



[[Karlsruhe Institute of Technology **Fakultät für Informatik**

**Institut für Anthropomatik (IFA)**

**Lehrstuhl für Interaktive Echtzeitsysteme (IES)**

Prof. Dr.-Ing. J. Beyerer

Diplomarbeit

**Analyse von Verfahren zur Kombination  
von Expertenwissen in Form von  
Wahrscheinlichkeitsverteilungen im  
Hinblick auf die verteilte lokale  
Bayes'sche Fusion**

Johannes Fährdrich

1. Dezember 2009 – 28. Mai 2010

Betreut von:

Prof. Dr.-Ing. Jürgen Beyerer

Betreuende Mitarbeiterin: Dipl.-Math. Jennifer Sander

Karlsruher Institut für Technologie (KIT)

Fakultät für Informatik

Institut für Anthropomatik (IFA)

Lehrstuhl für Interaktive Echtzeitsysteme (IES)

Thema: Analyse von Verfahren zur Kombination von Expertenwissen in Form von Wahrscheinlichkeitsverteilungen im Hinblick auf die verteilte lokale Bayes'sche Fusion

Autor: Johannes Fährdrich

Johannes Fährdrich

Marienstr. 36

76137 Karlsruhe

E-Mail: [faehndrich@gmail.com](mailto:faehndrich@gmail.com)

Ich versichere hiermit wahrheitsgemäß, die Arbeit bis auf die dem Aufgabensteller bereits bekannte Hilfe selbständig angefertigt, alle benutzten Hilfsmittel vollständig und genau angegeben und alles kenntlich gemacht zu haben, was aus Arbeiten anderer unverändert oder mit Abänderung entnommen wurde.

.....  
Karlsruhe, 28.05.2010

(Johannes Fähndrich)

# Danksagung

An dieser Stelle möchte ich mich bei meiner Betreuerin Jennifer Sanders für ihre Geduld und ständige Unterstützung sowie die vielen inhaltlichen Anregungen bedanken, ohne die diese Arbeit so nicht möglich gewesen wäre. Des weiteren möchte ich mich bei all denen bedanken, die an diese Arbeit durch Diskussionen oder Formulierungsvorschläge mitwirkten.

# Kurzzusammenfassung

Bayes'sche Informationsfusion kann einen hohen Rechenaufwand erzeugen. Zum Zweck einer Verkürzung der Berechnungszeit wird in dieser Arbeit eine verteilte lokale Bayes'sche Informationsfusion betrachtet. Die Größen von Interesse werden durch eine Lokalisierung in Mengen aufgeteilt, die Bereiche von erhöhtem Interesse beinhalten. Auf diesen Mengen kann ein Agentensystem die Fusion verteilt durchführen. Dadurch, dass die Agenten des Agentensystems sich auf diese lokale Menge der Größen von Interesse beschränken, wird die Wahrscheinlichkeit für die Elemente dieser Menge wahrscheinlicher als sie es global gesehen sind. Die Ergebnisse der verteilten Berechnung müssen so zusammengeführt werden, dass möglichst das gleiche Ergebnis wie bei einer nicht verteilten Bayes'schen Informationsfusion entsteht. Dieses Problem ist mit dem einer Konsensfindung in einer Gruppe vergleichbar. Dabei muss aus den A-posteriori-Wahrscheinlichkeitsverteilungen der Gruppenmitglieder, eine Wahrscheinlichkeitsverteilung erzeugt werden, welche die *Ground Truth* ausreichend gut annähert.

Eine Pooling-Methode ist eine Methode der Entscheidungstheorie. Diese Pooling-Methoden vereinen die Wahrscheinlichkeitsverteilung, welche die Meinungen/Schätzungen der Gruppenmitglieder repräsentieren. Die entsprechende Fachliteratur wurde nach Pooling-Methoden und Möglichkeiten durchsucht, wie diese bewertet werden können. In der Entscheidungstheorie geschieht dies anhand von Eigenschaften, welche die Pooling-Methoden erfüllen können. Dieser Ansatz wurde hier für eine verteilte lokale Bayes'sche Informationsfusion fortgesetzt. Es werden Eigenschaften von Pooling-Methoden auf ihre Relevanz hinsichtlich einer verteilten lokaler Bayes'schen Informationsfusion untersucht. Auf Basis dieser Einsichten werden Vor- und Nachteile von Pooling-Methoden, (am Beispiel des Linear Opinion Pools, des Logarithmic Opinion Pools und des Supra Bayes Ansatz) im Kontext von lokaler Bayes'scher Informationsfusion beleuchtet. Somit wird für diese Beispiele durch Analyse und Bewertung der Eigenschaften eine Aussage über deren Eignung für eine verteilte lokale Bayes'sche Informationsfusion abgeleitet.



# Inhaltsverzeichnis

<b>Kurzzusammenfassung</b>	<b>ii</b>
<b>Inhaltsverzeichnis</b>	<b>v</b>
<b>Abbildungsverzeichnis</b>	<b>vii</b>
<b>Tabellenverzeichnis</b>	<b>ix</b>
<b>1. Einführung</b>	<b>1</b>
1.1. Problemstellung . . . . .	4
1.2. Stand der Wissenschaft . . . . .	6
1.3. Aufbau der Arbeit . . . . .	8
<b>2. Grundlagen</b>	<b>10</b>
2.1. Informationsfusion . . . . .	10
2.1.1. Frequentistische Statistik . . . . .	12
2.1.2. Fuzzy-Logik . . . . .	12
2.1.3. Künstliche Neuronale Netze . . . . .	13
2.1.4. Bayes'sche Statistik . . . . .	14
2.1.5. Dempster-Shafer-Evidenztheorie (DSE) . . . . .	15
2.2. Bayes'sche Fusion . . . . .	18
2.3. Lokale Konfiguration . . . . .	20
2.4. Alternative lokale Konfigurationen . . . . .	26
<b>3. Pooling Methoden</b>	<b>29</b>
3.1. Einführung in Pooling-Methoden . . . . .	29
3.2. Der Linear Opinion Pool und der generelle Linear Opinion Pool . . . . .	30
3.3. Logarithmic Opinion Pool . . . . .	34
3.4. Supra-Bayes Ansatz . . . . .	37

<b>4. Eigenschaften von Pooling Methoden</b>	<b>39</b>
4.1. Allgemeine Eigenschaften . . . . .	39
4.2. Beschränkungssätze . . . . .	48
4.3. Eigenschaften eines LinOP bzw. GenLinOP . . . . .	52
4.3.1. Wertung der Eigenschaften eines LinOP und GenLinOP für die lokale Bayes'sche Fusion . . . . .	57
4.4. Eigenschaften eines LogOP . . . . .	60
4.4.1. Wertung der Eigenschaften eines LogOP für die lokale Bayes'sche Fusion . . . . .	64
4.5. Analyse des Supra-Bayes Ansatzes für die lokale Bayes'sche Fusion . . . . .	65
<b>5. Lösungsvorschläge für die lokale Bayes'sche Fusion</b>	<b>67</b>
5.1. Globales a-priori-Wissen . . . . .	67
5.2. Gemeinsame Referenz . . . . .	68
5.3. Weitere Verfahren . . . . .	71
5.4. Linear Opinion Pool . . . . .	73
5.5. Logarithmischer Opinion Pool . . . . .	76
5.6. Supra Bayesian . . . . .	80
<b>6. Zusammenfassung und Ausblick</b>	<b>82</b>
<b>Literaturverzeichnis</b>	<b>86</b>
<b>A. Mathematische Grundlagen</b>	<b>92</b>
A.1. Stochastik Grundlagen . . . . .	92
A.2. Unsicherheit . . . . .	93
A.3. sigma-Algebren . . . . .	93
A.4. Scoring rules . . . . .	94
<b>B. Daten der Beispiele</b>	<b>95</b>





# Abbildungsverzeichnis

1.1. Luftbild für Fahrzeugerkennungsbeispiel . . . . .	6
1.2. Beispiel einer Straßenkarte als weiter Information [San10]. . . . .	7
1.3. Beispiel für eine Lokalisierung . . . . .	8
2.1. Fusionsschema . . . . .	11
2.2. Diskretes Fusionsbeispiel . . . . .	25
3.1. Ergebnis eines LinOP . . . . .	33
3.2. Auswirkung der Gewichte bei einem LogOP . . . . .	35
3.3. Beispiel für einen LogOP . . . . .	35
3.4. Ein Vergleich zwischen LogOP und LinOP . . . . .	36
5.1. Ergebnis eines LinOP mit gleichen Gewichten für alle Gruppenmitglieder. .	70
5.2. Die Wahrscheinlichkeiten der Referenz ( $z^1$ ), die alle Gruppenmitglieder be- trachten. . . . .	70
5.3. Ergebnis einer Informationsfusion mit Gewichtung mittels einer gemeinsa- men Referenz. . . . .	71

# Tabellenverzeichnis

2.1. Dargestellt sind die Likelihoodfunktionen $p(d_i z^j)$ . . . . .	24
3.1. Beispiel für Meinungen von zwei Gruppenmitgliedern. . . . .	32
4.1. Darstellung der Eigenschaften mit Markierung ob der LinOP/GenLinOP diese Eigenschaft erfüllt. . . . .	53
4.2. Beispiel für Meinungen von zwei Gruppenmitglieder mit einer Einstimmigkeit bezüglich der Wahrscheinlichkeit von $z^2$ . . . . .	54
4.3. Beispiel für einstimmige Meinungen von zwei Gruppenmitglieder. . . . .	55
4.4. Beispiel für Meinungen von zwei Gruppenmitgliedern. . . . .	56
4.5. Beispiel für Meinungen von zwei Gruppenmitgliedern, wobei $A$ und $B$ unabhängig sind. . . . .	56
4.6. Beispiel für Meinungen von zwei Gruppenmitgliedern . . . . .	57
4.7. Darstellung der Eigenschaften mit Markierung, ob der LogOP diese Eigenschaft erfüllt. . . . .	60
4.8. Beispiel für Meinungen von zwei Gruppenmitgliedern, wobei eine Einstimmigkeit bei $z^2$ auftritt. . . . .	62
4.9. Beispiel für Meinungen von zwei Gruppenmitgliedern. . . . .	62
4.10. Ergebnis des LogOP. . . . .	63
5.1. Beispiel für Meinungen von zwei Gruppenmitgliedern, wobei $Agent_1$ $z^3$ und $Agent_2$ $z^1$ nicht betrachtet. . . . .	76
5.2. Beispieldaten für den Beweis, dass der modLogOP die EB nicht erfüllt. . . . .	77
5.3. Ergebnis des Gegenbeispiels dafür, dass der modLogOP die EB nicht erfüllt. . . . .	78
B.1. Wahrscheinlichkeitsverteilung der Beobachtungen $z^1, \dots, z^{20}$ der zehn Informationsquellen $d1, \dots, d5$ für das Beispiel in Abschnitt 5.2. . . . .	96

B.2. Wahrscheinlichkeitsverteilung der Beobachtungen $z^1, \dots, z^{20}$ der Zehn Informationsquellen $d_6, \dots, d_{10}$ für das Beispiel in Abschnitt 5.2. . . . . .	97
B.3. Dargestellt sind die bedingten Wahrscheinlichkeiten $p(d_i z^j)$ . . . . .	98
B.4. Dargestellt sind Ergebnisse der mit der gemeinsamen Referenz gewichteten Wahrscheinlichkeiten. . . . .	98
B.5. Dargestellt sind Ergebnisse des LinOP mit gleichen Gewichten für alle Gruppenmitglieder . . . . .	99

# 1. Einführung

Stehen mehrere Informationsquellen bezüglich einer *Größe von Interesse (Property of Interest PoI)* zur Verfügung, dann kann die Information aller Quellen die Ungewissheit bezüglich der PoI mehr verringern, als eine der Quellen alleine. Dies ist beispielsweise der Fall, wenn ein Fahrzeug automatisch gesteuert werden soll. Ein Sensorsystem wie zum Beispiel eine Kamera hat Vor- und Nachteile. Werden zusätzlich weitere Sensortypen verwendet, nehmen wir an ein Radar und ein Laser, dann können Fehler eines Sensortyps durch einen anderen Sensor ausgeglichen werden. Ist beispielsweise die Information der Kamera wegen Gegenlicht verwaschen, können die anderen Sensoren wichtige Information bezüglich möglicher Hindernisse auf der Fahrbahn liefern.

Die Information der Informationsquellen ist immer mit Unsicherheit behaftet. Mit Unsicherheit behaftete Information über eine PoI so zur Informationsfusion zu verwenden, dass die Unsicherheit bezüglich der PoI minimiert wird, ist ein Ziel<sup>1</sup> der lokalen Bayes'schen Informationsfusion. Dazu muss die Information als Wahrscheinlichkeit dargestellt werden. Diese wird hier als *Grad des Dafürhaltens (Degree of Belief (DoB))* interpretiert.

Die Berechnung der Informationsfusion soll lokal realisiert werden. Die Lokalisierung ist eine Einschränkung auf bestimmte Bereiche von Interesse. Dazu werden die betrachteten PoIs in Teilmengen unterteilt. Jede Teilmenge kann separat untersucht werden. So können die Untersuchungen verteilt berechnet werden.

Die untersuchenden Elemente, Sander und Beyerer [San06] nennen sie Agenten, werden hier als eine Gruppe betrachtet. Diese Gruppe arbeitet zusammen, um eine Schätzung der PoI zu finden. Damit die Parallelität zu den Lösungsansätzen aus der Entscheidungstheorie deutlich wird, werden die Agenten im Folgenden als Gruppenmitglieder bezeichnet. Durch die Lokalisierung kann die Fusion verteilt berechnet werden. Dies kann einen Speedup ermöglichen. Durch die Konzentration auf die PoI von Interesse, wird der Rechenaufwand weiter verringert [San09b].

---

<sup>1</sup>Ein weiteres Ziel ist beispielsweise die Berechenbarkeit.

Sander und Beyerer ziehen Parallelen zu einer kriminalistischen Ermittlung. Dabei entsprechen die von ihnen beschriebenen Fusionsagenten den verschiedenen Ermittlern, die Informationsquellen überprüfen, um Beweise für die Schuld eines Täters zu sammeln. Dieser Ablauf aus dem Alltag eines Ermittlers, kann hier auf die Gruppenmitglieder übertragen werden.

Was die hier betrachtete Problemstellung von der eines Standard Gruppenentscheids [Fre85] unterscheidet, ist, dass die Gruppenmitglieder eine lokale Sicht aufweisen. Lokal bedeutet hier, dass die Gruppenmitglieder nicht alle die gleichen, sondern unterschiedliche Bereiche der PoI betrachten, um redundante Berechnungen zu vermeiden. Ein weiterer Unterschied ist, dass eine *Ground Truth* existiert. Während bei einem klassischen Gruppenentscheid der Entscheidungstheorie meistens nicht klar ist, was die optimale Lösung für den Konsens ist, kann in den hier betrachteten Szenarien die „wahre“ Wahrscheinlichkeitsverteilung bestimmt werden. Diese „wahre“ Wahrscheinlichkeitsverteilung entspricht der eines einzelnen Bayes'schen Gruppenmitglieds, welches gesamt  $Z$  betrachtet.

Dadurch ist es fraglich, ob Verfahren, die für einen Standard Gruppenentscheid beschrieben wurden, auch für einen lokalen Kontext anwendbar sind. Ziel ist es, Methoden zur Kombination von Schätzungen oder Meinungen in Form von Wahrscheinlichkeitsverteilungen auf ihre Eigenschaften und deren Nutzen bezüglich des hier verwendeten Verfahrens, der verteilten lokalen Bayes'schen Fusion, zu untersuchen. Dabei werden Methoden zur Findung eines Konsens untersucht, die nicht für den Kontext einer lokalen Fusion entworfen wurden. Ob eine solche Methode für eine lokale Fusion anwendbar ist, wird im Rahmen dieser Arbeit geklärt. Des Weiteren werden Eigenschaften dieser Methode analysiert, um zu klären, welche dieser Eigenschaften für oder gegen die Verwendung in einem lokalen Kontext sprechen.

Ein Konsens der Gruppenmitglieder ist hier eine Wahrscheinlichkeitsverteilung über den Bereich der PoI, den die Gruppenmitglieder insgesamt betrachteten. Ein Konsens ist also meistens keine Verschmelzung von Schätzungen der Gruppenmitglieder, da jedes Gruppenmitglied auf einer anderen Teilmenge der PoI arbeitet.

Soll Information aus mehreren Informationsquellen in eine Informationsfusion einfließen, um die Unsicherheit bezüglich einer PoI zu verringern, muss ein Verfahren gewählt werden, welches die folgenden Anforderungen erfüllt:

- Die Fähigkeit zur Transformation: Wird diese Eigenschaft erfüllt, dann kann das

---

Verfahren mit Informationen auf verschiedenen Abstraktionsniveaus umgehen.

- Die Fähigkeit zur Fusion erlaubt es, die Information aus den Informationsquellen zu einer Aussage zu verschmelzen.
- Die Fähigkeit der Fokussierung. Diese Fähigkeit erlaubt es, sich mit der Fusion auf einen lokalen Kontext, der von erhöhtem Interesse ist, zu konzentrieren.

Erst dadurch wird die Fusion für reale Problemstellungen berechenbar [Bey07].

Die erste Frage ist: Wie kann man verschiedene Wahrscheinlichkeitsverteilungen zu einer Konsenswahrscheinlichkeitsverteilung zusammenfassen? Der Begriff *Amalgamation* (im übertragenen Sinne: Verbindung, Vereinigung, Vermischung, Verschmelzung) scheint hier passend, da die verschiedenen Wahrscheinlichkeitsverteilungen zu einer einzigen zusammengeführt werden.

Die zweite Frage ist: Warum wird solch eine Amalgamation der Wahrscheinlichkeiten benötigt? Dazu führt French [Fre85] drei Aspekte auf:

- Erstens das *Experten Problem*: Die Gruppe besteht aus einer Menge von Experten, welche um ihre Meinung gebeten werden. Diese Expertenmeinungen können dann benutzt werden, um eine Entscheidung zu treffen.
- Zweitens das *Gruppenentscheid-Problem*: Dabei ist die Gruppe an einer Entscheidung interessiert. Die Gruppenmitglieder wollen dabei ihre Meinung in einem Gruppenentscheid amalgamieren.
- Drittens das *Textbuch-Problem*: Hier sollen die Gruppenmitglieder ihre Meinung publizieren, damit andere sie verwenden können. French [Fre85] ist der Ansicht, dass hier das eigentliche Gruppenentscheidungsproblem nicht gegeben ist. Jedoch bleibt es das gleiche Problem, ob die Gruppe für sich die Meinungen amalgamieren, oder ob diese Aufgabe einer anderen Instanz überlassen wird: Wie werden aus mehreren Meinungen eine? In dieser Arbeit wird das Experten Problem betrachtet, wie der Abschnitt 1.1 erläutert. Des weiteren beschränkt sich diese Arbeit im Wesentlichen auf die Betrachtung von diskreten Wahrscheinlichkeitsverteilungen.

### 1.1. Problemstellung

Aufgabe ist es, eine Bayes'sche Fusion von lokalen Wahrscheinlichkeitsverteilungen verteilt durchzuführen. Um die Information, auf denen die Fusion berechnet wird, mathematisch greifbar zu machen, wird ihnen eine Struktur gegeben. Die allgemeine Konfiguration ist wie folgt definiert: Information aus verschiedenen Informationsquellen wird als DoB für ein Ereignis interpretiert. Diese DoB werden so zusammengeführt, dass sie mit dem Vorwissen eine A-posteriori-Wahrscheinlichkeitsverteilung für die betrachteten Ereignisse ergeben.

Durch die lokale Bayes'sche Informationsfusion lassen sich Rechen- und Speicheraufwand der Bayes'schen Fusionsmethodik signifikant verringern. Dabei wird die zur Verfügung stehende Information nach Auffälligkeiten (Spuren) abgesucht, entlang derer dann lokal „ermittelt“ wird. Hierzu wird die eigentliche Fusion auf die Spuren und die zugehörigen lokalen Kontexte beschränkt. Lokale Bayes'sche Informationsfusion kann in Anlehnung an kriminalistische Ermittlungen mittels einer agentenbasierten Fusionsarchitektur verteilt umgesetzt werden. Zur verteilten Realisierung lokaler Bayes'scher Fusion bedarf es im Allgemeinen Methoden, mit denen sich lokale A-posteriori-Wahrscheinlichkeitsverteilungen, die mit verschiedenen Auffälligkeiten korrespondieren, möglichst verlust- und artefaktfrei im Hinblick auf das Fusionsergebnis kombinieren lassen. Das Fusionsergebnis kann dabei je nach Aufgabenstellung die unsicherheitsbehaftete Informationsbeschreibung der Zielgröße in Form der A-posteriori-Wahrscheinlichkeitsverteilung oder eine hieraus abgeleitete Schätzung der Zielgröße sein.

Diese beschriebene Kombination lokaler A-posteriori-Wahrscheinlichkeitsverteilungen entspricht in gewisser Hinsicht einer Entscheidungsfindung in einer Gruppe wo Expertenmeinungen in Form von Wahrscheinlichkeitsverteilungen zu kombinieren sind. Zu den geläufigsten Verfahren zur Entscheidungsfindung in einer Gruppe gehören der Linear Opinion Pool, der Logarithmic Opinion Pool sowie der Supra-Bayes-Ansatz. Da diese Techniken jedoch im Allgemeinen zur Kombination von Wahrscheinlichkeitsverteilungen über dem gleichen Zielgrößenraum eingesetzt werden, stellt sich die Frage inwieweit sie direkt für die lokale Bayes'sche Informationsfusion nutzbar sind.

Ein Teil dieser Arbeit besteht deshalb aus einer Analyse der in der Literatur gebräuchlichen Verfahren zur Entscheidungsfindung in einer Gruppe im Hinblick auf ihre Eignung für die verteilte lokale Bayes'sche Informationsfusion. Gegebenenfalls soll untersucht werden inwieweit gebräuchliche Verfahren abgewandelt werden können, um sie für die verteilte lo-



kale Bayes'sche Informationsfusion nutzbar machen zu können. Die erzielten theoretischen Ergebnisse sollen an Beispielen verdeutlicht werden.

Um die Problemstellung zu verdeutlichen, soll hier ein Beispiel helfen: Die Aufgabe sei, Fahrzeugtypen spezieller Autos zu klassifizieren. Interessiert sind wir an der Koordinate, wo das Fahrzeug steht, am Typ, spezieller um welches Modell einer Marke es sich handelt, und an der Ausrichtung des Fahrzeugs. Für die Erkennung stehen mehrere Informationsquellen zur Verfügung: Straßenkarten, Luftbilder und Aussagen von Menschen. Denkbar sind noch weitere Informationsquellen, wie beispielsweise Statistiken über die Häufigkeit von Fahrzeugtypen in gewissen Regionen. Die Information dieser Informationsquellen soll so zur Informationsfusion verwendet werden, dass möglichst viele Fahrzeuge richtig klassifiziert werden. Um die DoB zu erzeugen, muss jede Information aus den Informationsquellen in eine Wahrscheinlichkeit übersetzt werden. So kann beispielsweise die Straßenkarte so interpretiert werden, dass die Häufigkeit von Fahrzeugen auf einer Straße höher ist als abseits von Straßen. Die daraus resultierende Wahrscheinlichkeitsverteilung nimmt an den Positionen der Straßen und Parkplätze höhere Werte an als auf sonstigem Gelände. In Abbildung 1.1 ist ein simuliertes Luftbild zu sehen, welches als ein Bild einer Kamera als Informationsquelle interpretiert werden kann. Da diese Informationen auf verschiedenen Abstraktionsniveaus vorliegen, ist hier die Fähigkeit zur Transformation wichtig. Abbildung 1.2 zeigt eine zu diesem Bildausschnitt passende Straßenkarte, welche in eine DoB Repräsentation transformiert wurde. Bei der Abbildung 1.2 wurde eine Gleichverteilung der Wahrscheinlichkeit über der Straße angenommen. Die Ränder der Straße sind von der Wahrscheinlichkeit her nicht feste Schnitte, sondern sie wurden durch eine Normalverteilung verwischt. Für eine genauere Beschreibung siehe [San10]. Die Informationsquellen können in diesen Fällen unkorreliert und damit unabhängig voneinander sein. Der Grad der Korrelation ist ungewiss, da die Art, wie diese Informationen entstanden sind nicht erklärt wird. Das soll anhand von zwei Beispielen erläutert werden: Im ersten Fall wurde die Straßenkarte aus dem Luftbild gewonnen. Die Aussage des Menschen wurde ebenfalls aus dem Bild entnommen, es wurden zum Beispiel die darin vorhandenen Fahrzeuge gezählt. All diesen Informationen liegen dieselben Daten, nämlich das Bild, zu Grunde. Damit sind sie stark korreliert. Im Zweiten Fall sind die Quellen der Informationen unterschiedlich. Die Straßenkarte wurde durch ein Vermessungsamt erstellt, die Statistik kommt aus der Zulassungsstelle. Dann liegen den Informationen verschiedene Daten zu Grunde, die damit fast vollständig unkorreliert sind.

Die Fusion über dem gesamten betrachteten Bereich ist zu aufwendig. Daher werden durch



**Abbildung 1.1.:** Beispiel eines Luftbildes für auf dem Fahrzeuge erkannt werden sollen, aus [San10].

eine Vorverarbeitung Teilmengen der zu betrachteten Ereignisse erzeugt, welche von höherem Interesse sind als andere. Abbildung 1.3 zeigt für das Beispiel des Luftbildes, wie solche Bereiche (grüne Markierungen) aussehen könnten. Zu beachten ist, dass bei dieser Einschränkung der PoI auf die lokalen Bereiche schon Fehler auftreten können. So wurden Bereiche in Abbildung 1.3 markiert, in denen kein Fahrzeug enthalten ist oder es wurde ein Fahrzeug (linke Straße, Mitte) übersehen. Diese Bereiche können nun von verschiedenen Agenten betrachtet werden. Für eine optimale Parallelisierung kann jedem Gruppenmitglied ein Bereich zugewiesen werden, in dem dieses die Fusion lokal durchführt. Die Ergebnisse dieser verteilten Berechnungen müssen amalgamiert werden, damit eine weitere Verarbeitung möglich ist. Die Frage, ob Methoden aus der Entscheidungstheorie für diese Amalgamation verwendet werden können, ist zentraler Bestandteil dieser Arbeit.

## 1.2. Stand der Wissenschaft

Ein großer Teil dieser Arbeit besteht aus einer Literaturübersicht. Daher werden die hier vorgestellten Veröffentlichungen nur kurz angesprochen. Eine Einführung in die verteilte



**Abbildung 1.2.:** Beispiel einer Straßenkarte als weitere Information [San10].

lokale Bayes'sche Informationsfusion ist in [San08, San09a] gegeben. Das Zusammenführen der Ergebnisse der verteilten lokalen Bayes'schen Informationsfusion wird in dieser Arbeit als eine Art Gruppenentscheid interpretiert. Die Verwendung von Methoden aus der Entscheidungstheorie scheint damit denkbar. Swain beispielsweise untersucht in [Swa92] mehrere Methoden aus der Entscheidungstheorie für eine Klassifikation von geografischen Daten aus verschiedenen Sensoren.

Viele Autoren beschreiben Aspekte der Probleme, wie sie in [Fre85] aufgeführt werden. Dies entspricht jedoch nicht der hier betrachteten Aufgabe, da der lokale Ansatz fordert, dass nicht alle Gruppenmitglieder die gleichen Ereignisse betrachten.

Eine Übersicht über das Thema Pooling-Methoden ist in [Gen86b, Jac95] gegeben. Stone [Sto61] prägte den Begriff *Linear Opinion Pool*. Viele Eigenschaften von Pooling-Methoden werden von Pennock and Wellman in [Pen05] für graphische Modelle beschrieben. McConway analysiert den Linear Opinion Pool in [McC81] auf seine Eigenschaften. Zwei dieser Eigenschaften sind die *Marginalization Property* und die *External Bayesianity*. Eine Klasse von Pooling-Methoden, welche die Marginalization Property erfüllen, wird in [McC81, Gen84d] beschrieben. Genest [Gen84b, Gen86a] beschäftigt sich ebenfalls mit



Abbildung 1.3.: Beispiel eines Luftbildes auf dem relevante Bereiche hervorgehoben wurden.

Pooling-Methoden, welche die External Bayesianity erfüllen. Weitere Pooling-Methoden wurden von [Fre80, Fre81] im Bereich der Entscheidungstheorie beschrieben. Eine dieser Pooling-Methoden ist der *Logarithmic Opinion Pool* [Gen84a]. Wie die Gewichte bei einem Logarithmic Opinion Pool gewählt werden können, wird in [Hes98, Acz80] beschrieben.

### 1.3. Aufbau der Arbeit

In Kapitel 2 wird in die Informationsfusion eingeführt. Dabei wird eine Methodik beschrieben, wie aus verschiedenen Informationsquellen Information zusammengeführt werden können. Um die Wahl der Fusionsmethodik zu erläutern, werden fünf Arten der Fusion betrachtet und gegeneinander abgegrenzt.

- Frequentistische Statistik
- Fuzzy-Logik
- Künstliche Neuronale Netze
- Bayes'sche Statistik

- Dempster-Shafer-Evidenztheorie

Dann folgt eine tiefere Einführung in die Bayes'sche Informationsfusion, gefolgt von der Definition eines lokalen Szenarios.

Kapitel 3 beschreibt die Methoden, die zur Zusammenführung der Schätzungen der Gruppenmitglieder betrachtet werden. Nach einer allgemeinen Einführung werden die drei bekanntesten Methoden (Linear/Logarithmic Opinion Pool und der Supra-Bayes Ansatz) beschrieben.

Danach werden in Kapitel 4 die Eigenschaften solcher Methoden untersucht. Zunächst werden die Eigenschaften allgemein eingeführt, um dann drei Pooling-Methoden auf die Eignung für einen Einsatz in einer lokalen Bayes'schen Fusion zu untersuchen. Dann werden allgemeingültige Regeln diskutiert, welche die Kombination dieser Eigenschaften betreffen. Je mehr von diesen Eigenschaften erfüllt sind, ohne dabei mit den Regeln der Kombinationsmöglichkeiten der Eigenschaften in einen Konflikt zu kommen, desto besser repräsentiert das Ergebnis der verwendeten Methode die Wahrscheinlichkeitsverteilungen der Gruppenmitglieder, was das Ergebnis interpretierbar macht. Die darauf folgenden Kapitel dienen der Zusammenfassung und der kritischen Betrachtung, wie gut die betrachteten Methoden lokal anwendbar sind, sowie einem Ausblick über offene Fragen in der Problemstellung.

## 2. Grundlagen

In diesem Kapitel soll die Informationsfusion, speziell die Bayes'sche Informationsfusion besprochen werden. Zur Übersicht werden zunächst die gängigsten methodischen Ansätze zur Informationsfusion vorgestellt. Dies soll auch der Motivations dienen, warum die Bayes'sche Informationsfusion verwendet wird. Danach wird genauer auf die Bayes'sche und die lokale Bayes'sche Fusion eingegangen.

