeit ist auf das Intervall  $[0, 1]$  abzubilden, wie auch das Maß der Qualität, was mit den funktionalen Zugehörigkeiten aus Abbildung 17 durch ihre Schnitte nicht möglich wäre. Es ergibt sich aus der Zugehörigkeit zu einer der unscharfen Zahlen aus  $\tilde{C}$  die nicht Null ist, direkt ein Maß des kontextuellen Sugeno-Systems, mit dem das Qualitätsmaß des qualitativen Sugeno-Systems direkt verglichen werden kann.

Es wird jetzt noch eine Menge der Eingabe benötigt, die die Resultate der beiden Systeme bewirkt. Um wieder eine bessere Visualisierung zu erreichen, wird die Menge dieser Eingabedaten möglichst klein gehalten. Für diese Menge ergibt sich nun die Ausgabe des qualitativen (grüne Linie) und des kontextuellen Systems (blaue Linie), deren Graphen zusammen in der Abbildung 29 auf der linken Seite dargestellt sind. Es ist gut ersichtlich, dass eine gewisse Korrelation der beiden Graphen vorhanden ist, wobei der

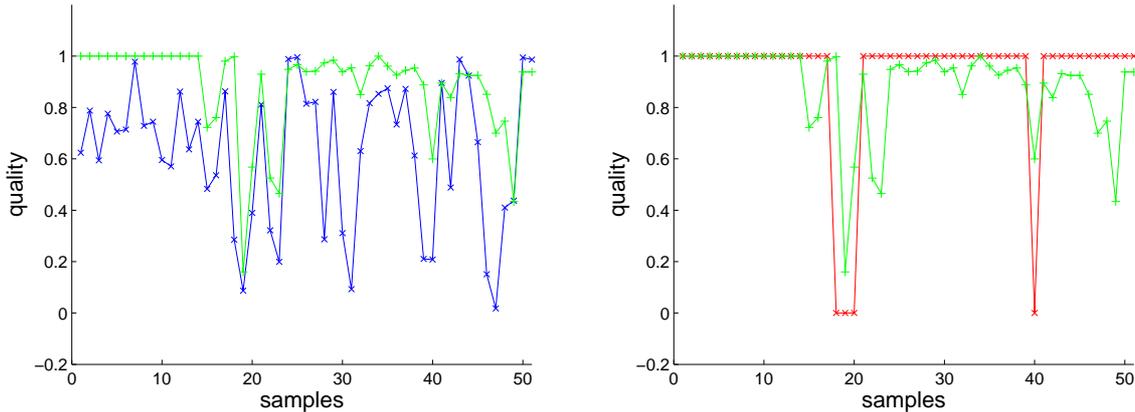


Abbildung 29: Graphen für das Maß der Zugehörigkeit (blau), zu den unscharfen Zahlen, und Maß der Qualität (grün) in einer Abbildung (links); Graphen für tatsächliche Richtigkeit (rot) und Maß der Qualität (grün) in einer Abbildung (rechts)

Graph der kontextuellen Ausgabe ein wesentlich sprunghafteres Verhalten besitzt, als der Graph der qualitativen Ausgabe. Für die ersten 14 normalisierten Ausgabedaten des qualitativen Sugeno-Systems ist ein nahezu konstanter Wert von 1 zu beobachten. Diese Ausgabedaten sind auf den kontextuellen Zustand 'liegend' zurückzuführen, was, wie schon in Kapitel 4.4.3 festgestellt wurde, zum besten Maß der Qualität führt. Der Graph des kontextuellen Systems ist für diese ersten 14 Werte sehr viel variierender und dennoch recht nahe bei Eins.

Der weitere Verlauf beider Graphen ist zuerst auf Daten des kontextuellen Zustandes 'bewegt' und dann auf den Daten des Zustandes 'schreibend' zurückzuführen. Dieser Verlauf der Graphen erscheint dem Betrachter besonders stark korrelierend, was nun mathematisch für die gesamten Graphen nachgewiesen werden soll.

Um die Korrelation zweier Funktionen, die durch ihre äquidistanten Funktionswerte definiert werden, zu berechnen, ist folgende Definition der mathematischen Korrelation  $d_{corr}$  gegeben:

$$\begin{aligned}
 d_{corr}(\vec{x}, \vec{y}) &:= \frac{\sum_{i=0}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=0}^n (x_i - \bar{x})^2 \sum_{i=0}^n (y_i - \bar{y})^2}} \\
 &= \frac{\langle (\vec{x} - \bar{x}), (\vec{y} - \bar{y}) \rangle}{\|\vec{x} - \bar{x}\| \|\vec{y} - \bar{y}\|} \quad (\text{Definition 4.4.4.7})
 \end{aligned}$$

Das übliche Funktionssymbol  $\mu$  das für die Funktion zur Berechnung eines Mittelwertes benutzt wird, ist in dieser Arbeit schon für funktionale Zugehörigkeiten verwendet worden, weshalb Konstanten  $\bar{x}$  deklariert wurden, die nach Definition 4.4.4.8 zu berechnen sind.

$$\bar{x} := \frac{1}{n} \sum_{i=0}^n x_i \quad (\text{Definition 4.4.4.8})$$

Die Funktion  $d_{corr}$  bildet auf das Intervall  $[-1, 1]$  ab. Diese Werte dieses Intervalls sind noch zu interpretieren. Die Korrelation  $d_{corr}(\vec{x}, \vec{y}) = 1$  bedeutet für die Vektoren  $\vec{x}$  und  $\vec{y}$  eine totale Korrelation,

was für zwei Graphen, die durch die Stützpunkte der Vektoren definiert werden, eine gleiche Schwingung indiziert. Eine gleiche Schwingung bedeutet, dass wenn für zwei aufeinander folgende Datenpunkte  $x_i$  und  $x_{i+1}$  eine Steigung auftritt, auch für die zwei Datenpunkte  $y_i$  und  $y_{i+1}$  eine Steigung auftritt. Das Gleiche gilt für zwei Datentupel wenn sie gleich oder fallend sind.

Eine totale Antikorrelation ist bei dem Wert der Korrelation von  $d_{corr}(\vec{x}, \vec{y}) = 0$  für die Vektoren  $\vec{x}$  und  $\vec{y}$  der Stützpunkte der Fall. Diese nicht Korrelation bedeutet, dass die Verläufe der beiden Graphen nichts miteinander zu tun haben.

Sind für zwei aufeinander folgende Datenpunkte  $x_i$  und  $x_{i+1}$  die steigend sind, die korrespondierenden Datenpunkte  $y_i$  und  $y_{i+1}$  fallend, für alle  $i \in \{0, \dots, n\}$ , dann sind die Graphen total antikorreliert. Dies gilt auch für  $x_i$  und  $x_{i+1}$  fallend und  $y_i$  und  $y_{i+1}$  steigend. Der Wert für die Korrelationsfunktion ist dann  $d_{corr}(\vec{x}, \vec{y}) = -1$  für die Vektoren  $\vec{x}$  und  $\vec{y}$  der Stützpunkte.

Im Fall der Graphen für die beiden Systeme der Kontexterkenkung und der Qualitätsanalyse kann eine mathematische Korrelation berechnet werden. Für den Satz an Daten, der zu den Graphen in Abbildung 29 auf der linken Seite geführt hat, ist die Korrelation mit  $d_{corr}(\vec{x}, \vec{y}) = 0,6$  relativ gut und deutlich einer maximalen Korrelation zuzuordnen. Mit diesem Ergebnis kann der Rückschluss auf das System der Qualitätsanalyse getroffen werden, dass es den Fehler der Klassifikation des kontextuellen Systems repräsentiert. Die geringere Varianz der Werte für die Qualitätsanalyse deutet auf die Einbeziehung des kontextuellen Zustandes in die Analyse hin, die ein stabileres Maß der Qualität bewirkt. Diese Abhängigkeit vom kontextuellen Zustand besitzt das Sugeno-System zur Kontexterkenkung nicht. Ein rekurrentes neuronales Fuzzy Netz könnte diese Eigenschaft beinhalten, was jedoch in dieser Arbeit nicht behandelt wird.

Der zweite Graph in Abbildung 29 auf der rechten Seite zeigt wieder die Kurve der Werte der Qualitätsanalyse (grün), jedoch dem Graphen der tatsächlichen Richtigkeit oder Falschheit (rot) gegenübergestellt. Die Korrelation mit dieser roten Kurve bedeutet den Fehler des qualitativen Sugeno-Regelsystems, der durch das Training zu reduzieren versucht wurde. Das Ergebnis der mathematischen Berechnung der Korrelation ist  $d_{corr}(\vec{x}, \vec{y}) = 0,5$ . Das ist schlechter als das der vorherigen Graphen, indiziert aber doch noch eine gewisse Korrelation. Das qualitative Sugeno-System wurde zwar auf den Fehler der Klassifikation trainiert, was jedoch nicht das einzige Resultat der Qualitätsanalyse ist. Es wurde festgestellt, dass die nicht eindeutige Klassifikation des kontextuellen Sugeno-Systems vom qualitativen System repräsentiert wird. Die Gewichtung dieser beiden Repräsentanten ist durch die Anzahl an Epochen die das qualitative Sugeno-Regelsystem trainiert wird und die Auswahl der Daten des Trainings zu erreichen. Um die Gewichtung zu bestimmen ist jedoch die Zielsetzung festzulegen, denn in einem Fall kann es wichtiger sein das Verhalten des kontextuellen Systems möglichst genau zu repräsentieren, im anderen Fall soll möglichst genau die falsche Klassifikation gefunden werden.

Bisher wurde nur auf ein Sugeno-System zur Kontexterkenkung eingegangen. Im nächsten Abschnitt wird versucht die Ausgaben von Qualitätsanalyse und kontextuellem neuronalen Netz in Verbindung zu bringen. Es wird analysiert ob ein ähnliches Verhalten, wie für eine unscharfes Regelsystem, beobachtet werden kann.

#### 4.4.5 Kontextuelle Neuronale Netze $N_C$ und unscharfe Qualitätsanalyse $S_Q$

In Kapitel 4.3 wurde neben einem Sugeno-System auch ein vorwärts gerichtetes neuronales Netz (FFNN) zur kontextuellen Klassifikation vorgestellt. Da die Qualitätsanalyse durch ein Sugeno-System berechnet wird, ist es fraglich, in wie fern das Resultat der qualitativen Analyse durch die Wahl eines Sugeno-Systems zur Kontexterkenkung beeinflusst wird. Es scheint nahe liegend zu sein, dass ein unscharfes Regelsystem ein anderes repräsentieren kann. Um auch die funktionale Adaptivität des qualitativen Sugeno-Systems bei anderen nicht unscharfen Algorithmen zur Kontexterkenkung nachzuweisen, wird nun in diesem Abschnitt auf die Qualitätsanalyse eines FFNN eingegangen.

Das neuronale Netz, das durch das qualitative System in seinem Fehler beurteilt werden soll, wurde in Kapitel 4.3.2 beschrieben. Der Satz an Daten, der jetzt zur statistischen Analyse der qualitativen und kontextuellen Ausgaben dient, ist von den Daten des Trainings beider Systeme unabhängig.

Ein reines neuronales Netz, nicht eine der hybriden Formen die in dieser Arbeit vorgestellt wurden, kann sehr genau an Daten des Trainings angepasst werden. Diese Anpassung kann jedoch auf unabhängigen Daten ein sehr schlechtes Resultat liefern. Es soll nun die Frage beantwortet werden ob eine Qualitätsanalyse anhand der Eingabedaten und des erkannten Kontextes die schlechte Klassifikation von für das

neuronalen Netz unbekanntem Daten erkennen kann.

In der Abbildung 30 ist auf der linken Seite das Resultat der Qualitätsanalyse für die Daten des Trainings des FFNN und des Sugeno-Regelsystems zu sehen. Da das neuronale Netz schon nach wenigen Epochen des Trainings gegen eine sehr gute Klassifikation konvergiert, ist es nicht verwunderlich, dass die rich-

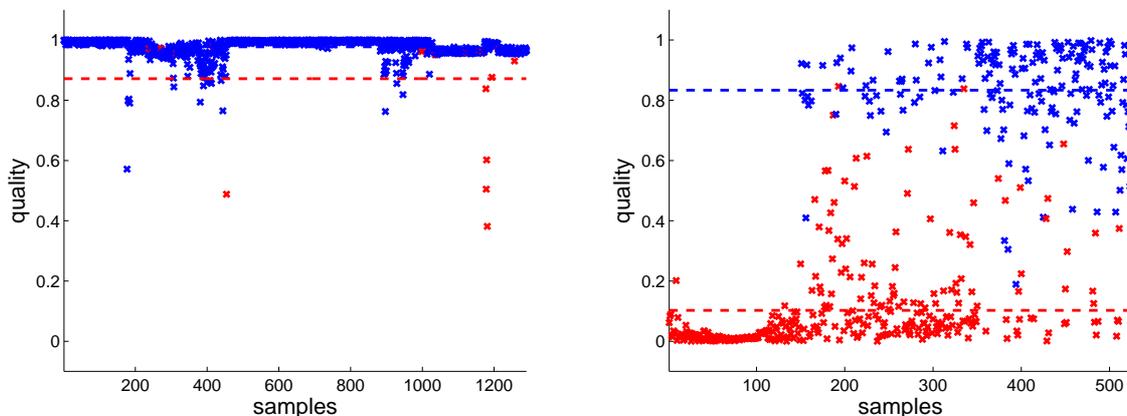


Abbildung 30: Maß der Qualität für bekannte Daten des FFNN (links) und unbekanntem unabhängigen Daten (rechts), mit Kennzeichnung der richtigen (blau) und falschen kontextuellen Klassifikation (rot)

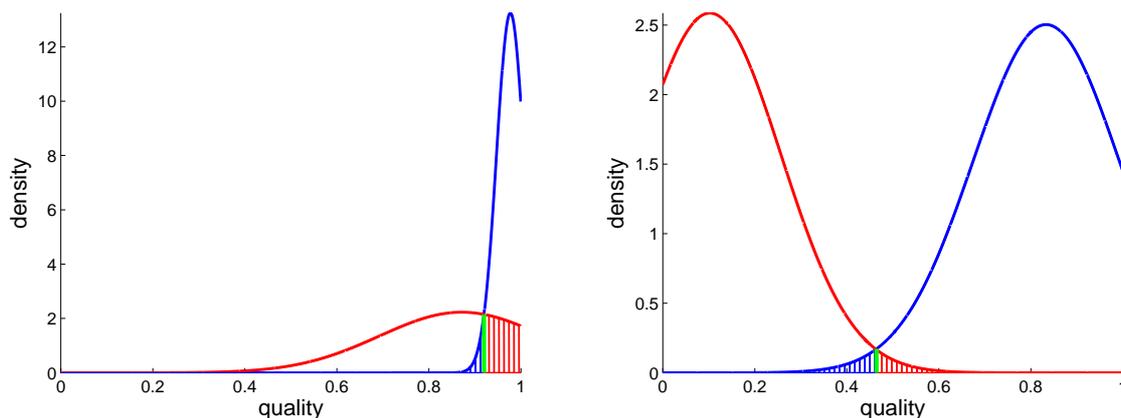


Abbildung 31: Dichtefunktionen der Normalverteilungen für kontextuell richtig (blau) und falsch (rot) klassifizierte Qualitätsmaße, mit Kennzeichnung der Schnittmengen (rot/blau schraffiert) und eines möglichen Schwellwertes (grün); Dichtefunktionen für dem FFNN bekannte Daten (links) und für unbekanntem unabhängigen Daten (rechts)

tig klassifizierte Datenpunkte (blau) überwiegen. Dem qualitativen Sugeno-System fehlen aufgrund der selben Daten des Trainings die Faktoren für eine schlechte Qualität. Aus dieser Grafik ist also nicht viel mehr heraus zu lesen, als dass mit den Daten des Trainings keine Aussage bezüglich der Qualität über ein FFNN getroffen werden kann. Die zweite Grafik in Abbildung 30 auf der rechten Seite zeigt nun das Qualitätsmaß für Daten, die unabhängig von den Daten des Trainings sind. Hier zeigt sich dass die Klassifikation des FFNN für unbekanntem Daten sehr schlecht ist. Diese schlechte Klassifikation wird jedoch von der Qualitätsanalyse sehr gut erkannt, denn fast alle falsch klassifizierte Daten besitzen eine Qualität nahe bei Null. Auch der Abstand zwischen den Mittelwerten für richtig und falsch klassifizierte Daten ist sehr nahe am Maximum. Für diese Testdaten wäre also mit der Qualitätsanalyse aus einer schlechten Klassifikation noch die richtig erkannten kontextuellen Zustände heraus zu filtern.