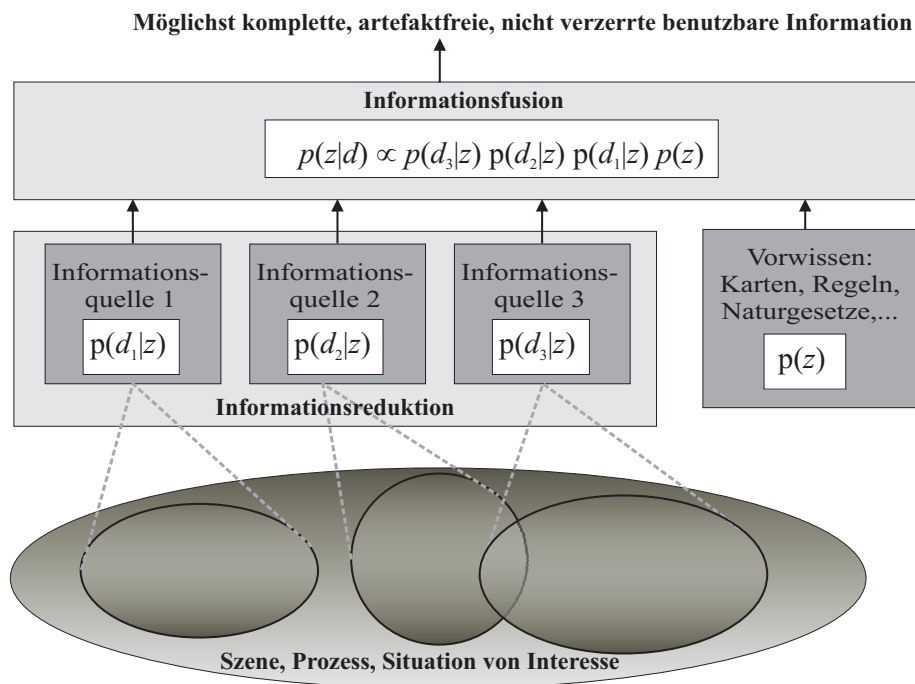
### 2.1. Informationsfusion

Wissen über die Welt zu gewinnen, ist schon lange ein Ziel der Menschen. Früher waren es Entdecker und Geologen wie Humboldt, die jeden Zentimeter, den sie bereisten, vermaßen, jede Pflanze untersuchten und jede Beobachtung notierten, um damit den menschlichen Horizont zu erweitern. Um Informationen über die Welt zu gewinnen, können aus verschiedenen Informationsquellen - bei Humboldt waren dies hauptsächlich die Sinne - verwendet werden. Informationsquellen liefern Information, die üblicherweise verarbeitet und kombiniert werden muss, um neues Wissen zu erzeugen oder Wissen zu präzisieren. Das menschliche Gehirn ist für die wissenschaftliche Disziplin der Informationsfusion ein großes Vorbild. Seine Leistung, die fünf Sinne des Menschen so genau zu nutzen, dass komplexe Abläufe, wie beispielsweise Fussballspielen, möglich sind, kann bislang nicht vollständig maschinell nachgeahmt werden.

Einhergehend mit dem technischen Fortschritt werden immer mehr Sensoren verwendet. Sie finden ihren Einsatz fast überall in unserem Alltag. Mit diesen Sensoren werden Messungen durchgeführt. Die Messergebnisse werden entweder direkt angezeigt, wie das beispielsweise bei der Außentemperaturmessung im Auto der Fall ist, oder weiterverarbeitet. Zur späteren Weiterverarbeitung können die Information aufgezeichnet werden. Die Aufzeichnung von Information macht beispielsweise eine moderne Wettervorhersage erst effektiv. Für spezielle Aufgaben werden spezielle Sensoren verwendet, die verschiedene

Funktionsweisen implementieren.

Bei der Informationsfusion wird Information aus verschiedenen Quellen so verknüpft, dass neues und/oder präziseres Wissen über die interessierenden Größen erzeugt wird. Angenommen, es sei von Interesse, was die passende Bekleidung für den heutigen Tag sein könnte. Dann wäre es sinnvoll, nicht nur die Temperatur zu messen, sondern auch die Windstärke und die Luftfeuchtigkeit. Diese drei Messergebnisse könnten dann zusammen eine Entscheidungshilfe generieren, die maßgeblich beeinflusst, welche Bekleidung zu wählen ist. Abbildung 2.1 stellt das Prinzip der Informationsfusion schematisch dar. Die Informationsquellen erfassen Teile des „Universums“ bzw. des relevanten Weltausschnitts. Die Informationsbeiträge werden mit dem Vorwissen fusioniert und ergeben neues und/oder präzisiertes Wissen über das interessierende Objekt. Die Informationsfusion bewerkstelligt



**Abbildung 2.1.:** Abbildung einer schematischen Informationsfusion aus [San10].

also genau das Zusammenführen von Information aus den verschiedenen Informationsquellen vgl. [War07]. Wie genau die vorliegende Information bei der Informationsfusion repräsentiert, interpretiert und verarbeitet wird, hängt vom zugrunde liegenden methodischen Ansatz ab. Die gängigsten methodischen Ansätze sollen hier kurz vorgestellt werden:

### 2.1.1. Frequentistische Statistik

Der Informationsbeitrag  $d_i$  aus einer Informationsquelle wird in der klassischen Statistik als Realisierung einer Zufallsgröße interpretiert. Anhand dieser Information soll die interessierende Größe  $z$  in Form einer deterministischen Konstanten geschätzt werden. Dabei liegt eine empirische frequentistische Interpretation der Wahrscheinlichkeit zugrunde. Die Wahrscheinlichkeit einer Beobachtung  $d$  ist definiert als der Grenzwert von dessen relativer Häufigkeit, d.h.

$$p(d) = \lim_{n \rightarrow \infty} \left( \frac{n_d}{n} \right).$$

Dabei ist  $n_d$  die absolute Häufigkeit der Beobachtung  $d$  in  $n$  unabhängigen, identischen Zufallsexperimenten. Dies bedeutet in Worten, dass die Wahrscheinlichkeit von  $d$  größer wird, je öfter  $d$  als Ergebnis des Zufallsexperiments erscheint.

Ein Beispiel soll dies verdeutlichen: Sei  $z$  die Distanz der Erde zur Sonne am heutigen Tag. Die Messungen, die durchgeführt werden, sind immer fehlerbehaftet. Angenommen wir fragen jeden, den wir im Laufe des Tages treffen, nach einer Schätzung für  $z$ . Dann sind die Antworten die Beobachtungen  $d_i$ . Die bedingte Wahrscheinlichkeit  $p(d|z)$  wird als Funktion von  $z$  aufgefasst, siehe Abschnitt A.1. So könnte z.B. angenommen werden, dass die Aussagen sich normalverteilt um den „wahren“ Wert von  $z$  häufen.  $z$  kann dann durch eine sogenannte *Maximum-Likelihood-Schätzung* angenähert werden [Bey99, Blo98]:

$$z_{\text{ML}} = \arg \max_z \{p(d|z)\}$$

### 2.1.2. Fuzzy-Logik

Die Fuzzy-Logik gründet sich auf eine Verallgemeinerung des Mengenbegriffs. Mit dem Erweiterungsprinzip werden Zugehörigkeiten zu einer Menge auf unscharfe Zugehörigkeitsfunktionen abgebildet. Dies hat das Ziel, eine unscharfe Wissensrepräsentation zu erhalten, in der sich Unschärfe (Vagheit) explizit ausdrücken lässt. Durch eine Zugehörigkeitsfunktion kann ein Ereignis so mehreren Klassen angehören, die nicht scharf trennbar sind. Diese Zugehörigkeitsfunktion ordnet dem Ereignis einen Grad der Zugehörigkeit zu jeder dieser Klassen zu. Durch die weitgehend freie Wahl der Zugehörigkeitsfunktion ist die Fuzzy-Logik sehr subjektiv. Seien beispielsweise die Beobachtungen Temperaturen von Wasser. Innerhalb welcher Temperaturintervalle das Wasser „kalt“, „warm“ oder „heiß“ ist, kann nicht zweifelsfrei bestimmt werden. Mit einer Zugehörigkeitsfunktion lässt sich eine Tem-



peratur von beispielsweise  $43^{\circ}\text{C}$  auf die Mengen „kalt“, „warm“ und „heiß“ abbilden. So könnten  $43^{\circ}\text{C}$  zu 0,1 kalt, zu 0,6 warm und zu 0,3 heiß sein. Mittels der Fuzzy-Logik lässt sich gut menschliches Wissen mit Unschärfe repräsentieren. Um den Unterschied zwischen Unsicherheit und Unschärfe zu demonstrieren, sei folgendes Beispiel angegeben: Ein Topf enthält Kohlen unterschiedlicher Temperatur. Beim Ziehen eines Kohlenstücks ist es *unsicher*, ob ein heisses oder ein kaltes Kohlenstück gezogen wird. Die Unterteilung, wann eine Kohle heiss und wann kalt ist, ist *unscharf*.

Da Menschen oft linguistische Aussagen unscharf einordnen, wird mittels der Fuzzy-Logik menschliches Wissen gut repräsentiert. Problematisch ist die Fuzzy-Theorie jedoch insbesondere aufgrund der hohen Freiheit bei der Wahl der Zugehörigkeitsfunktionen und der Operationen zur Kombination von Zugehörigkeitsfunktionen. Resultate dieser Kombination können problematische Ergebnisse erzeugen [Eck06].

### 2.1.3. Künstliche Neuronale Netze

Künstliche neuronale Netze (KNN) bilden das zentrale Nervensystem von Tieren und Menschen nach. Wie dabei die Biologie abgebildet wird, kann in [Roj93] (Kapitel 1) nachgelesen werden. Dabei werden durch verschiedene Topologien eine Eingabeschicht von Neuronen auf eine Ausgabeschicht von Neuronen abgebildet (vgl. [Gal93]). Dazu werden die Neuronen des KNN z.B. mit ihren Aktivierungsfunktionen und deren Verbindungen simuliert. KNN eignen sich gut, wenn es nicht möglich ist, einen Algorithmus für die Erzeugung der Ausgabe aus den Eingabedaten anzugeben. Einem KNN kann durch verschiedene Trainingsmethoden anhand von Trainingsbeispielen ein entsprechendes Verhalten beigebracht werden [Bra95]. Das Ergebnis eines Trainings ist ein trainiertes KNN. Dieses kann stark variieren, je nach dem welche Trainingsmethode, Topologie und Darstellung des Netzes gewählt wurde. KNN lassen sich jedoch sehr schwer interpretieren, da sie sehr viele Neuronen umfassen können. Zum besseren Verständnis werden viele Anwendungsbeispiele in [Hel90] besprochen.

Wegen der fehlenden Interpretierbarkeit ist es kaum möglich, Vorwissen mit in die Berechnungen eines KNN einfließen zu lassen. Bis jetzt ist es noch nicht möglich, so große KNN, wie sie in der Natur auftreten können, zu simulieren. Jedoch kann ein kleines KNN relativ komplizierte Funktionen gut approximieren. So kann z.B. jede boolesche Funktion mit einer endlichen Anzahl an Eingaben durch ein Multi-Layer-Perceptron (MLP) modelliert werden [Gal93].

### 2.1.4. Bayes'sche Statistik

Im Gegensatz zur klassischen Statistik wird bei der Bayes'schen Statistik auch die interessierende Größe  $z$  als Zufallsgröße interpretiert. Sie basiert auf dem Satz von Thomas Bayes, mit dem sich aus der A-priori-Wahrscheinlichkeitsverteilung  $p(z)$  und der Likelihoodfunktion  $p(d|z)$  die A-posteriori-Wahrscheinlichkeitsverteilung  $p(z|d)$  berechnen lässt [Ber04]:

$$p(z|d) = \frac{p(d|z)p(z)}{p(d)} \propto p(d|z)p(z).$$

Wird ein  $z$  bestimmt, für das  $p(z|d)$  maximal wird, so wird das Ergebnis  $z_{\text{MAP}}$  ein *Maximum-A-Posteriori-Schätzwert* genannt:

$$z_{\text{MAP}} = \arg \max_z p(z|d).$$

Die Bayes'sche Statistik hat den Vorteil, dass sich a-posteriori eine Wahrscheinlichkeitsverteilung für  $z$  bei gegebenen Informationsbeiträgen  $d$  angeben lässt, in die auch Vorwissen bezüglich  $z$  in adäquater Weise eingeht. Die Beobachtungen fließen in die A-posteriori-Wahrscheinlichkeit nur mittels der Likelihood mit ein. Ein weiterer Vorteil der Bayes'schen Statistik ist, dass *subjektive Wahrscheinlichkeiten* verwendet werden können. Die frequentistische Interpretation von Wahrscheinlichkeit kann dies beispielsweise nicht [Lin83].

Wahrscheinlichkeit wird hier als *Grad des Dafürhaltens (Degree of Belief, DoB)* interpretiert [Bey99]. Mit jeder Beobachtung werden Evidenzen für Hypothesen gesammelt. Für einen tieferen Einblick sei der Artikel von Lindley in [Ait91] empfohlen. Eine Hypothese könnte zum Beispiel sein, dass die interessierende Größe  $z$  einen bestimmten Wert annimmt. Die A-priori-Wahrscheinlichkeitsverteilung  $p(z)$  kann als DoB für die unterschiedlichen Werte von  $z$ , bevor Ereignisse beobachtet wurden, verstanden werden. Die von den Informationsquellen gelieferten Beobachtungen „erhärten“ oder „verweichlichen“ den Verdacht, dass eine bestimmte Hypothese richtig ist. Der Bruch  $\frac{p(d|z)}{p(d)}$  kann als Stärke des Einflusses interpretiert werden, den ein „Beweis“ in Form einer Beobachtung  $d$  auf die BoD der Hypothese bezüglich des Wertes von  $z$  hat. Ist die Wahrscheinlichkeit, dass die Daten  $d$  beobachtet werden, klein und ist die Wahrscheinlichkeit groß, dass  $d$  beobachtet wird, wenn  $z$  einen bestimmten Wert annimmt, so ist der Faktor auch groß. Daraus resultiert eine hohe A-posteriori-Wahrscheinlichkeit für den betrachteten Wert von  $z$  - vorausgesetzt, dass dieser Wert auf Basis des Vorwissens nicht extrem unwahrscheinlich ist [Ber04].

Ein Beispiel soll das beschriebene Vorgehen verdeutlichen: Bei einem medizinischen Test können zwei Ausgänge beobachtet werden. Der Test soll ein positives Resultat ergeben, wenn ein Patient an der zu testenden Krankheit erkrankt ist und im anderen Fall ein negatives Resultat liefern. Dies gelingt in der Praxis nicht immer. Ein *false positive* ist, wenn ein Patient nicht an der Krankheit erkrankt ist, der Test jedoch positiv ausfällt. Mit der Bayes'schen Methodik kann man nun berechnen, wie groß die Wahrscheinlichkeit für ein *false positive* ist. Bei einer Krankheit die 0,1 Prozent der Bevölkerung haben, ist die a-priori-Wahrscheinlichkeit dafür, dass eine Person wirklich erkrankt ist,  $p(\text{Krank}) = 0,001$ . Nehmen wir nun an, der Test würde in 5 Prozent der Fälle, in denen der Patient nicht krank ist, ein positives Ergebnis liefern. Wenn der Patient krank ist, liefert der Test in 99 Prozent der Fälle ein positives Ergebnis. In Formeln gilt also:  $p(\text{Test positiv} | \text{nicht Krank}) = 0,05$  und  $p(\text{Test positiv} | \text{Krank}) = 0,99$ . Die Wahrscheinlichkeit für einen positiven Test berechnet sich zu  $p(\text{positiv}) = p(\text{Test positiv} | \text{Krank})p(\text{Krank}) + p(\text{Test positiv} | \text{nicht Krank})p(\text{nicht Krank}) \approx 0,05094$ . Mit der Bayes'schen Formel ergibt sich dann:

$$p(\text{Krank} | \text{Test positiv}) = \frac{p(\text{Test positiv} | \text{Krank})p(\text{Krank})}{p(\text{positiv})} \approx 0,019.$$

Die Wahrscheinlichkeit für ein false positive ist nun  $1 - p(\text{krank} | \text{Test positiv}) \approx 0,981$ . Bei ca. 98 Prozent der positiven Testergebnisse handelt es sich also um ein false positive. Dies liegt daran, dass so wenige Menschen die Krankheit haben. Ein derartiges Resultat wird anti-intuitiv genannt, da häufig die Annahme, der Test würde in 5 Prozent der Fälle, in denen der Patient nicht krank ist, ein positives Ergebnis liefern, fälschlicherweise als false positive interpretiert wird.

Die Bayes'sche Informationsfusion wird in Abschnitt 2.2 näher betrachtet.

### 2.1.5. Dempster-Shafer-Evidenztheorie (DSE)

Die Evidenztheorie wurde erstmals von Dempster in den späten sechziger Jahren in Harvard beschrieben. Ungefähr zehn Jahre später beschäftigte sich Glenn Shafer in seinem Buch „A Mathematical Theory of Evidence“ [Sha76] mit den Ideen von Dempster. Die daraus resultierende Theorie wird oft als Erweiterung der Wahrscheinlichkeitstheorie betrachtet. Nach Dempster lässt sich die Unsicherheit mittels eines Intervall darstellen. Das Unsicherheitsintervall ist in der Regel bei wenig Information groß und wird mit

zunehmendem Informationsstand immer kleiner [Eck06]. Dempster benutzt die Struktur  $(\Omega, M(), \text{bel}(), \text{pl}())$ . Dabei sind die Elemente aus  $\Omega$  die Menge der Elementarereignisse, die z.B. die Lösung eines Problems darstellen. Ziel der DSE ist es, mit der vorhandenen Information unsicherheitsbehaftete Aussagen über die Elemente aus  $\Omega$  zu treffen. Die Information erzeugt eine Menge von Glauben, die vollständig auf  $\Omega$  verteilt wird. Nicht allen Elementen aus  $\Omega$  muss die gleiche Menge Glauben zugeordnet werden. Glauben ist dabei eine Menge an „Evidenz“ oder „Beweisen“, die für ein Element aus  $\Omega$  spricht.  $M$  ist die Verteilung des Glaubens über  $\Omega$ . Diese wird auch als Basismaß bezeichnet. Grundgedanke dabei ist, die Menge an Glauben, die von der vorhanden Information zugelassen wird, über die Potenzmenge  $\mathcal{P}(\Omega)$  zu verteilen. Dabei muss der Glaube nicht zwingend auf Elementarereignisse verteilt werden. Das Basismaß  $M$  ist im Allgemeinen kein Wahrscheinlichkeitsmaß.  $M$  erfüllt per Definition folgende Eigenschaften:

$$\begin{aligned} M(\omega) &\in [0, 1], \\ M(\emptyset) &= 0, \\ \sum_{\omega \subseteq \Omega} M(\omega) &= 1, \end{aligned} \tag{2.1}$$

für alle Teilmengen  $\omega \in \mathcal{P}(\Omega)$ .  $M(\omega)$  bezeichnet dabei die Menge an Glauben, die exakt  $\omega$  zugeordnet wird. Da  $M$  Glauben auch exakt an nichtatomare Ereignisse  $\omega \in \mathcal{P}(\Omega)$  verteilen kann, ist Ungewissheit gut modellierbar [War07]. Das Basismaß  $M$  induziert eine Glaubensfunktion. Die Glaubensfunktion, die den Mindestgrad des Dafürhaltens beschreibt, wird  $\text{bel}(\cdot) : \mathcal{P}(\Omega) \rightarrow [0, 1]$  genannt. Die Funktion  $\text{bel}(\omega)$  kann als Wahrscheinlichkeit interpretiert werden, dass Glaube unter der gegebenen Information bewiesenermaßen  $\omega$  zugeordnet werden kann. Sie wird wie folgt definiert:

$$\text{bel}(\omega) = \sum_{o: o \subseteq \omega} M(o).$$

In Worten ist dies die Summe des Glaubens aller Ereignisse, die  $\omega$  implizieren. Somit gibt die Glaubensfunktion die Menge an Glauben an, die mit Sicherheit  $\omega$  zugeordnet wird. Existiert keine Information, so kann anhand der DSE „vollkommene Unwissenheit“ explizit

angegeben werden. Dann ist:

$$bel(\Omega) = 1, \tag{2.2}$$

$$bel(\omega) = 0, \tag{2.3}$$

für jede echte Teilmenge  $\omega \subset \Omega$ . Dies entspricht nicht einer Gleichverteilung des Glaubens über alle Elementarereignisse, wie dies z.B. in der bayes'schen Statistik durch den Maximum-Entropie-Ansatz realisiert wird. [Hei08]

Die Plausibilitätsfunktion wird durch  $pl(\cdot)$  beschrieben. Die Plausibilitätsfunktion  $pl(\omega)$  kann als Summe der Wahrscheinlichkeiten für alle Elemente von  $\omega$  interpretiert werden, mit denen  $\omega$  möglich ist [Koh95]. Formal wird die Plausibilitätsfunktion wie folgt definiert:

$$pl(\omega) = \sum_{o: o \cap \omega \neq \emptyset} M(o),$$

wobei  $o$  die Mengen sind, deren Schnitt mit  $\omega$  nicht leer ist.  $pl(\omega)$  gibt das Maximum an Glauben an, das möglicherweise  $\omega$  zugeordnet werden kann. Dempster fasst  $bel(\omega)$  und  $pl(\omega)$  als untere und obere Schranke für die tatsächliche Wahrscheinlichkeit  $p(\omega)$  von  $\omega$  auf:

$$bel(\omega) \leq p(\omega) \leq pl(\omega).$$

Es wird klar, dass die klassische Wahrscheinlichkeit genau dann, wenn  $bel(\omega)$  gleich  $pl(\omega)$  ist, ein Spezialfall des Wahrscheinlichkeitsintervalls ist. Lindley diskutiert in [Lin82] jedoch, dass für die Darstellung von Unsicherheit Wahrscheinlichkeiten im klassischen Sinne „unausweichlich“ sind. Lindley zieht in seinem Artikel [Lin87] anhand von *scoring rules* (siehe Abschnitt A.4) den Schluss, dass der Mittelpunkt des Unsicherheitsintervalls bessere Werte erreicht als das Intervall. Dies stellt in Frage, ob ein Wahrscheinlichkeitsintervall einen Vorteil gegenüber einer klassischen Wahrscheinlichkeit darstellt.

In diesem Abschnitt wurde das Gebiet Informationsfusion einführend besprochen. Es wurden verschiedene methodische Ansätze vorgestellt. In dieser Diplomarbeit besteht die Aufgabe darin, Verfahren zu begutachten, mittels derer sich Teilergebnisse einer verteilten Bayes'schen Fusion zusammenführen lassen. Bayes'sche Fusion basiert auf der Bayes'schen Statistik.

## 2.2. Bayes'sche Fusion

Bei der Bayes'schen Fusionsmethodik handelt es sich (im Wesentlichen) um einen der methodischen Ansätze zur Informationsfusion, die in Abschnitt 2.1 eingeführt wurden. Der Bayes'sche Ansatz wurde ausgewählt, weil ein fundierter mathematischer Hintergrund existiert. Durch die Interpretation von Wahrscheinlichkeit als DoB kann jede Art der Ungewissheit ausgedrückt werden. Zusätzlich lassen sich mit der Bayes'schen Fusionsmethodik unterschiedliche skalierte Größen und Beiträge von heterogenen Informationsquellen zwanglos simultan verarbeiten [Bey07]. Bei der Bayes'schen Fusion werden durch die DoB-Betrachtung nur Wahrscheinlichkeiten als Informationsträger zugelassen. Die von den Informationsquellen gelieferte Information und das Vorwissen müssen wahrscheinlichkeitstheoretisch mittels DoB-Wahrscheinlichkeitsverteilung formuliert werden.

Anhand des Theorems von Thomas Bayes können Likelihoodfunktionen und die A-priori-Wahrscheinlichkeitsverteilung (WV) multiplikativ zu einer A-posteriori-Wahrscheinlichkeitsverteilung verknüpft werden:

$$\underbrace{p(z|d)}_{A\text{-Posteriori-WV}} = \frac{p(d|z)p(z)}{p(d)} \propto \underbrace{p(d|z)}_{\text{Likelihood-Funktion}} \underbrace{p(z)}_{A\text{-Priori-WV}} \quad (2.4)$$

Die A-posteriori-Wahrscheinlichkeitsverteilung  $p(z|d)$  lässt sich bis auf einen konstanten Faktor durch eine Gewichtung der Likelihoodfunktion  $p(d|z)$  durch die a-priori Wahrscheinlichkeitsverteilung  $p(z)$  erzeugen [Ber04]. Die Daten  $d = (d_1, \dots, d_S)$  stellen die Informationsbeiträge aus  $S \in \mathbb{N}$  Informationsquellen dar. Die Kombination der  $z^i$  zu  $Z := (z^1, \dots, z^m)$   $m \in \mathbb{N}$  sind die Größen, deren Wahrscheinlichkeit geschätzt werden soll. Die  $z^i$  mit  $i \in \{1, \dots, m\}$  setzen sich aus den *Properties of Interest* (PoI) zusammen:  $z := (z_1, \dots, z_k)$ ,  $k \in \mathbb{N}$  ist ein Vektor, in dem die PoI zusammengefasst werden. Die Wahrscheinlichkeitsverteilung der Beobachtung, welche die Skalierung vornimmt und damit die A-posteriori-Wahrscheinlichkeitsverteilung auf das Intervall  $[0,1]$  normiert, ist

$$p(d) = \int \cdots \int p(d|z_1, \dots, z_k) p(z_1, \dots, z_k) dz_1 \dots dz_k.$$

Dies ist die Auftrittswahrscheinlichkeit der Beobachtung  $d^1$ . Die Berechnung von  $p(z|d)$  für jeden möglichen Wert von  $z$  ist sehr aufwendig. Der Aufwand ist oft so groß, dass eine

---

<sup>1</sup>Für diskrete  $z_i$  muss hier aufsummiert werden. Maßtheoretisch gesehen ist die Summe jedoch nur ein Spezialfall des Integrals vlg. [San09b].

Berechnung für reale Beispiele nicht durchgeführt werden kann. Der Aufwand liegt nach [San09b] in der Klasse

$$O\left(\sqrt[k]{\prod_{i=1}^k Val(Z_i)}\right)$$

wobei  $Val(Z_i)$  die Anzahl der möglichen Werte der Komponenten von  $z$  sind. Die Berechnung von  $p(z|d)$  ist also selbst für einfache Beispiele sehr rechenaufwendig.

Es können zwei Fälle unterschieden werden. Im ersten Fall sind alle Informationen der Informationsquellen zu Beginn der Berechnung gleichzeitig vorhanden. Die Informationen können z.B. aufgezeichnet vorliegen. Dann kann, wie oben erklärt, das Fusionsergebnis berechnet werden. In dieser Arbeit wird, wie in den Arbeiten von Sander et al. (vgl. [San09b]), davon ausgegangen, dass alle Beobachtungen vor Beginn der Fusion abgeschlossen sind. Im zweiten Fall sind die Informationen nicht zu Beginn der Berechnung vollständig vorhanden. Die Informationsquellen liefern weiter neue Beobachtungen, die in die Fusion mit einfließen sollen. Die Agenten durchlaufen die Informationsquellen nacheinander, um das Fusionsergebnis zu berechnen vgl. [San09b]. Die Fusion nach (2.4) mit der Information  $d$  wird dann nur mit einem Teil der Information berechnet. Unter der Annahme der bedingten Unabhängigkeit der Daten, gegeben  $z$ , kann sequenziell fusioniert werden. Bedingt unabhängig bedeutet in diesem Fall, dass:

$$p(z|d_i, d_j) \propto p(d_i|z)p(d_j|z)p(z)$$

für alle  $i, j \in \{1, \dots, S\}, i \neq j$ , gilt. Unter dieser Bedingung liefert die Bayes'sche Fusion das gleiche Resultat, egal ob alle Information gleichzeitig fusioniert wurde, oder ob die Fusion sequenziell durchgeführt wurde. Die Formel für die sequenzielle Fusion (zum Zeitpunkt, wenn die Information der Informationsquelle  $1, \dots, i \leq S - 1$  bereits verarbeitet wurden und die Information der Informationsquelle  $i + 1$  als neue Beobachtung dazukommt) sieht wie folgt aus:

$$p(z|d_1, \dots, d_{i+1}) = \frac{p(d_{i+1}|z)p(z|d_1, \dots, d_i)}{p(d_{i+1}|d_1, \dots, d_i)} \propto p(d_{i+1}|z)p(z|d_1, \dots, d_i). \quad (2.5)$$

Das Ergebnis des  $i$ -ten Schritts in dieser Sequenz wird also als A-priori-Wahrscheinlichkeitsverteilung für den  $i + 1$ -ten Schritt interpretiert. Wegen der Kommutativität der Multiplikation ist die Reihenfolge, in der die Information fusioniert wird, egal [San09b]. Dies hat den Vorteil, dass die Informationsquellen separat in eine Wahrscheinlichkeitstheorie

retische Repräsentation überführt werden können. Anstatt  $p(d_1, \dots, d_S|z)$  zu berechnen kann  $p(d_1|z), \dots, p(d_S|z)$  berechnet werden.

Eine *Konfiguration* soll hier beschreiben, wie die dem Wahrscheinlichkeitsraum zu Grunde liegende Menge und  $\sigma$ -Algebra definiert sind. Dabei kann beispielsweise  $Z$  endlich oder abzählbar sein. Des weiteren wird in einer Konfiguration beschrieben, wie die Bayes'sche Informationsfusion lokalisiert wird. Dies kann beispielsweise durch eine Vergrößerung oder durch eine Einschränkung geschehen.

### 2.3. Lokale Konfiguration

Um auf das eingangs erwähnte Beispiel aus der Kriminalistik zurück zu kommen, entspricht die Grundmenge  $Z$  der  $z$  gewissermaßen der gesamten Menschheit. Wurde ein Verbrechen verübt, so ist eine Berechnung der Wahrscheinlichkeit, wie verdächtig a-posteriori ein Element  $z \in Z$  ist, also  $p(z|d)$ , nicht für jedes  $z \in Z$  durchführbar. Eine solche Vorgehensweise würde bedeuten, dass am Anfang der Ermittlungen jeder Mensch als potenzieller Täter gesehen würde. Dies geschieht natürlich nicht. Deshalb werden die Ermittlungen auf plausible Verdächtige konzentriert. Dies geschieht durch eine Art Vorverarbeitung, die zu einer Verringerung der Komplexität führt. Je mehr relevante Information bei der Vorverarbeitung verwendet wird, desto besser sollte die Menge der Verdächtigen eingeschränkt werden können. So eine Vorverarbeitung entdeckt eine Menge verdächtiger Elemente  $U$  aus  $Z$ . Das Komplement zu  $U$  ist  $\bar{U} := Z \setminus U$ , also all die Elemente, die für die Untersuchung wahrscheinlich uninteressant sind. Ziel lokaler Bayes'scher Fusion ist es, die a-posteriori-Wahrscheinlichkeitsverteilung im Detail nur für die Elemente von  $U$  zu bestimmen. Die Wahrscheinlichkeit für  $\bar{U}$  wird allenfalls kumuliert betrachtet. Sie ist dann im kontinuierlichen Fall:

$$p(\bar{U}|d) = \int_{z \in \bar{U}} p(z|d) dz,$$

und im diskreten Fall:

$$p(\bar{U}|d) = \sum_{z \in \bar{U}} p(z|d).$$

Um zu entscheiden welches der Elemente aus  $Z$  a-posteriori wahrscheinlicher als andere ist, also welchem Element  $z$  die Daten  $d_1, \dots, d_S$  eine hohe A-posteriori-Wahrscheinlichkeit zuordnen, müssen die verschiedenen Werte von  $p(z|d)$  verglichen werden. Dazu können



z.B. die Quotienten der A-posteriori-Wahrscheinlichkeiten verschiedener Elemente aus  $Z$  verglichen werden.

$$\frac{p(z^i|d)}{p(z^j|d)} = \frac{p(d|z^i) p(z^i)}{p(d|z^j) p(z^j)} \quad (2.6)$$

Anhand des Likelihood-Verhältnisses können zwei Elemente aus  $Z$  darauf hin verglichen werden, wie plausibel die Beobachtungen ein Element erscheinen lassen [San08]. Dabei kann die Bayes'sche Fusion als eine Art Beweise zu sammeln, interpretiert werden. Sind die  $d_i$  bedingt unabhängig im Bezug auf  $z$ , dann kann Gleichung (2.6) auch wie folgt geschrieben werden [San08]:

$$\frac{p(z^j|d)}{p(z^k|d)} = \prod_{i=1}^S \frac{p(d^i|z^j) p(z^j)}{p(d^i|z^k) p(z^k)}$$

Ein Element  $z$  wird dann als "clue" (Verdachtsmoment) bezeichnet, wenn es in die Menge  $U$  aufgenommen wird. Diese „clues“ werden bei der Fusion detailliert betrachtet. Sander und Beyerer schlagen in [San08] vor, eine "relative Likelihoodfunktion"  $p_{max}(d|\cdot) : Z \rightarrow [0, 1]$  einzuführen, die alle Elemente aus  $Z$  mit dem plausibelsten Element vergleicht und so eine Ordnung auf  $Z$  erzeugt.  $p_{max}$  ist wie folgt definiert:

$$p_{max}(d|z) := \frac{p(d|z)}{\max_{z^* \in Z} p(d|z^*)}$$

Diese relative Likelihood-Funktion beschreibt, um wie viel unplausibler ein bestimmter Wert von  $z$  gegenüber dem plausiblen Element  $z^*$  ist.  $\bar{U}$  wird dann durch eine obere Schranke für die relative Likelihoodfunktion definiert:

$$\bar{U} := \left\{ z \mid p_{max}(d|z) \leq \left( \frac{1}{k} \right)^{f_S} \right\},$$

mit  $d := (d_0, \dots, d_S)$ , wobei  $d_0$  das a-priori-Wissen repräsentiert.

$U$  ist dann  $Z \setminus \bar{U}$ . Hierbei ist  $f_S$  eine Konstante, deren Wert davon abhängt, wie stark bedingt unabhängig gegeben  $z$  die Informationsquellen sind. Werden heterogene Informationsquellen verwendet, kann davon ausgegangen werden, dass diese allenfalls sehr schwach bedingt abhängig sind. Dann sollte  $f_S$  nahe bei 1 sein. Sind die Informationsquellen jedoch vollkommen bedingt abhängig, dann sollte  $f_S$  den Wert  $S$  annehmen [San08]. Sind die Informationen der  $S$  Quellen identisch, wird die Wahrscheinlichkeitsgrenze nach oben verschoben. Kommt bei jeder neuen Beobachtung neue unverrauschte Information hinzu,

so kann die Wahrscheinlichkeitsschranke bei  $\frac{1}{k}$  bleiben. Der Bruch  $\frac{1}{k}$  stellt die Größe der Umgebung dar.

Alle  $z \in \bar{U}$  können mathematisch als ein Element betrachtet oder vollständig ignoriert werden. Werden sie vollständig ignoriert, dann ist  $p(\bar{U}) = 0$  und die Lokalisierung wird *Einschränkung* genannt. Wird die  $\sigma$ -Algebra des Wahrscheinlichkeitsraums vergrößert, dann wird die Fokussierung *Marginalisierung* genannt. Die Marginalisierung wird in Abschnitt 2.4 genauer besprochen. Die einzelnen Umgebungen  $U_c$  sind voneinander unabhängig, wenn die  $U_c$  eine Partition von  $U$  bilden. Damit eignen sie sich für eine *Parallelisierung*. *Parallelisierung* ist eine verteilte, gleichzeitige Berechnung. Durch die parallele Berechnung kann ein Zeitersparnis erreicht werden. Die Bereiche  $U_c$  können dann als *Arbeitspakete* an Agenten vergeben werden. Diese können dann auf verschiedener Hardware ausgeführt werden um einen *Speedup* zu erzielen. Ein *Speedup* ist eine Verkürzung der Rechenzeit ohne eine Reduktion des Rechenaufwands. Meistens wird der Rechenaufwand bei einer verteilten Berechnung durch z.B. Kommunikationsaufwand erhöht. Wie der Rechenaufwand verringert werden kann, ist Thema des nächsten Abschnitts. Durch eine Lokalisierung wird die Fusion fokussiert [Bey07]. Hierbei muß man sich mangels eines allgemeingültigen Distanz-Begriffes von dem Begriff der Lokalität als räumliche Nähe trennen. Diese eingeschränkte Fusion ist begrenzt auf die lokale Welt  $(U \times D, p_U(z, d))$  statt auf die globale Welt  $(Z \times D, p(z, d))$ . Es gilt  $U \subseteq Z$  und  $p_U(\cdot) = p(\cdot|U)$  mit:

$$p_U(z) = \begin{cases} \frac{p(z)}{\int p(z) \mathrm{d}z}, & z \in U \\ 0, & z \notin U \end{cases}, \quad (2.7)$$

$$p_U(z|d) = \begin{cases} \frac{p(z|d)}{\int p(z|d) \mathrm{d}z}, & z \in U \\ 0, & z \notin U \end{cases} \quad (2.8)$$

$(U \times D, p_U(z, d))$  bedeutet, dass die gleiche Information  $D$  beobachtet wird, diese jedoch nur für die PoI, die in  $U$  liegen, durch die eingeschränkte Wahrscheinlichkeitsfunktion  $p_U(z, d)$  eine Aussage liefert<sup>2</sup>. Allen Elementen in  $\bar{U}$  kann keine Auftrittswahrscheinlichkeit mehr zugeordnet werden. Dies bedeutet, dass  $\sum_{u \in U} p(u) = 1$  ergibt. Die Wahrscheinlichkeit, die bei einer globalen Sicht  $\bar{U}$  zugeordnet wurde, wird in der eingeschränkten Konfiguration auf  $U$  umverteilt. Klar ist dadurch (außer  $p(\bar{U}) = 0$ ), dass allen Elementen in  $U$  lokal eine höhere Wahrscheinlichkeit zugeordnet wird, als dies global der Fall wäre. Lokalisierung als Einschränkung bedeutet in diesem Zusammenhang die vollständige Konzentration auf

---

<sup>2</sup>In Gleichung (2.7) muss das Integral für diskrete  $z$  durch eine Summe ersetzt werden.

die Werte von  $Z$ , die plausibel genug erscheinen. Die Werte, aus denen sich die Ereignisse zusammen setzen, die nicht von Interesse sind, werden Nuisance-Parameter genannt. Nuisance-Parameter können durch Herausintegrieren bei kontinuierlichen bzw. Heraussummieren bei diskreten Merkmalen, entfernt werden[Ber04]. Eine weitere Möglichkeit Nuisance-Parameter zu entfernen, stellt die Bildung des A-posteriori Erwartungswerts dar [Bey07]. Eine alternative Art eine Lokalisierung zu beschreiben wird in Abschnitt 2.4 besprochen. Durch eine Lokalisierung reduziert sich der Rechenaufwand der Fusion drastisch [San06]:

$$O \left( \sqrt[t]{\prod_{i=1}^t Val(U_i)} \right),$$

wobei  $Val(U_i) \ll Val(Z_i)$  mit  $z = (z_1, \dots, z_t) \in U$  ist. Die Größe  $Val(U_i)$  beschreibt die Menge der Werte, die  $z \in U_i$  annehmen kann. Sie ist kleiner als bei der globalen Betrachtung, da möglicherweise Nuisance-Parameter herausgenommen wurden. Wie bei der Kriminalistik wird um die „verdächtigsten“ Elemente ermittelt. Alle Elemente die nicht zu dieser Gruppe  $U$  der Verdächtigen gehören, werden nicht mehr betrachtet.

Der Fehler, der durch die Umverteilung der Wahrscheinlichkeitsmasse entsteht, wird für die Reduktion des Rechenaufwands in Kauf genommen. Diesem Fehler entgegen zu wirken ist ein gewünschter Effekt einer Pooling-Methode. Das verteilte Rechnen der Bayes'schen Fusion und die Probleme, die dabei auftreten, sind Thema des nächsten Abschnitts. Nun folgt ein Beispiel, das verdeutlicht, wie stark sich der soeben besprochene lokale Überschätzungsfehler bei einer lokalen Betrachtung der Wahrscheinlichkeitsverteilung auswirken kann. Diese Situation kann mit der eines Ermittler verglichen werden, der sich auf eine kleine Menge von Verdächtigen konzentriert. Wenn das Wissen über andere Verdächtige fehlt, ist es möglich, dass einer dieser Verdächtigen - auch wenn er im Vergleich zu anderen nicht sehr verdächtig ist - eine hohe Wahrscheinlichkeit vom Ermittler erhält, das Verbrechen verübt zu haben. Dies gilt, wenn zum Beispiel der Ermittler gezwungen wird, sich für einen Täter zu entscheiden. Um diesen Effekt mathematisch zu verdeutlichen, wurde ein einfaches Beispiel gewählt, bei dem  $|Z| = 7$  gilt und sieben verschiedene Informationsquellen betrachtet werden. Die probabilistische Repräsentation der von den Informationsquellen gelieferten Information sieht wie in Tabelle 2.1 aus: Die Tabelle 2.1 stellt die bedingten Wahrscheinlichkeiten  $p(d_i|z^j)$  dar. Je größer  $p(d_i|z^j)$ , desto wahrscheinlicher ist die Beobachtung  $d_i$  der  $i$ -ten Informationsquelle unter der Hypothese, dass  $z$  den Wert  $z^j$  annimmt,  $i, j \in \{1, \dots, 7\}$ . In diesem Beispiel werden zwei Gruppenmitglieder betrachtet, welche die Fusion auf verschiedenen Bereichen von  $Z$  durch-

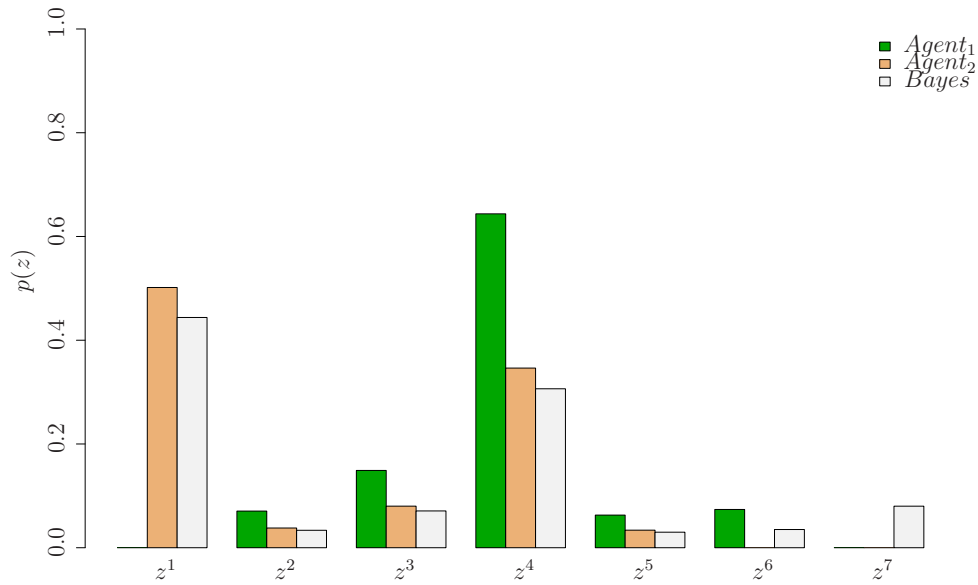
	d1	d2	d3	d4	d5	d6	d7
$z^1$	0.40	0.20	0.08	0.10	0.06	0.22	0.04
$z^2$	0.20	0.01	0.40	0.05	0.16	0.20	0.02
$z^3$	0.04	0.12	0.04	0.15	0.25	0.03	0.25
$z^4$	0.06	0.08	0.15	0.20	0.27	0.10	0.06
$z^5$	0.10	0.19	0.20	0.05	0.01	0.40	0.03
$z^6$	0.09	0.33	0.05	0.01	0.10	0.04	0.45
$z^7$	0.11	0.07	0.08	0.44	0.15	0.01	0.15

**Tabelle 2.1.:** Dargestellt sind die Likelihoodfunktionen  $p(d_i|z^j)$

führen. *Agent*<sub>1</sub> arbeitet auf  $\{z^2, \dots, z^6\} =: U_1$  *Agent*<sub>2</sub> auf  $\{z^1, \dots, z^5\} =: U_2$ . Die Wahrscheinlichkeiten der von ihm nicht betrachteten Ereignisse, welche in  $\bar{U}_i$  zusammengefasst werden, setzt jeder Agent auf Null. Als Vergleich wurde hier zusätzlich die globale Bayes'sche Fusion durchgeführt. Das Ergebnis der drei Fusionsvarianten stellt Abbildung 2.2 dar. Abbildung 2.2 zeigt den Unterschied zwischen dem Ergebnis der globalen Bayes'schen Fusion (Bayes), die über dem gesamten Raum  $Z$  berechnet wurde und den Ergebnissen der beiden Agenten, deren Fusionsergebnisse aufgrund der lokalen Betrachtung verzerrt wurden. Es wurde die Wahrscheinlichkeitsmasse von  $\bar{U}_i$  auf  $U_i$  aufgeteilt. Gut zu sehen ist, wie die Agenten global sehr kleine Wahrscheinlichkeiten lokal überschätzen. Für *Agent*<sub>2</sub> ist  $z^1$  das wahrscheinlichste Element aus  $U_2$ . Da *Agent*<sub>2</sub> nur die Verdächtigen  $z^1, \dots, z^5$  kennt, kann er nicht entscheiden wie verdächtig  $z^1$  wirklich ist. Wurde die komplette Bayes'sche Fusion berechnet, so sind die Wahrscheinlichkeiten der  $z^i$  bezüglich der Beobachtungen, exakt bekannt. Für die Reduktion des Rechenaufwandes bei der lokalen Fusion müssen Schätzfehler in Kauf genommen werden. Es ist aus Abbildung 2.2 jedoch ersichtlich, dass ein Agent die Verhältnisse zwischen den Wahrscheinlichkeiten der  $z^i$ , die er betrachtet, richtig einschätzt. Die Agenten schätzen die Wahrscheinlichkeiten von  $z^2$  bis  $z^4$  in den Verhältnissen untereinander richtig, jedoch immer proportional zu  $p(z|d)$  zu hoch ein. Weil  $p(U_1|d)$  relativ klein ist, schätzt *Agent*<sub>1</sub> die Wahrscheinlichkeit für  $z^4$  so groß ein, dass gilt:  $p_{U_2}(z^1|d) \leq p_{U_1}(z^4|d)$ . Müssten die Agenten nun entscheiden welches  $z^i$  das Wahrscheinlichste ist, so würde  $z^4$  gewählt<sup>3</sup> werden, obwohl  $z^1$  global am Wahrscheinlichsten ist. Es gibt eine obere und untere Schranke für die globale a-posteriori-

---

<sup>3</sup>Eine Wahl ist hier beispielsweise ein gewichtetes arithmetisches Mittel, beidem die Gruppenmitglieder gleich gewichtet werden.



**Abbildung 2.2.:** Diskretes Fusionsbeispiel Unterschied zwischen globalem und lokalen Fusionsergebnissen

Wahrscheinlichkeitsverteilung  $p(z|d)$ . Die obere Schranke liegt bei:

$$p(z|d) \leq \frac{p(z|d)}{p(U|d)} = p_U(z|d), z \in U.$$

Ist  $U = Z$ , dann ist das Ergebnis der auf  $U$  lokalisierten Bayes'schen Fusion gleich  $p(z|d)$  ohne Fehler, da ganz  $Z$  betrachtet wurde, also  $p(U|d) = 1$ . Je kleiner  $p(U|d)$ , desto größer wird die obere Schranke für die globale a-posteriori-Wahrscheinlichkeitsverteilung. Genauso ist die obere Schranke für größere Werte der globalen Wahrscheinlichkeitsverteilung  $p(z|d)$  höher. Ist  $p(z|d)$  also groß und  $p(U|d)$  klein, dann kann der Schätzfehler für die globale a-posteriori-Wahrscheinlichkeitsverteilung relativ groß sein. In Abbildung 2.2 wird jedes Ereignis  $z^i$  als gleich wahrscheinlich vorausgesetzt. Diese Annahme wird als *nicht informatives a-priori-Wissen* bezeichnet. In diesem Fall ist  $p(U_i|d) = \frac{3}{7}$ . Die untere Schranke für die globale Relevanz kann wie folgt motiviert werden:  $p(z|d)$  kann berechnet werden als  $p_U(z|d)p(U|d)$ . Die lokal berechnete Wahrscheinlichkeitsverteilung für  $z$  gegeben  $d$  wird mit  $p(U|d)$ , der Relevanz von  $U$  gegeben  $d$ , gewichtet.  $p(U|d)$  kann nach unten abgeschätzt

werden durch:

$$p(U|d) = \frac{\sum_{z \in U} p(d|z)p(z)}{p(d)} \quad (2.9)$$

$$\begin{aligned} &= \frac{\sum_{z \in U} p(d|z)p(z)}{\sum_{z \in U} p(d|z)p(z) + \sum_{z \in \bar{U}} p(d|z)p(z)} \\ &\geq \frac{\sum_{z \in U} p(d|z)p(z)}{\sum_{z \in U} p(d|z)p(z) + \epsilon[1 - p(U|d)]} \\ &= \frac{p(U|d)}{p(U|d) + \epsilon[1 - p(U|d)]}. \end{aligned} \quad (2.10)$$

$\epsilon$  ist dabei die Grenze, ab welcher die Elemente von  $Z$  anhand der Schranke für die Likelihoodfunktion nicht mehr zur relevanten Menge  $U$  gehören, d.h. es gilt  $p(d|z) \leq \epsilon$  für  $z \in \bar{U}$  [San09a].

Bei einer Einschränkung als Lokalisierung wird die Fusion auf einen relevant erscheinenden Bereich eingeschränkt. Dies ermöglicht die Berechnung der Fusion für kompliziertere Aufgaben, verursacht jedoch einen Fehler in der Fusion.

## 2.4. Alternative lokale Konfigurationen

Es gibt mehrere Arten, eine Lokalisierung zu realisieren. Die in Abschnitt 2.3 vorgestellte Einschränkung ist eine der beiden Arten, die in dieser Arbeit betrachtet werden. Im folgenden Abschnitt wird die zweite Art der Lokalisierung, die Marginalisierung oder Vergrößerung, besprochen. Ein Bayes'sches Modell wird maßtheoretisch durch die Struktur  $(Z \times D, \sigma(\mathcal{Z}, \mathcal{D}), p(z, d))$  beschrieben. In der Menge  $Z \times D$  sind nur atomare Elemente enthalten. Für das Beispiel der Fahrzeugerkennung ist  $Z$  aus Tupeln der Form  $\{Position, Typ, Orientierung\}$ , d.h. den möglichen Werten der PoI, zusammengesetzt.  $D$  ist die Menge der möglichen Beobachtungen, also die potentiell mögliche Information der Informationsquellen. Die Wahrscheinlichkeitsverteilung  $p : \sigma(\mathcal{Z}, \mathcal{D}) \rightarrow [0, 1]$  erfülle die Kolmogorow'schen Axiome. Sie weist den Elementen von  $\sigma(\mathcal{Z}, \mathcal{D})$  einen Wert in  $[0, 1]$  zu. Die mengentheoretische Strukturen  $\mathcal{Z}, \mathcal{D}$  sind dabei  $\sigma$ -Algebren, die definieren, welchen Ereignissen eine Wahrscheinlichkeit durch  $p$  zugeordnet werden kann. Durch die Eigenschaften einer  $\sigma$ -Algebra (siehe Abschnitt A) wird auch nicht-elementaren Ereignissen

eine Wahrscheinlichkeit zugeordnet. Ist zum Beispiel  $Z := \{z^1 = ((1, 2), VW, Nord), z^2 = ((2, 3), BMW, Nord), z^3 = ((2, 3), VW, Ost)\}$  dann kann  $\sigma(z^1)$  z.B.  $\{Z, \{z^1\}, \{z^2 \vee z^3\}, \emptyset\}$  sein. Dadurch ist es möglich, einem Element wie  $\{z^2 \vee z^3\}$  eine Wahrscheinlichkeit zuzuordnen. Nun ist es möglich, eine Vergrößerung bezüglich  $Z$  zu definieren. Dies geschieht, indem verschiedene Elemente aus der feingranularsten  $\sigma$ -Algebra  $\sigma(\mathcal{Z}, \mathcal{D})$  weggelassen werden. Dadurch wird eine Lokalisierung auf die Elemente erreichbar, die von Interesse sind.

Der Unterschied zur Einschränkung aus Abschnitt 2.3 ist, dass für die Ereignisse, die nicht betrachtet werden, die Wahrscheinlichkeitsverteilung nicht auf Null gesetzt wird. Dadurch bleibt die Wahrscheinlichkeit der Elemente von  $\sigma(\mathcal{U}, \mathcal{D})$  gleich wie vor der Vergrößerung. Der Fehler, der durch die Umverteilung von Wahrscheinlichkeitsmasse bei der Einschränkung entsteht, fällt weg. Es wird eine Menge gebildet, die alle nicht betrachteten Elemente enthält, wie dies im obigen Beispiel 2.12 der Fall ist.

Es kann eine Art Mischung aus diesen beiden Konfigurationen beschrieben werden. Dabei wird die Konfiguration, bei der eine Einschränkung lokalisiert, mit Zusatzwissen auf eine quasi Marginalisierung erweitert. Dazu muss  $p(U_i)$  bekannt sein. Die Verzerrung, welche durch die Umverteilung der Wahrscheinlichkeitsmasse entstanden ist, kann damit aufgehoben werden. Jedes Gruppenmitglied erzeugt eine Wahrscheinlichkeitsverteilung über dem von ihm betrachteten  $U_i$ . Das bedeutet das  $\sum_{z \in U_i} p_i(z) = 1$  ist. Genauer ist damit  $p_i(z) = \frac{p(z)}{p(U_i)}$ . Diese Verzerrung kann mit dem Zusatzwissen  $p(U_i)$  entzerrt werden. Dazu muss jedes Gruppenmitglied nur seine Wahrscheinlichkeitsverteilung über  $U_i$  mit  $p(U_i)$  multiplizieren. Damit wird die gleiche Wahrscheinlichkeitsverteilung wie bei der alternativen Konfiguration mit einer Marginalisierung erzeugt. Diese Konfiguration soll *Einschränkung mit Zusatzwissen* genannt werden. Während die Marginalisierung keine Aussage darüber trifft, wie die Wahrscheinlichkeit auf  $\bar{U}$  verteilt ist, kann mit der Maximum-Entropie Methode bei der Einschränkung mit Zusatzwissen eine Gleichverteilung über  $\bar{U}$  angenommen werden. Dadurch wird eine neue Wahrscheinlichkeitsverteilung jedes Gruppenmitglieds erzeugt:

$$p_i^*(z) = \begin{cases} p(z), & \text{falls } z \in U_i \\ \frac{p(\bar{U}_i)}{|\bar{U}_i|}, & \text{sonst} \end{cases} \quad (2.11)$$

Auf  $U_i$  ist die Wahrscheinlichkeitsverteilung unverzerrt, sonst wird eine Gleichverteilung angenommen. Ein Beispiel soll dies verdeutlichen: Sei  $U := \{z^1 = ((1, 2), VW), z^2 = ((1, 2), Audi)\}$ . Wobei die erste Komponente die x- und y-Kordinate ist, die zweite Kom-

ponente der Fahrzeugtyp eines Autos. Die lokalen Wahrscheinlichkeiten könnten dann z.B. wie folgt aussehen:

$$\begin{aligned} p(z^1|d) &= 0,6, \\ p(z^2|d) &= 0,25, \\ p(\bar{U}|d) &= 0,15, \end{aligned} \tag{2.12}$$

Die Beobachtungen  $d$  lassen schließen, dass an dem Punkt (1,2) zu 60 Prozent Wahrscheinlichkeit ein VW und zu 25 Prozent ein Audi zu beobachten ist. Bei dem obigen Beispiel wurde  $U$  aus einzelnen  $z$  zusammen gesetzt, wobei diese die Elemente repräsentieren die von erhöhtem Interesse sind. Dies ist stark vereinfacht. Normalerweise erhält man um ein verdächtiges Element  $c$  (für das englische Wort clue) eine Umgebung  $U_c$  mit Elementen die ebenfalls verdächtig sind:

$$U_c := \{z | d(c, z) \leq \epsilon\}, \tag{2.13}$$

$d(\cdot, \cdot)$  ist hierbei eine passende Metrik. Alternativ kann durch

$$\bar{U} := \left\{ z \mid \frac{p(d|z)}{\max_{z^* \in Z} p(d|z^*)} \leq \epsilon \right\} = \{z | p_{\max}(d|z) \leq \epsilon\}, \tag{2.14}$$

eine Schranke für die Relevanz aller Elemente, die nicht von Interesse zu sein scheinen, definiert werden. Diese Elemente werden dann in  $\bar{U}$  zusammengefasst [San09b]. Im nicht verteilten Ansatz werden alle  $U_c$  zusammengefasst zu  $U := \bigcup U_c$ .

Mit den in diesem Abschnitt 2 vorgestellten Werkzeugen ist die in dieser Arbeit gestellte Aufgabe einer Analyse zugänglich. Anhand der Bayes'schen Informationsfusion kann gegebene Information optimal ausgenutzt und kombiniert werden. Mit ihr ist es möglich, Vorwissen mit in die Informationsfusion einfließen zu lassen. Die vorgestellten Strukturen erlauben eine Einschränkung sowie eine Vergrößerung, um eine Lokalisierung der Bayes'schen Informationsfusion zu erreichen. Lokale Bayes'sche Modelle können verteilt bearbeitet werden. Die Ergebnisse sind jedoch global unvollständig oder mit einem Fehler behaftet. Der Fehler tritt bei der Einschränkung auf und wurde mit einer oberen und unteren Schranke abgesteckt. Im Folgenden werden ausgewählte Verfahren, mit denen verteilt berechnete lokale Fusionsergebnisse zu einer globaleren Sicht zusammengeführt werden können, besprochen.



## 3. Pooling Methoden

### 3.1. Einführung in Pooling-Methoden

Pooling-Methoden werden verwendet, um die Konsensfindung in Gruppen zu modellieren. Bei den Mitgliedern der Gruppe, den sogenannten Agenten, kann es sich um menschliche Experten oder um Software handeln. Die Information ist dabei in Form von Wahrscheinlichkeitsverteilungen vorhanden. Die Gruppenmitglieder erzeugen eine Wahrscheinlichkeitsverteilung über den von ihnen betrachteten PoI. Diese Wahrscheinlichkeitsverteilungen der Agenten können mit Hilfe einer Pooling-Methode zu einem Gruppenentscheid zusammengeführt werden. Ein *Entscheider* (Decision Maker) kann als ein  $n + 1$ -tes Gruppenmitglied interpretiert werden. Dies ist beispielsweise die Person, die an dem Gruppenentscheid interessiert ist.

Die Meinungen der Gruppenmitglieder können sich unterscheiden. Dafür gibt es mehrere Erklärungen. Wahrscheinlichkeiten werden *objektiv* genannt, wenn die unterschiedlichen Individuen bei gleichen Beobachtungen  $d$  und gleichem Vorwissen auf die gleiche Wahrscheinlichkeitsverteilung schließen. Gruppenmitgliedern können verschiedene Beobachtungen für ihre Meinungsbildung zur Verfügung stehen. Dann können die Wahrscheinlichkeiten objektiv sein, obwohl die Meinungen der Gruppenmitglieder unterschiedlich sind. Liegen den Gruppenmitgliedern die gleichen Beobachtungen vor, jedoch verschiedenes Vorwissen, dann kann die Meinungsbildung trotz verschiedener resultierender Wahrscheinlichkeitsverteilungen objektiv sein. Werden unterschiedliche Zusatzannahmen wie beispielsweise eine Einschränkung des Wertebereichs der Größen von Interesse gemacht, dann kann eine objektive Meinungsbildung ebenfalls verschiedene Ergebnisse liefern.

Eine weitere Erklärung dafür, warum Gruppenmitglieder bei der Meinungsbildung verschiedene Wahrscheinlichkeitsverteilungen erzeugen, liefert die Interpretation der Wahrscheinlichkeit als *subjektive* Wahrscheinlichkeit. Dabei begründet sich ein Unterschied in den Wahrscheinlichkeitsverteilungen zusätzlich durch die Subjektivität der Gruppenmit-

glieder. Hier wird also der Grund für die Ungleichheit nicht auf verschiedenes Vorwissen oder unterschiedliche Beobachtungen begrenzt. Dies repräsentiert eher eine menschliche Aussage, da die Erfahrungen und das Wissen eines Menschen schlecht greifbar sind.

Bei der Wahl einer Pooling-Methode wird zwischen zwei Ansätzen unterschieden. Beim axiomatischen Ansatz wählt man die Pooling-Methode anhand gewünschter Eigenschaften aus. Die Wahl wird oft durch Sätze erleichtert, die hier *Beschränkungssätze* (siehe 4.2) genannt werden. Der zweite Ansatz ist analytisch: Man überprüft, welche Eigenschaften eine bestimmte Pooling-Methode erfüllt. Dieser Ansatz wird in Abschnitt 4.3, 4.4 und 4.5 anhand des Linear Opinion Pools, des Logarithmic Opinion Pool und des Supra-Bayes Ansatzes veranschaulicht.

## 3.2. Der Linear Opinion Pool und der generelle Linear Opinion Pool

Ein Opinion Pool wurde von Stone [Sto61] als eine Funktion zur mathematischen Formulierung einer Entscheidungsfindung in einer Gruppe eingeführt. Dabei geben  $n$  Mitglieder einer Gruppe Schätzungen über ein Ereignis  $A \in \mathcal{Z}$  ab. Sie werden durch Wahrscheinlichkeitsverteilungen  $p_1(A), \dots, p_n(A)$  dargestellt. Dabei ist  $p_i(A)$  die Schätzung des  $i$ -ten Gruppenmitglieds bezüglich des Ereignisses  $A$ . Die Wahrscheinlichkeit wird als DoB interpretiert.