Um diese Feststellung zu verifizieren, wurde ein weiterer Satz von Daten zum Test der Qualitätsanalyse benutzt. Diese Daten sind wieder unabhängig von Daten des Trainings und anhand ihres Qualitätsmaßes in Abbildung 32 einzusehen. Das qualitative Sugeno-System, dessen Ausgabe wieder auf das Intervall

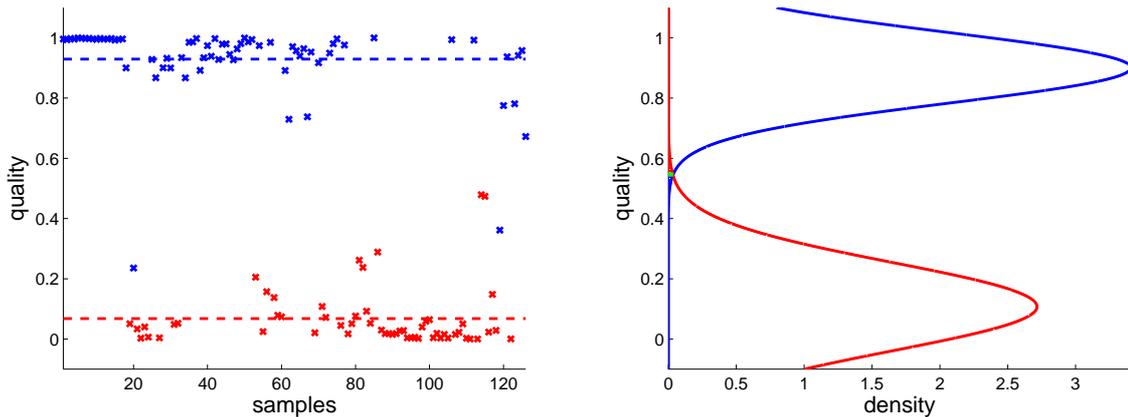


Abbildung 32: Qualitätsmaß bezüglich der kontextuellen Klassifikation eines FFNN, mit richtiger (blau) und falscher Klassifikation (rot) und Mittelwerten (rot/blau gestrichelte Linien), für einen weiteren unbekanntem unabhängigen Datensatz; Dichtefunktionen (rot/blau) und Schwellwert (grün) der Qualitätsmaße (rechts) für die Qualitäten dieses Datensatzes

[0, 1] normalisiert wurde, trennt die richtig von den falsch klassifizierten Daten nahezu perfekt. Die Daten die durch das FFNN falsch klassifiziert wurden, besitzen alle eine Qualität kleiner 0,5. Nur zwei richtig klassifizierte Datenpaare befinden sich unterhalb dieser Schwelle.

Durch diese zweite Bestätigung kann eine erfolgreiche Charakterisierung der Qualitätsanalyse für FFNN festgestellt werden. Dieser zweite Algorithmus zur kontextuellen Klassifikation lässt einen Rückschluss auf alle klassifizierenden Systeme der qualitativen Analysierbarkeit zu, denn mit dem Beweis der universellen Approximation durch unscharfe Regelsysteme von Wang [Wan98] und der Analyse der qualitativen Sugeno-Systeme kann der induktive Schluss bestätigt werden.

Es ist noch zu klären in wie weit die Vollständigkeit der Daten des Trainings für die qualitativen Sugeno-Systeme von Bedeutung ist. Im nächsten Abschnitt soll die Frage beantwortet werden, auf welche Qualität unbekannte kontextuelle Zustände abgebildet werden.

#### 4.4.6 Auswirkung der Trainingsdaten auf das resultierende System zur Qualitätsanalyse $S_Q$

Wurde bisher ein Sugeno-System  $\hat{S}_{Qlin}$  zur Qualitätsanalyse erstellt, so waren die Daten, auf Basis deren das subtraktive Clustering die Struktur erstellte und das korrespondierende ANFIS trainiert wurde, vollständig. In diesen Daten des Trainings waren alle kontextuellen Zustände, die das kontextuelle System erkennen kann, zu gleichen Teilen vorhanden. Das Resultat des Trainings mit einem solchen Datensatz resultierte in einem ausgewogenen System, denn jeder kontextuelle Zustand wurde entsprechend seiner Erkennbarkeit in seinem Qualitätsmaß quantisiert. Doch diese Vollständigkeit eines Trainingsdatensatzes kann nicht immer vorausgesetzt werden, da zu Entwicklungszeiten nicht alle Zustände bekannt sein müssen. Durch diese Möglichkeit der unvollständigen Datensätze tritt die Frage auf, auf welche Qualität ein solcher unbekannter kontextuelle Zustand durch das qualitative Sugeno-System abgebildet wird. Es ist auch zu klären wie die restlichen Qualitäten der bekannten kontextuellen Zustände beeinflusst werden.

In einem ersten Versuch wurden sämtliche Daten zum Training des qualitativen Systems weggelassen, die in dem Zustand 'liegend' entstanden sind. Das normalisierte Resultat der Qualitätsanalyse für einen Datensatz des Testes, der wieder alle kontextuellen Zustände enthält, ist in Abbildung 33 auf der linken Seite dargestellt. Wie festzustellen ist, wurden alle Daten des unbekanntem Zustandes 'liegend' in einer Qualität von Null erkannt. Dieses Resultat ist sehr positiv, da jede andere Qualität keinen Rückschluss auf einen falschen oder unbekanntem Zustand ermöglichen würde. Wird die Funktion des Sugeno-Systems näher betrachtet, so ist dieses Ergebnis jedoch nicht verwunderlich, denn die strukturelle Identifikation des Systems wird durch ein subtraktives Clustering bewerkstelligt. Dieses Clustering bildet für den Identifikator des unbekanntem Zustandes keinen Cluster aus, und somit existiert in diesem Bereich keine Zugehörigkeitsfunktion. Die Regeln die eine Verknüpfung der funktionalen Zugehörigkeiten mittels einer

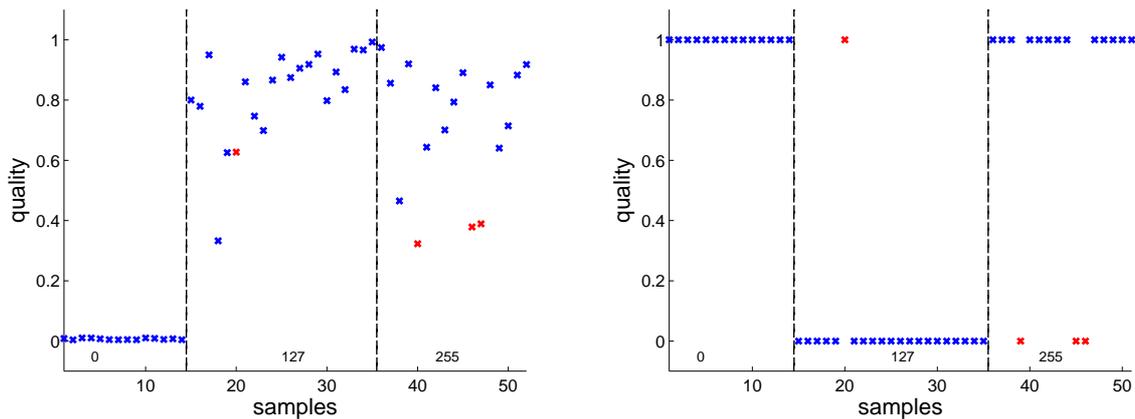


Abbildung 33: Werte der Qualität von qualitativen Sugeno-Systemen denen zum einen der Zustand 'liegend $\hat{=}$ 0' (links) und zum anderen 'bewegt $\hat{=}$ 127' (rechts) beim Training vorenthalten wurde

Produkt-Norm realisieren, werden für einen unbekanntem Zustand mit einem Faktor nahe bei Null sehr klein, was wiederum die gesamte normalisierte gewichtete Summe eines Sugeno-Regelsystems gegen Null konvergieren lässt.

Die dem qualitativen Sugeno-System bekannten Zustände werden in ihrer Qualität wie zuvor bei den mit kompletten Daten trainierten Systemen beurteilt. Diese Feststellung kann wieder in der Abbildung 33 auf der linken Seite für die Zustände 'bewegt' (127) und 'schreibend' (255) eingesehen werden.

In einem zweiten Versuch werden alle Daten, die in dem Zustand 'bewegt' entstanden sind, aus dem Datensatz zum Erstellen und Training des qualitativen Sugeno-Regelsystems, eliminiert. Wie in der Abbildung 33 auf der rechten Seite festgestellt werden kann, sind die Werte der Qualität absolut, sie sind entweder Null oder Eins. In der zustandsabhängigen Analyse wurde zuvor festgestellt, dass ein kontextueller Klassifikator die Zustände einer Bewegtheit von einer nicht Bewegtheit gänzlich unterscheiden kann. Wird also einer der dynamischen Zustände wie 'bewegt' und 'schreibend' entfernt, so kann mit nahezu vollständiger Sicherheit die Richtigkeit der Klassifikation entschieden werden. Der Grund dafür liegt wieder in der Methode des subtraktiven Clusterings, das an der Stelle des unbekanntem kontextuellen Identifikators keine Zugehörigkeitsfunktion platziert. Die normalisierte gewichtete Summe konvergiert wieder gegen Null und somit die Ausgabe des qualitativen Sugeno-Systems. Diese Feststellung wird durch die falsch klassifizierte Zustände (rote Kreuze) des kontextuellen Systems bestätigt, da diese für den bekannten Zustand gegen Eins streben und für den unbekanntem gegen Null.

Ein qualitatives System, das aufgrund unvollständiger Daten mit einem unbekanntem kontextuellen Zustand konfrontiert wird, besitzt eventuell Aussagekraft über die bekannten Zustände, nicht jedoch über den unbekanntem. Der unbekanntem Zustand wird immer mit einem Wert der Qualität nahe bei Null beurteilt, weshalb er in einer weiteren Verarbeitung keinen oder einen sehr geringen Einfluss besitzt. Wie in der Analyse für den eliminierten Zustand 'bewegt' festgestellt wurde, kann auch die Qualität der bekannten Zustände verändert werden. In dem Fall des Stift-Artefaktes bewirkt die Veränderung eine absolute Kategorisierung, die positiv zu werten ist. Werden jedoch kontextuelle Systeme in ihrer Qualität beurteilt, die mehr als zwei Zustände besitzen die sich einen Raum der Werte von Sensoren teilen, so kann das Ergebnis wesentlich negativer sein. Aus diesen Gründen wird ein vollständiger Datensatz des Trainings gefordert.

## 5 Kontexterkenkung und Qualitätsanalyse zweiter Ebene

Die weitere Verarbeitung der berechneten Qualitäten ist ohne die Betrachtung der Kontexte nicht möglich. Es wird eine Vorschrift der weiteren Verarbeitung benötigt, die sich auf der Art der Fusion oder Aggregation von kontextuellen Identifikatoren begründet. Eine automatisierte Bestimmung eines Systems zur Qualitätsanalyse, das nicht nur mit kontextuellen Informationen und verarbeiteten Werten aus Sensoren gespeist wird, ist sehr komplex, da die komplette Permutation der eingehenden Zustände bewertet werden muss. Für jedes Eingabedatum der Schlussfolgerung aus kontextuellen und qualitativen Identifikatoren muss der Entwickler eine Beurteilung der Richtigkeit zur Verfügung stellen. Dies ist sehr aufwendig und zuweilen sogar unmöglich, da zu Entwicklungszeiten nicht alle Permutationen der Eingabe und Ausgabe bekannt sein müssen.