Um einen Gruppenentscheid zu erreichen, wird eine Pooling-Methode benötigt, welche die Schätzungen der Gruppenmitglieder zusammenführt. Stone schlägt dafür ein gewichtetes arithmetisches Mittel vor:

$$K(p_1, \dots, p_n)(A) = \sum_{i=1}^n w_i p_i(A), \quad (3.1)$$

mit den Gewichten  $w_i$ , für die  $\sum_{i=1}^n w_i = 1$  und  $w_i \geq 0$  gilt. Bacharach [Bac75] bezeichnet diese Pooling-Methode als *Linear Opinion Pool (LinOP)*.

Genest verallgemeinert den LinOP so, dass er eine größere Klasse von Pooling-Methoden umfasst<sup>1</sup> [Gen84d]. Dieser wird hier *generalisierter Linear Opinion Pool (GenLinOP)* ge-

---

<sup>1</sup>Die Klasse enthält alle Pooling-Methoden, die die Marginalization Property von McConway erfüllt.

nannt. Diese Pooling-Methoden haben die Form:

$$K(p_1, \dots, p_n)(A) = \sum_{i=1}^n w_i p_i(A) + \left[ 1 - \sum_{i=1}^n w_i \right] R(A), \quad (3.2)$$

mit Gewichten  $w_1, \dots, w_n \in [-1, 1]$  und einem beliebigen Wahrscheinlichkeitsmaß  $R$  auf  $\mathcal{Z}$ , so dass gilt:  $\forall J \subseteq \{1, \dots, n\} : \left| \sum_{j \in J} w_j \right| \leq 1$ .

Die Gewichte drücken aus, wie viel Einfluss die Wahrscheinlichkeitsverteilung  $p_i(A)$  auf den Gruppenentscheid  $K(p_1, \dots, p_n)(A)$  hat.

Die Erweiterung des LinOP auf den GenLinOP ist bei gewissen Anwendungen durchaus sinnvoll. Stellen wir uns beispielsweise vor, dass der Gruppenentscheid nicht nur durch die Gruppenmitglieder erzeugt wird. Es könnte a-priori durch Vorwissen eine grobe Schätzung der Wahrscheinlichkeitsverteilung vorliegen. Diese soll nun mit den Meinungen der Gruppenmitglieder vereinigt werden. In diesem Fall kann das Vorwissen als  $R(A)$  mit in den GenLinOP einfließen. Die Gewichte der Gruppenmitglieder können dann beispielsweise auf 0,8 aufsummieren, was das Vorwissen mit einem Gewicht von 0,2 mit in den GenLinOP einfließen lässt.

Die Pooling-Methode kann z.B. demokratisiert werden, indem der Entscheider alle Gewichte  $w_i = \frac{1}{n}$  setzt. Damit fließen alle Schätzungen in die Entscheidungsfindung gleichermaßen mit ein. Gegenteilig dazu können alle Gewichte bis auf eines auf Null gesetzt werden, was einer *Diktatur* entspricht. Durch die Wahl der Gewichte können alle Linearkombinationen der Schätzungen erzeugt werden. Die Idee dabei ist, dass Gruppenmitglieder entsprechend ihrer Expertise bezüglich der zu treffenden Entscheidung gewichtet werden [McC81].

Reiffa [Rai70] erklärt anhand eines Beispiels, wie die Gruppenmitglieder agieren, wenn nicht klar ist, wann die Pooling-Methode angewandt wird. Hängen die Gewichte von  $p_i(A)$  ab, d.h. werden diese gewählt, nachdem die Wahrscheinlichkeitsverteilungen der Gruppenmitglieder bekannt sind, dann können die Gruppenmitglieder ihre Wahrscheinlichkeiten so daran anpassen, dass sie möglichst großen Einfluss auf den Gruppenentscheid haben, da der Zusammenhang zwischen den Meinungen und den Gewichten bekannt sein kann [Gen90]. Das heißt, sie können ihre Schätzungen so manipulieren, dass diese ihren Einfluss auf den Gruppenentscheid maximieren. Somit würde die Wahrscheinlichkeitsverteilung im Ergebnis verzerrt.

Wenn das Ergebnis von Gleichung (3.1) eine Wahrscheinlichkeitsverteilung sein soll, müs-

sen die Gewichte die dargestellten Anforderungen erfüllen.

Der GenLinOP muss mit den hier gegebenen Bedingungen nicht immer eine Wahrscheinlichkeitsverteilung ergeben. Je nach Wahl der Gewichte und der Wahrscheinlichkeitsverteilung  $R$  können negative Werte im Ergebnis erscheinen. Dafür hier ein Beispiel:

	<i>Agent</i> <sub>1</sub>	<i>Agent</i> <sub>2</sub>
$z^1$	0.9	0.3
$z^2$	0.1	0.7
$z^3$	0.0	0.0

**Tabelle 3.1.:** Beispiel für Meinungen von zwei Gruppenmitgliedern.

Sei  $R(z^1) = 0, R(z^2) = 0$  und  $R(z^3) = 1$  und die Gewichte seien  $w_1 = -1$  und  $w_2 = 1$ . Zu beachten ist, dass diese Gewichte die Bedingungen des GenLinOP erfüllen. Dann hat der GenLinOP folgende Ergebnisse:  $GenLinOP(z^1) = -0.6$ ,  $GenLinOP(z^2) = 0.6$  und  $GenLinOP(z^3) = 1$ . Es soll hier betont werden, dass die Wahl des Intervalles  $[0, 1]$  als Wertebereich für eine Wahrscheinlichkeitsfunktion willkürlich ist [Gen84a]. Aufgrund dieser Konvention ist das Ergebnis eines GenLinOP nicht immer eine Wahrscheinlichkeitsverteilung. Zum besseren Verständnis soll hier ein Beispiel eines LinOP erläutert werden: Die zwei Gruppenmitglieder in diesem Beispiel betrachten die  $z \in U_i$  in ihrer lokalen Umgebung. Dabei wurden die dieselben beiden Gruppenmitglieder betrachtet wie im Beispiel 2.2. Die Wahrscheinlichkeitsverteilungen der beiden Gruppenmitglieder wurden mit einem LinOP mit den Gewichten  $w_i = \frac{1}{2}$  kombiniert. Abbildung 3.1 zeigt das Ergebnis des LinOP der beiden Gruppenmitglieder und einer Bayes'schen Fusion über alle  $z \in Z$ .

Als Referenz ist in Abbildung 3.1 das Ergebnis einer Bayes'schen Informationsfusion (Bayes) ohne Fehler (Überschätzung durch die Einschränkung auf  $U_i$ ) und Informationsverlust (dadurch, dass  $z^7$  von den Gruppenmitglieder nicht betrachtet wurde) mit eingezeichnet. Die Fehler der Schätzungen der Gruppenmitglieder sind im Ergebnis des LinOP wiederzuerkennen. So wird beispielsweise  $z^4$  von beiden Gruppenmitgliedern als zu wahrscheinlich eingeschätzt. Dies spiegelt sich im Ergebnis des LinOP wieder, hier wird  $z^4$  ebenfalls überschätzt. Zu sehen ist auch, dass  $z^1$  von *Agent*<sub>2</sub> als zu wahrscheinlich eingeschätzt wurde. Im Ergebnis des LinOP ist die Einschätzung jedoch zu niedrig. Dies liegt daran, dass *Agent*<sub>1</sub> das Ereignis  $z^1$  nicht betrachtet. Wegen der lokalen (eingeschränkten) Betrachtung ordnet *Agent*<sub>1</sub> dem Ereignis  $z^1$  eine Wahrscheinlichkeit von Null zu. Dies wird bei der Berechnung des LinOP mit berücksichtigt und zieht damit das Ergebnis des LinOP bei  $z^1$  nach unten.

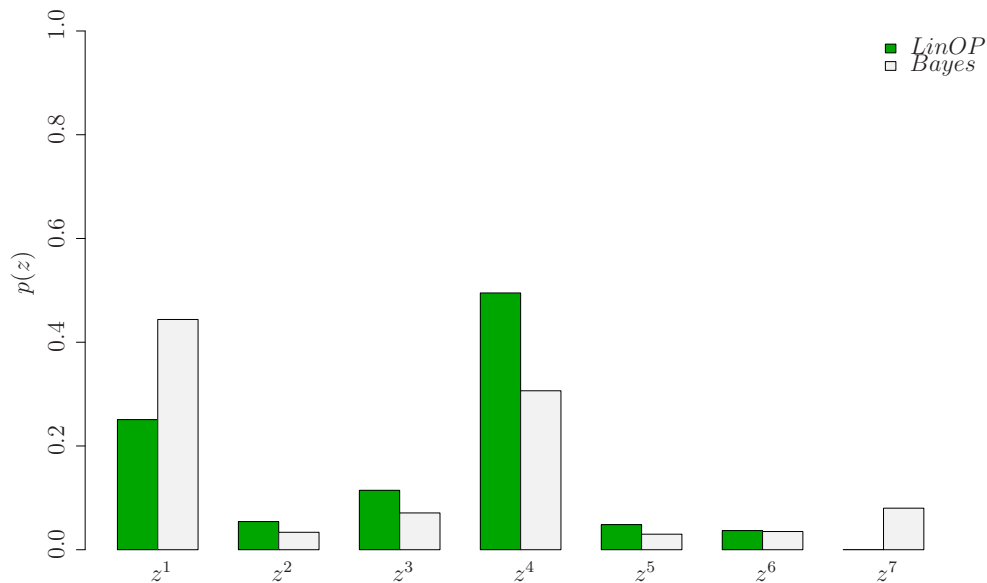


Abbildung 3.1.: Ergebnis eines LinOP

Ein wichtiges Problem des LinOP ist die Wahl der Gewichte  $w_i$ . Diese bestimmen die Zusammensetzung des Gruppenentscheid. Die Wahl der Gewichte beeinflusst die Eigenschaften und das Verhalten eines LinOP. Um zu verstehen, wie die Gewichte in einem Anwendungsfall zu wählen sind, muss eine Interpretation der Gewichtung festgelegt werden. Da der LinOP in Abschnitt 4.3 auf seine Eigenschaften und sein Verhalten überprüft wird, sollen hier zum besseren Verständnis einige Interpretationen der Gewichte des LinOP beschrieben werden.

Es ist nicht immer sinnvoll, dass die Gewichte auf 1 aufsummieren. Nach Winkler [Win68] gilt, dass, wenn alle Experten die gleiche Information haben, dann sollte  $\sum w_i = 1$  sein. Bei vollkommen dekorrelierten Schätzungen der Gruppenmitglieder folgert er, dass die Gewichte  $w_i = 1, i \in \{1, \dots, n\}$  mit der Summe von  $\sum w_i = n$  sein sollten. Dies ist für die lokale Bayes'sche Informationsfusion interessant, da die Gruppenmitglieder nicht die gleichen Ereignisse betrachten.

Winkler [Win68] und Swain [Swa92] fassen vier Möglichkeiten der Wahl der Gewichte zusammen:

- Alle Gruppenmitglieder erhalten das gleiche Gewicht.

- Die Gewichte sind proportional zur Expertise des Gruppenmitglieds.
- Die Gewichte werden durch *scoring rules* bestimmt.
- Jedes Gruppenmitglied gewichtet sich selbst.

### 3.3. Logarithmic Opinion Pool

Eine weitere Pooling-Methode ist der Logarithmic Opinion Pool (LogOP). Im Gegensatz zum LinOP, der ein gewichtetes arithmetisches Mittel darstellt, ist der LogOP ein gewichtetes geometrisches Mittel<sup>2</sup>

$$K(p_1(z), \dots, p_n(z)) = \frac{\prod_{i=1}^n [p_i(z)]^{w_i}}{\sum_{z \in Z} \prod_{i=1}^n [p_i(z)]^{w_i}} \quad (3.3)$$

mit den Gewichten  $w_i \in [0, 1]$  und  $\sum_{i=1}^n w_i = 1$ . Dass die Gewichte positiv sind und sich zu Eins aufsummieren ist nicht zwingende Bedingung [Gen86b]. Hier wird sie jedoch gefordert, um Fälle wie  $w_i = 0, \forall i \in \{1, \dots, n\}$  oder andere Gewichte aus zu schießen, bei denen der Nenner des LogOP Null wird. Zu beachten ist, dass für die Schätzungen  $0 \leq p_i(z) \leq 1$  gilt. Die Gewichte wirken hier exponentiell auf die Meinungen der Gruppenmitglieder ein. Abbildung 3.2 illustriert anhand eines Beispiels, wie sich die Gewichte auswirken. Die Gewichte haben dabei einen größeren Einfluss als bei der multiplikativen Verknüpfung mit der Meinung wie bei einem LinOP. Ein kleineres Gewicht schränkt damit den Einfluss eines Gruppenmitglieds auf den Konsens stärker ein. Abbildung 3.3 zeigt das Resultat der Kombination der lokalen Schätzung beider Agenten aus Beispiel 2.2 mittels eines LogOP. Als Referenz wurde das Ergebnis einer bayes'schen Fusion über ganz  $Z$  mit eingezeichnet. Die Gewichte wurden jeweils auf  $\frac{1}{2}$  gesetzt. Da die Meinungen der Agenten multipliziert werden, stellt eine Schätzung der Form  $p_i(A) = 0$  ein Veto für das Ereignis  $A$  da. Dies ist am Beispiel von  $z^1, z^6$  und  $z^7$  in Abbildung 3.3 zu sehen. Dieser Effekt wird oft als Nachteil des LogOP aufgefasst. Der LogOP kann laut [Win68, Swa92] als *natural-conjugate* interpretiert werden. Dies können wir uns einfach klar machen: Werden natural-conjugate Wahrscheinlichkeitsverteilungen der gleichen Familie multipliziert, gehört das Ergebnis

---

<sup>2</sup>Hier werden die atomaren Ereignisse der zugrundeliegenden  $\sigma$ -Algebra in Bezug auf die PoI betrachtet. Da der LogOP nicht (wie der LinOP) die Marginalisation Property erfüllt, können die nicht-atomaren Ereignisse  $A \in \mathcal{Z}$  nicht direkt betrachtet werden.

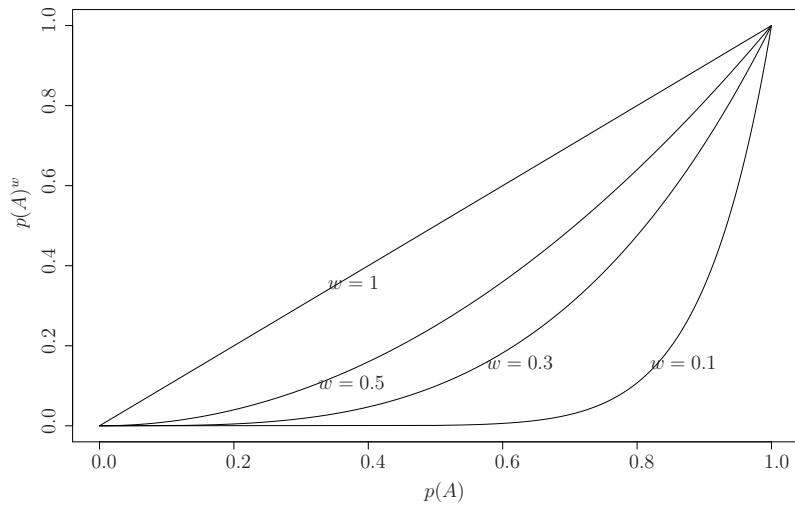


Abbildung 3.2.: Auswirkung der Gewichte bei einem LogOP

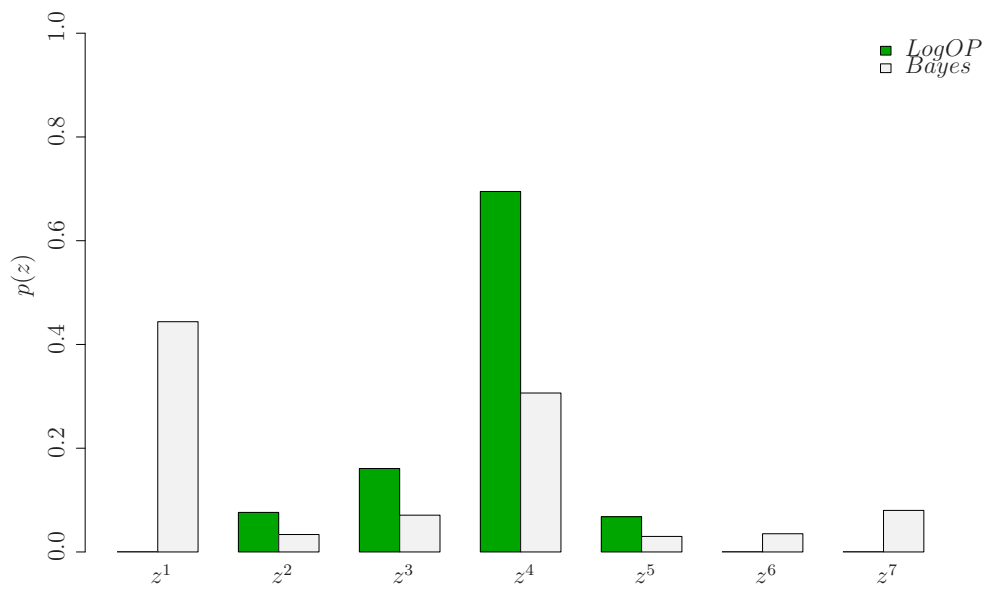
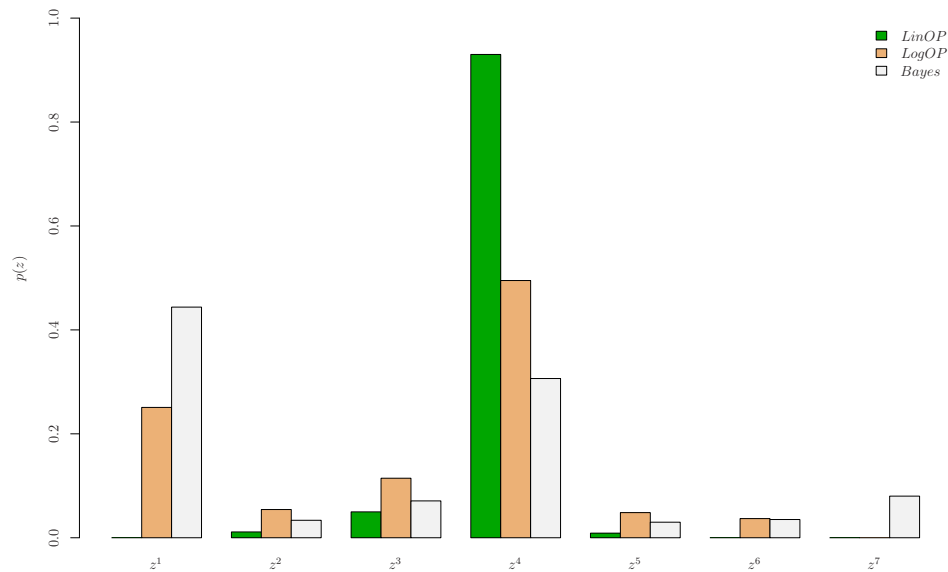


Abbildung 3.3.: Fusionsergebnis der beiden Agenten aus Beispiel 2.2 mittels eines LogOP



**Abbildung 3.4.:** Ein Vergleich zwischen LogOP und LinOP

wieder zu der gleichen Familie von Wahrscheinlichkeitsverteilungen. Gehören die Wahrscheinlichkeitsverteilungen der Gruppenmitglieder alle der gleichen natural-conjugate Familie von Wahrscheinlichkeitsverteilung an, dann ist das Ergebnis eines LogOP auch in der natural-conjugate Familie, da der LogOP die Wahrscheinlichkeitsverteilungen der Gruppenmitglieder multipliziert.

Winkler zeigt in [Win68], dass sich das Ergebnis eines natural-conjugate (hier der LogOP) von dem eines LinOP dadurch unterscheidet, dass die Varianz reduziert wird und dadurch, dass das Ergebnis der gleichen Familie von Wahrscheinlichkeitsverteilungen angehört wie die Schätzungen der Gruppenmitglieder. Ist die Familie der Wahrscheinlichkeitsverteilung der Schätzungen der Gruppenmitglieder beispielsweise unimodal<sup>3</sup>, dann ist das Ergebnis des LogOP auch unimodal.

Werden zum Beispiel  $n$  Normalverteilungen  $\mathcal{N}(\mu_i, \sigma_i)$  mit Mittelwert  $\mu_i$  und Varianz  $\sigma_i$  kombiniert, dann ist der LogOP eine Normalverteilung  $\mathcal{N}(\mu_K, \sigma_K)$  wobei der Mittelwert

---

<sup>3</sup>Unimodal bedeutet für eine Funktion, dass sie bis zu einem Wert streng monoton steigend und ab diesem Wert streng monoton fallend ist. Diese Eigenschaft wird auch eingipflig genannt. Sie gilt nur für Zustandsräume auf denen eine Ordnung definiert ist.



und die Varianz durch

$$\mu_K = \frac{\sum_{i=1}^n w_i^* \mu_i}{\sum_{i=1}^n w_i^*}$$

$$\sigma_K = \left( \sum_{i=1}^n w_i^* \right)^{-1}$$

gegeben sind. Für die Gewichten gilt  $w_i^* = w_i/\sigma_i^2$  [Kas09]. Für weitere Beispiele sei auf [Coo91] verwiesen.

Der Unterschied zwischen einem LinOP und einem LogOP ist in 3.4 illustriert. Sowohl der LinOP als auch der LogOP kombinieren die Ergebnisse der Schätzungen der beiden Agenten aus dem Beispiel 2.2. Ebenso ersichtlich ist, dass beim Ergebnis des LogOP drei Vetos auftreten. Das eine bei  $z^1$ , weil *Agent*<sub>1</sub> dieses nicht betrachtet und eines bei  $z^6$ , da *Agent*<sub>2</sub> dort ein Veto einlegt und ein drittes bei  $z^7$ , das gemeinschaftlich von beiden Agenten gehalten wird. Beim Ergebnis des LinOP bleibt nur das gemeinschaftlich gehaltene Veto<sup>4</sup> erhalten.

### 3.4. Supra-Bayes Ansatz

Der Supra-Bayes'sche (SB) Ansatz wurde von Morris [Mor77] vorgestellt. Dabei werden die Meinungen der Gruppenmitglieder als Beobachtungen aufgefasst. Dies kann wie folgt interpretiert werden.

Der Entscheider erstellt einen Likelihood  $p_{n+1}(p_1(z), \dots, p_n(z), d)$  für das Ereignis  $z$  aus den Meinungen der Gruppenmitglieder. Er kann nun mittels der bayes'schen Formel sein Vorwissen mit dem Likelihood zu einer A-posteriori-Wahrscheinlichkeitsverteilung fusionieren. Die Pooling-Methode nimmt dann folgende Form an:

$$\begin{aligned} SB(p_1, \dots, p_n)(z|d) &= p_{n+1}(z|p_1(z|d), \dots, p_n(z|d), d) \\ &= \frac{p_{n+1}(p_1(z|d), \dots, p_n(z|d), d|z)p_{n+1}(z|d)}{p_{n+1}(p_1(z|d), \dots, p_n(z|d), d)} \\ &\propto p_{n+1}(p_1(z|d), \dots, p_n(z|d), d|z)p_{n+1}(z|d) \end{aligned} \tag{3.4}$$

Hier sieht der Entscheider zusätzlich zu den Wahrscheinlichkeitsverteilungen der Grup-

<sup>4</sup>Diese Eigenschaft wird *Zero Preservation Property* genannt.

penmitglieder die Beobachtungen  $d$ . Dies ermöglicht, dass eines der Gruppenmitglieder die Rolle des Entscheiders übernimmt. Durch das Erstellen des Likelihood ordnet der Entscheider (das  $n + 1$ -te Gruppenmitglied) jedem der  $n$  Gruppenmitglieder implizit eine Glaubwürdigkeit zu.

Durch die A-priori-Wahrscheinlichkeitsverteilung und durch die implizite Gewichtung der  $n$  Gruppenmitglieder kann Vorwissen mit in die Pooling-Methode einfließen. Die verwendeten Daten (die Schätzungen der Gruppenmitglieder) sind mit Unsicherheit behaftet. Die Unsicherheit entspricht hier dem Fehler der Schätzungen der Gruppenmitglieder.

Nach Givens et al. [Giv99] und Jacobs [Jac95] ist das erstellen des Likelihoods durch den Entscheider die größte Schwierigkeit des SB-Ansatzes. Der Entscheider muss eine Likelihood der Meinungen der Gruppenmitglieder bezüglich des Ereignisses  $z$  erstellen. Der zu erstellende Likelihood sollte die Korrelation der Schätzungen und die Expertise der Gruppenmitglieder beschreiben. Dies ist ähnlich schwierig wie die Wahl der Gewichte  $w_i$  für die Pooling-Methoden LinOP und LogOP. Der Entscheider kann zusätzlich durch die Wahl seines Vorwissens in Form der A-priori-Wahrscheinlichkeitsverteilung  $p_{n+1}(z)$  den Gruppenentscheid beeinflussen. Kleine Fehler in der Wahl dieser A-priori-Wahrscheinlichkeitsverteilungen und des Likelihood haben großen Einfluss auf die Wahrscheinlichkeitsverteilung des Gruppenentscheids [Giv99].

Sind die Schätzungen der Gruppenmitglieder bedingt unabhängig bezüglich  $z$ , so kann der Likelihood nach [DeG91, Gel95] wie folgt erzeugt werden:

$$p_{n+1}(p_1(z|d), \dots, p_n(z|d), d|z) = p_{n+1}(p_1(z|d), \dots, p_n(z|d)|z)$$

In diesem Fall vereinfacht sich der SB-Ansatz auf:

$$SB(p_1, \dots, p_n)(z) = \frac{p_{n+1}(p_1(z|d), \dots, p_n(z|d)|z)p_{n+1}(z)}{p_{n+1}(p_1(z|d), \dots, p_n(z|d))}$$

## 4. Eigenschaften von Pooling Methoden

### 4.1. Allgemeine Eigenschaften

McConway [McC81] und Genest [Gen86b] beschreiben einen axiomatischen Ansatz, bei dem eine passende Pooling Methode ausgewählt wird, indem bestimmt wird, welche Eigenschaften von der Pooling-Methode gewünscht werden. Manche der Eigenschaften tragen zur Interpretierbarkeit der Ergebnisse bei, andere beschreiben, wie Gewichte gewählt werden müssen, damit das Ergebnis sinnvoll bleibt. Dass die Eigenschaften nicht frei wählbar sind und sich manchmal gegenseitig ausschliessen, wird anhand von Beschränkungssätzen gezeigt.

Alle folgenden Aussagen sind unter den Annahme, dass  $\mathcal{Z}$  mindestens zwei nicht leere und disjunkte Elemente enthält, deren Vereinigung eine echte Teilmenge von  $Z$  ist [Wag82]. Die  $\sigma$ -Algebra  $\mathcal{Z}$  heisst dann *tertiär*. Die erste betrachtete Eigenschaft ist die von McConway in [McC81] vorgestellte *Marginalization Property*:

**Marginalization Property (MP):**

$$K(p_1(A), \dots, p_n(A))^{\mathcal{U}} = K(p_1(A)^{\mathcal{U}}, \dots, p_n(A)^{\mathcal{U}}). \quad (4.1)$$

Die linke Seite dieser Gleichung entspricht der Marginalisierung des Gruppenentscheids auf die Sub- $\sigma$ -Algebra  $\mathcal{U}$  von  $\mathcal{Z}$ . Die rechte Seite entspricht dem Gruppenentscheid der auf  $\mathcal{U}$  marginalisierten Wahrscheinlichkeiten. Die MP fordert von der Pooling-Methode, dass sie mit der Marginalisierung beliebig in der Reihenfolge der Anwendung vertauschbar sein muss. Während man auf der linken Seite der Gleichung noch alle  $A \in \mathcal{Z}$  betrachtet und danach auf  $\mathcal{U}$  marginalisiert, kann man auf der rechten Seite nur die  $A \in \mathcal{U}$  betrachten. Die rechte Seite impliziert eine Reduktion des Rechenaufwandes, da davon ausgegangen werden kann, dass  $|\mathcal{U}| \ll |\mathcal{Z}|$ .

Details der Marginalisierung werden im Abschnitt 2.4 besprochen.

Um dies auf ein Beispiel aus der Kriminalistik zurückzuführen, kann die MP wie folgt beschrieben werden: Bei einer Ermittlung kann der Degree of Belief (DoB) für die Schuld eines Individuums  $z$  durch das Sammeln von Beweisen erhöht oder verringert werden. Am Ende einer Ermittlung sollte der DoB für die Schuld des Täters möglichst nahe an 1 liegen, damit es zu einer Verurteilung kommt. Soll für den Gruppenentscheid der Richter oder der Jury die MP gelten, dann gilt: Zum Zeitpunkt der Verurteilung darf es keinen Unterschied mehr machen, ob das Urteil für jedes atomare Element von  $\mathcal{Z}$  getroffen wurde und danach diese zu einer sub- $\sigma$ -Algebra marginalisiert wurden oder umgekehrt. Angenommen in einem Mordprozess gibt es eine Gruppe von Verdächtigen  $Z = \{Eve, Alice, Bob\}$ . Von einem Ermittler wurden entlastende Beweise für Eve gefunden. Bei dem Prozess genügt es dem Anwalt von Eve, zu beweisen, dass Alice oder Bob schuldig sind. Dabei ist es egal, welcher von beiden tatsächlich der Schuldige ist. Was zählt ist nur, dass Eve unschuldig ist. Dies ist eine Marginalisierung auf die  $\sigma$ -Algebra erzeugt von  $\{Alice \vee Bob\}$ . Gilt die MP dann ist es egal, ob die Ermittler Beweise gegen Alice oder gegen Bob finden oder Beweise gegen beide.

Eine Pooling-Methode, die nicht die MP erfüllt, hängt also von den gesamten Schätzungen  $p_i(A), \forall A \in \mathcal{Z}$  und von  $\mathcal{Z}$  ab und dadurch auch von weiteren Ereignissen zusätzlich zu  $A$  [McC81]. Eine wünschenswerte Eigenschaft wäre, dass der Gruppenentscheid nur von  $A \in \mathcal{Z}$  und  $p_i(A), A \in \mathcal{U}$  abhängt. Dies Forderung wird in der *Weak Setwise Function Property* formuliert [McC81, Gen84d].