Auch für die weiteren Schlussfolgerungen, aus vorhandenen Informationen bezüglich kontextueller Zustände, stehen verschiedene Algorithmen zur Verfügung. Der Nachweis, dass die Repräsentation des Fehlers eines kontextuellen Algorithmus durch ein unscharfes Regelsystem geleistet werden kann, wurde schon im letzten Kapitel für Systeme der ersten Ebene der Kontexterstellung gebracht. In dem ergründeten und analysierten Fall der Verarbeitung von Sensordaten war kein Wissen über das kontextuelle Systemverhalten und problematisch zu kategorisierende Sensorwerte vorhanden. Dies ist in der weiteren Verarbeitung von Kontextinformationen nicht so zutreffend, da im allgemeinen zwei Schlussmöglichkeiten unterschieden werden können, der fusionierende und der aggregative Schluss. Es wurde schon in Kapitel 2.2.3 theoretisch erörtert, wie eine Fusion oder eine Aggregation sich auf ein Qualitätsmaß auswirken könnte. Diese Annahmen beziehen sich auf die Eingabe, die unter gewissen Bedingungen eine Ausgabe bewirkt, welche zunächst unabhängig von der Art des Algorithmus ist, der zum Schluss verwendet wird. Es wurde angenommen, dass eine Fusion aus gleichen Kontexten auf einen gleichen Kontext eine umso größere Bestätigung des geschlossenen Kontextes bewirkt, desto mehr Kontexte dieses gleichen Typs vorhanden sind. Aus diesem Grund sollten sich die einzelnen Qualitäten auch bestärken. Wird aus einzelnen kontextuellen Informationen auf einen neuen kontextuellen Zustand geschlossen, so vermehrt sich die

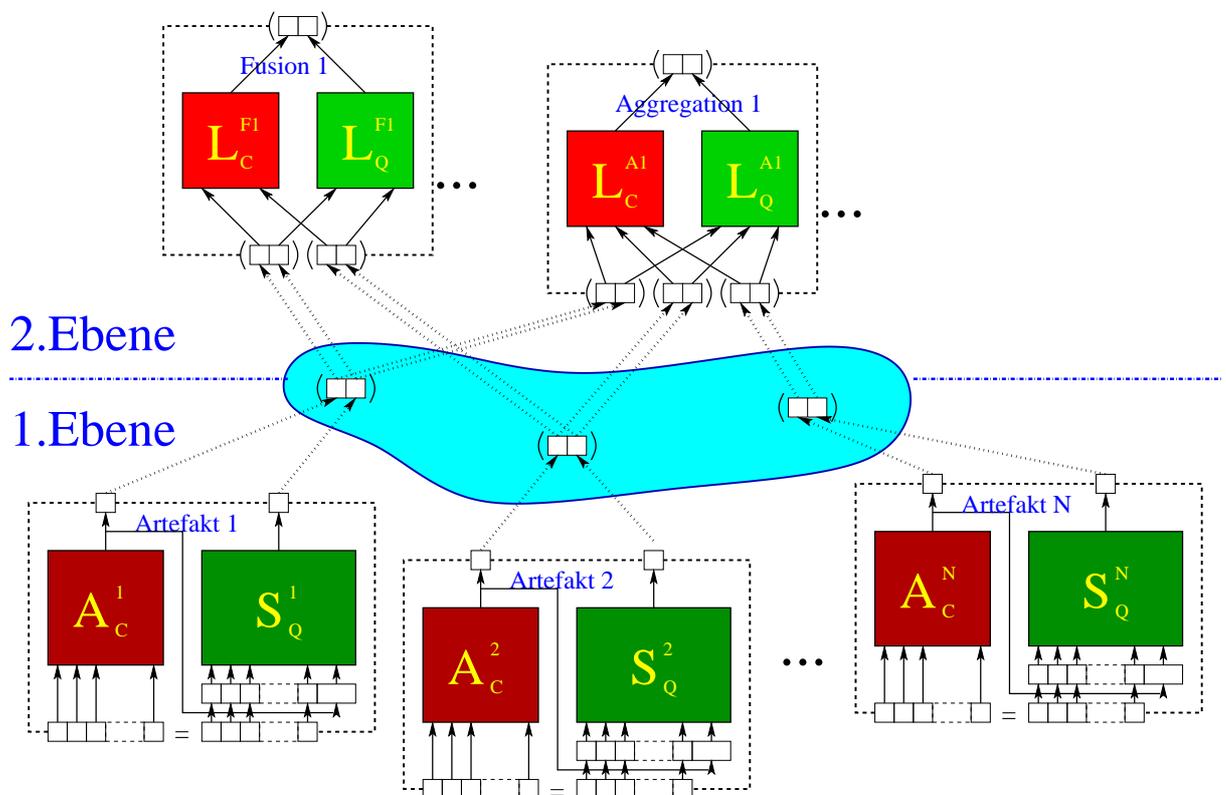


Abbildung 34: Darstellung der Verbindung von Systemen der Generierung (Artefakt 1-N) und der Ableitung (Fusion 1 und Aggregation 1), die durch einen Raum (hellblauer Bereich) von kontextuellen/qualitativen Tupeln verbunden sind, wobei die kontextuelle Verarbeitung  $L_C^{F1}$  und  $L_C^{A1}$  rot und die qualitative  $L_Q^{F1}$  und  $L_Q^{A1}$  grün gekennzeichnet ist

---

Ungenauigkeit durch die Anzahl der Eingaben. Deshalb bewirkt eine Aggregation eines neuen Kontextes eine Verminderung der Gewissheit der einzelnen kontextuellen Zustände und somit eine Abschwächung der Qualität des resultierenden Kontextes. Diese theoretischen Folgerungen beruhen auf statistisch unabhängigen Eingaben, was für statistisch abhängige nicht der Fall wäre.

Wird das Tupel der kontextuellen und der zugehörigen qualitativen Information als unscharfes Paar erachtet, so könnte die Schlussfolgerung aus einer solchen Eingabe im Ganzen durch eine unscharfe Logik erfolgen. Diese Methodik erfordert jedoch eine Festlegung des kontextuellen schließenden Systems, was bei der Fokussierung dieser Arbeit auf ein Qualitätsmaß zu restriktiv wäre. Es soll dem Entwickler freigestellt werden, welche Systeme zur Verarbeitung von Kontexten verwendet werden, ohne die Verwendbarkeit der hier vorgestellten Qualitätsanalyse zu verhindern.

Durch die nun getroffenen Ansprüche an eine Verarbeitung von Qualitätsmaßen kann nun der allgemeine Aufbau für Schlüsse aus Qualitäten definiert werden. In der Abbildung 34 sind zwei Systeme zum Schluss aus Kontexten und Qualitäten dargestellt, eines für einen fusionierenden und eines für einen aggregativen Schluss, wobei die Schlüsse für Kontext und Qualität in getrennten Bereichen bearbeitet werden. Die Verbindung mit Systemen der ersten Ebene, die auf einen kontextuellen Zustand und dessen Qualität aus vorverarbeiteten Sensordaten schließen, erfolgt über einen Tupel-Raum. Dieser Tupel-Raum, der in der Abbildung als hellblaue Wolke dargestellt ist, wird als nicht speichernd definiert, da diese Vereinfachung für die Ergründung von Qualitäten der zweiten Ebene ausreicht. Werden also mehrere Kontexte mit deren zugehörigen Qualitäten für Schlussfolgerungen verwertet, so müssen diese zur gleichen Zeit auftreten.

Das Schließen aus Qualitäten auf eine resultierende Qualität, wird im ersten Abschnitt formal definiert. Es werden n-dimensionale Bedingungen angenommen und mit Hilfe von unscharfen Operatoren fusioniert und aggregiert.

Im zweiten Abschnitt dieses Kapitels wird ein weiteres Artefakt beschrieben, das zur praktischen Evaluierung der Schlüsse benötigt wird. Dieses Hand-Artefakt bildet mit dem im letzten Kapitel beschriebenen Stift-Artefakt die Grundlage für Grafiken, die zur statistischen Analyse verwendet werden.

Die eigentliche statistische Analyse der Verbindung von Stift- und Hand-Artefakt ist im letzten Abschnitt beschrieben. Für diese Analyse wurden zwei Sätze an Daten für beide Artefakte erstellt, die eine zufällige und lange Abfolge von kontextuellen Zuständen der beiden Artefakte beschreibt.

## 5.1 Allgemeiner Aufbau von Systemen zweiter Ebene

Die Systeme der zweiten Ebene sind wieder in zwei Teile unterteilt, ein Teil übernimmt die Ableitung der Kontexte und der andere die der Qualitäten. Der Algorithmus der Kontexterkennung ist wieder nicht bekannt, kann jedoch dadurch beschrieben werden, ob er eine Fusion oder eine Aggregation ausführt. Entsprechend dieser semantischen Beschreibung kann die Verarbeitung der Qualitäten definiert werden, was durch eine abgewandelte unscharfe Logik geschieht. Diese Abwandlung äußert sich durch die spezielle Eingabe der unscharfen Schlussfolgerung, da schon bekannt ist, welcher kontextuelle Zustand die Qualität hervorgerufen hat. Die Eingabe der Schlussfolgerung ist daher nicht ein Element eines ursprünglichen Universums, das erst unscharf interpretiert wird, sondern alleinig die 'Unschärfe'. Es wird also nur der Grad der Zugehörigkeit, welcher bisher als Qualität bezeichnet wurde, als Eingabe der unscharfen qualitativen Schlussfolgerung benutzt. Es ist für diese Schlussfolgerung irrelevant, zu welcher unscharfen Menge dieser Grad eine Zugehörigkeit ausdrückt, sondern nur welche Art der Schlussfolgerung verwendet wird.

Die Grundlage der Verarbeitung und die Eingabe des gesamten Systems, ist wieder ein Vektor, der an der Schnittstelle des Systems übergeben wird. Dieser Vektor  $\vec{v}_{CQ}$  (Abbildung 35, innerhalb unterer gestrichelter Linie) enthält die Tupel aus kontextuellen und qualitativen Identifikatoren, die das Teilsystem der Kontextverarbeitung bearbeiten kann. Diese Tupel werden für die einzelnen Teilsysteme aufgeteilt, das System zur qualitativen Schlussfolgerung  $L_Q$  erhält die Qualitäten und das Kontextsystem  $L_C$  die kontextuellen Identifikatoren. Es ist hierbei zu beachten, dass das kontextuelle System  $L_C$  nur eine Abstraktion eines eigentlichen Algorithmus ist, die die Art der Schlussfolgerung repräsentiert. Die Art der

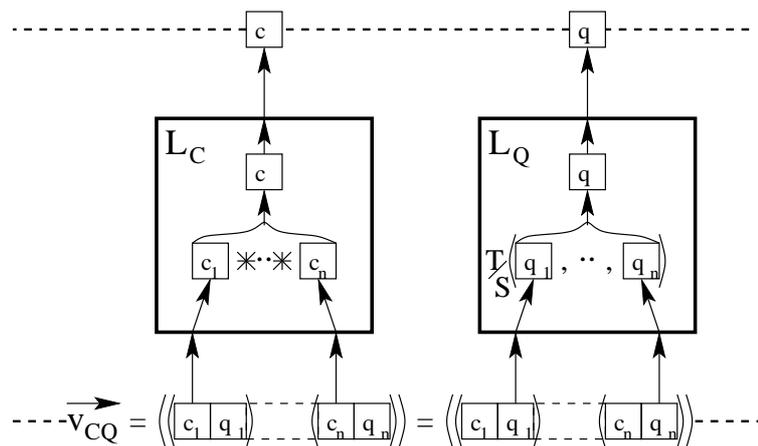


Abbildung 35: Aufbau von Systemen der Zweiten Ebene, mit System für die kontextuelle Schlussfolgerung  $L_C$  und qualitative Schlussfolgerung  $L_Q$

Schlussfolgerung für das kontextuelle System legt den Operator (Abbildung 35, mit \* gekennzeichnet) der Verarbeitung fest, welcher wiederum den Operator der unscharfen qualitativen Ableitung bestimmt. Das in den beiden Schlussfolgerungen entstandene Tupel aus abgeleiteter kontextuelle und qualitativer Information, wird an der oberen Schnittstelle (Abbildung 35, obere gestrichelte Linie) des gesamten Systems zur Verfügung gestellt. Eine weitere Verarbeitung durch andere Systeme kann durch Ablegen der Information in dem zuvor beschriebenen Tupel-Raum ermöglicht werden, diese Möglichkeit wird hier jedoch nicht betrachtet, da sich die Resultate der weiteren Schlussfolgerungen mit den hier betrachteten decken würden.

In den zwei weiteren Abschnitten in diesem Kapitel werden nun die zwei Arten der Schlussfolgerung aus Qualitäten beschrieben. Es wird ergründet, wie die Art der kontextuellen Schlussfolgerung die spezielle unscharfe Schlussfolgerung der Qualitäten definiert.

### 5.1.1 Fusionierender Schluss aus Kontexten und Qualitäten

Der fusionierende Schluss aus mehreren Qualitäten wird durch die logische Abstraktion des kontextuellen Algorithmus definiert, welche ein Modell spezifiziert. Werden mehrere kontextuelle Informationen zur Bestätigung dieser Information herangezogen, so entspricht dies einer logischen Konjunktion. Es ist

eigentlich zu vermuten, dass eine Fusion aufgrund einer Disjunktion vollzogen wird, die die einzelnen Bestandteile der Fusion verknüpft, was jedoch in diesem Modell, der Abstraktion eines realen Algorithmus, nicht der Fall ist. In der klassischen Logik müssen alle Variablen des konjunktiven Terms wahr sein, damit die gesamte Aussage wahr wird. Das hier verwendete Modell kann nur Artefakte zu einer bestätigenden Fusion verarbeiten, die auch den kontextuellen Zustand der Fusion besitzen. Besitzen nicht alle Artefakte den zu fusionierenden Zustand, so können sie nicht durch diese Fusion gegenseitig bestätigt werden. In der unscharfen Verarbeitung existieren beliebig viele Stufen zwischen wahr und falsch, weshalb die Anzahl der bestätigenden Qualitäten der Eingabe, das Resultat der gefolgerten Qualität positiv gewichtet. Die qualitative Fusion basiert auf dem Modell der kontextuellen Fusion, und wird nur berechnet wenn das kontextuelle Modell wahr ist. Die qualitative Fusion wird mit einem unscharfen Operator berechnet (Kapitel 3.1.3), der das erweiterte Äquivalent der Disjunktion ist.

### Abstraktion der kontextuellen Fusion

Die logische Abstraktion einer beliebigen algorithmischen Fusion von  $n$  gleichen kontextuellen Identifikatoren  $x_i$ , wird in einem logischen Term aus Prädikaten wie folgt definiert:

$$t_c = \exists x_1 P_{\underline{c}}(x_1, c) \wedge \exists x_2 P_{\underline{c}}(x_2, c) \wedge \dots \wedge \exists x_n P_{\underline{c}}(x_n, c) \quad (\text{Definition 5.1.1.1})$$

Das Prädikat  $P_{\underline{c}}$  testet die Eingabe einer bestimmten Herkunft, auf die syntaktische Gleichheit zu dem gesuchten kontextuellen Identifikator  $c$  der Fusion. Die Herkunft beschreibt das gewünschte Artefakt, das Bestandteil der Fusion ist, durch die Dimension  $i$  im Vektor  $\vec{\nabla}_{CQ} = ((c_1, q_1), \dots, (c_n, q_n))$  der Eingabe. Wenn ein Artefakt  $i$  der Fusion zu dem Zeitpunkt der Schlussfolgerung einen kontextuellen Zustand  $c_i$  liefert, der nicht durch diesen Term abgedeckt wird, so ist die Prüfung auf syntaktische Gleichheit erfolglos und die Interpretation des Prädikats  $\mathcal{I}(C^{\text{Artefakt}_i}, \beta)(P_{\underline{c}}(c_i, c)) = \text{false}$  unter der Belegung  $\beta(x_i) = c_i$ . Die Interpretation  $\mathcal{I}(C^{\text{Artefakt}_1} \times \dots \times C^{\text{Artefakt}_N}, \beta)(t_c)$  des Terms  $t_c$ , der in der Definition 5.1.1.1 festgelegt wurde, besitzt das kartesische Produkt  $C^{\text{Artefakt}_1} \times \dots \times C^{\text{Artefakt}_N}$  als Menge aller möglichen Belegungen  $\beta$ . Dieses kartesische Produkt besteht aus den Universen der kontextuellen Identifikatoren der Artefakte  $i = 1, \dots, n$ . Das Resultat der Interpretation  $\mathcal{I}$  des Terms  $t_c$  mit der Belegung  $\beta(x_1), \dots, \beta(x_n)$ , die dem aktuellen Vektor der Eingabe  $\vec{\nabla}_{CQ}$  entspricht, ist der zu fusionierende Wahrheitswert  $\text{true}$ , wenn alle der Prädikate wahr sind.

Die Funktion, die bei einem Wahrheitswert  $\text{true}$  der Interpretation  $\mathcal{I}$  des Terms  $t_c$ , auf den kontextuellen Identifikator  $c$  der Fusion abbildet, ist wie folgt definiert:

$$L_C^F(c_1, \dots, c_n) := \begin{cases} c & , \text{ wenn } \mathcal{I}(C^{\text{Artefakt}_1} \times \dots \times C^{\text{Artefakt}_N}, \beta)(t_c) = \text{true} \\ \emptyset & , \text{ sonst.} \end{cases} \quad (\text{Definition 5.1.1.2})$$

Ist der Term  $t_c$  unter der Belegung  $\beta(x_1) = c_k, \dots, \beta(x_n) = c_m$  falsch, so wird durch die Funktion  $L_C^F$  auf die leere Menge  $\emptyset$  abgebildet. Diese Funktion ist, wie schon erwähnt, nur eine Modellierung des realen Algorithmus zur kontextuellen Fusion, da die Adaption an einen normalen Algorithmus sehr umfangreich und deshalb hier kaum zu verwirklichen ist. Da diese Funktion  $L_C^F$  nicht rückgekoppelt wird und als zusätzliche Eingabe für den Algorithmus zur qualitativen Fusion dient, wird mit dieser Abstraktion kein Fehler gemacht. Die einzigen Informationen des realen fusionierenden Algorithmus, die bekannt sein müssen sind, welcher Kontext zu fusionieren ist und welche Artefakte zur Fusion beitragen.

### Definition der qualitativen Fusion

Das Verhalten der abstrahierten kontextuellen Fusion wird nun unscharf für die Qualität erweitert. Es wird der aufgrund des kontextuellen Modells die Fusion der Qualitäten definiert, nun mit unscharfen Operatoren und in einer aussagenlogischen Form, was in folgendem Term resultiert:

$$t_q = S(x_1, x_2, \dots, x_n) = S(x_1, S(x_2, S(\dots S(x_{n-1}, x_n)))) \quad (\text{Definition 5.1.1.3})$$

Wie in Definition 5.1.1.3 eingesehen kann, sind die einzelnen Variablen der Qualitäten durch eine  $S$ -Norm verknüpft, dem unscharfen Äquivalent zum scharfen UND-Operator. Die  $S$ -Norm ist induktiv durch die binäre  $S$ -Norm auf  $n$ -Dimensionen erweitert. Wie schon in Kapitel 3.1.3 beschrieben wurde, existieren viele Möglichkeiten eine  $S$ -Norm zu definieren, es wird jedoch mit der allgemeinen Bezeichnung zunächst nicht auf eine spezielle  $S$ -Norm eingegangen. Eine Wahl bleibt dem Entwickler überlassen, je nachdem welche Ansprüche an das Resultat der Fusion gestellt werden.

Wie schon bei der Abstraktion der kontextuellen Fusion, wird auch bei der Qualitativen Fusion der

Zwischenschritt über die Interpretation  $\mathcal{I}$  des Terms  $t_q$  gegangen. Die Interpretation  $\mathcal{I}$  belegt die Variablen  $x_1, \dots, x_n$  mit den Qualitäten des Vektors  $\vec{v}_{CQ} = ((c_1, q_1), \dots, (c_n, q_n))$ , die aus dem Universum der Qualität des jeweiligen Artefaktes entstammen. Im Gegensatz zur kontextuellen Eingabe sind die Mengen der Qualitäten auf das Intervall  $[0, 1]$  festgelegt, da wieder die normalisierte und nicht klassifizierte Ausgabe der Qualitätsanalysen verwendet wird. Die Interpretation  $\mathcal{I}$  des Terms  $t_q$  ist somit wie folgt definiert:

$$\mathcal{I}([0, 1]^n, \beta)(S(q_1, \dots, q_n)) \in [0, 1] \quad (\text{Definition 5.1.1.4})$$

Die Funktion  $L_Q^F$  ist, wie schon bei der kontextuellen Fusion, die Schlussfolgerung der qualitativen Fusion. Sie gibt im Fall der Fusion von Qualitäten nur den Wert der Interpretation  $\mathcal{I}$  des Terms  $t_q$  zurück, was in formeller Schreibweise wie folgt definiert ist:

$$L_Q^F(q_1, \dots, q_n) := \begin{cases} \mathcal{I}([0, 1]^n, \beta)(S(q_1, \dots, q_n)) & , \text{ wenn } L_C^F(c_1, \dots, c_n) \neq \emptyset \\ \emptyset & , \text{ sonst.} \end{cases} \quad (\text{Definition 5.1.1.5})$$

Wurde das Ergebnis der Qualitätsanalyse der ersten Ebene nicht nur auf das Intervall  $[0, 1]$  normalisiert, sondern auch einer bestimmten Qualitätsklasse  $[q_k] \in Q_K$  zugeordnet, so müssen die für die Fusion verwendeten Artefakte alle auf die gleiche Menge von Qualitätsklassen abbilden. Sind die Mengen der Qualitätsklassen nicht gleich, so können die Identifikatoren der jeweiligen Klassen nicht in einer Fusion miteinander verrechnet werden, da die Vergleichbarkeit nicht gegeben ist. Werden Qualitäten der gleichen Mengen miteinander verrechnet, so ist die Funktion  $L_Q^F$  wie folgt definiert:

$$L_Q^F(q_{1k}, \dots, q_{nk}) := \begin{cases} \mathcal{I}(Q_K^n, \beta)(S(q_{1k}, \dots, q_{nk})) & , \text{ wenn } L_C^F(c_1, \dots, c_n) \neq \emptyset \\ \emptyset & , \text{ sonst.} \end{cases} \quad (\text{Definition 5.1.1.6})$$

Die einzelnen Qualitätsklassen  $[q_k]$  werden hierbei durch ihre Identifikatoren  $q_{ik}$ , des jeweiligen Artefaktes  $i$ , repräsentiert und verrechnet.

Diese Methode der Fusion von Qualitäten ist durch die Definition aufgrund der abstrahierten kontextuellen Fusion mit einem Verlust an Information behaftet. Es ist jedoch nicht bekannt wie ein realer kontextueller Algorithmus die Fusion durchführen würde, weshalb dieser Fehler zu vernachlässigen ist. Die größere Bedeutung hat die universelle Einsetzbarkeit des Qualitätsmaßes, das für die Erstellung und Verrechnung keine Informationen benötigt.

### 5.1.2 Aggregativer Schluss aus Kontexten und Qualitäten

Der aggregative Schluss aus Kontexten schließt aus beliebigen kontextuellen Zuständen auf einen neuen Zustand. Auch im Fall der Aggregation wird der Schluss aus Qualitäten über den Schluss aus Kontexten definiert. Bei dieser Möglichkeit des Schlusses gehen jedoch mehr Informationen durch die Abstraktion des kontextuellen Algorithmus verloren. Eine Aggregation kann z.B. die einzelnen kontextuellen Zustände unterschiedlich gewichten, weshalb ein solcher aggregativer Schluss mehr auf bestimmte Aspekte der Eingabe abgerichtet ist und auf andere weniger. Dieser Sachverhalt muss hier jedoch aus genannten Gründen vernachlässigt werden.

#### Abstraktion der kontextuellen Aggregation

Die Aggregation eines neuen Kontextes beruht wieder auf bestimmten kontextuellen Identifikatoren, welche von Artefakten  $i$  stammen. Diese Identifikatoren sind in dem Vektor der Eingabe  $\vec{v}_{CQ} = ((c_1, q_1), \dots, (c_n, q_n))$  gegeben, wobei die die Tupel der Paare aus Kontext und Qualität getrennt werden. Diese kontextuellen Identifikatoren bilden die Belegung des Terms  $t_c$ , der die Abstraktion des aggregativen Schlusses bestimmt. Der Term  $t_c$  ist wieder ein Konstrukt aus Prädikaten  $P_{\pm}$ , die jedoch jetzt über eine Konjunktion wie folgt verknüpft werden:

$$t_c = \exists x_1 P_{\pm}(x_1, c_1) \wedge \exists x_2 P_{\pm}(x_2, c_2) \wedge \dots \wedge \exists x_n P_{\pm}(x_n, c_n) \quad (\text{Definition 5.1.2.7})$$

Das Prädikat  $P_{\pm}(x_i, c_k)$  entscheidet die syntaktische Gleichheit für eine Belegung  $\beta(x_i) = q_i$  der Variablen  $x_i$  und des zu prüfenden kontextuellen Zustandes  $c_i$ . Die Identifikatoren  $c_i$  der kontextuellen Zustände sind Elemente der Menge aller Zustände  $C = \{c_1, \dots, c_n, c\}$ , die durch den Term  $t_c$  verarbeitet werden und den Zustand  $c$  der durch die Aggregation abgeleitet wird.

Die Interpretation  $\mathcal{I}$  liefert wieder den Wahrheitswert *true* oder *false* des Terms  $t_c$ . Die Universen  $C^{\text{Artefakt}_i}$  der Eingabe werden durch die in der Aggregation zu verarbeitenden Artefakte  $i = 1, \dots, n$

bestimmt. Die Funktion  $L_C^A$  liefert den zu aggregierenden kontextuellen Zustand  $c$  aufgrund des Wahrheitswerts, der durch die Interpretation  $\mathcal{I}$  bestimmt wird, wie folgt:

$$L_C^A(c_1, \dots, c_n) := \begin{cases} c & , \text{ wenn } \mathcal{I}(C^{\text{Artefakt}_1} \times \dots \times C^{\text{Artefakt}_n}, \beta)(t_c) = \text{true} \\ \emptyset & , \text{ sonst.} \end{cases} \quad (\text{Definition 5.1.2.8})$$

Wenn der Term  $t_c$  unter der Belegung  $\beta(x_1) = c_k, \dots, \beta(x_n) = c_m$ , mit  $c_k \in C^{\text{Artefakt}_1}, \dots, c_m \in C^{\text{Artefakt}_n}$ , falsch ist, so wird die leere Menge  $\emptyset$  durch die Funktion  $L_C^A$  zurück gegeben. Auch diese Funktion  $L_C^A$  ist nur eine Abstraktion des realen Algorithmus des aggregativen Schlusses, die das Modell für die qualitative Aggregation bildet.

### Definition der qualitativen Aggregation

Die Erweiterung des scharfen Terms der kontextuellen Abstraktion auf einen unscharfen Term, ist die Grundlage des aggregativen Schlusses aus Qualitäten. Die Eingabe sind die qualitativen Identifikatoren  $q_i$  des Vektors  $\vec{v}_{CQ} = ((c_1, q_1), \dots, (c_n, q_n))$  der kontextuellen und qualitativen Tupel, die die Belegung  $\beta(x_i) = q_i$  des Terms  $t_q$  bilden. Der Term, der die Aggregation bestimmt, ist wie folgt definiert:

$$t_q = T(x_1, x_2, \dots, x_n) = T(x_1, T(x_2, T(\dots T(x_{n-1}, x_n)))) \quad (\text{Definition 5.1.2.9})$$

Eine bestimmte  $T$ -Norm, die das unscharfe Äquivalent zu der scharfen Konjunktion bildet, wurde nicht festgelegt. Die Wahl einer spezifischen  $T$ -Norm bleibt dem Entwickler überlassen, da diese seinen Ansprüchen genügen muss. Auch die allgemeine  $T$ -Norm aus Definition 5.1.1.3 ist wieder induktiv aus der binären Norm erweitert.

Die Interpretation  $\mathcal{I}$  des Terms  $t_q$  auf den Intervallen  $[0, 1]$  der Qualität und der Belegung  $\beta(x_1) = q_1, \dots, \beta(x_n) = q_n$  der Variablen  $x_i$  mit  $i = 1, \dots, n$ , ist wie folgt abzuleiten:

$$\mathcal{I}([0, 1]^n, \beta)(T(q_1, \dots, q_n)) \in [0, 1] \quad (\text{Definition 5.1.2.10})$$

Die Funktion, die aufgrund der Interpretation  $\mathcal{I}$  den zu aggregierenden Kontext  $c$  oder, im Falle einer Interpretation des Terms  $t_q$  von *false*, die leere Menge  $\emptyset$  zurück geben soll, ist wie folgt definiert:

$$L_Q^A(q_1, \dots, q_n) := \begin{cases} \mathcal{I}([0, 1]^n, \beta)(T(q_1, \dots, q_n)) & , \text{ wenn } L_C^A(c_1, \dots, c_n) \neq \emptyset \\ \emptyset & , \text{ sonst.} \end{cases} \quad (\text{Definition 5.1.2.11})$$

Ist der qualitative Identifikator nicht ein normalisiertes Element des Intervalls  $[0, 1]$ , sondern ein Element einer Menge von Qualitätsklassen  $[q_k] \in Q_K$ , so ist die Funktion  $L_Q^A$  wie folgt definiert:

$$L_Q^A(q_{1k}, \dots, q_{nl}) := \begin{cases} \mathcal{I}(Q_K^n, \beta)(T(q_{1k}, \dots, q_{nl})) & , \text{ wenn } L_C^A(c_1, \dots, c_n) \neq \emptyset \\ \emptyset & , \text{ sonst.} \end{cases} \quad (\text{Definition 5.1.2.12})$$

Es wird wieder gefordert, dass die Mengen der Qualitätsklassen für jedes Artefakt, das zu dem aggregativen Schluss beiträgt, gleich sind. Ist diese Gleichheit nicht vorhanden, so können die Elemente der Menge der Qualitätsklassen nicht in einer Aggregation abgeleitet werden.