#### **Weak Setwise Function Property (WSFP):**

Es gibt eine Funktion  $F$  so dass gilt:

$$K(p_1(A), \dots, p_n(A)) = F(A, p_1(A), \dots, p_n(A)), \forall A \in \mathcal{U}.$$

Die WSFP impliziert, dass der Gruppenentscheid der Schätzungen  $p_i(A)$  nur von dem Wert  $A \in \mathcal{Z}$  und den Schätzungen  $p_i(A)$  abhängt. Dies kommt der Anforderung an eine Jury gleich, unvoreingenommen jeden Angeklagten gerecht<sup>1</sup> zu behandeln. Genauer bedeutet das, dass keine zusätzliche Information mit in die Meinungsbildung einfließen darf. Der endgültige Entschluss einer Jury sollte also nur von der Meinung der einzelnen Jury-

---

<sup>1</sup>Gerecht soll nicht nur heißen gleich, sondern auch dem Fall angepasst.

mitglieder und dem zu beurteilenden Fall abhängen. Dies ist nicht zu verwechseln mit der Bildung der Einzelmeinung der Mitglieder der Gruppe. Dabei ist es unmöglich und auch nicht wünschenswert, dass die Meinungen und das Vorwissen irgendeines Gruppenmitglieds nicht in seine Meinungsbildung mit einfließen.

Pooling-Methoden, die die MP erfüllen, erfüllen auch WSFP. Der GenLinOP (siehe (3.2)) erfüllt also MP und WSFP. Um das Verständnis dieser Eigenschaften zu fördern, wird nun gezeigt, dass  $MP \Leftrightarrow WSFP$  ist.

**Satz 1** *Die MP ist äquivalent zur WSFP [McC81].*

Beweis: Für den Beweis muss gezeigt werden, dass gilt:

- i  $WSFP \Rightarrow MP$
- ii  $WSFP \Leftarrow MP$

Hier ist  $U$  die Teilmenge von  $Z$ , auf die sich in der lokalen Fusion konzentriert wird. Zu i: Erfüllt eine Pooling-Methode die Eigenschaft WSFP, ist die Menge der nicht betrachteten Ereignisse  $\bar{U}$  für den Gruppenentscheid bedeutungslos. Spezieller könnte die Konsensbildung äquivalent bezüglich jeder sub- $\sigma$ -Algebra  $\mathcal{U} \subseteq \mathcal{Z}$  durchgeführt werden, solange das Ereignis von Interesse darin enthalten ist. Klar wird dadurch, dass eine Funktion, die WSFP erfüllt auch MP erfüllt. Es gilt also  $WSFP \Rightarrow MP$ .

Zu ii: Angenommen, die Pooling-Methode  $K$  erfüllt MP und  $U \subseteq Z$ . Dann gibt es drei Fälle, die betrachtet werden müssen:

- $U = Z$ : In diesem Fall sind alle Elemente von  $Z$  auch in  $U$  enthalten. Es gibt keine Elemente innerhalb  $\bar{U}$ , die nicht relevant für den Gruppenentscheid  $K$  sein könnten. Damit hängt  $K$  nur von  $A \in \mathcal{U}$  und den Wahrscheinlichkeiten  $p(A)$ ,  $A \in \mathcal{U}$  ab.
- $U = \emptyset$ : Wird kein Ereignis  $A \in \mathcal{U}$  betrachtet, so kann der Gruppenentscheid auch nicht davon abhängen. Diese beiden Punkte sind verständlich, wenn wir uns klar machen, dass eine Pooling-Methode auf einer Wahrscheinlichkeitsverteilung operiert und dass ihr Ergebnis wieder eine Wahrscheinlichkeitsverteilung ist.
- $U \subset Z$ : Ist  $U$  eine echte Teilmenge von  $Z$ , dann kann eine  $\sigma$ -Algebra  $\mathcal{U} = \{\emptyset, U, \bar{U}, Z\}$  definiert werden. Diese ist die kleinste  $\sigma$ -Algebra, in der  $U$  enthalten ist. Damit enthält jede  $\sigma$ -Algebra  $\mathcal{Z}^*$ , die  $U$  enthält, auch  $\mathcal{U}$ . Zu beachten ist, dass  $\mathcal{U}$  alleine von  $U$  definiert wird. Siehe  $\sigma$ -Operator in Appendix A. Spezieller ist jede Wahrscheinlich-

keitsverteilung über  $\mathcal{U}$  durch  $p_i(A), A \in \mathcal{U}$  gegeben. Da  $K$  die MP erfüllt, gilt nach Gleichung (4.1):

$$K(p_1^{\mathcal{Z}^*}, \dots, p_n^{\mathcal{Z}^*})(A) = K(p_1^{\mathcal{Z}^*}, \dots, p_n^{\mathcal{Z}^*})(A), \quad (4.2)$$

für alle Sub- $\sigma$ -Algebren  $\mathcal{Z}^*$  von  $\mathcal{Z}$  mit  $A \in \mathcal{Z}^*$ . Auf der linken Seite der Gleichung 4.2 findet die Marginalisierung auf  $\mathcal{Z}^*$  erst statt, nachdem die Konsensverteilung über ganz  $\mathcal{Z}$  berechnet wurde. Da nur die  $A \in \mathcal{Z}^*$  betrachtet werden, gilt:

$$K(p_1, \dots, p_n)(A) = K(p_1^{\mathcal{Z}^*}, \dots, p_n^{\mathcal{Z}^*})(A), \quad \forall A \in \mathcal{Z}^*.$$

Spezieller gilt dies auch für die kleinste  $\sigma$ -Algebra  $\mathcal{U}$ , die  $U$  enthält:

$$K(p_1, \dots, p_n)(A) = K(p_1^{\mathcal{U}}, \dots, p_n^{\mathcal{U}})(A), \quad \forall A \in \mathcal{U}.$$

Die rechte Seite hängt nur noch von  $A \in \mathcal{U}$  und  $p_1(A)^{\mathcal{U}}, \dots, p_n(A)^{\mathcal{U}}$  ab, womit der Beweis für WSFP  $\Rightarrow$  MP vollständig ist.  $\square$

Damit ist gezeigt, dass MP äquivalent ist zu WSFP.

Die Eigenschaft (MP oder WSFP) kann noch verstärkt werden. Wird gefordert, dass der Gruppenentscheid nur noch von den Schätzungen  $p_i(A)$  abhängt und nicht mehr von dem betrachteten Ereignis  $A$ , dann heißt die Eigenschaft *Strong Setwise Function Property*:

### Strong Setwise Function Property (SSFP):

Es existiert eine Funktion  $F : [0, 1] \times \dots \times [0, 1] \rightarrow [0, 1]$ :

$$K(p_1, \dots, p_n)(A) = F(p_1(A), \dots, p_n(A)), \quad \forall A \in \mathcal{Z}. \quad (4.3)$$

Ersichtlich ist, dass die Pooling-Methode nicht von  $A$  abhängen darf, sondern nur noch von den Schätzungen  $p_i(A)$ . Dies könnte mit einem Experten verglichen werden, der Spuren wie Fingerabdrücke oder DNA-Spuren eines Tatortes untersucht. Dabei sollte er sich nur auf diese Information verlassen, losgelöst von dem Fall, um den es sich handelt.

Eine Pooling-Methode die SSFP erfüllt, ist nach [McC81] ein LinOP. Dies wird genauer in Satz 4 erläutert.

Die Forderung der SSFP ist sehr restriktiv, sie lässt zum Beispiel nicht die gesamte Klasse

von geometrischen Mitteln wie

$$K(p_1, \dots, p_n)(A) \propto \prod_{i=1}^n p_i(A)^{w_i},$$

zu [Gen87]. Genest [Gen86a, Gen84c] schlägt eine gelockerte SSFP vor, die geometrische Mittel zulässt. Dazu ist folgende Erweiterung nötig: Das Gleichheitszeichen in Gleichung (4.3) wird durch ein Proportional-Zeichen ersetzt. Pennock et al. [Pen05] benennt diese schwächere Form der SSFP zur *Proportional Dependence on States (PDS)* um:

$$K(p_1, \dots, p_n)(A) \propto F(p_1(A), \dots, p_n(A))$$

Die Proportionalitätskonstante hängt dabei nicht von den Elementen von  $\mathcal{Z}$  ab, aber von den Wahrscheinlichkeiten der atomaren Elemente aus  $\mathcal{Z}$ . Die PDS wird von Pennock et al. als eine „Unabhängigkeit von irrelevanten Ereignissen“ beschrieben. Pooling-Methoden von der Form der Gleichung (3.3) erfüllen z.B. die PDS, aber nicht die SSFP [Gen84c].

Mit den bisher eingeführten Eigenschaften werden die Formen von Pooling-Methoden charakterisiert. Es wird jedoch nicht ausgesagt, wie die betrachteten Wahrscheinlichkeitsverteilungen  $p_i(z)$  mit in den Gruppenentscheid einfließen. Wie der Gruppenentscheid von den Wahrscheinlichkeitsverteilungen  $p_i(z)$  beeinflusst werden kann, legen die folgenden Eigenschaften fest.

### Zero Preservation Property (ZPP):

$$p_1(A) = \dots = p_n(A) = 0 \Rightarrow K(p_1, \dots, p_n)(A) = 0$$

Wenn alle Gruppenmitglieder der Meinung sind, dass die Auftrittswahrscheinlichkeit des Ereignisses  $A \in \mathcal{Z}$  gleich Null ist, dann sollte die Pooling-Methode dies auch widerspiegeln [Ber86]. Nach Genest [Gen84d] wird diese Eigenschaft *Zero Preservation Property* genannt. (McConway nennt sie in [McC81] „Zero Probability Property“.)

In der Kriminalistik lässt sich die ZPP folgendermaßen interpretieren: Liefert kein Ermittler Beweise für die Schuld und liefert kein Vorwissen einen Grund für die Annahme der Schuld des Angeklagten, dann sollte dieser nicht verurteilt werden, da der DoB für die Schuld dieses Angeklagten nahe bei Null ist. Die ZPP ergänzt die WSFP so, dass die SSFP genau dann für eine Pooling-Methode gilt, wenn auch die WSFP und die ZPP gelten. Die

ZPP beschreibt einen Spezialfall der *Unanimity*.

#### **Unanimity (UNAM):**

Pennock et al. formulieren die UNAM in [Pen05] wie folgt:

$$p_1(A) = \dots = p_n(A) \Rightarrow K(p_1, \dots, p_n)(A) = p_1(A), \forall A \in \mathcal{Z}.$$

Wenn alle Gruppenmitglieder die gleiche Wahrscheinlichkeitsverteilung über  $\mathcal{Z}$  erzeugen, muss die Pooling-Methode die korrespondierende eindeutige Wahrscheinlichkeitsverteilung übernehmen. Bei einer Jury eines Bundesgerichts der Vereinigten Staaten muss z.B. die Verurteilung einstimmig erfolgen. Da Jurymitglieder nur zwischen Schuld ( $p(\text{Schuld}) = 1$ ) und Unschuld ( $p(\text{Schuld}) = 0$ ) entscheiden können, muss die UNAM Eigenschaft erfüllt sein. Verallgemeinert kann die Einstimmigkeit (deutsch für Unanimity) wie folgt formuliert werden [Sei09] :

$$c_1 \leq p_i(A) \leq c_2, \forall i \in \{1, \dots, n\} \Rightarrow c_1 \leq K(p_1, \dots, p_n)(A) \leq c_2, \forall A \in \mathcal{Z}$$

Die von Pennock et al. beschriebene Einstimmigkeit ist ein Spezialfall mit  $c_1 = c_2$ . Meist wird auch gefordert, dass ein Jurymitglied nicht den Entscheider so beeinflussen kann, dass er seine Meinung vollständig übernimmt. Diese Forderung wird in der *Non Dictatorship* Eigenschaft formuliert.

#### **Non Dictatorship (ND):**

ND bedeutet, dass die Pooling-Methode keine Diktatur sein darf. Eine Diktatur ist, wenn alle Schätzungen der Gruppenmitglieder bis auf eine ignoriert werden. Dies wäre bei einem externen Entscheider, z.B. einem Richter der Fall, der eine Gruppe von Experten einsetzt, bei seiner Meinungsbildung jedoch alle bis auf einen Experten ignoriert. Pennock et al. [Pen05] beschreiben dies wie folgt: „Mit der UNAM wird sichergestellt, dass ein Multi-Agenten-Problem nicht auf ein Ein-Agenten-Problem reduziert wird.“



**Independence Preservation Property (IPP):**

Laddaga formuliert in [Lad77] zunächst die IPP. Pooling-Methoden, welche die IPP erfüllen, erhalten die von allen Gruppenmitgliedern einstimmig gehaltenen Unabhängigkeiten von Ereignissen [Pen05]. Zwei Ereignisse  $A, B$  sind unabhängig, wenn alle Gruppenmitglieder diese als unabhängig betrachten, so dass für alle  $i \in \{1, \dots, n\}$  gilt:  $p_i(A \cap B) = p_i(A)p_i(B)$ . Dann ist diese Unabhängigkeit einstimmig. Wird diese Unabhängigkeit erhalten, dann gilt nach [Gen87] für die Pooling-Methode:

$$K(p_1(A \cap B), \dots, p_n(A \cap B)) = K(p_1, \dots, p_n)(A)K(p_1(B), \dots, p_n(B)),$$

Sind alle Gruppenmitglieder der Meinung, dass das Eintreten des einen Ereignisses nicht vom Eintreten des anderen beeinflusst wird und umgekehrt, dann fordert die IPP, dass auch die Konsensverteilung diese Unabhängigkeiten repräsentiert. Die IPP ist eine Eigenschaft, die aussagt, wie gut ein Konsens in Übereinstimmung mit den Wahrscheinlichkeitsverteilungen steht, aus denen er erzeugt wurde. Genest et al. argumentieren in [Gen87] gegen die Forderung der IPP, indem sie zeigen, dass IPP und WSFP gleichzeitig nur von einer Diktatur erfüllt werden können, wenn  $Z$  mehr als vier Elemente enthält. In [Pen05] wird die Aussage auf die PDS erweitert. Diese Erkenntnis wird von Wagner das „Unmöglichkeitstheorem“ genannt. Genest et al. [Gen87] diskutieren, dass die Forderung der IPP in vielen praktischen Anwendungen nicht sinnvoll ist. Dies ist beispielsweise der Fall, wenn für einen externen Entscheider  $n$  Experten eine Schätzung über die Wahrscheinlichkeitsverteilungen zweier Ereignisse  $A, B$  abgeben. Dann wird durch die IPP im Fall der einstimmigen Unabhängigkeit induziert, dass die zusammengeführte Schätzung der  $n$  Experten bezüglich der Wahrscheinlichkeit des Auftretens eines Ereignisses  $B$  sich für den Entscheider nichts ändert, wenn  $A$  beobachtet wurde und umgekehrt. Beurteilt der Entscheider die Experten jedoch nach ihrer Leistung bei vorherigen Schätzungen, dann verändert er unter Umständen die Wahrscheinlichkeit für  $B$ , wenn  $A$  beobachtet wurde. Daraus folgt, dass für Pooling-Methoden, die die Qualität von Experten iterativ berücksichtigt, die Gewichte nur für abhängige Ereignisse verändert werden dürften, da sonst die IPP nicht mehr erfüllt wäre. Pennock et al. definieren in [Pen05] eine *Event Independence Preservation Property*, bei der nicht wie bei der IPP alle Unabhängigkeiten erhalten bleiben müssen.

**Event Independence Preservation Property (EIPP):**

Die EIPP ist für eine Pooling-Methode erfüllt, wenn für alle Elemente  $A, B$  der  $\sigma$ -Algebra  $\mathcal{Z}$  gilt:

$$p_i(A|B) = p_i(A), i \in \{1, \dots, n\} \Rightarrow K(p_1(A|B), \dots, p_n(A|B)) = K(p_1, \dots, p_n)(A).$$

Wenn zwei Ereignisse  $A, B \in \mathcal{U}$  für den Experten  $i$  unabhängig von einander sind, gilt  $p_i(A|B) = p_i(A)$ ,  $i \in \{1, \dots, n\}$ . Die Ereignisse, die in  $\mathcal{U}$  enthalten sind, werden von Pennock et al. [Pen05] *primär* genannt. Erfüllt eine Pooling-Methode die EIPP, dann erhält sie die Unabhängigkeit zwischen Elementen von Interesse aus  $\mathcal{U}$ . Es müssen nicht alle Unabhängigkeiten erhalten werden, sondern nur die direkten zwischen den betrachteten Elementen von Interesse. Ein Beispiel aus [Pen05] soll verdeutlichen, was Elemente von Interesse sind: Sei  $Z = \{z_1, \dots, z_7\}$ , dann kann die  $\sigma$ -Algebra  $\mathcal{Z}$  über  $Z$  die Elemente  $A = z_3 \wedge \bar{z}_7$  und  $B = z_4 \wedge \bar{z}_2 \vee z_7$  enthalten. Diese Art von Unabhängigkeit zwischen  $A$  und  $B$  wird von der EIPP nicht erhalten, ausser es handelt sich dabei um die Ereignisse von Interesse.

Pennock et al. zeigen, dass ein LinOP die EIPP nicht erfüllt. Sie räumen jedoch die Möglichkeit ein, dass die EIPP und die PDS von einer Pooling-Methode erfüllt werden kann, die keine Diktatur ist.

**External Bayesianity (EB):**

Madansky, Genest et al. und Pennock et al. [Mad64, Gen86a, Pen05] definieren eine Pooling-Methode dann als EB, wenn sie folgende Bedingung erfüllt:

$$K \left( \frac{p(d|z)p_1(z)}{\sum_{z \in Z} p(d|z)p_1(z)}, \dots, \frac{p(d|z)p_n(z)}{\sum_{z \in Z} p(d|z)p_n(z)} \right) = \frac{p(d|z)K(p_1, \dots, p_n)(z)}{\sum_{z \in Z} p(d|z)K(p_1, \dots, p_n)(z)}.$$

Dabei ist  $d$  neue Information, die der Gruppe zur Verfügung gestellt wird. Der Likelihood  $p(d|z)$  ist der gemeinsame Likelihood der Gruppe bezüglich der Beobachtungen  $d$ . Dieser Likelihood  $p(d|z)$  muss so gewählt werden, dass für alle  $i \in \{1, \dots, n\}$  gilt  $0 < \sum p(d|z)p_i(z) < \infty$  [Gen86a].

EB beschreibt eine Eigenschaft in Bezug auf die Kommutativität von Pooling-Methoden

und die Miteinbeziehung neuer Information mittels bayes'scher Fusion. Wenn die gemeinsame neue Information  $d$  in Form des Likelihood  $p(d|z)$  verfügbar ist, kann die Konsensverteilung mittels der Bayes'scher Formel aktualisiert werden. Mit dem gleichen Ergebnis können die einzelnen Gruppenmitglieder vor Berechnung der Konsensverteilung ihre Schätzungen individuell aktualisieren. Wird die EB erfüllt, so ist die Reihenfolge von Aktualisierung und Anwendung der Pooling-Methode egal. Externe Bayes'sche Pooling-Methoden sind eine wichtige Klasse von Pooling-Methoden, die z.B. den Logarithmischen Opinion Pool enthält.

Der LogOP erfüllt - im Gegensatz zum LinOP - die EB. Raiffa [Rai70] beschreibt anhand eines Beispiels, wie eine Gruppe agieren kann, wenn die Pooling-Methode EB nicht erfüllt. In diesem Beispiel lässt er die Gruppenmitglieder entweder einen Kompromiss a-priori oder a-posteriori durchführen. Sein Beispiel kombiniert die Wahrscheinlichkeitsverteilungen der Gruppenmitglieder mit einem LinOP. Die Ergebnisse der beiden Kompromisse sind unterschiedlich, da der LinOP die EB nicht erfüllt. Raiffa kommt zu dem Schluss, dass in solchen Fällen die Gruppe bessere a-priori die Wahrscheinlichkeitsverteilungen kombiniert.

Eine Pooling-Methode, die EB erfüllt, ist nicht gleich einem SB Ansatz, da eine Supra-Bayes'sche Pooling-Methode nicht die EB erfüllen muss [Jac95]. Bei einer Pooling-Methode, welche die EB erfüllt, ist der Gruppenentschluss für einen Außenstehenden ein bayes'scher Entschluss, unter der Bedingung neuer Beobachtungen.

### Relative Propensity Consistency (RPC):

Vorgestellt von Genest, Weerahandi und Zidek in [Gen84a] bedeutet die RPC, dass die Propensität<sup>2</sup> der Verhältnisse von Schätzungen unter der Pooling-Methode erhalten bleiben. Dafür werden für  $z^1, z^2, z^3, z^4 \in Z$  zwei Funktionen definiert:

$$RP(z^1, z^2) = \left( \frac{p_1(z^1)}{p_1(z^2)}, \dots, \frac{p_n(z^1)}{p_n(z^2)} \right)$$

und

$$H(z^1, z^2) = \frac{K(p_1, \dots, p_n)(z^1)}{K(p_1, \dots, p_n)(z^2)}$$

<sup>2</sup>Mit der Propensität wird hier die Tatsache bezeichnet, dass das Verhältnis nicht umschlägt. Ist beispielsweise ein Verhältnis größer Eins, dann kann es sich verändern aber, nicht kleiner Eins werden.

Die Pooling-Methode  $K$  erfüllt die RPC, wenn für beliebige Ereignisse  $z^1, z^2, z^3, z^4 \in Z$  gilt:

$$RP(z^1, z^2) \geq RP(z^3, z^4) \Rightarrow H(z^1, z^2) \geq H(z^3, z^4).$$

Dabei ist  $RP(z^1, z^2) \geq RP(z^3, z^4)$  per Definition dann erfüllt, wenn  $\frac{p_i(z^1)}{p_i(z^2)} \geq \frac{p_i(z^3)}{p_i(z^4)}$  für alle  $i \in \{1, \dots, n\}$  gilt. Halten alle Gruppenmitglieder ein beliebiges Ereignis  $z^1$  für mindestens so wahrscheinlich wie ein anderes Ereignis  $z^2$ , dann fordert die RPC insbesondere, dass  $z^1$  auch nach der Anwendung der Pooling-Methode mindestens so wahrscheinlich wie  $z^2$  bleibt.

Im folgenden Abschnitt werden Beziehungen zwischen Eigenschaften von Pooling-Methoden betrachtet. Diese sind bei der Wahl einer geeigneten Pooling-Methode hilfreich.

## 4.2. Beschränkungssätze

Nicht alle erstrebenswerten Eigenschaften einer Pooling-Methode können in der Regel gleichzeitig erfüllt werden. Wie in [Pen05, McC81, Gen84d] beschrieben, lassen sich Beschränkungssätze herleiten.

Dies Sätze beschränken die möglichen Pooling-Methoden auf eine bestimmte Klasse von Pooling-Methoden, welche die geforderten Eigenschaften erfüllt.

Genest zeigt in [Gen84d] folgenden Satz für Pooling-Methoden, welche die MP erfüllen:

**Satz 2** *Ist  $Z$  tertiär und genügt die Pooling-Methode der MP, so handelt es sich um einen GenLinOP wie bei Gleichung (3.2) beschrieben.*

Sind bei einem GenLinOP alle Gewichte Null, dann ist  $GenLinOP(p_1(z), \dots, p_n(z)) = R(z)$  und damit unabhängig von den Schätzungen der Gruppenmitglieder. Sind die Gewichte nicht negativ und gilt  $\sum_{i=1}^n w_i = 1$ , so reduziert sich der GenLinOP auf den LinOP (3.1) [Gen84d]. Wird also die MP gefordert, dann ist die Pooling-Methode praktisch immer von der Form (3.2). Genauso impliziert die RPC praktisch immer einen LogOP. Dies zeigt im wesentlichen der folgende Satz [Gen84a]:

**Satz 3** *Ist  $Z$  tertiär und gilt die RPC, dann ist  $K(p_1, \dots, p_n)(z)$ ,  $z \in Z$  von folgender Form:*

$$K(p_1, \dots, p_n)(z) = Q(p_1(z), \dots, p_n(z))c(z) \prod_{i=1}^n p_i(z)^{w_i},$$

mit Gewichten  $w_i \geq 0$ .  $Q$  und  $c$  sind positive Konstanten.  $Q$  hängt von den Schätzungen der Experten und  $c$  hängt vom betrachteten Ereignis  $z$  ab.

Der Fall, bei dem die Wahrscheinlichkeiten  $p_i(z)$  den Wert Null annehmen können, wurde von Genest ausgeschlossen. Genest geht davon aus, dass die Meinungen  $p_i(z)$  der Gruppenmitglieder größer Null sind.

Enthält  $\mathcal{Z}$  genau zwei disjunkte Elemente, dann wird die RPC von einem LinOP erfüllt [Gen84a].

Genest et al. [Gen84a] stellen fest, dass eine Pooling-Methode, welche die RPC erfüllt, auch das *Dominanz-Prinzip von Schmitt*<sup>3</sup> erfüllt.

Die SSFP impliziert praktisch immer einen LinOP, wie der folgende Satz aus [McC81] zeigt:

**Satz 4** *Ist  $\mathcal{Z}$  tertiär und gilt die SSFP, dann ist die Pooling-Methode ein LinOP.*

Bei Satz 4 ist zu beachten, dass eine Pooling-Methode eine Wahrscheinlichkeitsverteilung als Ergebnis liefern muss. Ist dies nicht der Fall, können auch andere Pooling-Methoden die SSFP erfüllen.

Werden von einer Pooling-Methode MP und ZPP gefordert, so ist dies äquivalent dazu, die SSFP zu fordern [McC81]. Aus Satz 4 lässt sich folgern, dass die möglicherweise vorhandenen Gewichte der Pooling-Methode unabhängig von den betrachteten Ereignissen sein müssen [Gen84d, McC81]. Dies ist leicht einzusehen, da eine Pooling-Methode, welche die SSFP erfüllt, nur von den Schätzungen der Gruppenmitglieder bezüglich des betrachteten Ereignisses abhängt, nicht aber von dem betrachteten Ereignis selbst.

Die EB charakterisiert nach [Gen84b] einen LogOP.

**Satz 5** *Ist  $\mathcal{Z}$  tertiär und  $Z$  abzählbar und erfüllt die Pooling-Methode EB und eine zur PDS analoge Variante der WSFP, dann ist sie von der Form:*

$$K(p_1, \dots, p_n)(z) = \frac{g(z) \prod_{i=1}^n p_i(z)^{w_i}}{\sum_{z \in Z} g(z) \prod_{i=1}^n p_i(z)^{w_i}}$$

<sup>3</sup>Das Dominanz-Prinzip von Schmitt verlangt, dass  $K(p_1, \dots, p_n)(z^1) \geq K(p_1, \dots, p_n)(z^2)$  wenn  $p_i(z^1) \geq p_i(z^2)$ ,  $\forall i \in \{1, \dots, n\}$

#### 4. Eigenschaften von Pooling Methoden

---

Wobei  $g : Z \rightarrow (0, \infty)$  eine Likelihood ähnliche Funktion ist und die Gewichte  $w_i$  nichtnegativ sind und zu Eins aufsummieren.

Nach [Gen86a] wird durch die EB und die PDS im kontinuierlichen Fall eine allgemeinere Form des LogOPs als im Satz 5 charakterisiert. Genest verweist in [Gen86a, Gen84b] für den Fall, dass die vor wie in Satz 5 sind auf [McC78].

Nach Genest [Gen84c] gilt:

**Satz 6** *Ist  $Z$  tertiär und erfüllt die Pooling-Methode die SSFP und die EB, dann ist die Pooling-Methode eine Diktatur.*

Nach Satz 4 charakterisiert die SSFP einen LinOP, wenn  $Z$  tertiär ist. Soll ein LinOP in diesem Fall die EB erfüllen, dann sind alle Gewichte bis auf eines gleich Null. Genest bemerkt in [Gen84c, Gen84b], dass folgender Satz gilt, wenn man die SSFP durch die PDS ersetzt:

**Satz 7** *Ist  $Z$  tertiär, dann erfüllt der LogOP die EB und die PDS.*

Für den Fall, dass zusätzlich zur SSFP die IPP gefordert wird, gilt laut Genest [Gen84d] und Wagner [Wag84] der folgender Satz:

**Satz 8** *Ist  $Z$  tertiär und die Pooling-Methode erfüllt sowohl die SSFP wie auch die IPP, dann ist die Pooling-Methode eine Diktatur.*

Wird von der Pooling-Methode die IPP und anstatt der SSFP die MP gefordert, wird also auf die ZPP verzichtet<sup>4</sup>, dann gilt laut Genest [Gen84d] der folgenden Satz:

**Satz 9** *Ist  $Z$  tertiär und gilt die MP sowie die IPP und gilt nicht die ZPP, dann ist die Pooling-Methode von der Form (3.2) mit Gewichten  $w_i = 0$  und  $R$  nimmt Werte aus  $\{0, 1\}$  an.*

Damit reduziert sich der GenLinOP (siehe Gleichung (3.2)) auf die Funktion  $R$ .  $R$  ist eine beliebige binäre<sup>5</sup> Wahrscheinlichkeitsverteilung über  $Z$ , unabhängig von den Schätzungen

---

<sup>4</sup>Es gilt, dass die Forderung der SSFP äquivalent zu der Forderung der MP und der ZPP ist [McC81].  
Siehe Abschnitt 4.1

<sup>5</sup>Binär soll hier bedeuten, dass die Funktion nur die Werte Null und Eins annimmt.

der Gruppenmitglieder. Eine solche Funktion wird *dogmatische Funktion* genannt. Dies ist laut Genest kaum eine Verbesserung zu einer Diktatur [Gen84d].

Hier wurde der Satz 3.1 von Genest [Gen84d] zum besseren Verständnis auf zwei Sätze aufgeteilt. Dadurch soll klar werden, welche Rolle die ZPP in diesem Fall spielt.

Die Kombination der Eigenschaften SSFP und IPP schränken also die Wahl der Pooling-Methode sehr ein. Deshalb werden im folgenden die SSFP durch PDS und die IPP durch die EIPP ersetzt. Damit kann nach [Pen05] folgender Satz als Abwandlung von Satz 8 interpretiert werden:

**Satz 10** *Ist  $\mathcal{Z}$  tertiär und gelten die EIPP und die PDS, dann existieren Konstanten  $w_1, \dots, w_n$  und  $c$ , so dass die Pooling-Methode folgende Form annimmt:*

$$K(p_1, \dots, p_n)(A) = \left[ \sum_{i=1}^n w_i p_i(A) \right] + c. \quad (4.4)$$

Wird zusätzlich zu der EIPP und der PDS noch die UNAM gefordert, dann kann laut Penock et al. [Pen05] folgender Satz formuliert werden:

**Satz 11** *Ist  $\mathcal{Z}$  tertiär und gilt die EIPP, die PDS und die UNAM, dann handelt es sich bei der Pooling-Methode um eine Diktatur.*

Durch die Konstante  $c$  in der Gleichung (4.4) kann die UNAM nicht mehr gelten, außer sie ist gleich Null.