Die Aggregation von Qualitäten ist auch wie bei der Fusion wieder fehlerhaft, da sie nicht auf Basis des Verhaltens eines realen Algorithmus zur kontextuellen Aggregation definiert wurde, sondern eines Modells von Prädikaten. Es kann nicht bestimmt werden, wie groß dieser Fehler ist, da dies von der Art des realen und nicht vorhandenen Algorithmus abhängt. Die Verwendung einer Logik zur Aggregation von neuen Kontexten ist jedoch eine sehr nahe liegende Methodik, die auch gegenüber dem hier verwendeten Modell der Abstraktion keinen Fehler hervorruft. Die Verfügung eines universellen Qualitätsmaßes, dessen Fehler bei der Aggregation und der Fusion gegenüber dem Mehrwert der Auswertung von Qualitäten zu vernachlässigen ist, steht hier im Vordergrund. Dieser Mehrwert wurde im Zusammenhang mit der statistischen Analyse in Kapitel 4.4 ergründet.

## 5.2 Das Hand-Artefakt

Wird eine fusionierender oder aggregativer Schluss aus Kontexten in einer mehrstufigen Erkennung angestrebt, so benötigt man zu dem vorhandenen Stift-Artefakt ein weiteres, um die Schlüsse zu ermöglichen. In diesem Abschnitt wird ein Hand-Artefakt vorgestellt, das gleiche und unterschiedliche kontextuelle Zustände besitzt, wie das ausführlich beschriebene Stift-Artefakt. Die Beschreibung und Analyse dieses neuen Hand-Artefakts wird recht kurz gehalten, da die essentiellen Bestandteile mit dem Stift-Artefakt übereinstimmen und schon mit diesem erklärt wurden.

In der Abbildung 36 auf der linken Seite können die einzelnen Bestandteile des Hand Artefaktes eingesehen werden, das aus einem Partikel 2/29, einem Sensor-Board 1/92, einer Batterie als Stromversorgung und einer Halterung zur Befestigung an der Hand zusammensetzt ist. Die Montage des Hand-Artefaktes

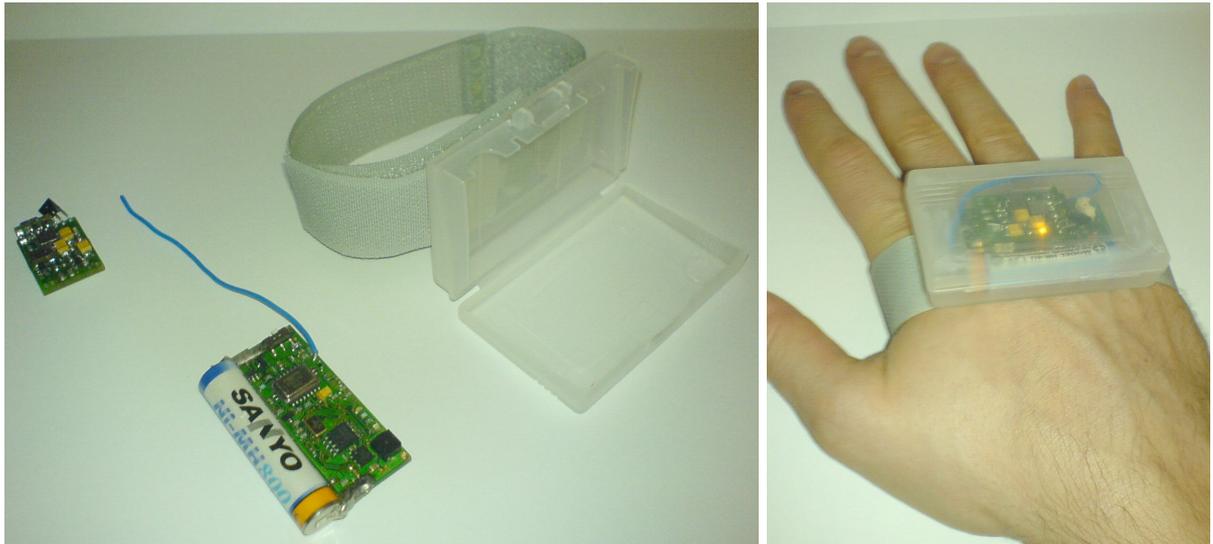


Abbildung 36: Bestandteile des Hand-Artefakts in Einzelteilen (links) und montiert auf einer Hand (rechts)

an einer Hand ist in Abbildung 36 auf der rechten Seite abgebildet. Die kontextuellen Zustände werden wieder über die Interpretation der Beschleunigung über die x-, y- und z-Achse erkannt.

In den nächsten Abschnitten wird zuerst der Aufbau des Systems zur Kontexterkenkung beschrieben. Dieses unterscheidet sich in seiner Konstruktion von dem Sugeno-System, das zur Bestimmung von kontextuellen Zuständen des Stift-Artefaktes eingesetzt wurde. In einem weiteren Abschnitt in diesem Kapitel wird näher auf das System zur Qualitätsanalyse des kontextuellen Algorithmus eingegangen, das dem analytischen System des kontextuellen neuronalen Netz des Stift-Artefaktes ähnelt. Der letzte Abschnitt in diesem Kapitel beinhaltet eine kurze statistische Analyse der spezifischen Qualitäten de Hand-Artefaktes.

### 5.2.1 Erkennung von kontextuellen Zuständen mit dem Hand-Artefakt

Das Hand-Artefakt wurde zur Erkennung von vier verschiedenen kontextuellen Zuständen konstruiert, den Zuständen 'ruhig' mit dem Identifikator 0, 'klatschend' mit 85, 'schüttelnd' mit 170 und 'schreibend' dem der Identifikator '255' zugeordnet wurde. Der Zustand 'ruhig' deckt sich, wie an dem Identifikator 0 erkannt werden kann, mit dem Zustand 'liegend' des Stift-Artefaktes, genauso wie die Zustände 'schreibend'.

Wird wieder die Zugehörigkeit zu einem Element der Menge der kontextuellen Zustände  $C = \{0, 85, 170, 255\}$  über eine Menge von unscharfen Dreiecks-Zahlen  $\tilde{C} = \{\tilde{0}, \tilde{85}, \tilde{170}, \tilde{255}\}$  entschieden, so werden die Intervallgrenzen anhand der Schnitte der funktionalen Zugehörigkeiten entschieden. Die linearen funktionalen Zugehörigkeiten sind für diese vier unscharfen Zahlen in Abbildung 37 abgebildet. Da der Bereich der kontextuellen Identifikatoren gleich geblieben ist, jedoch jetzt vier Zustände darin Platz finden, sind die Intervalle, die die Zugehörigkeit zu einem kontextuellen Zustand bestimmen, kleiner geworden. Da die Projektion eines kontextuellen Sugeno-Systems auf diese Intervalle schon in dem Fall des Stift-Artefaktes

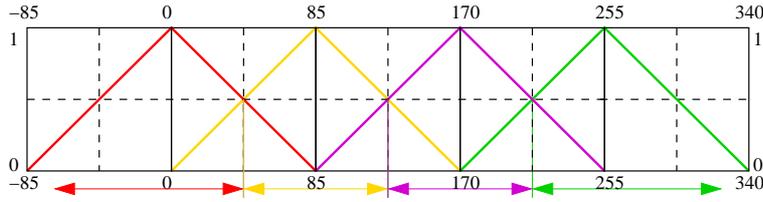


Abbildung 37: Funktionale Zugehörigkeiten für die unscharfen Zahlen der Menge  $\tilde{C} = \{\tilde{0}, \tilde{85}, \tilde{170}, \tilde{255}\}$ , deren Schnittpunkte die Intervalle festlegen, die die Größe der Kontextklassen bestimmen

mit drei kontextuellen Zuständen problematisch war, wurden weitere Kriterien der Eingabe benötigt, die eine Unterscheidung der Zustände ermöglichen. Auch die höhere Anzahl der Zustände der Bewegtheit, bewirkt eine erschwerte Unterscheidung der Zustände. Wie schon bei dem neuronalen Netz zur Erkennung der Zustände des Stift-Artefaktes, wurde auch für die kontextuelle Sugeno-Logik des neu entwickelten Hand-Artefaktes der Vektor der Eingabe, der die Standardabweichungen für alle Achsen enthält, um die Mittelwerte der Beschleunigung für jede Achse ergänzt. Dieser 6-dimensionale Vektor  $\nabla_C^{FuzzyHand} := (v_1, v_2, v_3, v_4, v_5, v_6)$  der Eingabe bewirkt, mit der Methode der Gitter Partitionierung, ein kontextuelles Sugeno-Regelsystem  $\hat{\mathbf{S}}_{C_{lin}}^{FuzzyHand}$  mit einer hohen Anzahl von Regeln. Diese erhöhte Anzahl bewirkt nicht nur ein komplexeres Regelsystem, sondern auch ein korrespondierendes ANFIS mit mehr Neuronen, das zum Training des Systems verwendet wird. Um eine komplexere Architektur des ANFIS und des Regelsystems zu vermeiden, wurde im Zusammenhang mit den qualitativen Regelsystemen eine Methode der Reduktion eingesetzt, das subtraktive Clustering. Da das Hand-Artefakt ein neu entworfenes System ist und nicht auf die Arbeiten von Estévez [Est05][Est06] zurückgeht, wurde das subtraktive Clustering zur Identifikation des kontextuellen Sugeno-Systems  $\hat{\mathbf{S}}_{C_{lin}}^{FuzzyHand}$  eingesetzt.

Die weitere Spezifikation des kontextuellen Sugeno-Regelsystems  $\hat{\mathbf{S}}_{C_{lin}}^{FuzzyHand}$  ist mit dem Sugeno-System des Stift-Artefaktes identisch. Zuerst wird wieder ein Satz von Daten zum Erstellen und Training des Systems bestimmt, der die Eingabedaten und die gewünschte Ausgabe für jedes Datenpaar enthält. Das subtraktive Clustering liefert für diesen Datensatz die Anzahl der funktionalen Zugehörigkeiten für jede Dimension der Eingabe und somit auch die Anzahl der Regeln. Mit dieser strukturellen Information kann das ANFIS aufgebaut werden. In der ersten Iteration der Methode der kleinsten Quadrate werden die Parameter der initialen linearen Konsequenzen bestimmt. Die Feinabstimmung der Parameter des ANFIS erfolgt wieder über das hybride Training, das das Backpropagation und die Methode der kleinsten Quadrate vereint. Das hybride Training stoppt beim Auffinden eines guten Minimums. Das resultierende ANFIS bestimmt das Sugeno-Regelsystem  $\hat{\mathbf{S}}_{C_{lin}}^{FuzzyHand}$ , das im Rahmen des Hand-Artefaktes zur Erkennung von kontextuellen Zuständen eingesetzt wird.

Die Ausgabe des kontextuellen Sugeno-Systems muss noch klassifiziert werden, was durch die Zuordnung zu den zuvor bestimmten Intervallen der kontextuellen Identifikatoren  $C = \{0, 85, 170, 255\}$  geschieht, da der Ausgabewert in genau einem dieser Intervalle liegt. Die Intervalle wurden durch die Schnittpunkte der unscharfen Dreiecks-Zahlen  $\tilde{C} = \{\tilde{0}, \tilde{85}, \tilde{170}, \tilde{255}\}$  bestimmt, die in der Abbildung 37 über ihre funktionalen Zugehörigkeiten dargestellt sind. Der identifizierte kontextuelle Identifikator ist Ausgabe des kontextuellen Systems  $\mathbf{S}_C^{FuzzyHand}$  und wieder Eingabe des korrespondierenden qualitativen Systems  $\mathbf{S}_Q^{FuzzyHand}$ , das im nächsten Abschnitt beschrieben wird.

### 5.2.2 Spezifikation der Qualitätsanalyse für das Hand-Artefakt

Das analytische System zur Bestimmung einer Qualität zu einem Datum der Eingabe und einem erkannten Kontext ist dem der Qualitätsanalyse für das kontextuelle neuronale Netz, das im Zusammenhang mit dem Stift-Artefakt benutzt wurde, ähnlich. Der einzige Unterschied liegt in der erhöhten Anzahl der kontextuellen Zustände und den damit verbundenen höheren Anzahl von Identifikatoren. Diese größere Anzahl wird durch das subtraktive Clustering erkannt und resultiert in einer größeren Anzahl von Clustern. Eine höhere Zahl an Clustern bedeutet mehr funktionale Zugehörigkeiten für jede Dimension der Eingabe.

Es wird wieder der gleiche Satz an Daten für das Erstellen und Training des qualitativen ANFIS einge-

setzt, der schon für das kontextuelle Sugeno-System verwendet wurde. Dieser Satz an Daten ergibt durch das subtraktive Clustering für jede der sieben Dimensionen der Eingabe sieben funktionale Zugehörigkeiten, die wieder durch das subtraktive Clustering auf Gauss-Funktionen limitiert sind. Da jede funktionale Zugehörigkeit über alle sieben Dimensionen für einen Cluster steht, jeder Cluster eine Regel repräsentiert und jede Regel eine lineare funktionale Konsequenz bedingt, besteht das ANFIS aus sieben Neuronen der Eingabe- und der Zwischenschicht. Die Parameter der linearen Konsequenzen sind mit der Methode der kleinsten Quadrate zu bestimmen, wobei auch die Feinabstimmung über diese Methode im hybriden Trai-

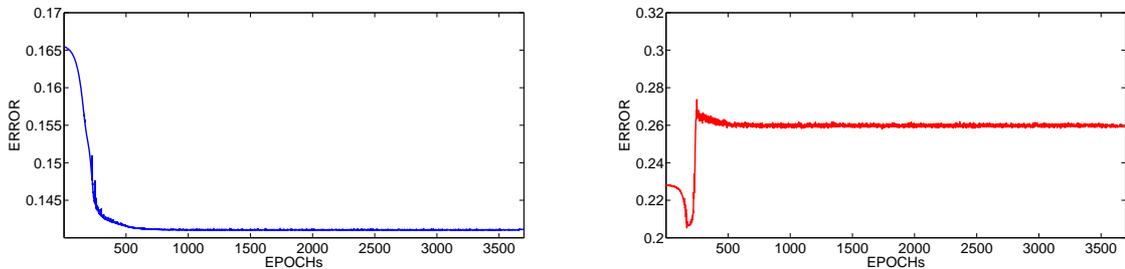


Abbildung 38: Fehler der Ausgabe des qualitativen Sugeno-Systems gegenüber gewünschter Ausgabe von Daten des Trainings (links) und des Testes (rechts)

ning erfolgt. Das Backpropagation trainiert die Parameter der Gauss'schen Zugehörigkeitsfunktionen, bis ein möglichst globales Minimum des Fehlers für den gesamten Satz an Daten gefunden wurde. Ein globales Minimum wurde ungefähr ab der Epoche  $\approx 3600$  nicht mehr überschritten, was auch in der Abbildung 38 auf der linken Seite für die Trainingsdaten eingesehen werden kann. Die Entwicklung des Fehlers, für einen unabhängigen Satz an Testdaten, ist in der gleichen Abbildung auf der rechten Seite zu sehen. Die Testdaten zeigen ein Minimum bei der Epoche 170, die auch die Epoche des Endes des Trainings bedeutet.

Das nach dem hybriden Training vorhandene System wird zur Qualitätsanalyse bezüglich der Kontexte des Stift-Artefaktes eingesetzt, wobei die Ausgabe des Sugeno-Systems  $\hat{\mathbf{S}}_{Q_{lin}}^{FuzzyHand}$  nur auf das Intervall  $[0, 1]$  normalisiert wird, was die Ausgabe der Qualitätsanalyse  $\mathbf{S}_Q^{FuzzyHand}$  ergibt, und nicht, wie auch in Kapitel 4.1.5 beschrieben wurde, einer Klasse der Qualität zugeordnet wird. Die normalisierten Qualitäten für die unabhängigen Daten des Testes, werden im nächsten Abschnitt kurz analysiert.

### 5.2.3 Analyse der Qualitäten

Die allgemeinen Eigenschaften, die das Qualitätsmaß besitzt, wurden ausführlich in Kapitel 4.4 statistisch analysiert. In diesem Abschnitt soll nur kurz die Wirksamkeit der qualitativen Analyse  $\mathbf{S}_Q^{FuzzyHand}$  für das Stift-Artefakt nachgewiesen werden.

Die Ergebnisse der Qualitätsanalyse sind für die Daten des Testes in der Abbildung 39 auf der linken Seite einzusehen, wieder mit den kontextuell richtig klassifizierten Datenpaaren in blau und den falsch klassifizierten in rot. Die Dichtefunktionen der Normalverteilungen sind in der gleichen Abbildung auf der rechten Seite abgebildet, die Dichte für die Richtigen in blau und die Falschen in Rot. Der Schnittpunkt der Dichtefunktionen bildet wieder den Schwellwert  $s$  für die Wahrscheinlichkeiten. Die Wahrscheinlichkeit dass eine Qualität  $q$  wirklich eine richtige kontextuelle Klassifikation  $c = true$  identifiziert, ist mit  $\mathbb{P}(c = true|q > s) = 0.6226$  gegeben. Eine Wahrscheinlichkeit unter dem Schwellwert  $s$  anhand der Qualität  $q$  wirklich einen falsch klassifizierten Kontext  $c = false$  zu erkennen, wird mit  $\mathbb{P}(c = false|q < s) = 0.6226$  bewertet. Die Wahrscheinlichkeit der falschen Positiven ist  $\mathbb{P}(c = false|q > s) = 0.0122$  und der falschen Negativen ist  $\mathbb{P}(c = true|q < s) = 0.3652$ . Diese Wahrscheinlichkeiten zeigen eine gut mögliche Kategorisierung der Kontexte durch das Qualitätsmaß und somit einen Erfolg der qualitativen Analyse für das kontextuelle System des Hand-Artefaktes.

Die beiden Systeme, zur Qualitätsanalyse und zur kontextuellen Klassifikation des Hand-Artefaktes, bilden nun mit den unscharfen Systemen des Stift-Artefaktes die Grundlage der Verarbeitung auf höheren Schichten. Im nächsten Abschnitt werden die Grundlagen erläutert, wie die Fusion und die Aggregation von Kontexten, und somit der Qualitäten, bewerkstelligt werden können.

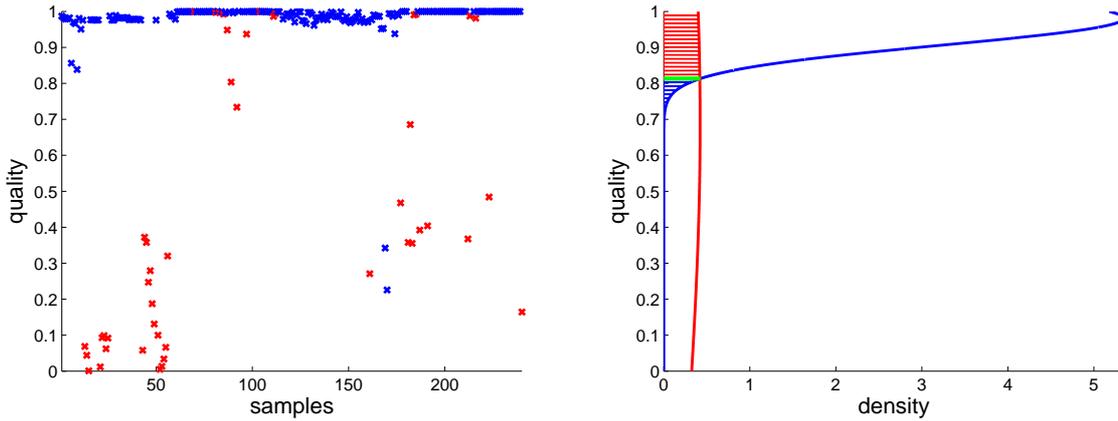


Abbildung 39: Maß der Qualität für das kontextuelle Sugeno-System und einen Satz an Testdaten (links), mit kontextuell richtig klassifizierten Daten (blau) und falsch klassifizierten (rot); Dichtefunktionen für die Verteilung der Qualitäten (rechts), mit Dichte für falsch (rot) und richtig (blau) klassifizierte Daten

### 5.2.4 Schlüsse aus Stift- und Hand-Artefakt

Die Verbindung des Stift- und des Hand-Artefaktes erfolgt über den nicht speichernden Tupel-Raum, durch den die Tupel aus Kontext und Qualität übergeben werden. Die kontextuelle Fusion und Aggregation werden über das Modell der abstrahierten Prädikatenlogik ausgeführt, da keine speziellen Algorithmen für den Schluss vorhanden sind und auch die Evaluation des Qualitätsmaßes im Vordergrund steht. Der Schluss der Fusion und der Aggregation findet auf der zweiten Ebene statt, welcher nur erstellte kontextuelle und qualitative Informationen und keine Daten von Sensoren verarbeitet.

Der Aufbau der Fusion und der Aggregation und die Verbindung mit den Systemen der ersten Ebene, dem Stift- und dem Hand-Artefakt, ist in der Abbildung 40 zu sehen. Die Eingabe der Systeme des

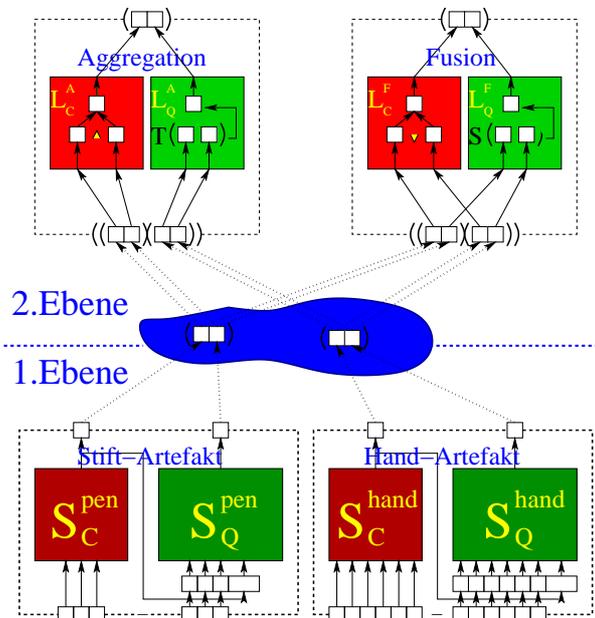


Abbildung 40: Aufbau von Stift- und Hand-Artefakt, mit Verbindung zu Systemen zweiter Ebene über einen Tupel-Raum (blau); Systeme zweiter Ebene des fusionierenden und aggregativen Schlusses

Schlusses ist ein Vektor  $\vec{\nabla}_{CQ} = ((c_{pen}, q_{pen}), (c_{hand}, q_{hand}))$  für die Fusion und die Aggregation, der die Tupel aus Kontext und Qualität der beiden Artefakte enthält. Diese beiden Tupel werden aufgespalten, wobei der jeweilige kontextuelle Schluss die kontextuellen und der qualitative Schluss die qualitativen

Identifikatoren erhält.

Die Funktionen  $L_C^A$  und  $L_Q^A$ , die den aggregativen Schluss aus Kontexten und Qualitäten ableiten, sind wie folgt definiert:

$$L_C^A(c_{pen}, c_{hand}) := \begin{cases} 300 & , \text{ wenn } \mathcal{I}(C^{pen} \times C^{hand}, \beta)(P_{\pm}(c_{pen}, 0) \wedge P_{\pm}(c_{hand}, 85)) = true \\ \emptyset & , \text{ sonst.} \end{cases}$$

$$L_Q^A(q_{pen}, q_{hand}) := \begin{cases} \mathcal{I}([0, 1]^n, \beta)(T_P(q_{pen}, q_{hand})) & , \text{ wenn } L_C^A(c_{pen}, c_{hand}) \neq \emptyset \\ \emptyset & , \text{ sonst.} \end{cases}$$

Die Funktion  $L_C^A$  der kontextuellen Aggregation schließt aus einem Stift-Artefakt im Zustand 'liegend' mit dem Identifikator 0 und einem Hand-Artefakt im Zustand 'klatschend' mit 85, auf einen neuen kontextuellen Zustand 'Vortrag beendet' mit dem Identifikator 340. Ist dieser neue Zustand 'Vortrag beendet' aggregiert worden, so werden die Qualitäten  $q_{pen}$  und  $q_{hand}$  in der Aggregation der Funktion  $L_Q^A$  der neuen Qualität durch eine  $T_P$ -Norm (Kapitel 3.1.3) verrechnet. Es wurde die Produkt-Norm gewählt, da diese die Qualitäten multipliziert und dadurch eine kleinere Qualität erreicht, als die kleinste Qualität der Eingabe. Dieses Verhalten entspricht der hier gewünschten Verschlechterung der Qualität durch die Aggregation.

Der fusionierende Schluss ist über die Funktionen  $L_C^F$  und  $L_Q^F$  wie folgt definiert:

$$L_C^F(c_{pen}, c_{hand}) := \begin{cases} 255 & , \text{ wenn } \mathcal{I}(C^{pen} \times C^{hand}, \beta)(P_{\pm}(c_{pen}, 255) \wedge P_{\pm}(c_{hand}, 255)) = true \\ \emptyset & , \text{ sonst.} \end{cases}$$

$$L_Q^F(q_{pen}, q_{hand}) := \begin{cases} \mathcal{I}([0, 1]^n, \beta)(S_P(q_{pen}, q_{hand})) & , \text{ wenn } L_C^F(c_{pen}, c_{hand}) \neq \emptyset \\ \emptyset & , \text{ sonst.} \end{cases}$$

In der Fusion aus Hand- und Stift-Artefakt bestätigen sich die kontextuellen Zustände 'schreibend' mit dem Identifikator 255. Das Modell des abstrahierten Prädikatenlogik testet dabei ob beide Artefakte den gleichen zu fusionierenden kontextuellen Identifikator liefern, was einer logischen Konjunktion entspricht. Befindet sich nur eines der Artefakte in diesem Zustand, so kann der kontextuelle Zustand nicht bestätigt werden und behält die Qualität, die dieses Artefakt für den Zustand erkannt hat. Dieses Tupel kann also nicht in der Fusion verarbeitet werden, sondern wird direkt aus dem Tupel-Raum ausgelesen. Die Fusion der qualitativen Identifikatoren erfolgt, wenn das Stift- und das Hand-Artefakt beide den gleichen zu fusionierenden Zustand besitzen. Die resultierende Qualität wird mit der probabilistischen Summe  $S_P$  (Kapitel 3.1.3) berechnet, die eine bessere Qualität als die einzelnen Artefakte liefert. Die Fusion die hier verwendet wird soll eine Verbesserung erreichen, wenn sich die Artefakte bestätigen, was diese  $S_P$ -Norm erfüllt.

Die definierte Fusion und Aggregation, des Stift- und des Hand-Artefaktes, ist die Grundlage der Schlussfolgerungen, die im nächsten Abschnitt analysiert werden. Es wird aufgrund einer zufälligen Zustandsabfolge eine Qualität der Fusion und der Aggregation berechnet und diese Resultate analysiert.

### 5.3 Analyse von Resultaten des Qualitätsmaßes zweiter Ebene

Die Analyse von fusionierenden und aggregativen Schlüssen aus Qualitäten benötigt einen großen Datensatz von qualitativen Werten für das Stift- und das Hand-Artefakt. Diese Daten sollten die Qualitäten für eine möglichst zufällige Abfolge von kontextuellen Zuständen beinhalten, damit eine gewisse Anzahl von Zuständen, die fusioniert oder aggregiert werden sollen, zu gleichen Zeitpunkten auftreten. Für die statistische Analyse ist die Information darüber erforderlich, ob die durch die Kontexterkenkung erkannten Zustände auch dem realen kontextuellen Zustand entsprechen. Durch diese Information kann wieder für den Schluss gekennzeichnet werden, ob auch die Fusion oder die Aggregation der Qualität die richtigen von den falschen Klassifikationen trennbar macht. Es kann durch Normalverteilungen die Wahrscheinlichkeit berechnet werden, ob Qualitäten über einem gewissen Schwellwert wirklich richtige Kontexte kennzeichnen. Diese Berechnung von Wahrscheinlichkeiten anhand eines Schwellwertes, wurde schon im Zusammenhang mit der statistischen Analyse für das Stift-Artefakt (Kapitel 4.4) verwendet. Die Analyse der Fusion und der Aggregation beruht auf Offline-Daten, die für die Artefakte nicht synchron gesammelt wurden.