In Satz 4 wurde gezeigt, dass die Forderung der SSFP in einem LinOP resultiert. Der Satz 5 zeigt, welche Form einer Pooling-Methode annimmt wenn sie die EB und ein Variante der PDS erfüllt. Klar ist, dass SSFP und EB nicht gleichzeitig gelten können, außer bei einer Diktatur. Die SSFP kann als Kombination der Eigenschaften MP und ZPP interpretiert werden. Fraglich ist nun, welche dieser Eigenschaften in Kombination eine Diktatur erzwingen. Die EB und die modifizierte PDS zusammen charakterisieren einen LogOP, welcher eine extreme Variante der ZPP erfüllt. Die ZPP wird nicht nur erfüllt, sondern es genügt, wenn ein Gruppenmitglied ein Ereignis für unmöglich hält, damit die Konsenswahrscheinlichkeitsverteilung dies auch tut. Folglich kann die Inkompatibilität nur noch durch die MP verursacht werden.

Wie in Satz 8 gesehen, können SSFP und IPP gleichzeitig nur von einer Diktatur er-

füllt werden. Wird hier die SSFP durch die PDS ersetzt, dann lässt sich folgender Satz formulieren [Gen87]:

**Satz 12** *Ist  $Z$  abzählbar und enthält mindestens fünf Ereignisse und erfüllt die Pooling-Methode PDS und IPP, dann ist sie von der Form:*

$$K(p_1, \dots, p_n)(A) = \sum_{i=1}^n w_i p_i(A) + c$$

Genest und Wagner zeigen weiterhin, dass, die Pooling-Methode aus Satz 12 eine Diktatur ist, wenn  $Z$  endlich ist.

In diesem Abschnitt wurden die Kombinationen der Eigenschaften von Pooling-Methoden aus Abschnitt 4.1 betrachtet. Werden die Eigenschaften in Kombinationen von einer Pooling-Methode gefordert, dann beschränken sie die Auswahl an Pooling-Methoden, erleichtern jedoch den Umgang mit ihnen. Pooling-Methoden, die eine gewisse Eigenschaft erfüllen, können zu einer Klasse zusammengefasst werden. Bei der Konstruktion einer Pooling-Methode, die für eine spezielle Konfiguration geeignet ist, kann anhand der geforderten Eigenschaften die entsprechende Klasse von Pooling-Methoden gewählt werden. Manche Kombinationen von geforderten Eigenschaften erlauben nur noch eine Diktatur als Pooling-Methode.

Die vorgestellten Beschränkungssätze zeigen, welche Klasse von Pooling-Methoden welche Bedingungen erfüllt. Der *LinOP* und externe Bayes'sche Pooling-Methoden - genauer der *LogOP* - werden im Folgenden genauer betrachtet. Es soll bestimmt werden, welche Eigenschaften erfüllt und welche sinnvollerweise für die in Abschnitt 2.3 vorgestellte lokale Konfiguration gefordert werden sollten.

### 4.3. Eigenschaften eines LinOP bzw. GenLinOP

In diesem Abschnitt werden Eigenschaften eines LinOPs besprochen. Jede Eigenschaft wird daraufhin untersucht, ob sie für eine, durch eine Einschränkung lokalisierte Konfiguration der bayes'schen Informationsfusion sinnvoll ist. In Situationen, bei denen es einen Unterschied macht, ob der LinOP ( siehe Gleichung (3.1)) oder der GenLinOP ( siehe Gleichung (3.2)) betrachtet wird, wird eine Fallunterscheidung gemacht. Da der LinOP ein Spezialfall des GenLinOP ist, müssen Eigenschaften, die für den LinOP nicht gelten,



Eigenschaft	LinOP	GenLinOP
MP	✓	✓
WSFP	✓	✓
SSFP	✓	
PDS	✓	
ZPP	✓	
UNAM	✓	
ND		
IPP		
EB		
RPC		

**Tabelle 4.1.:** Darstellung der Eigenschaften mit Markierung ob der LinOP/GenLinOP diese Eigenschaft erfüllt.

auch für den GenLinOP nicht gelten. Umgekehrt gilt: Gelten Eigenschaften für den GenLinOP, dann gelten sie auch für den LinOP. Die Tabelle 4.1 stellt dar, welche relevanten Eigenschaften von LinOP und welche vom GenLinOP erfüllt werden. Die Annahme, dass  $\mathcal{Z}$  tertiär ist, wird ab jetzt vorausgesetzt, da Probleme, bei denen dies nicht gilt, eine Lokalisierung als nicht nötig erscheinen lassen.

Wie von Genest [Gen84d] bewiesen, charakterisiert die *MP* einen GenLinOP. Das heißt: jede Pooling-Methode, welche die *MP* erfüllt, ist von der Form der Gleichung (3.2). Dies ist der Grund dafür, dass der GenLinOP wie in Gleichung (3.1) auf der beliebigen Menge  $A \in \mathcal{Z}$  beschrieben werden kann, anstatt nur auf solchen Mengen, die nur aus einem Element bestehen.

So wie die *MP* einen GenLinOP charakterisiert, charakterisiert die *SSFP* einen LinOP von der Form der Gleichung (3.1). Der LinOP erfüllt also die *SSFP*. Wie McConway [McC81] gezeigt hat, gilt die *SSFP* genau dann, wenn die *MP* und die *ZPP* für eine Pooling-Methode gelten.

Aus der *SSFP* folgt, dass die Gewichte  $w_i$  eines LinOP nicht von dem betrachteten Ereignis  $A$  abhängen können.

Da der LinOP die *SSFP* erfüllt, hat er auch die *PDS* Eigenschaft [Pen05].

Die *SSFP* kann auf zwei Arten interpretiert werden: Erstens beschreibt sie die Form, wie der Konsens gebildet werden muss. Der Konsens darf nur von den Meinungen der Gruppenmitglieder bezüglich des betrachteten Ereignisses abhängen. Zweitens setzt sich die *SSFP* aus der *ZPP* und der *MP* zusammen. Gilt für eine Pooling-Methode die *SSFP*,

	<i>Agent</i> <sub>1</sub>	<i>Agent</i> <sub>2</sub>
$z^1$	0.4	0.0
$z^2$	0.6	0.6
$z^3$	0.0	0.4

**Tabelle 4.2.:** Beispiel für Meinungen von zwei Gruppenmitglieder mit einer Einstimmigkeit bezüglich der Wahrscheinlichkeit von  $z^2$

dann erfüllt diese Pooling-Methode auch die MP und die ZPP.

Die Eigenschaft, welche die MP zur SSFP äquivalent macht, ist also die ZPP. Der LinOP erfüllt die SSFP und damit auch die ZPP [McC81]. Das bedeutet, wenn alle Gruppenmitglieder einem Ereignis einen DoB von Null zuordnen, tut dies der Konsens auch. Dies können wir leicht nachvollziehen: Da alle DoB bezüglich des betrachteten Ereignisses gleich Null sind, ändern die Gewichte, ganz egal wie sie gewählt werden, nichts daran, dass die Summe der DoB Null ist.

Im Gegensatz dazu gilt die ZPP bei einem GenLinOP nicht immer. Ob der GenLinOP die ZPP erfüllt, hängt davon ab, wie die Gewichte und die Funktion  $R$  gewählt werden. Hierfür ein Beispiel, bei dem die ZPP nicht gilt: Die Wahrscheinlichkeitsverteilungen der beiden Gruppenmitglieder ist in Tabelle 3.1 dargestellt. Sei  $R(z^1) = 0$ ,  $R(z^2) = 0$  und  $R(z^3) = 1$  und die Gewichte seien  $w_1 = \frac{1}{4}$  und  $w_2 = \frac{1}{4}$ . Zu beachten ist, dass diese Gewichte die Bedingungen des GenLinOP erfüllen, obwohl sie nicht zu Eins aufsummieren. Dann hat der GenLinOP folgende Ergebnisse:  $GenLinOP(p_1, p_2)(z^1) = 0.3$ ,  $GenLinOP(p_1, p_2)(z^2) = 0.2$  und  $GenLinOP(p_1, p_2)(z^3) = 0.5$ . Das Ereignis  $z^3$  wird damit nicht mehr als unmöglich betrachtet. Die einstimmige Wahrscheinlichkeit von  $p_1(z^3) = 0$  mit  $i \in \{1, 2\}$  bleibt nicht erhalten.

Klar ist, dass die UNAM von einem GenLinOP nicht immer erfüllt ist, da die ZPP nicht erfüllt ist.

Der LinOP im Gegenzug erfüllt die UNAM. Sind die Wahrscheinlichkeitsverteilungen der Gruppenmitglieder gleich, ändert die Anwendung des LinOP diese Wahrscheinlichkeitsverteilung nicht. Als Erweiterung der UNAM gilt, dass die Einstimmigkeit auch erhalten bleibt, wenn die Gruppe nur für ein  $z \in Z$  die gleiche Wahrscheinlichkeit bestimmt hat. Dazu ebenfalls ein Beispiel: Die Wahrscheinlichkeitsverteilungen der Agenten werden in Tabelle 4.2 dargestellt. Seien die Gewichte  $w_1 = 0,2$  und  $w_2 = 0,8$ . Dann liefert der LinOP:  $LinOP(p_1, p_2)(z^1) = 0,08$ ,  $LinOP(p_1, p_2)(z^2) = 0,6$  und  $LinOP(p_1, p_2)(z^3) = 0,32$ . Nun folgt ein Beweis, dass der LinOP die UNAM auch für einzelne Elemente erfüllt: An-

	<i>Agent</i> <sub>1</sub>	<i>Agent</i> <sub>2</sub>
$z^1$	0.0	0.0
$z^2$	0.6	0.6
$z^3$	0.4	0.4

**Tabelle 4.3.:** Beispiel für einstimmige Meinungen von zwei Gruppenmitglieder.

genommen alle Gruppenmitglieder haben für das Ereignis  $z$  die gleiche Wahrscheinlichkeit  $p_*(z)$  ermittelt. Der LinOP hat dann folgende Form:

$$\text{LinOP}(p_1, \dots, p_n)(z) = \sum_{i=1}^n w_i p_i(z) = \sum_{i=1}^n w_i p_*(z)$$

$p_*(z)$  hängt damit nicht mehr von  $i$  ab und kann aus der Summe herausgezogen werden.

$$\text{LinOP}(p_1, \dots, p_n)(z) = p_*(z) \sum_{i=1}^n w_i$$

Da die Gewichte des LinOP zu Eins aufsummieren, gilt:

$$\text{LinOP}(p_1, \dots, p_n)(z) = p_*(z)$$

Damit ist klar, dass der LinOP die UNAM auch für einzelne Ereignisse erfüllt.  $\square$

Die UNAM ist je nach Wahl der Wahrscheinlichkeitsverteilung  $R$  und der Gewichte  $w_i$  beim GenLinOP nicht immer erfüllt. In Beispiel von Tabelle 4.3 ist dies für  $R(z^1) = 1$ ,  $R(z^2) = 0$  und  $R(z^3) = 0$  mit den Gewichten  $w_1 = \frac{1}{4}$  und  $w_2 = \frac{1}{4}$  der Fall. Hier erhält man folgendes Ergebnis:  $\text{GenLinOP}(p_1, p_2)(z^1) = 0.5$ ,  $\text{GenLinOP}(p_1, p_2)(z^2) = 0.3$  und  $\text{GenLinOP}(p_1, p_2)(z^3) = 0,2$ . Dabei bleibt die Wahrscheinlichkeitsverteilung der Agenten nicht erhalten.

Ob ein LinOP die *ND* erfüllt, hängt von den gewählten Gewichten ab. Der LinOP kann wie in Abschnitt 3.2 besprochen zu einer Diktatur werden, wenn alle Gewichte außer einem auf Null gesetzt werden. Dieses eine Gewicht ist dann aufgrund der Bedingung, dass die Gewichte auf Eins aufsummieren, gleich Eins. Damit wird die Wahrscheinlichkeitsverteilung des Gruppenmitglieds, dessen Gewicht gleich Eins ist, ohne Änderung übernommen.

Der GenLinOP kann auf diese Weise ebenfalls zu einer Diktatur werden. Jedoch ist das nicht die einzige Möglichkeit: Werden die Gewichte alle gleich Null gewählt, dann diktiert die Wahrscheinlichkeitsverteilung  $R$ .

	<i>Agent</i> <sub>1</sub>	<i>Agent</i> <sub>2</sub>
$z^1$	0.12	0.05
$z^2$	0.18	0.05
$z^3$	0.28	0.45
$z^4$	0.42	0.45

**Tabelle 4.4.:** Beispiel für Meinungen von zwei Gruppenmitgliedern.

	<i>Agent</i> <sub>1</sub>	<i>Agent</i> <sub>2</sub>
$A$	0.3	0.1
$B$	0.4	0.5
$A \cap B$	0.12	0.05

**Tabelle 4.5.:** Beispiel für Meinungen von zwei Gruppenmitgliedern, wobei  $A$  und  $B$  unabhängig sind.

Wie in Satz 8 beschrieben, erfüllt ein LinOP die *IPP* nur als Diktatur. Der GenLinOP ist, wenn er die *IPP* erfüllt, nach Satz 9 und Satz 8 damit entweder eine Diktatur oder eine *dogmatische* Funktion. Für den LinOP hier ein Beispiel zur Verdeutlichung: In diesem Beispiel wird ein  $Z = \{z^1, z^2, z^3, z^4\}$  betrachtet. Das Ereignis  $A$  setze sich aus  $A = \{z^1, z^2\}$  und Ereignis  $B$  aus  $B = \{z^1, z^3\}$  zusammen. Die Wahrscheinlichkeiten für die Ereignisse  $A$  und  $B$  wurde in Tabelle 4.4 aufgeführt. Die Wahrscheinlichkeitsverteilungen der Gruppenmitglieder sind in der Tabelle 4.5 dargestellt. Das Ergebnis eines LinOP mit den Gewichten  $w_i = \frac{1}{2}$  ist dann für  $p(A) = 0,2$  und für  $p(B) = 0,45$ . Das Pooling-Ergebnis für  $A \cap B$  sollte dann  $0,09$  sein, für diesen LinOP ist es jedoch  $0,085$ . Damit ist klar, dass der LinOP - mit Ausnahme der Diktatur - die *IPP* nicht erfüllt. Der GenLinOP erfüllt die *IPP* dementsprechend auch nicht.

Pennock und Wellman [Pen05] zeigen, dass der LinOP die *EIPP* ebenfalls nicht erfüllt. Auch die direkten Unabhängigkeiten zwischen den Ereignissen von Interesse bleiben also nicht erhalten.

Wie im Satz 6 gezeigt, können die Eigenschaften *EB* und *SSFP* gemeinsam nur von einer Diktatur erfüllt werden. Da jeder LinOP die *SSFP* erfüllt, kann ein LinOP nur als Diktatur die *EB* erfüllen.

Satz 3 zeigt, dass der LinOP die *RPC* nur dann erfüllt, wenn genau zwei disjunkte Ereignisse in  $Z$  enthalten sind [Gen84a]. Da hier  $Z$  als tertiär angenommen wird, erfüllt der LinOP die *RPC* nicht.

$p(z^i)$	$Agent_1$	$Agent_2$
$z^1$	0.2	0.3
$z^2$	0.4	0.1
$z^3$	0.4	0.6

Tabelle 4.6.: Beispiel für Meinungen von zwei Gruppenmitgliedern

### 4.3.1. Wertung der Eigenschaften eines LinOP und GenLinOP für die lokale Bayes'sche Fusion

In dieser Arbeit werden zwei Arten der Lokalisierung betrachtet: die Einschränkung aus Abschnitt 2.3 und die Marginalisierung aus Abschnitt 2.4. Eine Eigenschaft wie die *MP* für die Marginalisierung kann auch für die Einschränkung formuliert werden. Dabei wird gefordert, dass die Pooling-Methode das gleiche Ergebnis liefert, egal ob vor ihrer Anwendung eingeschränkt wurde oder danach. Dass diese Eigenschaft für einen LinOP nicht gilt, soll anhand eines Beispiels verdeutlicht werden:

Es werden zwei Gruppenmitglieder betrachtet. Es muss das gleiche Ergebnis folgen, wenn zuerst die Wahrscheinlichkeitsverteilungen der Gruppenmitglieder mittels eines LinOP berechnet und dann auf ein  $U \subset Z$  eingeschränkt wird oder umgekehrt. Für zwei Gruppenmitglieder und  $Z = \{z^1, z^2, z^3\}$ ,  $U := \{z^1, z^2\}$  mit den Gewichten  $w_1 = w_2 = \frac{1}{2}$  seien die Meinungen der Gruppenmitglieder wie in Tabelle 4.6 dargestellt. Wird zuerst der LinOP berechnet und dann eingeschränkt, so erhält man die Wahrscheinlichkeiten  $p(z^1|U) = \frac{1}{2}$  und  $p(z^2|U) = \frac{1}{2}$ . Kombiniert man jedoch die eingeschränkten Verteilungen mittels eines LinOP, dann ergibt sich folgendes Resultat:  $LinOP(p_1, p_2)(z^1|U) = \frac{13}{24}$  und  $LinOP(p_1, p_2)(z^2|U) = \frac{11}{24}$ . Dadurch ist klar, dass die Einschränkung nicht mit der Anwendung der Pooling-Methode vertauschbar ist.

Die Lokalisierung findet in der hier vorgestellten Konfiguration schon im Vorfeld durch eine Vorverarbeitung statt. Während der Bayes'schen Fusion werden alle Ereignisse in  $\mathcal{U} \subseteq \mathcal{Z}$  betrachtet. Die Frage der Reihenfolge der Lokalisierung und der Kombination der Wahrscheinlichkeitsverteilungen der Gruppenmitglieder stellt sich daher nicht mehr. Die *MP* muss von einer Pooling-Methode, die hier verwendet werden soll, also nicht erfüllt werden. Wie French in [Fre85] richtig beobachtet, kann die *MP* als nicht wünschenswert betrachtet werden. Sind die Gruppenmitglieder beispielsweise Experten auf verschiedenen Fachgebieten, so kann es sinnvoll sein, wenn ein Experte auf der sub- $\sigma$ -Algebra seiner Expertise ein höheres Gewicht erhält als auf den restlichen betrachteten Ereignissen. Dies verhin-

dert jedoch die MP. Die MP kann auch als positive Eigenschaft für den verteilten lokalen Kontext interpretiert werden. Sie ermöglicht, dass die Pooling-Methode auf Ereignissen berechnet werden kann, ohne dass die atomaren Ereignisse betrachtet werden müssen.

Die *SSFP* hat für die lokale Bayes'sche Fusion keine besondere Bedeutung. Generell ist die *SSFP* eine gute Eigenschaft für die Berechenbarkeit, da zur Berechnung des Poolingergebnisses nur die Wahrscheinlichkeitsverteilungen der Gruppenmitglieder bezüglich des betrachteten Ereignisses bekannt sein muss. Es gibt also keine Abhängigkeit zu anderen Ereignissen. Des Weiteren schränkt die Forderung der *SSFP* an eine Pooling-Methode, kombiniert mit der Forderung, dass das Ergebnis der Pooling-Methode eine Wahrscheinlichkeit ist, die Wahl der zulässigen Pooling-Methoden auf einen *LinOP* ein. Kann auf eine Normalisierung verzichtet werden, dann kann beispielsweise ein nicht normalisierter *LogOP* (*nnLogOP*) die *SSFP* erfüllen. Der *nnLogOP* hat die Form:

$$nnLogOP(p_1, \dots, p_n)(z) = \prod_{i=1}^n p_i(z)^{w_i} \quad (4.5)$$

Je nachdem, wie die Konsenswahrscheinlichkeitsverteilung weiter verwendet wird, kann dies sinnvoll sein. Soll beispielsweise das wahrscheinlichste Ereignis gefunden werden, kann auf die Normierung verzichtet werden.

Da bei der lokalen Bayes'schen Fusion mit einer Einschränkung als Lokalisierung die Wahrscheinlichkeit der Ereignisse außerhalb von  $U$  von allen Gruppenmitgliedern auf Null gesetzt wird, sollte auch der Konsens der Ereignisse, die von keinem Agenten betrachtet wurden, die Wahrscheinlichkeit Null zuordnen. Die *ZPP* ist also für den lokalen bayes'schen Fusionsansatz notwendig, wenn in der verwendeten Konfiguration durch eine Einschränkung lokalisiert wurde. Für die lokale Konfiguration, bei der die  $U_i$  eine Partitionierung von  $U$  sind, ist die *UNAM* nicht von Bedeutung, da jedes  $z \in U$  genau von einem Gruppenmitglied betrachtet wird.

Die *UNAM* wird dann wichtig, wenn alle Gruppenmitglieder die gleichen Ereignisse betrachten. Dann könnte es vorkommen, dass die Wahrscheinlichkeitsverteilungen der Gruppenmitglieder gleich sind. Damit ist die Voraussetzung der *UNAM* erfüllt. Für eine lokale Fusion, bei der nicht alle Gruppenmitglieder die gleichen Ereignisse betrachten, wird die Voraussetzung der *UNAM* nicht erfüllt. Wird durch eine Einschränkung lokalisiert, dann sind die lokalen Wahrscheinlichkeitsverteilungen der Gruppenmitglieder verzerrt. In diesem Fall ist die *UNAM* nicht wünschenswert, da dies bedeuten würde, dass die Verzer-

rung auch im Konsens noch enthalten ist. Stattdessen ist es erwünscht, dass die Pooling-Methode diese Verzerrung aufhebt. Für den Fall, in dem durch eine Marginalisierung lokalisiert wird, ist die Forderung der UNAM sinnvoll, da die lokalen Wahrscheinlichkeitsverteilungen der Gruppenmitglieder nicht verzerrt sind. Betrachten zwei Gruppenmitglieder das gleiche Ereignis, müssen sie diesem Ereignis die gleiche Wahrscheinlichkeit zuordnen, wenn die Wahrscheinlichkeitsverteilung der Gruppenmitglieder objektiv sind. Diese sollte dann auch so durch den Konsens repräsentiert werden.

Die  $ND$  ist eine Anforderung an die Gewichte eines LinOP. Bei einer lokalen Konfiguration werden bei einer Diktatur alle Ereignisse ignoriert, die nicht von dem Gruppenmitglied betrachtet wurden, dessen Meinung das Gewicht Eins erhalten hat. Dies kommt einer Einschränkung auf den betrachteten Bereich  $U_i$  des Diktators  $i$  gleich. Dadurch werden alle Ereignisse in  $\bar{U}_i$  als unmöglich angenommen.

Bei einem lokalen Ansatz, bei dem der Schnitt der  $U_i$  leer ist, kann es keine gemeinschaftlich gehaltenen Unabhängigkeiten geben. Daher ist das Nichterhalten dieser, also der Verstoß gegen die  $IPP$  des LinOP, in diesem Fall kein Nachteil. Diese Erkenntnis lässt jedoch eine Wertung des Pooling-Ergebnisses zu, denn das Ergebnis des LinOP erhält die Beziehungen zwischen den betrachteten Ereignissen nicht.

Für eine lokale Fusion sind Pooling-Methoden interessant, die  $EB$  erfüllen, da die Gewichte wie folgt gewählt werden können: Durch die Einschränkung verzerrt jedes Gruppenmitglied seine lokale A-posteriori-Wahrscheinlichkeitsverteilung um den Faktor  $\frac{1}{p(U_i|d)}$ . Die Gruppenmitglieder betrachten dabei disjunkte  $U_i$ . Um diesen Faktor auszugleichen, könnten die Gewichte beim LinOP  $w_i = p(U_i|d)$  gewählt werden. Damit wäre das Ergebnis unverzerrt. Es gilt:

$$K(p_1, \dots, p_n)(z|d) = \sum_{i=1}^n w_i p_i(z|d) = \sum_{i=1}^n w_i \frac{p(z|d) \cdot \mathbf{1}_{\{z \in U_i\}}}{p(U_i|d)} = \sum_{i=1}^n p(U_i|d) \frac{p(z|d) \cdot \mathbf{1}_{\{z \in U_i\}}}{p(U_i|d)} \quad (4.6)$$

Wird nun gekürzt, erhält man die Summe der unverzerrten Wahrscheinlichkeiten. Der Verzerrung, die durch die Einschränkung auf die Bereiche  $U_i$  entstanden ist, wurde mit den Gewichten  $w_i = p(U_i|d)$  entgegengewirkt.

Wünschenswert ist also, die Gewichte  $w_i = p(U_i|d)$  zu wählen. Die Berechnungen dafür

Eigenschaft	LogOP
MP	
WSFP	
SSFP	
PDS	✓
ZPP	✓
UNAM	✓
ND	
IPP	
EB	✓
RPC	✓

**Tabelle 4.7.:** Darstellung der Eigenschaften mit Markierung, ob der LogOP diese Eigenschaft erfüllt.

sind jedoch aufwendig.  $p(U_i|d)$  berechnet sich wie folgt:

$$p(U_i|d) = \frac{p(d|U_i)p(U_i)}{p(d)}$$

Wenn die dargestellte Pooling-Methode die EB erfüllen würde, könnte man die Gewichte schon vor der eigentlichen Fusion festlegen, ohne das Ergebnis zu verändern.

$$p(U_i) = \sum_{z \in U_i} p(z)$$

Da die Gewichte nicht mehr nach den Beobachtungen bedingt werden, reduziert sich der Aufwand, der entsteht, um die Gewichte zu wählen. Die Berechnung von  $p(d) = \sum_{j=1}^n p(d|U_j)p(U_j)$  und von  $p(d|U_i)$  fallen weg. So würde für den Fall einer lokalen Bayes'schen Fusion mit einer Einschränkung als Lokalisierung eine optimale Wahl der Gewichte ermöglicht.

Für eine lokale Konfiguration ist die *RPC* nicht bedeutsam. Damit die *RPC* in einer lokalen Konfiguration von Bedeutung ist, müssen alle Gruppenmitglieder mindestens ein Ereignisse gemeinsam betrachten. Dies ist im Allgemeinen jedoch nicht der Fall.

#### 4.4. Eigenschaften eines LogOP

In nachstehender Tabelle 4.7 wird zur Übersicht dargestellt, welche der in Kapitel 4.1 vorgestellten Eigenschaften ein LogOP erfüllt.



Der LogOP erfüllt die Eigenschaft *MP* nicht. Wie in Satz 2 gezeigt wird, impliziert die *MP* einen GenLinOP (siehe Gleichung (3.2)) wenn  $Z$  tertiär ist. Da der LogOP nicht zu dieser Klasse von Pooling-Methoden gehört, erfüllt er auch nicht die *MP*.

Der LogOP erfüllt nicht die *SSFP*. Wie in Satz 4 gezeigt, impliziert die *SSFP* einen LinOP wenn  $Z$  tertiär ist. Wird jedoch auf die Normalisierung (den Nenner des LogOP) verzichtet, dann erfüllt dieser nnLogOP (siehe Gleichung (4.5)) die *SSFP*. Damit erfüllt der nnLogOP auch die *MP* und die *ZPP* [McC81]. Hinzu kommt, dass Satz 1 ebenfalls gilt, und der nnLogOP dadurch auch die *WSFP* erfüllt.

Der LogOP erfüllt die *PDS* [Gen84b, Pen05]. Durch die Proportionalität, welche die *PDS* von der *SSFP* unterscheidet, wird die Normalisierung des LogOP, welche das Erfüllen der *SSFP* verhindert hat, wieder möglich. Die Proportionalitätskonstante der *PDS* ist dabei der Nenner des LogOP.

Der LogOP erfüllt die *ZPP* auf eine drastische Weise. Wenn beim LogOP nur ein Gruppenmitglied die Wahrscheinlichkeit eines Ereignisses mit Null einschätzt, wird die Wahrscheinlichkeit des Konsens für dieses Ereignis gleich Null. Damit ist klar, dass der Konsens auch gleich Null ist, wenn alle Gruppenmitglieder die Wahrscheinlichkeit eines Ereignisses mit Null einschätzen.

Der LogOP erfüllt die *UNAM* [Pen05]. Dies können wir leicht nachvollziehen: Durch die Rechenregel der Multiplikation für Potenzen gilt:  $a^s \cdot a^r = a^{s+r}$ . Sind die Wahrscheinlichkeitsverteilungen aller Gruppenmitglieder gleich, dann hat der LogOP folgende Form:

$$\text{LogOP}(p_1, \dots, p_n)(z) = \frac{\prod_{i=1}^n p_1(z)^{w_i}}{\sum_{z \in Z} \prod_{i=1}^n p_1(z)^{w_i}}, \quad z \in Z$$

Da die Gewichte des LogOP zu Eins aufsummieren, gilt:

$$\text{LogOP}(p_1, \dots, p_n)(z) = \frac{p_1(z)^1}{\sum_{z \in Z} p_1(z)^1}$$

Im Nenner werden die Wahrscheinlichkeiten der Ereignisse in ganz  $Z$  aufsummiert. Diese Summe muss wegen der Eigenschaften einer Wahrscheinlichkeit gleich Eins sein. Daraus folgt:

$$\text{LogOP}(p_1, \dots, p_n)(z) = p_1(z), \quad \forall z \in Z$$

$p(z^i)$	<i>Agent</i> <sub>1</sub>	<i>Agent</i> <sub>2</sub>
$z^1$	0.2	0.1
$z^2$	0.5	0.5
$z^3$	0.3	0.4

**Tabelle 4.8.:** Beispiel für Meinungen von zwei Gruppenmitgliedern, wobei eine Einstimmigkeit bei  $z^2$  auftritt.

Die Wahrscheinlichkeitsverteilung des ersten Gruppenmitglieds kann durch die eines beliebigen Gruppenmitglieds ersetzt werden, da alle Wahrscheinlichkeitsverteilungen identisch sind. Da der LogOP jedoch die SSFP nicht erfüllt, erhält er die einheitlich gehaltene Wahrscheinlichkeit für einzelne Ereignisse nicht. Dazu hier ein Beispiel:

Zwei Gruppenmitglieder *Agent*<sub>1</sub> und *Agent*<sub>2</sub> betrachten ein  $Z = \{z^1, z^2, z^3\}$  aus drei Elementarereignissen. Die Gewichte des LogOP sind  $w_i = \frac{1}{2}, i \in \{1, 2\}$ . Die Agenten ordnen den Ereignissen die Wahrscheinlichkeiten wie in Tabelle 4.8 zu. Das Ergebnis des LogOP für das Ereignis  $z^2$  ist:  $LogOP(z^2) \approx 0,5062$ . Dies ist ungleich der gemeinsam gehaltenen Wahrscheinlichkeit  $p_i(z^2) = \frac{1}{2}, i \in \{1, 2\}$  der Gruppenmitglieder für das Ereignis  $z^2$ . Der LogOP erhält also die einheitlich gehaltene Wahrscheinlichkeitsverteilung für einzelne Ereignisse nicht.

Wie beim LinOP hängt beim LogOP das Erfüllen der *ND* von den Gewichten ab. Werden alle Gewichte bis auf eines auf Null gesetzt, wird der Konsens von dem Gruppenmitglied diktiert, dessen Gewicht nicht Null ist. Dies können wir uns leicht klar machen: Das Potenzgesetz für Null besagt:  $a^0 = 1 \forall a \in \mathbb{R}$ . Damit sehen wir, dass alle gewichteten Wahrscheinlichkeitsverteilungen der Gruppenmitglieder, die nicht vom Diktator sind, Eins ergeben und somit das Ergebnis des Produktes nicht verändern.