Die Qualitäten für eine zufällige Abfolge von kontextuellen Zuständen des Stift-Artefaktes sind in der Abbildung 41 dargestellt. Die einzelnen Zustände, die den jeweiligen Wert der Qualität hervorgerufen haben, sind in der Abbildung durch den kontextuellen Identifikator gekennzeichnet. Die falsch klassifi-

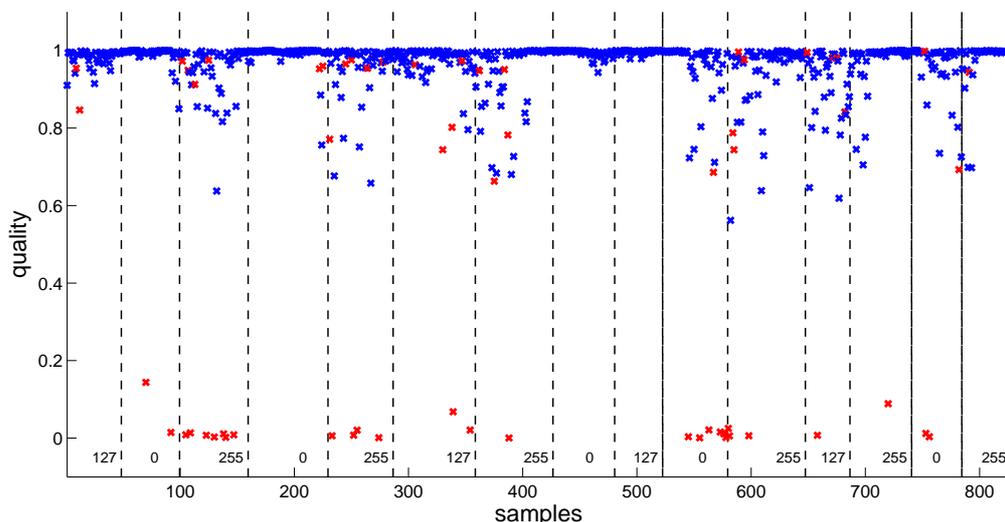


Abbildung 41: Maß der Qualität für einen zufälligen Wechsel der Zustände des Stift-Artefaktes; Kennzeichnung der richtig (blau) und falsch (rot) klassifizierten kontextuellen Zustände und die Markierung des Zustandes durch den Identifikator

zierten Zustände sind rot und die richtig klassifizierten blau markiert. Es ist zu sehen, dass eine große Anzahl der falsch klassifizierten Zustände durch ihre Qualität deutlich von den richtigen zu trennen sind. Der Nachweis über ein Maß der Wahrscheinlichkeit wird hier nicht erbracht, da dieser Nachweis schon in den statistischen Analysen des Stift-Artefaktes erbracht wurde.

Eine zufällige Abfolge von kontextuellen Zuständen des Hand-Artefaktes erzeugt die Werte der Qualität die in der Abbildung 42 abgebildet sind. Die Anzahl der Datenpaare ist für beide Artefakte gleich, da sonst nicht für jedes Datum des einen Artefaktes ein Datum des anderen zum Schluss zur Verfügung stehen würde. Die Intervalle der einzelnen kontextuellen Zustände sind für die beiden Artefakte unterschiedlich groß, damit eine zufällige Überschneidung der Intervalle für die Schlüsse auftritt.

Die Offline-Fusion und -Aggregation, die auf die kontextuellen und qualitativen Informationen angewandt wird, entspricht zunächst nicht einer eigentlichen Benutzung von Systemen der zweiten Ebene, was jedoch für die statistischen Analysen in diesem Kapitel nicht von Bedeutung ist. Durch die Länge der Datensätze und die zufällige Korrelation der kontextuellen Zustände, wird allerdings eine scheinbare reelle Benutzung erreicht.

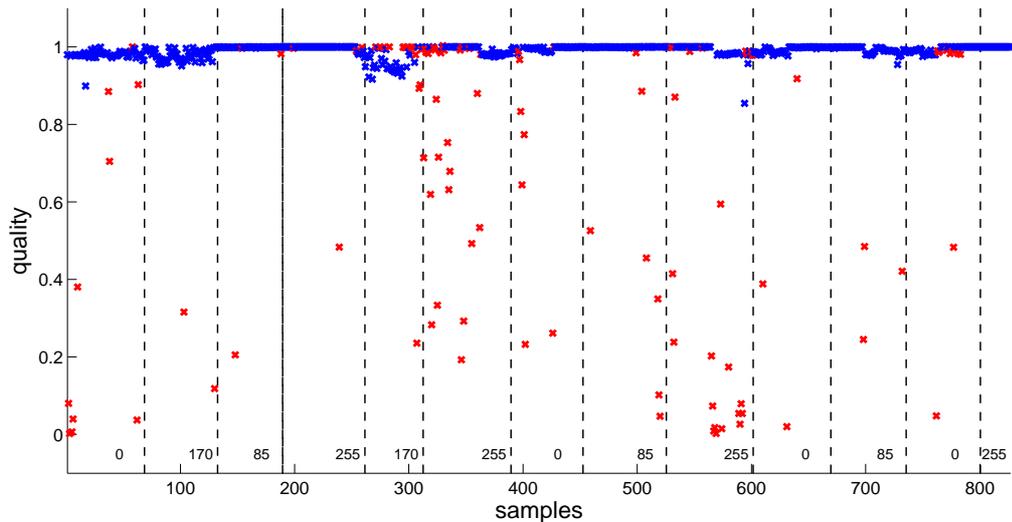


Abbildung 42: Maß der Qualität für einen zufälligen Wechsel der Zustände des Hand-Artefaktes; Kennzeichnung der richtig (blau) und falsch (rot) klassifizierten kontextuellen Zustände und die Markierung des Zustandes durch den Identifikator

In den nächsten Abschnitten werden nun die resultierenden Qualitäten der Fusion und der Aggregation analysiert. Die kontextuellen Zustände des Stift- und des Hand-Artefaktes, die aggregiert und fusioniert werden, und die Art wie der Vorgang stattfindet, wurden im Kapitel 5.2.4 definiert.

### 5.3.1 Fusionierende Schlüsse aus Kontexten und Qualitäten

In diesem Abschnitt werden Werte der Qualität analysiert, die bei der Fusion der Qualität des kontextuellen Zustandes 'schreibend' (255) des Stift- und des Hand-Artefaktes erreicht werden. Es ist für die Analyse bekannt, ob das jeweilige Artefakt diesen Zustand erkannt hat, wenn dieser Zustand real wirklich vorhanden war. Das Modell, einer abstrahierten realen Fusion, wird hier zur Fusion der der kontextuellen Identifikatoren verwendet, welches für die beiden Artefakte in Kapitel 5.2.4 definiert wurde. Die Qualitäten die fusioniert werden sollen, sind mit den realen Zuständen in den Abbildungen 41 und 42 dargestellt. Ist der jeweilige Datenpunkt der Qualität in den Abbildungen blau, so entspricht der Identifikator des Zustandes dem durch das Artefakt erkannten Zustand. Eine rote Kennzeichnung des Datenpunkts bedeutet die Erkennung eines anderen Zustandes, der nicht dem Identifikator des Bereiches entspricht. Es werden also die Qualitäten miteinander fusioniert, die zum gleichen Zeitpunkt (sample) in einem Bereich mit der Markierung 255 auftreten und für jedes Artefakt blau gekennzeichnet sind. Sind in Bereichen, die mit einem anderen kontextuellen Identifikator als 255 markiert sind, Datenpunkte zu gleichen Zeitpunkten vorhanden, die für beide Artefakte den erkannten falschen Zustand 'schreibend' beinhalten, so werden diese Datenpunkte auch in der Fusion verarbeitet.

Die Ergebnisse des fusionierenden Schlusses sind in der Abbildung 43 dargestellt. Es ist zu sehen, dass alle Werte der Qualität blau gekennzeichnet sind, was bedeutet, dass der erkannte kontextuelle Zustand dem realen entspricht und dies für mindestens ein Artefakt. Die Verknüpfung der tatsächlichen Richtigkeit, die als zusätzliche Information in der statistische Analyse verwendet wird, entspricht also einer logischen Disjunktion. Da keine Daten, die in der Fusion verarbeitet wurden, von beiden Artefakten fälschlich erkannt wurden, können keine Wahrscheinlichkeiten berechnet werden. Es stehen für die Dichtefunktionen keine Daten der falschen Klassifikation zur Verfügung. Dieser Sachverhalt ist jedoch für eine Fusion typisch, da es sich um eine gegenseitige Bestätigung handelt. Es ist in einem realen Gebrauch der Artefakte unwahrscheinlich, dass beide einen falschen kontextuellen Zustand erkennen. Die schwarzen Datenpunkte in der Abbildung sind die Werte der Fusion, die auf die leere Menge abgebildet wurden, bei denen also eines oder beide Artefakte einen anderen Zustand, als den der Fusion, erkannt haben.

Die Werte der gelungenen Qualität sind für die Fusion nahezu Eins, was sich in der verwendeten  $S$ -Norm

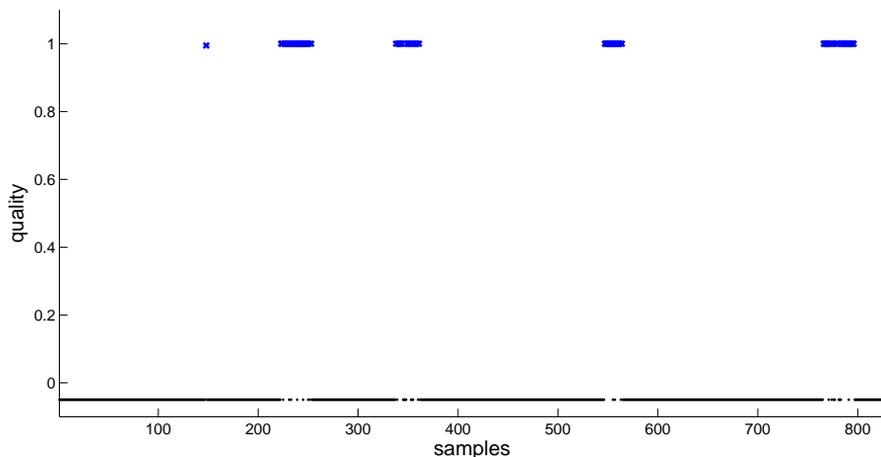


Abbildung 43: Werte der Qualität des fusionierenden Schlusses aus Stift- und Hand-Artefakt, mit Kennzeichnung der tatsächlichen Richtigkeit; eingehende Daten die nicht in dieser Fusion verarbeitet werden, sind durch schwarzen Punkte (-0,05) visualisiert

ergründet, die in einer höhere Qualität resultiert als die Eingangsgrößen. Diese generell hohe Qualität spiegelt jedoch den Mechanismus der Fusion wieder, denn es findet eine Bestätigung der einzelnen kontextuellen Zustände statt. Es muss also eine resultierende Qualität einen Wert nahe Eins haben, sonst wäre die Fusion in ihrer Semantik nicht korrekt.

Die Analyse der aggregativen Schlüsse, die im nächsten Abschnitt beschrieben wird, wird gegenteilige Resultate liefern. Diese ergründen sich in der Generierung eines neuen kontextuellen Zustandes und der verwendeten  $T$ -Norm, die diese gleichwertige Abhängigkeit von den Eingangsgrößen repräsentiert.

### 5.3.2 Aggregative Schlüsse aus Kontexten und Qualitäten

Der hier analysierte aggregative Schluss, welcher den kontextuellen Zustand 'klatschend' (85) des Hand-Artefaktes mit dem Zustand 'liegend' (0) des Stift-Artefaktes kombiniert, wurde in Kapitel 5.2.4 definiert. Es wird der neue Zustand 'Vortrag beendet' aggregiert, der zu gleichen Teilen von den beiden Artefakten abhängig ist. Die Verknüpfung der beiden Zustände der Aggregation ist eine logische Konjunktion, die auch für die Qualitäten auf ein unscharfes Äquivalent, die  $T_P$ -Norm erweitert wird. Die Information, ob die kontextuellen und qualitativen Zustandspaare auch tatsächlich richtig sind, und somit zu dem richtigen aggregierten kontextuellen Zustand führen, wird wieder als zusätzliche Information zur statistischen Analyse verwendet. Um aus den einzelnen Informationen über die richtige oder falsche Klassifikation der Artefakte die Korrektheit der Aggregation abzuleiten, werden die eingehenden Informationen konjugiert. Die Daten die zur Aggregation verwendet werden, entstanden wieder aus getrennten nicht synchronen Aufnahmen der Artefakt-Zustände, was auch für die statistische Analyse der aggregierten Qualitäten keine Relevanz besitzt.

Die Resultate der aggregierten Qualitätsmaße sind in der Abbildung 44 zu sehen, wobei die Konjunktion der richtigen kontextuellen Klassifikationen blau und die der falschen rot gekennzeichnet sind. Es ist deutlich zu erkennen, dass die richtig erkannten kontextuellen Zustände sich deutlich von den falschen durch ihre Qualität abheben. Ein Grund hierfür liegt in dem kaum falsch klassifizierten Zustand 'liegend' des Stift-Artefaktes, der Bestandteil der Aggregation ist. Die Qualität ist auch generell hoch für diesen Zustand, auch wenn er eigentlich eine falsche Klassifikation ist. Dies ist auch ein Grund dafür, dass eine andere  $S$ -Norm als die Produkt-Norm kaum bessere Qualitäten als diese liefert. In Kapitel 3.1.3 wurde erwähnt, dass die  $T_P$ -Norm die kleinste Norm ist, der Beweis kann in [HR05] nachgelesen werden.

Ein anderer Grund für die hohe Qualität der Aggregation aus richtigen kontextuellen Zuständen der Artefakte ist die verwendete Konjunktion, als Operator zur Bestimmung der Richtigkeit des Schlusses. Es müssen beide kontextuelle Identifikatoren dem tatsächlichen Zustand der Artefakte entsprechen, die dann auch, wie statistisch nachgewiesen wurde, mit einer hohen Wahrscheinlichkeit eine Qualität nahe Eins besitzen.

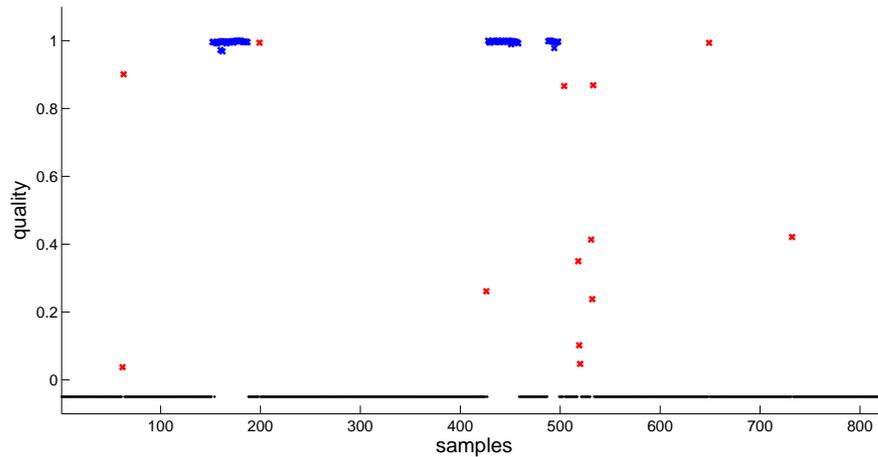


Abbildung 44: Werte der Qualität des aggregativen Schlusses aus Stift- und Hand-Artefakt, mit Kennzeichnung der tatsächlichen Richtigkeit; eingehende Daten die nicht in dieser Aggregation verarbeitet werden sind durch schwarzen Punkte (-0,05) visualisiert

Die Dichtefunktionen der Normalverteilungen, für richtig und falsch klassifizierte kontextuelle Zustände, sind in der Abbildung 45 dargestellt. Die Wahrscheinlichkeiten, die über die Integrale der abgeschnittenen Dichtefunktionen berechnet werden, zeigen auch eine klare Trennbarkeit der richtigen von den falschen aggregativen Schlüssen. Wird ein Qualitätsmaß aggregiert, das sich über dem Schwellwert (Abbildung

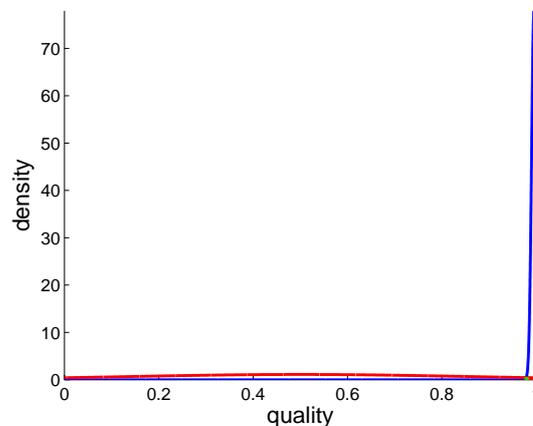


Abbildung 45: Dichtefunktionen für richtig (blau) und falsch Klassifizierte (rot) kontextuelle Zustände

45, grüne Linie) der Schnittpunkte befindet, bedeutet eine Wahrscheinlichkeit von  $\mathbb{P}(c = true|q > s) = 0,9099$ , dass dieser aggregierte Kontext wirklich richtig ist. Ein Qualitätsmaß unterhalb des Schwellwertes ist mit der Wahrscheinlichkeit  $\mathbb{P}(c = false|q < s) = 0,9099$  bezeichnend für eine falsche Aggregation. Die falschen positiven sind mit der Wahrscheinlichkeit von  $\mathbb{P}(c = false|q > s) = 0,0008$  und die richtigen Negativen mit  $\mathbb{P}(c = true|q < s) = 0,0893$  berechnet.

Es wurde somit auch für kontextuelle Schlüsse ein statistischer Nachweis gebracht, dass das Qualitätsmaß eine Trennbarkeit der richtigen von den falschen Klassifikationen ermöglicht. Die Aggregation eines neuen kontextuellen Zustandes resultiert in einer Minimierung der Qualität, wobei die Fusion eine Maximierung bewirkt.

---

## 6 Zusammenfassung und Ausblick

Die Theorie der Qualität des Kontextes existierte bereits, in dieser Arbeit wurde jedoch ein praktisches System definiert und evaluiert, das eine Verlässlichkeit dynamisch für jeden kontextuellen Zustand berechnen kann. Dieses qualitative System wurde so designed, dass es nicht nur auf Rechnern ausgeführt werden kann, die eine hohe Performanz besitzen, sondern selbst auf schwachen Sensorknoten berechnet werden kann. Eine Theorie wurde somit durch eine reales System bestätigt, das selbst zukünftigen Anforderungen entsprechen kann.

### 6.1 Zusammenfassung

Das Ziel dieser Arbeit war ein System zu entwickeln, das eine Verlässlichkeit einer kontextuellen Erkennung dynamisch berechnet. Dieses qualitative System sollte eine automatisierte Erstellung besitzen, die zur Verfügung gestellt und detailliert beschrieben wurde. Es wurde eine formalisierte Herleitung der Spezifikation des qualitativen Systems bereitgestellt, die den mathematisch und logischen Grundlagen entspricht.

Die Algorithmen, die die automatisierte Spezifikation des Aufbaus und der Parameter ermöglichen, wurden detailliert beschrieben und gegenüber anderen Verfahren abgehoben. Es wurden zunächst zwei kontextuelle Systeme bezüglich ihres Fehlers der Klassifikation adaptiert, wobei eine genaue Beschreibung der kontextuellen Systeme und des Vorgangs der Adaption vorliegt. Es wurde versucht die Besonderheiten der adaptiven Algorithmen herauszuarbeiten und für die Zwecke der Qualitätsanalyse zu optimieren.

Die Evaluation der erstellten qualitativen Systeme wurde aufgrund statistischer Methoden durchgeführt. Es wurden verschiedene Aspekte analysiert, die zum Teil zuvor theoretisch ergründet wurden. Es wurde versucht den Mehrwert des Qualitätsmaß hervorzuheben und für die Nutzbarkeit in ubiquitären Umgebungen zu bestätigen.

Ein abstraktes Modell der kontextuellen Schlussfolgerung diente zur Evaluierung des Qualitätsmaßes auf einer zweiten Ebene. Die Schlussfolgerungen die aus Qualitäten gezogen wurde, konnten auf Basis dieses abstrahierten Modells gemacht werden. Ein zweites Artefakt wurde mit dem schon zuvor benutzten und beschriebenen zur realen Evaluierung der Schlussfolgerungen benutzt.

Eine automatisierte Konstruktion eines qualitativen Systems wurde also theoretisch erörtert und praktisch evaluiert. Die Maße der Qualität für eine Verlässlichkeit eines kontextuellen Systems wurden auf zwei Ebenen analysiert.

Das vorgestellte System der Qualitätsanalyse von kontextuellen Zuständen ist völlig neuartig und mit keiner anderen Arbeit zu vergleichen. Es ist auf jedes System zur kontextuellen Zustandserkennung anwendbar, ohne primäres Wissen über den Algorithmus und die Sensorik zu benötigen. Das qualitative System ist universell in bestehende Systeme integrierbar, nahezu unabhängig von der Performanz des Rechners.

### 6.2 Ausblick

Das qualitative System wurde mit zwei Artefakten und zwei verschiedenen Algorithmen zur Kontexterkenkung eingesetzt. Die weiteren Möglichkeiten, um das qualitative System zu evaluieren, bestehen bei Verwendung weiterer Artefakte und Algorithmen. Eine komplexe ubiquitäre Umgebung mit vielen Artefakten und hoher Kommunikation könnte einen langzeit Test liefern, was auch schon teilweise in anderen Arbeiten des TecO's zur Anwendung kommt. Eine praktische Verwendung des qualitativen Systems in ubiquitären Umgebungen könnte zur Bestätigung und Anerkennung dieser Qualität des Kontextes dienen.

Die komplette Trennung von kontextueller Klassifikation und qualitativer Analyse könnte bei der Konstruktion von neuen Systemen überdacht werden. Den Vorteil einer unabhängigen Qualitätsanalyse gilt es bei einer kombinierten Systementwicklung zu wahren, was gewisse Probleme ergeben könnte. Eine Lösung könnten rekurrente neuronale Fuzzy Netze sein. Auch ein geschicktes Scheduling-Verfahren könnte die komplette Parallelität reduzieren helfen.

Die automatisierte Erstellung des Systems zur qualitativen Analyse basiert auf Algorithmen, die die beste

Auswahl der vorhandenen Methoden darstellt. Eine spezielle Anpassung verschiedener Methoden auf die Erstellung eines qualitativen Systems könnte das Resultat der Analyse noch verbessern. Die Verfahren des Clusterings könnten erweitert und untereinander kombiniert werden.

## Literatur

- [AD99] Gregory D Abowed and Anind K Dey. Towards a better understanding of context and context-awareness. *Panel at the 1st International Symposium on Handheld and Ubiquitous Computing*, pp 304-307, 1999.
- [Bez73] J C Bezdek. *Fuzzy mathematics in pattern classification*. PhD thesis, Center for Applied Mathematics, Cornell University, Ithaca, New York, USA, 1973.
- [BKS03] Thomas Buchholz, Axel Küpper, and Michael Schiffers. Quality of context: What it is and why we need it. *Workshop of the HP OpenView University Association*, 2003.
- [BLD95] P J Costa Branco, N Lori, and J A Dente. An autonomous approach to the mountain clustering method. *The Third International Symposium on Uncertainty Modeling and Analysis and Annual Conference of the North American Fuzzy Information Processing Society*, pp 649-654, 1995.
- [Chi96] Stephen Chiu. Method and software for extracting fuzzy classification rules by subtractive clustering. *IEEE Control Systems Magazine*, 1996, vol. pp. 461-465, 1996.
- [Chi97] Stephen Chiu. *Fuzzy Information Engineering: A Guided Tour of Applications*, chapter 9, Extracting Fuzzy Rules from Data for Function Approximation and Pattern Classification. John Wiley&Sons, 1997.
- [Deu02] P Deussen. Skriptum: Formale systeme. Universität Karlsruhe, 2002.
- [DH93] Deufelhard and Hohmann. *Numerische Mathematik 1, 2. Auflage*, chapter 5.4 Singulärwertzerlegung, S.151-156. 1993.
- [DP00] D Dubois and H Prade. Possibility theory in information fusion. *In Proc. of the 3<sup>rd</sup> IEEE International Conference on Information Fusion*, pp 6-19, 2000.
- [DS81] N R Draper and H Smith. *Applied Regression Analysis, Second Edition*. 1981.
- [Dun73] J C Dunn. A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters. *Journal Cybernetics*, 3, pp. 32-57, 1973.
- [Est05] Pablo Estévez. Step by step tutorial to neural networks and fuzzy logic in the teco-smart-it platform. Technical report, Telecooperation Office, University of Karlsruhe, 2005.
- [Est06] Pablo Estévez. Master's thesis, 2006.
- [EW04] Amr Ali Eldin and René Wagenaar. A fuzzy logic based approach to support users self-control of their private contextual data retrieval. *Proceedings of the 13th European Conference on Information System*, 2004.
- [GG89] I Gath and A B Geva. Unsupervised optimal fuzzy clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol 11(7), pp 773-781, 1989.
- [GK] E E Gustafson and W C Kessel. Fuzzy clustering with a fuzzy covariance matrix. *IEEE CDC, San Diego, California*, pp 761-766.
- [HKK97] Frank Höppner, Frank Klawonn, and Rudolf Kruse. *Fuzzy-Clusteranalyse, Verfahren für die Bilderkennung, Klassifikation und Datenanalyse*. VIEWEG, 1997.
- [HR05] Uwe D Hanebeck and Patrick Rößler. Unschärfe mengen. Lehrstuhl für Intelligente Sensor-Aktor-Systeme, Institut für Technische Informatik, 2005.
- [Jan93] Jyh-Shing Roger Jang. Anfis: Adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man and Cybernetics*, 1993, vol. 23 pp. 665-685, 1993.
- [JS95] Jyh-Shing Roger Jang and Chuen-Tsai Sun. Neuro-fuzzy modeling and control. *Proceedings of the IEEE*, 1995, 1995.
- [KL06] Younghee Kim and Keumsuk Lee. A quality measurement method of context information in ubiquitous environments. *International Conference on Hybrid Information Technology*, vol 2, 2006.

- 
- [Koh81] Teuvo Kohonen. Automatic formation of topological maps of patterns in a self-organizing system. *Proceedings of 2SCIA, Scand. Conference on Image Analysis*, pp 214-220, 1981.
- [KSSF03] Andreas Krause, Daniel P. Siewiorek, Asim Smailagic, and Jonny Farrington. *Unsupervised, Dynamic Identification of Physiological and Activity Context in Wearable Computing*. Munich University of Technology and Carnegie Mellon University and BodyMedia Inc., 2003.
- [KuUS00] H Kunle and G Aumann und U Schober. Skripten: Lineare algebra und analytische geometrie i+ii. Universität Karlsruhe, 2000.
- [LSK01] Jung W Lee, Seo H Son, and Soon H Kwon. Advanced mountain clustering method. *North American Fuzzy Information Processing Society*, vol 1, pp 275-280, 2001.
- [LY90] H L Larsen and R R Yager. Efficient computation of transitive closures. *Fuzzy Sets and Systems*, vol 38, pp 81-90, 1990.
- [MA75] E H Mamdani and S Assilina. An experiment in linguistic synthesis with a fuzzy logic controller. *International Journal of Man-Machine Studies*, vol 7(1), pp 1-13, 1975.
- [Män95] Manfred Männle. *Modellierung nichtlinearer Systeme mit unscharfen Regeln*, chapter 3.2.2 Identifikation mit Algorithmus RPROP. Diplomarbeit, 1995.
- [Män00] Manfred Männle. Schnelle takagi-sugeno fuzzy modellierung. *Proceedings of GMA/GI Computational Intelligence*, pp. 323-328, Baden-Baden, Germany, 2000.
- [OL02] Stanislaw Osowski and Tran Hoai Linh. Neuro-fuzzy tsk network for approximation. *Control and Cybernetics*, 2002, vol. pp., 2002.
- [Par06] <http://particle.teco.edu>. World Wide Web, Telecooperation Office, University of Karlsruhe, 2006.
- [PLYSK04] Jihyung Park, Seungsoo Lee, Kiwon Yeom, and Seokho Lee Sungju Kim. A context-aware system in ubiquitous environment: a research guide assistant. *Conference on Cybernetics and Intelligent Systems*, 2004.
- [RDN05] M. Razzaque, S. Dobson, and P. Nixon. Categorisation and modelling of quality in context information. In *Proc. of the IJCAI 2005 Workshop on AI and Autonomic Communications, Edinburgh, Scotland*, August 2005.
- [SK88] M Sugeno and G Kang. Structure identification of fuzzy model. *Fuzzy Sets and Systems*, 1988, vol 26(1), pp 15-33, 1988.
- [TS85] T Tagaki and M Sugeno. Fuzzy identification of systems and its application to modelling and control. *IEEE Trans. Syst. Man and Cybernetics*, 1985, vol SMC-15, no. 1, pp 116-132, 1985.
- [Tsu90] Y Tsukamoto. *Advances in Fuzzy Set Theory and Applications*, chapter An approach to fuzzy reasoning method, pp 137-149. 1990.
- [Wan98] L X Wang. *Adaptive Fuzzy Systems and Control*. Prentice-Hall, Englewood Cliffs, 1998.
- [Wei91] Mark Weiser. The computer of the 21st century. *Scientific American*, September 1991.
- [YF94] R. Yager and D. Filer. Generation of fuzzy rules by mountain clustering. *Journal on Intelligent Fuzzy Systems*, vol 2, pp 209-219, 1994.
- [Zel94] Andreas Zell. *Simulation Neuronaler Netze*. Oldenbourg, 1994.
- [Zim06] Tobias Zimmer. Qoc: Quality of context - improving the performance of context-aware applications. *Advances in Pervasive Computing*, pp 209-214, 2006.