Der LogOP erfüllt nicht die *IPP*. Dies soll nun bewiesen werden: Annahme ist, dass zwei Ereignisse  $A, B$  für alle Gruppenmitglieder unabhängig sind. Tabelle 4.9 stellt dabei die Meinungen der Gruppenmitglieder für  $Z = \{z^1, z^2, z^3, z^4, z^5\}$  dar. Sei  $A = \{z^1, z^2\}$  und

	<i>Agent</i> <sub>1</sub>	<i>Agent</i> <sub>2</sub>
$z^1$	0.36	0.25
$z^2$	0.24	0.25
$z^3$	0.24	0.25
$z^4$	0.1	0.2
$z^5$	0.06	0.05

**Tabelle 4.9.:** Beispiel für Meinungen von zwei Gruppenmitgliedern.

	$LogOP(p_1, p_2)(z)$
$z^1$	0.3042
$z^2$	0.2484
$z^3$	0.2484
$z^4$	0.1434
$z^5$	0.0555

**Tabelle 4.10.:** Ergebnis des LogOP.

$A = \{z^1, z^3\}$  unabhängige Ereignisse. Wird nun ein LogOP mit den Gewichte  $w_i = \frac{1}{2}$  für  $Z$  berechnet, liefert der LogOP Ergebnisse wie in Tabelle 4.10 dargestellt. Die Ergebnisse von  $A$  und  $B$  sind dann  $LogOP(p_1, p_2)(A) = LogOP(p_1, p_2)(z^1) + LogOP(p_1, p_2)(z^2) \approx 0,5526$  und  $LogOP(p_1, p_2)(B) = LogOP(p_1, p_2)(z^1) + LogOP(p_1, p_2)(z^3) \approx 0,5526$ . Das Ergebnis für  $A \cap B$  ist dann  $LogOP(p_1, p_2)(A \cap B) = LogOP(p_1, p_2)(z^1) \approx 0.3042$ . Dies entspricht jedoch nicht  $LogOP(p_1, p_2)(A) \cdot LogOP(p_1, p_2)(B) \approx 0.3054$ . Somit ist gezeigt, dass der LogOP ist IPP nicht erfüllt.

Das der LogOP die EB erfüllt, soll nun Bewiesen werden:

$$LogOP(p_1, \dots, p_n)(z|d) \propto \prod_{i=1}^n \underbrace{p_i(z|d)^{w_i}}_{\text{Fusion vor der Kombination der Meinungen}}$$

Wegen des Satzes von Bayes gilt dann:

$$\begin{aligned} LogOP(p_1, \dots, p_n)(z|d) &\propto \prod_{i=1}^n \left[ \frac{p_i(z)p(d|z)}{p_i(d)} \right]^{w_i} \\ &= \frac{\left[ \prod_{i=1}^n p_i(z)^{w_i} \right] \cdot \left[ \prod_{i=1}^n p(d|z)^{w_i} \right]}{\left[ \prod_{i=1}^n p_i(d)^{w_i} \right]} \end{aligned}$$

Da  $p_i(d)$  nicht von  $z$  abhängt, kann es als Proportionalitätskonstante aus der Gleichung genommen werden:

$$\begin{aligned} LogOP(p_1, \dots, p_n)(z|d) &\propto \left[ \prod_{i=1}^n p_i(z)^{w_i} \right] \cdot \left[ \prod_{i=1}^n p(d|z)^{w_i} \right] \\ &= \left[ \prod_{i=1}^n p_i(z)^{w_i} \right] \cdot p(d|z)^{\sum_{i=1}^n w_i} \end{aligned}$$

Da die Gewichte des LogOP zu Eins aufsummieren, gilt:

$$\begin{aligned} \text{LogOP}(p_1, \dots, p_n)(z|d) &\propto \left[ \prod_{i=1}^n p_i(z)^{w_i} \right] \cdot p(d|z) \\ &\propto \underbrace{\text{LogOP}(p_1, \dots, p_n)(z)p(d|z)}_{\text{Fusion nach der Kombination der Meinungen}} \end{aligned}$$

Damit ist gezeigt, dass der LogOP die EB erfüllt.

Der LogOP erfüllt die RPC. Dies soll nun bewiesen werden: Gilt  $\frac{p_i(z^1)}{p_i(z^2)} \geq \frac{p_i(z^3)}{p_i(z^4)} \forall i \in \{1, \dots, n\}$  für Ereignisse  $z^1, z^2, z^3, z^4$  aus  $\mathcal{Z}$ , dann gilt auch:

$$\prod_{i=1}^n \frac{p_i(z^1)}{p_i(z^2)} \geq \prod_{i=1}^n \frac{p_i(z^3)}{p_i(z^4)}$$

Da die Gewichte des LogOP alles positiv sind ändern sie nichts an der Ungleichung:

$$\prod_{i=1}^n \left[ \frac{p_i(z^1)}{p_i(z^2)} \right]^{w_i} \geq \prod_{i=1}^n \left[ \frac{p_i(z^3)}{p_i(z^4)} \right]^{w_i}$$

Dies kann nun wie folgt umgeformt werden:

$$\begin{aligned} \prod_{i=1}^n \left[ \frac{p_i(z^1)}{p_i(z^2)} \right]^{w_i} \geq \prod_{i=1}^n \left[ \frac{p_i(z^3)}{p_i(z^4)} \right]^{w_i} &\Leftrightarrow \prod_{i=1}^n \frac{p_i(z^1)^{w_i}}{p_i(z^2)^{w_i}} \geq \prod_{i=1}^n \frac{p_i(z^3)^{w_i}}{p_i(z^4)^{w_i}} \\ &\Leftrightarrow \frac{\prod_{i=1}^n p_i(z^1)^{w_i}}{\prod_{i=1}^n p_i(z^2)^{w_i}} \geq \frac{\prod_{i=1}^n p_i(z^3)^{w_i}}{\prod_{i=1}^n p_i(z^4)^{w_i}} \\ &\Leftrightarrow \frac{\text{LogOP}(p_1, \dots, p_n)(z^1)}{\text{LogOP}(p_1, \dots, p_n)(z^2)} \geq \frac{\text{LogOP}(p_1, \dots, p_n)(z^3)}{\text{LogOP}(p_1, \dots, p_n)(z^4)} \end{aligned}$$

Damit ist gezeigt, dass der LogOP die RPC erfüllt.

#### 4.4.1. Wertung der Eigenschaften eines LogOP für die lokale Bayes'sche Fusion

Der LogOP liefert in einer lokalen Konfiguration mit einer Einschränkung als Lokalisierung im Allgemeinen kein sinnvolles Ergebnis. Dies liegt daran, dass jedes Gruppenmitglied die Wahrscheinlichkeit für die Menge der Ereignisse außerhalb des betrachteten  $U_i$  auf Null setzt. Jedes Gruppenmitglied erhebt immer dann ein Veto, wenn das betrachtete Ereignis

nicht in dem von ihm betrachteten  $U_i$  liegt. Ist der Schnitt der  $U_i$  leer, dann ist das Ergebnis des LogOP immer Null. Damit macht der LogOP nur Sinn, wenn alle Agenten die selben Ereignisse betrachten, oder eine andere Art der Lokalisierung wie beispielsweise einer Einschränkung mit Zusatzwissen gewählt wird.

## 4.5. Analyse des Supra-Bayes Ansatzes für die lokale Bayes'sche Fusion

Wie French [Fre85] argumentiert, ist der SB Ansatz eine der Lösungen des hier betrachteten Problems, welches darin besteht, die Meinungen der Gruppenmitglieder zu verschmelzen. Wenn der Bayes'sche Ansatz nicht nur zur Vereinigung des a-priori Wissens mit dem Likelihood der Beobachtungen zu der A-posteriori-Wahrscheinlichkeitsverteilung der einzelnen Gruppenmitglieder verwendet wird, sondern auch bei einem Gruppenentscheid eingesetzt werden soll, dann ist der SB Ansatz eine konsistente Wahl.

Der Entscheider kann mittels des Likelihood  $p_{n+1}(p_1(z|d), \dots, p_n(z|d), d|z)$  und seinem Vorwissen  $p_{n+1}(z)$  beeinflussen, wie sich der Konsens der Gruppe zusammensetzt. Mit dem Likelihood kann der Entscheider die Gruppenmitglieder bewerten. Dies ist ein ähnliches Problem wie bei der Wahl der Gewichte des LinOP und LogOP.

Zur Vereinfachung soll hier angenommen werden, dass der Entscheider die Beobachtungen  $d$  nicht kennt. Er hat als Entscheidungsgrundlage also nur die Meinungen der Gruppenmitglieder und das betrachtete Ereignis. Das Vorwissen des Entscheiders kann als nicht-informativ angenommen werden. Damit steht dem Entscheider keine zusätzliche Information zu Verfügung. Der Entscheider ist in diesem Fall keine Instanz mit zusätzlichem Wissen, welche die Gruppenmitglieder bewerten oder das Vorwissen erweitern könnte. Die Annahme des nicht-informativen Vorwissens ist beispielsweise dann sinnvoll, wenn der Entscheider virtuell ist. Beispielsweise könnte ein Gruppenmitglied oder die ganze Gruppe die Rolle des Entscheiders übernehmen.

Es kann beispielsweise angenommen werden, dass die Meinungen der Gruppenmitglieder bedingt unabhängig sind. In diesem Fall kann der Likelihood z.B. so gewählt werden (siehe Abschnitt 3.4):

$$p_{n+1}(p_1(z|d), \dots, p_n(z|d)|z) = \prod_{i=1}^n p_i(z|d)$$

Mit diesem Likelihood enthält sich der Entscheider bei der Bewertung der Gruppenmitglieder und übernimmt deren Meinungen ohne sie zu gewichten. Diese Wahl des Likelihood ist dann Sinnvoll, wenn der Entscheider keine Information darüber besitzt, wie er die Gruppenmitglieder gewichten sollte. Der SB hat mit diesem Likelihood die Form:

$$SB(p_1, \dots, p_n)(z) = \frac{p_{n+1}(p_1(z|d), \dots, p_n(z|d)|z)p_{n+1}(z)}{p_{n+1}(p_1(z|d), \dots, p_n(z|d))}$$

Damit wird die Annahme gemacht, dass der Entscheider nur die Wahrscheinlichkeitsverteilungen der Gruppenmitglieder kennt. Sein Pooling ist damit nicht von den Beobachtungen beeinflusst. Wegen der Annahme das die Wahrscheinlichkeitsverteilungen der Gruppenmitglieder bedingt unabhängig sind, kann der SB-Ansatz weiter vereinfacht werden:

$$SB(p_1, \dots, p_n)(z) \propto \prod_{i=1}^n p_{n+1}(p_i(z|d)|z)p_{n+1}(z)$$

Hat der Entscheider ein nicht-informatives Vorwissen und kann die Beobachtungen nicht verwenden, dann sollte er die Gruppenmitglieder auch nicht bewerten. Damit vereinfacht sich der SB-Ansatz auf:

$$SB(p_1, \dots, p_n)(z) \propto \prod_{i=1}^n p_i(z|d)$$

Das Ergebnis des SB-Ansatz hängt stark mit der Wahl des Likelihood und des Vorwissens des Entscheiders ab.

Um für Probleme in der Praxis einen Likelihood zu bestimmen, wird eine (bei frequentistischer Betrachtungsweise) ausreichend große Menge von Beobachtungen benötigt. Beim SB Ansatz sind die Beobachtungen (die Daten) die Wahrscheinlichkeitsverteilung der Gruppenmitglieder. Diese Beobachtungen können beispielsweise durch Experimente erzeugt werden. Der Aufwand kann abhängig von der Aufgabe jedoch sehr hoch sein.

## 5. Lösungsvorschläge für die lokale Bayes'sche Fusion

In Kapitel 3 wurden drei verschiedene Pooling-Methoden betrachtet. Diese Pooling-Methoden entstammen der Entscheidungstheorie und wurden hier für eine Anwendung in einer verteilten lokalen Konfiguration genutzt. Durch die Lokalisierung ist die Anwendung dieser Pooling-Methoden nicht mehr intuitiv nachvollziehbar. Der Begriff eines Konsens wurde in dieser Arbeit mit einer anderen Bedeutung belegt. Einen Konsens zu finden bedeutet in einer Gruppe, welche eine verteilte lokale Fusion bearbeitet, dass eine Wahrscheinlichkeitsverteilung über alle betrachteten Bereiche  $U_i$  gefunden wird, so dass die lokalen Wahrscheinlichkeitsverteilungen der Gruppenmitglieder möglichst gut repräsentiert werden. Dabei sollte diese Wahrscheinlichkeitsverteilung möglichst identisch einer Wahrscheinlichkeitsverteilung sein, die durch einen nicht verteilte und nicht lokale Bayes'sche Informationsfusion entstanden wäre. Wie diese Methoden aus der Entscheidungstheorie zu dem Problem der lokalen Bayes'schen Fusion passen, soll nun erörtert werden.

Der grundlegende Unterschied in beiden betrachteten Konfigurationen ist, ob auf  $U$  bzw. die  $U_i$  eingeschränkt, oder ob zur Lokalisierung vergrößert wird.

### 5.1. Globales a-priori-Wissen

Liegt das a-priori-Wissen global vor, dann ist  $p(z)$  bekannt. Ein  $p(z)$  wird nicht-informatives Vorwissen genannt, wenn gilt:  $p(z) = \frac{1}{|Z|}$ . Die A-posteriori-Wahrscheinlichkeitsverteilung kann somit bis auf einen Proportionalitätsfaktor lokal berechnet werden:

$$p_i(z|d) \propto \underbrace{p(d|z)p(z)}_{=: p_i^{nn}(z|d)}, \quad z \in U_i.$$

Nun können diese nicht normalisierten Ergebnisse der Gruppenmitglieder vereint werden. Dazu werden diese  $p_i^{nn}(z|d)$  ohne zu normalisieren übernommen. Erst das gesamte Ergebnis über ganz  $U$  wird dann normalisiert, so dass wieder eine Wahrscheinlichkeitsverteilung entsteht. Dies ergibt eine bezüglich  $U$  unverzerrte Lösung. Dass mehrere Gruppenmitglieder einen gleichen Bereich ansehen ist kein Problem, da sie dort auf die gleichen Ergebnisse kommen.

Die Forderung, dass die Gruppenmitglieder eine Wahrscheinlichkeitsverteilung als Ergebnis ihrer Fusion erzeugen, macht eine andere Kombination nötig. Dazu kann  $p_i^{nn}(z|d)$  wie folgt verwendet werden:

$$p(U_i|d, U) \propto \sum_{z \in U_i} p_i^{nn}(z|d)$$

berechnet werden. Die Gruppenmitglieder müssen dabei, zusätzlich zu der von ihnen erstellen Wahrscheinlichkeitsverteilung,  $p(U_i|d, U)$  mit liefern. Diese  $p(U_i|d, U)$  eignen sich gut als Gewichte für einen LinOP. Mit diesen Gewichten liefert der LinOP eine Lösung die im Bezug auf  $U$  unverzerrt ist.

In den meisten Fällen ist jedoch  $p(z)$  nicht bekannt und es muss eine andere Methodik gesucht werden, diese Gewichte zu bestimmen. In [San08] wird eine weitere Lösungsmöglichkeit vorgeschlagen: Eine *gemeinsame Referenz*.

## 5.2. Gemeinsame Referenz

Eine gemeinsame Referenz ist eine Menge  $R \subseteq Z$ , die in allen Umgebungen  $U_i$  enthalten ist. Ist dies nicht der Fall, kann ein beliebiges Ereignis  $z^r \in \bar{U}$  allen  $U_i$  hinzugefügt werden. Dieses dient als gemeinsame Referenz. Von nun an wird diese Referenz mit  $R$  bezeichnet bzw.  $z^r$ , wenn es sich um ein Elementarereignis handelt. Wie in [San08] beschrieben, kann durch eine gemeinsame Referenz der Schätzfehler der Gruppenmitglieder in ein Verhältnis zueinander gebracht werden. Ohne globales Vorwissen  $p(z)$  sieht die Fusion eines Gruppenmitglieds wie folgt aus:

$$p_i(z|d) \propto p(d|z)p_i(z).$$

Im Fall der Einschränkung ist  $\sum_{z \in U_i} p_i(z) = 1$ . Ist  $p_i(z)$  ein nicht-informatives Vorwissen dann ist  $p_i(z) = \frac{1}{|U_i|}$ , da das Gruppenmitglied als Vorwissen lediglich die Bedingung  $z \in U_i$  nutzt. Die hierdurch entstandene Einschränkung hat die Umverteilung der Wahrscheinlich-



keitsmasse auf  $U_i$  zur Folge. Dadurch kann nicht wie bei globalem Vorwissen  $p(U_i|d, U)$  als Gewicht für die Pooling-Methode verwendet werden, (siehe Abschnitt 5.1). Um dennoch eine Aussage bezüglich  $U$  treffen zu können, kann die gemeinsame Referenz verwendet werden.

Es gilt:

$$p_i(z) = \frac{p(z)}{p(U_i)}$$

Die Verzerrung der Wahrscheinlichkeit für  $z^r$  ist damit repräsentativ für ganz  $U_i$ . Es gilt:

$$\text{Verhältnis}(i, j) = \frac{p_i(z^r|d)}{p_j(z^r|d)} = \frac{\frac{p(z^r|d)}{p(U_i|d)}}{\frac{p(z^r|d)}{p(U_j|d)}} = \frac{p(U_j|d)}{p(U_i|d)} \quad (5.1)$$

Dabei ist  $i$  beliebig, aber fest und  $j = 1, \dots, n$ . Somit lässt sich die auf  $U := \bigcup_{i=1}^n U_i \setminus \{z^r\}$  eingeschränkt A-posteriori-Wahrscheinlichkeitsverteilung  $p^*(z|d) = \frac{p(z|d)}{p(U|d)}$  berechnen. Falls  $U_j \cap U_k = z^r$ ,  $j, k \in \{1, \dots, n\}$ , so gilt:

$$p^*(z|d) \propto \sum_{j=1}^n \text{Verhältnis}(i, j) p_j(z|d)$$

Um diese Idee zu verdeutlichen, wird hier ein Beispiel betrachtet: Fünf Agenten betrachten fünf nicht disjunkte Mengen von Ereignissen. Es wurde eine gemeinsame Referenz  $z^r$  bestimmt, die in  $\bar{U}$  liegt (hier ist dies  $z^1$ ). Jeder Agent bekommt zu seinem  $U_i$  die gemeinsame Referenz hinzugefügt. Der Agent betrachtet die Referenz wie jedes andere Ereignis. Die Wahrscheinlichkeitsverteilungen der Agenten werden danach, wie oben beschrieben, mit dem Verhältnis(i,j) gewichtet. Dabei ist  $i$  gleich Eins gewählt worden. Alle Meinungen der Gruppenmitglieder werden also mit dem des ersten Gruppenmitglieds ins Verhältnis gesetzt. Das Ergebnis eines LinOP ohne die Gewichtung durch die oben beschriebenen Verhältnisse, sondern mit gleichen Gewichte für alle Gruppenmitglieder, sieht wie in Abbildung 5.1 aus. Die gemeinsame Referenz wurde von den Gruppenmitgliedern lokal wie in Abbildung 5.2 bewertet.

Abbildung 5.3 zeigt das Ergebnis der Gewichtung der Agenten mittels dem aus der gemeinsamen Referenz entstandenen Verhältnisse (siehe (5.1)). Das Ergebnis dieser Gewichtung wurde Normalisiert, damit eine Wahrscheinlichkeitsverteilung entsteht. Die in Abbildung 5.3 dargestellt Ergebnis ist gleich dem einer Informationsfusion mit einer Einschränkung

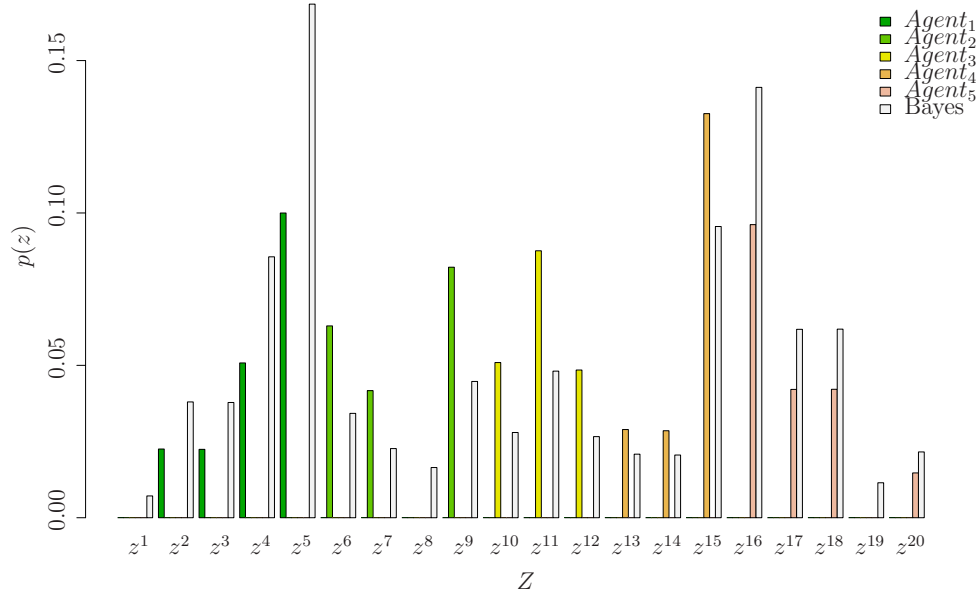


Abbildung 5.1.: Ergebnis eines LinOP mit gleichen Gewichten für alle Gruppenmitglieder.

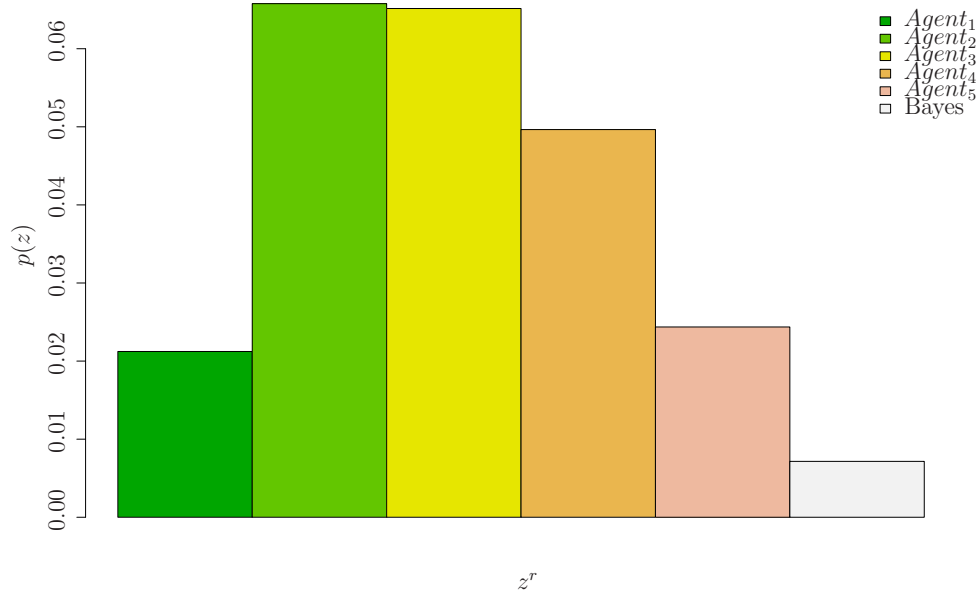
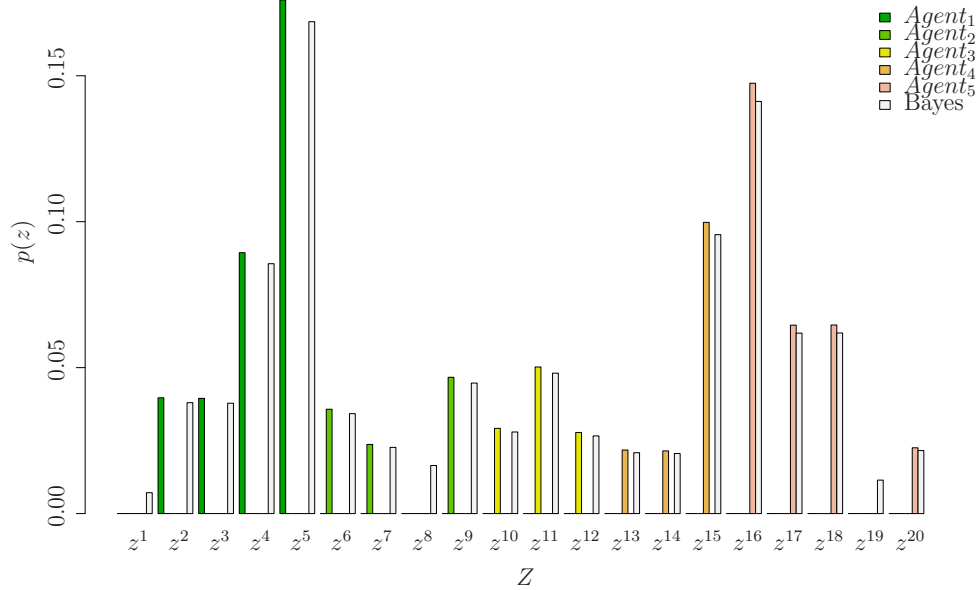


Abbildung 5.2.: Die Wahrscheinlichkeiten der Referenz ( $z^1$ ), die alle Gruppenmitglieder betrachten.



**Abbildung 5.3.:** Ergebnis einer Informationsfusion mit Gewichtung mittels einer gemeinsamen Referenz.

auf  $U$ . In Abbildung 5.3 ist ersichtlich, dass immer noch eine Verzerrung der Wahrscheinlichkeiten der Gruppenmitglieder vorliegt. Die zu dieser Abbildung passenden Werte sind gerundet in Tabelle B.4 dargestellt. Die Gruppenmitglieder untereinander stehen jedoch im richtigen Verhältnis. Sie überschätzen sich also alle um den gleichen Faktor. Würde beispielsweise das Ereignis  $z^{16}$  vor der Normalisierung mit der gemeinsamen Referenz noch als wahrscheinlicher bewertet als die Ereignisse  $z^{15}$  und  $z^5$ , so ist dieser Fehler nach der Normalisierung behoben.

### 5.3. Weitere Verfahren

Ist durch die Vorverarbeitung zusätzliche Information vorhanden, können unter gewissen Umständen Gewichte für eine Pooling-Methode daraus abgeleitet werden. Im Falle der Marginalisierung bleiben die Wahrscheinlichkeiten der betrachteten Ereignisse erhalten. Das  $i$ -te Gruppenmitglied berechnet für jedes  $z \in U_i$  gerade dessen globale A-posteriori-Wahrscheinlichkeit und zusätzlich noch  $p(U_i|d) = \sum_{z \in U_i} p(z|d)$ ,  $i = 1, \dots, n$ . Sind bei einer

Einschränkung zusätzlich zu den (verzerrten) lokalen A-posteriori-Wahrscheinlichkeitsverteilungen  $p_i(z|d)$  der Gruppenmitglieder auch noch die A-posteriori-Wahrscheinlichkeiten  $p(U_i|d)$  bekannt, so kann zur Kombination der  $p_i(z|d)$  ein LinOP mit den Gewichten  $w_i = p(U_i|d)$  verwendet werden. Man beachte, dass diese Situation genau genommen äquivalent zu einer Situation ist, in der mit einer Marginalisierung gearbeitet wird. Für den Fall der Marginalisierung ist also ein Lösungsansatz vorhanden, auch wenn die Berechnung von  $p(U_i|d)$ , abhängig von der Größe der  $U_i$ , aufwendig sein kann. Beim zweiten Fall wird  $p(U_i)$  als bekannt angenommen. Dies könnte beispielsweise der Fall sein, wenn eine Schätzung über die Wahrscheinlichkeitsmasse, die sich auf  $U_i$  verteilt, bekannt ist. In Abschnitt 4.3.1 wurde weiterhin besprochen, dass diese Information besonders gewinnbringend nutzbar ist, wenn die Pooling-Methode die EB erfüllt.

Im Fall der Einschränkung mit Zusatzwissen ist es möglich, eine unverfälschte Wahrscheinlichkeitsverteilung auf den  $U_i$ , wie in Gleichung (2.11) beschrieben, zu erzeugen. Hier ist eine Pooling-Methode unnötig, da die Wahrscheinlichkeitsverteilung des Gruppenmitglieds, das  $U_i$  betrachtet, einfach übernommen werden kann.

Für die Fälle, bei denen das globale a-priori Wissen, die gemeinsame Referenz oder  $p(U_i|d)$  gegeben ist, werden hier nicht weiter vertieft betrachtet, da für sie schon ein Lösungsansatz beschrieben wurde. Die drei betrachteten Pooling-Methoden (LinOP, LogOP und SB-Ansatz) sollen trotzdem auf deren Anwendbarkeit für die Fälle der Einschränkung mit Zusatzwissen und erweiternde Annahmen untersucht werden, um fest zu stellen, ob eine Verfälschung durch sie entsteht, oder ob durch sie auch eine optimale Lösung erzeugt werden kann. Der Schwerpunkt dieser Arbeit liegt auf der Betrachtung der Konfiguration mit einer Einschränkung. Hier gibt es zwei Fälle zu unterscheiden:

Der erste Fall, bei dem die  $U_i$  eine Partitionierung von  $U$  sind, ist wegen des möglichst großen Speedup durch die Lokalisierung eine berechtigte Konfiguration.

Der zweite Fall, bei dem die  $U_i$  nicht disjunkt sind, lässt sich auf den ersten Fall zurückführen: Ist der Schnitt von  $U_i$  und  $U_j$  mit  $i \neq j$  mit  $i, j \in \{1, \dots, n\}$  nicht leer, dann können diese beiden Mengen zu einer vereint werden. Aus dem Schnitt der beiden Mengen kann ein Ereignis als gemeinsame Referenz<sup>1</sup> gewählt werden. Dadurch kann eine neue Untermenge  $U_{ij}$  bestimmt werden. So lässt sich eine Menge von Untermengen von  $U$  erzeugen, deren Schnitt leer ist. Die Beschränkung auf einen disjunkten Fall, bei dem  $p(U_i|d)$  nicht

---

<sup>1</sup>Die gemeinsame Referenz dient dazu, die beiden Agenten in ein Verhältnis zueinander zu setzen. Für mehr Details siehe Abschnitt 5.2.

bekannt ist, ist also ausreichend.

## 5.4. Linear Opinion Pool

Das gewichtete arithmetische Mittel (LinOP) als Pooling-Methode kann lokal gut verwendet werden. Das Hauptproblem ist die Suche nach den Gewichten, welche die Verzerrung entzerren, die durch die Einschränkung erzeugt wurde. Der LinOP erzeugt eine mathematisch exakte Lösung für eine lokale Fusion, wenn  $p(U_i|d)$  bekannt ist (siehe Gleichung 4.6).

Der LinOP erzeugt jedoch kein optimales Ergebnis bei einer direkten Anwendung mit einer Konfiguration einer Einschränkung mit Zusatzwissen (siehe Gleichung (2.11)). Wird der LinOP als Pooling-Methode verwendet, dann kann hier ummodelliert werden.

Dann hat der LinOP folgende Form:

$$\begin{aligned} \text{LinOP}(p_1, \dots, p_n)(z) &= \sum_{i=1}^n w_i p_i^*(z) \\ &= \sum_{i=1}^n w_i p(z) I_{\{z \in U_i\}} + \sum_{i=1}^n w_i \frac{p(\bar{U}_i)}{|\bar{U}_i|} I_{\{z \in \bar{U}_i\}} \\ \text{mit } z &\in U_{i0}, i0 \in \{1, \dots, n\} \end{aligned}$$

Dabei ist die Funktion  $I_{\{z \in U_i\}}$  eine Indikatorfunktion:

$$I_{\{z \in U_i\}} := \begin{cases} 1, & z \in U_i \\ 0, & \text{sonst} \end{cases}$$

Der LinOP setzt sich hier aus zwei Summanden zusammen. Der erste Summand ist die unverzerrte Wahrscheinlichkeit  $p(z)$  von  $z$ , jedoch gewichtet mit der Summe der Gewichte der Gruppenmitglieder, welche  $z$  betrachten. Der zweite Summand ist die Summe der gewichteten Gleichverteilungen über  $\bar{U}_i$  der Gruppenmitglieder, die  $z$  nicht betrachten. Das Ergebnis des LinOP ist in diesem Fall durch den zweiten Summanden und durch die Gewichtung der unverzerrten Wahrscheinlichkeit  $p(z)$  von  $z$  verfälscht. Dieses kann durch eine bessere Wahl der Gewichte nicht mehr verbessert werden.

Ein weiterer Ansatz ist, das Vorwissen als global nicht-informativ anzunehmen. Mittels der Maximum-Entropie-Methode kann  $p(z) = \frac{1}{|Z|}$  angenommen werden. Für die lokale Betrachtung kann dann dementsprechend  $p(z) = \frac{1}{|U|}$  angenommen werden. Wie in Ab-

schnitt 5.1 beschrieben gibt es für den Fall von globalen a-priori Wissen eine Lösung. Die Bayes'sche Informationsfusion der Gruppenmitglieder hat dann folgende Form:

$$\begin{aligned}
 p_i(z|d) &= \frac{p(z)p(d|z)}{p(d)} \\
 &= \frac{p(z)p(d|z)}{\sum_{z \in U_i} p(z)p(d|z)} \\
 &= \frac{\frac{1}{|Z|}p(d|z)}{\sum_{z \in U_i} \frac{1}{|Z|}p(d|z)} \\
 &= \frac{p(d|z)}{\sum_{z \in U_i} p(d|z)}
 \end{aligned} \tag{5.2}$$

Dementsprechend hat der LinOP die Form:

$$\begin{aligned}
 LinOP(p_1, \dots, p_n)(z) &= \sum_{i=1}^n w_i \left[ \frac{p(d|z)}{\sum_{z \in U_i} p(d|z)} \right] \\
 &= \sum_{i=1}^n \underbrace{\frac{w_i}{\sum_{z \in U_i} p(d|z)}}_{=: \tilde{w}_i} p(d|z) \\
 &= \sum_{i=1}^n \tilde{w}_i p(d|z)
 \end{aligned} \tag{5.3}$$

Damit repräsentiert das Ergebnis des LinOP den Likelihood, welcher von den Beobachtungen für  $z$  erzeugt wird, da das Vorwissen aus der Berechnung herausfällt. Es kann analog argumentiert werden, wenn die Gruppenmitglieder  $|Z|$  nicht kennen und damit beispielsweise  $p(z) = \frac{1}{|U_i|}$  annehmen.

Werden  $p(U_i) = p(U_j)$  und  $|U_i| = |U_j|$ ,  $\forall i, j$  als einschränkende Annahmen hinzugenommen, dann hat der LinOP folgende Form:

$$LinOP(p_1, \dots, p_n)(z) = p(z) \sum_{i=1}^n w_i 1_{\{z \in U_i\}} + (1 - \sum_{i=1}^n w_i 1_{\{z \in U_i\}}) \frac{p(\bar{U}_{i0})}{|\bar{U}_{i0}|}, z \in U_{i0}, i0 \in \{1, \dots, n\}$$

Damit ist die Verfälschung zwar einfacher beschreibbar geworden, sie verschwindet al-

lerdings nicht. Diese Annahme erscheint dann sinnvoll, wenn, wie in dem Beispiel der Fahrzeugdetektion in der Szene des Luftbildes, die Bereiche um eine Auffälligkeit immer gleich groß gewählt werden können. Dann sind die zu betrachtenden möglichen Positionen ( $|U_i|$ ) des in diesem Bereich enthaltenen Fahrzeugs fast gleich.

Zusammenfassend lässt sich über den LinOP folgendes sagen:

- + Der LinOP kann ein optimales Ergebnis liefern.
- + Für die Verwendung des LinOP spricht, dass er intuitiv nachvollziehbar ist.
- + Die Gewichte sind gut interpretierbar.
- + Der LinOP erfüllt die SSFP. Damit muss zur Berechnung von  $LinOP(p_1, \dots, p_n)(z)$  nur die Wahrscheinlichkeitsverteilungen der Gruppenmitglieder in diesem  $z$  betrachtet werden.
- + Der LinOP erfüllt die MP. Dies kann als Vorteil interpretiert werden, da zur Berechnung des LinOP nicht mehr zwangsweise die atomaren Ereignisse betrachtet werden müssen.
- Der LinOP erfüllt nicht die EB. Das bedeutet, dass bei neuer Information entweder zentral gerechnet werden muss oder die Gewichte neu bestimmt werden müssen.
- Der LinOP erhält keine Unabhängigkeiten. Dadurch werden von den Gruppenmitglieder als unabhängige gesehene Ereignisse im Konsens abhängig.
- Die optimalen Gewichte sind außer in Sonderfällen schwer berechenbar.

## 5.5. Logarithmischer Opinion Pool

Das gewichtete geometrische Mittel (LogOP) als Pooling-Methode zu verwenden, kann sinnvoll sein. Der LogOP muss jedoch für die Anwendung in einer lokalen Bayes'sche Fusion modifiziert werden. Er hat dann nämlich unter Umständen keine Wahrscheinlichkeitsverteilung als Lösung. Beispielsweise ist in einer lokalen Konfiguration, bei der auf die  $U_i$  eingeschränkt wird und die  $U_i$  ein Partition von  $U$  bilden, das Ergebnis des LogOP überall gleich Null.

Die Modifikation besteht darin, dass ein Gruppenmitglied nur innerhalb des von ihm betrachteten Bereiches  $U_i$  ein Veto aussprechen darf. Für alle Ereignisse in  $\bar{U}_i$  wird seine Schätzung ignoriert. Ein solcher LogOP wird hier *modifizierter LogOP* (modLogOP) genannt.

$$\text{modLogOP}(p_1, \dots, p_n)(z) = \frac{\prod_{i=1}^n p_i(z)^{w_i}}{\sum_{z \in U} \prod_{i=1}^n p_i(z)^{w_i}}, \quad w_i \geq 0, \quad \sum_{i=1}^n w_i = 1 \quad (5.4)$$

Zur Erläuterung folgt ein Beispiel: Es seien die Ereignisse von Interesse in  $Z = \{z^1, z^2, z^3\}$  zusammengefasst. Zwei Gruppenmitglieder  $Agent_1$  und  $Agent_2$  betrachten jeweils zwei Ereignisse.  $Agent_1$  betrachtet  $U_1 = \{z^1, z^2\}$  und  $Agent_2$  betrachtet  $U_2 = \{z^2, z^3\}$ . Die Wahrscheinlichkeitsverteilungen der Agenten sind in Tabelle 5.1 zu sehen.

	$Agent_1$	$Agent_2$
$z^1$	0.5	0.0
$z^2$	0.5	0.2
$z^3$	0.0	0.8

**Tabelle 5.1.:** Beispiel für Meinungen von zwei Gruppenmitgliedern, wobei  $Agent_1$   $z^3$  und  $Agent_2$   $z^1$  nicht betrachtet.

Werden diese Meinungen mit den Gewichten  $w_i = \frac{1}{2}, i \in \{1, 2\}$  mit einem LogOP kombiniert, wird  $z^2$  die Wahrscheinlichkeit Eins zugeordnet, da dies das einzige Ereignis ist, welches beide Agenten betrachten und somit auch das einzige, gegen das kein Veto von einem der Gruppenmitglieder erhoben wurde. Werden die Meinungen der Gruppenmitglieder jedoch mit den gleichen Gewichten und dem modLogOP kombiniert, dann erhalten wir folgendes Resultat:  $\text{modLogOP}(Agent_1, Agent_2)(z^1) \approx 0.3687, \text{modLogOP}(Agent_1, Agent_2)(z^2) \approx$



0.1649 und  $\text{modLogOP}(Agent_1, Agent_2)(z^3) \approx 0.4664$ .

Dieses Resultat lässt sich wie folgt interpretieren: Das Ereignis  $z^3$  ist in der Konsensverteilung das wahrscheinlichste Ereignis, da ihm alle Gruppenmitglieder die  $z^3$  betrachtet haben, eine hohe Wahrscheinlichkeit zuordnen. Da nur  $Agent_1$  das Ereignis  $z^1$  betrachtet, dieses jedoch nicht für so wahrscheinlich hält wie  $Agent_2$  das Ereignis  $z^3$ , ist es am zweitwahrscheinlichsten. Das Ereignis  $z^2$  ist im Konsens das unwahrscheinlichste Ereignis, da  $Agent_2$  es für unwahrscheinlich hält. Die Ereignisse  $z^1$  und  $z^3$  werden von allen Gruppenmitgliedern (die sie betrachten, die anderen Gruppenmitglieder werden ja ignoriert) als wahrscheinlicher gesehen als  $z^2$ . Da  $Agent_1$  den Ereignissen  $z^1$  und  $z^2$  die gleiche Wahrscheinlichkeit zuordnet wird durch  $Agent_2$  die Wahrscheinlichkeit von  $z^2$  im Konsens nach unten verschoben. Durch die Normalisierung wird  $\text{modLogOP}(Agent_1, Agent_2)(z^2)$  dann unwahrscheinlicher als das Mittel der beiden Gruppenmitglieder.

Durch diese Modifikation wird der modLogOP in einer lokalen Konfiguration anwendbar. Die Interpretation der Gewichte ist bei einem modLogOP dieselbe wie bei einem LinOP. Sie stellt eine Art Wertung für die Expertise eines Gruppenmitglieds dar. Die Modifikation vom LogOP zum modLogOP kann auch durch eine Modifikation der Gewichte erreicht werden. Dabei werden die Gewichte  $w_i(z)$  auf Null gesetzt, wenn das betrachtete Ereignis nicht von dem Agenten betrachtet wird. Damit würden die Gewichte von den betrachteten Ereignissen abhängen.

	$Agent_1$	$Agent_2$		$p(d z^i)$
$z^1$	0.3	0.0	$z^1$	0.3
$z^2$	0.7	0.0	$z^2$	0.2
$z^3$	0.0	0.0	$z^3$	0.0
$z^4$	0.0	0.6	$z^4$	0.4
$z^5$	0.0	0.4	$z^5$	0.1

(a) Beispiel für Meinungen von zwei Gruppenmitgliedern, wobei die  $U_i$  disjunkt sind.

(b) Der Likelihood, der von den Beobachtungen  $d$  erzeugt wird.

**Tabelle 5.2.:** Beispieldaten für den Beweis, dass der modLogOP die EB nicht erfüllt.

Der modLogOP liefert zwar für den lokalen Fall ein Ergebnis, das etwas sinnvoller ist als

## 5. Lösungsvorschläge für die lokale Bayes'sche Fusion

---

das eines LogOP, er hat jedoch auch Nachteile. Der modLogOP erfüllt nicht die EB, was der größte Vorteil des LogOP für die verteilte lokale Anwendung ist. Um dies zu beweisen, soll hier ein Gegenbeispiel gezeigt werden:

Zwei Gruppenmitglieder,  $Agent_1$  und  $Agent_2$  betrachten jeweils zwei Ereignisse aus  $Z = \{z^1, z^2, z^3, z^4, z^5\}$ .  $Agent_1$  betrachtet  $U_1 = \{z^1, z^2\}$  und  $Agent_2$  betrachtet  $U_2 = \{z^4, z^5\}$ . Die Wahrscheinlichkeitsverteilungen der Agenten werden in Tabelle 5.2(a) dargestellt.

	$modLogOP(z^i d)$
$z^1$	0.2309
$z^2$	0.2879
$z^3$	0.0
$z^4$	0.3417
$z^5$	0.1395

	$modLogOP(z^i d)$
$z^1$	0.2332
$z^2$	0.2374
$z^3$	0.0
$z^4$	0.4397
$z^5$	0.0897

(a) Das Ergebnis eines modLogOP wenn zuerst nach  $d$  bedingt wurde und dann der modLogOP angewandt wurde.

(b) Das Ergebnis eines modLogOP wenn zu zuerst der modLogOP angewandt wurde und dann nach  $d$  bedingt.

**Tabelle 5.3.:** Ergebnis des Gegenbeispiels dafür, dass der modLogOP die EB nicht erfüllt.

Die Gewichte des modLogOP seien  $w_i = \frac{1}{2}, i \in \{1, 2\}$ . Die Beobachtungen  $d$  erzeugen einen Likelihood für die betrachteten Ereignisse, wie in Tabelle 5.2(b) dargestellt.

Bedingen die beiden Agenten zuerst nach den Beobachtungen und wird anschließend der modLogOP angewendet, ergibt sich näherungsweise ein Ergebnis wie in Tabelle 5.3(a) dargestellt.

Wird jedoch zuerst der modLogOP angewendet und danach nach den Beobachtungen bedingt, ergibt sich näherungsweise ein Ergebnis wie in Tabelle 5.3(b) dargestellt. Die Ergebnisse aus Tabelle 5.3(a) und 5.3(b) stimmen nicht überein. Daher erfüllt der modLogOP die EB nicht. Der Vorteil, vor der Fusion mit der Likelihood der Beobachtungen das Pooling durch zu führen, den der LogOP gegenüber des LinOP hatte, ist damit nicht auf den modLogOP übertragbar. Auch die Gewichte können nicht vor der Bedingung durch die Beobachtungen gewählt werden.

Beim Vorliegen einer Einschränkung mit Zusatzwissen ist der LogOP direkt anwendbar. Wie beim LinOP bleibt das Ergebnis des LogOP jedoch verzerrt, auch wenn  $p(U_i)$  bekannt ist, da der LogOP dann folgende Form annimmt:

$$\begin{aligned}
 \text{LogOP}(p_1^*, \dots, p_n^*)(z) &= \frac{\prod_{i=1}^n p_i^*(z)^{w_i}}{\sum_{z \in U} \prod_{i=1}^n p_i^*(z)^{w_i}} \\
 &= \prod_{i=1}^n p_i^*(z)^{w_i} \cdot \frac{\prod_{i=1}^n p_i^*(z)^{w_i}}{\sum_{z \in U} \prod_{i=1}^n p_i^*(z)^{w_i}} \\
 &= \frac{\prod_{i=1}^n p(z)^{w_i} \prod_{i=1}^n \left( \frac{p(\bar{U}_i)}{|U_i|} \right)^{w_i}}{\sum_{z \in U} \left[ \prod_{i=1}^n p(z)^{w_i} \prod_{i=1}^n \left( \frac{p(\bar{U}_i)}{|U_i|} \right)^{w_i} \right]} \\
 &= \frac{p(z)^{\sum_{i=1}^n w_i} \prod_{i=1}^n \left( \frac{p(\bar{U}_i)}{|U_i|} \right)^{w_i}}{\sum_{z \in U} \left[ p(z)^{\sum_{i=1}^n w_i} \cdot \prod_{i=1}^n \left( \frac{p(\bar{U}_i)}{|U_i|} \right)^{w_i} \right]}
 \end{aligned}$$

Der Zähler des LogOP setzt sich also aus zwei Faktoren zusammen. Der Erste stellt die gewichtete (ohne Gewichtung unverzerrte) Wahrscheinlichkeit von  $z$  dar. Der Zweite stellt das Produkt der gewichteten Gleichverteilung  $\frac{p(\bar{U}_i)}{|U_i|}$  der restlichen Gruppenmitglieder, in deren Umgebung  $U_i$  das betrachtete  $z$  nicht enthalten ist, dar.

In Folgenden wird die Annahme gemacht, dass  $p(U_i) = p(U_j)$  und  $|U_i| = |U_j|$  gilt. Sind die  $U_i$  disjunkt und werden alle Gewichte des LogOP als  $w = \frac{1}{n}$  gewählt, dann kann der

LogOP wie folgt formuliert werden:

$$\begin{aligned}
 \text{LogOP}(p_1, \dots, p_n)(z) &= \frac{\prod_{i=1}^n p_i^*(z)^w}{\sum_{z \in U} \prod_{i=1}^n p_i^*(z)^w} \\
 &= \frac{p(z)^w \cdot \left(\frac{p(\bar{U}_1)}{|U_1|}\right)^{1-w}}{\sum_{z \in U} \left[ p(z)^w \cdot \left(\frac{p(\bar{U}_1)}{|U_1|}\right)^{1-w} \right]} \\
 &= \frac{p(z)^w}{\sum_{z \in U} p(z)^w}
 \end{aligned}$$

Dieses Ergebnis ist durch die Gewichtung von  $p(z)$  mit dem Gewicht  $w$  verfälscht.

Eine Möglichkeit hier eine unverzerrte Wahrscheinlichkeitsverteilung über  $U$  zu erzeugen besteht offensichtlich darin, die Gewichte des LogOP alle gleich Eins zu wählen. Man beachte, dass diese Wahl der Gewichte gemäß der Definition eines LogOP nicht zulässig ist. In diesem Fall erscheint dies jedoch eine sinnvolle Möglichkeit.

Zusammenfassend kann über den LogOP folgendes gesagt werden:

- + Da der LogOP die EB erfüllt, ist er für die lokale Bayes'sche Fusion eine attraktive Pooling-Methode.
- Gegen die Verwendung des LogOP spricht, dass wenn es ein Veto gibt, wenn ein Gruppenmitglied ein Ereignis für unmöglich ( $p_i(z) = 0$ ) hält.
- Es ist eine Modifikation des LogOP nötig, um ihn in einer lokalen Konfiguration mit Einschränkung als Lokalisierung einsetzen zu können.
- Der modLogOP erfüllt nicht die EB.

## 5.6. Supra Bayesian

Wie in Kapitel 4.5 beschrieben ist es schwer, in der Praxis ein Likelihood für die Meinungen der Gruppenmitglieder zu erstellen. In Kapitel 4.5 wurde ein Likelihood vorgestellt, durch dessen Verwendung die Konsenswahrscheinlichkeit exakt die Meinungen der Gruppenmitglieder darstellen kann. Dieser Likelihood kann so interpretiert werden, dass der Entscheider nur die Meinungen der Gruppenmitglieder betrachtet, nicht aber die Beob-

achtungen. Dann kann er ohne weiteres Vorwissen die Gruppenmitglieder nicht werten. Somit kann der Likelihood wie folgt gewählt werden:

$$p_{n+1}(p_1(z|d), \dots, p_n(z|d)|z) = \prod_{i=1}^n p_i(z|d).$$

Dabei tritt die gleiche Überbewertung der Wahrscheinlichkeit Null wie beim LogOP auf. Ist ein  $p_i(z|d) = 0$ , wird der gesamte Likelihood Null. Schränkt jeder Agent auf sein  $U_i$  ein, so ist der Likelihood im Fall der Partitionierung immer Null. Für den Fall, in dem die  $U_i$  disjunkt sind, ist der SB-Ansatz mit diesem Likelihood in einer Konfiguration mit einer Einschränkung als Lokalisierung nicht verwendbar.

Ist  $p(U_i|d)$  bekannt, kann mit  $p_i^*(z|d)$  gerechnet werden, (siehe Gleichung (2.11)). Der Likelihood könnte beispielsweise wie nachfolgend gewählt werden:

$$p_{n+1}(p_1^*(z|d), \dots, p_n^*(z|d)|z) = \prod_{\substack{i=1 \\ z \in U_i}}^n p_i^*(z|d)$$

Der SB hat dann die Form

$$SB(p_1^*, \dots, p_n^*)(z|d) = \frac{\prod_{\substack{i=1 \\ z \in U_i}}^n p_i(z|d)p_{n+1}(z)}{\sum_{z \in U} \prod_{\substack{i=1 \\ z \in U_i}}^n p_i(z|d)p_{n+1}(z)}.$$

Für disjunkte  $U_i$  und ein nicht-informatives Vorwissen des Entscheiders ergibt sich dann:

$$SB(p_1^*, \dots, p_n^*)(z|d) = \frac{p(z|d)}{\sum_{z \in U} p(z|d)}$$

Unter diesen Bedingungen liefert der SB also eine auf  $U$  unverzerrte Lösung.

Zusammenfassend lässt sich über den SB-Ansatz sagen:

- + Der SB-Ansatz kann eine optimale Lösung liefern.
- Der Likelihood ist schwer zu wählen.
- Je nach Wahl des Likelihood macht der SB-Ansatz für die lokale Verwendung keinen Sinn.

## 6. Zusammenfassung und Ausblick

Bei einer lokalen Bayes'schen Informationsfusion werden aus allen möglichen PoI Bereiche von Interesse, mittels einer Vorverarbeitung extrahiert. Im Fall einer Einschränkung wird die Fusion ausschließlich auf diesen Bereichen durchgeführt. Diese lokale Betrachtung hat eine Aufwandsreduktion zur Folge, da weniger PoI betrachtet werden. Jeder Bereich wird dann von einem sogenannten Agenten betrachtet. Die Agenten können verteilt ausgeführt werden, was einen Speedup zur Folge haben kann. Dies kann mit einer kriminalistischen Ermittlung verglichen werden, bei der verschiedene Mitarbeiter (die Agenten) disjunkte Bereiche der Verdächtigen untersuchen. Am Ende einer Untersuchung müssen die Ergebnisse dieser Mitarbeiter dann zusammengeführt werden, um ein Urteil bilden zu können. Die Fusionsagenten können, wie in dem Beispiel aus der Kriminalistik, als eine Gruppe (Team) aufgefasst werden, welche zusammenarbeitet um die Fusionsaufgabe zu lösen.

Da jeder Agent nur seinen lokalen Bereich betrachtet, ergeben sich Fehler in seinen Berechnungen der A-posteriori-Wahrscheinlichkeitsverteilung. Durch das Einschränken der Wahrscheinlichkeitsverteilung auf den von ihm betrachteten Bereich, erhöht jeder Agent die Wahrscheinlichkeit der von ihm betrachteten Ereignisse.

Die verteilten Ergebnisse zu einer globalen Sicht zusammen zu führen, ist Aufgabe einer Pooling-Methode. Die Fachliteratur wurde nach möglichen Pooling-Methoden der Entscheidungstheorie durchsucht. Drei Pooling-Methoden, der LinOP, der LogOP und der SB-Ansatz, wurden näher auf ihre Nutzbarkeit für eine verteilte lokale Bayes'sche Informationsfusion untersucht.

Um die Ergebnisse einer Pooling-Methode zu bewerten, werden in der Fachliteratur Eigenschaften definiert, welche eine Pooling-Methode erfüllen kann oder soll. Auf diese Eigenschaften können die verschiedenen Pooling-Methoden überprüft werden.

In dieser Arbeit wurde unter anderem untersucht, wie gut sich Pooling-Methoden aus der Entscheidungstheorie für das Zusammenführen der durch eine verteilten lokalen Bayes'schen

---

Informationsfusion erzeugte Wahrscheinlichkeitsverteilungen, verwenden lassen. Dadurch, dass die verteilt berechneten Ergebnisse wie Meinungen von Gruppenmitglieder interpretiert wurden, können theoretisch gut untersuchte Pooling-Methoden aus dem Forschungsbereich Entscheidungstheorie benutzt werden, um die verteilt berechneten Ergebnisse der lokalen Bayes'schen Informationsfusion zu einer globalen Wahrscheinlichkeitsverteilung zusammen zu führen.

Werden die Eigenschaften nach ihrer Brauchbarkeit für den verteilten lokalen Einsatz gewertet, können in vielen Fällen auch die Pooling-Methoden, die diese Eigenschaften erfüllen können, ebenfalls gewertet werden. Es ist jedoch, wie am Beispiel des LogOP festgestellt, nicht ausreichend nur diese Eigenschaften zu überprüfen. Das Ergebnis der Pooling-Methode muss auch mit der *Ground Truth* verglichen werden.

Ein Beispiel für eine Eigenschaft, die in einer verteilten lokalen Konfiguration mit einer Einschränkung als Lokalisierung nicht wünschenswert ist, ist die UNAM. Bei ihr bleiben die Fehler der Wahrscheinlichkeitsverteilung, die von der Gruppe gemeinschaftlich gemacht werden, erhalten.

Die EB beispielsweise ist für eine lokale verteilte Informationsfusion eine sinnvoll zu wählende Eigenschaft. Sie fordert von der Pooling-Methode, dass die Reihenfolge der Anwendung der Pooling-Methode und der Informationsfusion mit den Beobachtungen unerheblich ist. Dadurch ist es möglich die Gewichte einer Pooling-Methode zu wählen, bevor die Beobachtungen in die Inferenz mit einfließen. So könnte beispielsweise  $p(U_i)$  anstatt  $p(U_i|d)$  als Gewichte gewählt werden, was eine Aufwandsreduktion in der Berechnung der Gewichte zur Folge hat. Zusätzlich kann, bei neuen Beobachtungen, direkt das Ergebnis der Pooling-Methode mittels einer Bayes'schen Informationsfusion aktualisiert werden. Es müssen nicht mehr alle Gruppenmitglieder die neuen Beobachtungen betrachten, die Informationsfusion lokal durchführen und diese Ergebnisse dann wieder zusammengeführt werden. Dies kann gegebenenfalls von Vorteil sein. Der LogOP erfüllt die EB. Er erzeugt jedoch bei disjunkten  $U_i$  ein Ergebnis, das überall gleich Null ist.

Im Unterschied zu einem klassischen Gruppenentscheid kann in dem hier betrachteten Kontext die *Ground Truth* bestimmt werden. Somit kann ein Ergebnis einer Pooling-Methode dadurch bewertet werden, wie weit es sich von der „wahren“ Wahrscheinlichkeitsverteilung unterscheidet. Dabei wurde herausgefunden, dass eine Einschränkung mit Zusatzwissen bei einer direkten Anwendung der betrachteten Pooling-Methoden, keine Verbesserung gegenüber der Einschränkung erbringt.

Das Zusammenführen der Ergebnisse der Gruppenmitglieder zu einer Wahrscheinlichkeitsverteilung entspricht nicht dem, was in der Entscheidungstheorie ein Gruppenentscheid ist. Dies liegt daran, dass für einen optimalen Speedup die Gruppenmitglieder bei einer verteilten lokalen Informationsfusion disjunkte Ereignisse betrachten. Im Gegensatz dazu betrachten die Gruppenmitglieder in der Entscheidungstheorie alle das gleiche Ereignis und versuchen ihre Meinungen, über dieses Ereignis, zu einem Konsens zu vereinen. Nicht alle Pooling-Methoden aus der Entscheidungstheorie lassen sich auf den lokalen Kontext übertragen.

Ein LogOP beispielsweise ist im verteilten lokalen Kontext nicht anwendbar und muss modifiziert werden, damit er auch bei disjunkten  $U_i$  eine Wahrscheinlichkeitsverteilung als Ergebnis liefert. Der SB-Ansatz ist sehr kompliziert. Die Auswahlmöglichkeiten für den Likelihood des Entscheiders und dessen Vorwissen sind so groß, dass in dieser Arbeit keine allgemeingültige Analyse möglich war. Unter den Bedingungen, dass die  $p(U_i|d)$  bekannt sind, dass das Vorwissen des Entscheiders nicht informativ ist und dass die  $U_i$  disjunkt sind, liefert der SB-Ansatz mit einem geeigneten Likelihood ein unverzerrtes Ergebnis, welches gleich dem nicht verteilt berechneten Ergebnis ist. Der LinOP kann jedoch bei Verwendung von geeigneten Gewichten ( $w_i = p(U_i|d)$ ) eine Lösung erzeugen, die gleich der eines einzelnen Bayes'schen Agenten ist.

Die Bewertung der Ergebnisse solcher Pooling-Methoden ist nicht immer einfach. In dieser Arbeit konnte nur unter sehr restriktiven Annahmen Pooling-Methoden bestimmt werden, die ein Ergebnis liefern, welches mit dem nicht verteilt berechneten Ergebnis übereinstimmen. Der LinOP, welcher bei optimalen Gewichten eine mathematisch korrekte Lösung liefern könnte, benötigt für die Wahl dieser Gewichte einen sehr hohen Aufwand. Der Aufwand entsteht, da für die Gewichte die globale A-priori-Wahrscheinlichkeiten der Bereiche der Agenten berechnet werden müssen.

Dass unter gewissen Bedingungen der LinOP und der SB-Ansatz aus der Entscheidungstheorie lokal anwendbar sind, sollte als Motivation verstanden werden, weitere solcher Pooling-Methoden auf die Eignung für die verteilte lokale Bayes'sche Informationsfusion zu überprüfen.

In dieser Arbeit wurde die Betrachtung auf das Vorliegen diskreter Wahrscheinlichkeitsverteilungen beschränkt. Zukünftige Arbeiten könnten den Fall von kontinuierlichen Wahrscheinlichkeitsverteilungen untersuchen. Die hier betrachteten Gruppenmitglieder bestehen aus identischen Agenten, da Vorwissen und sämtliche Beobachtungen zur Verfügung



---

stehen. Es sind Szenarien denkbar, bei denen die einzelnen Agenten nur einen Teil der Beobachtungen zur Verfügung steht, oder die Agenten verschiedene Aufgaben übernehmen oder verschiedene Aspekte des gleichen Ereignisses untersuchen. Die Agenten hätten dann verschiedene Expertisen und würden beispielsweise verschiedene Informationsquellen betrachten oder verschiedene Klassen von PoI detektieren. Beispielsweise könnte ein Agent für das Erkennen von Personen in Bildern erstellt worden sein, während ein Anderer Fahrzeuge detektiert und ein Dritter die Art des Untergrunds bestimmt.

Des Weiteren wurde in dieser Arbeit jede Art der Interaktion zwischen den Gruppenmitgliedern nicht betrachtet, was bei der Analogie der kriminalistischen Untersuchung so gut wie undenkbar erscheint. Die Gruppenmitglieder könnten beispielsweise ihre Meinungen austauschen, um die eigene Meinung, im Angesicht der Meinungen der anderen Gruppenmitglieder, anzupassen [DeG74]. Bacharach betrachtet in [Bac79] beispielsweise, wie ein Bayes'scher Dialog zu einem Konsens führen kann. Dieser stellt eine Art Diskussion zwischen den Gruppenmitgliedern dar, wobei jedes Gruppenmitglied seine Meinung entsprechend den Meinungen der anderen Gruppenmitglieder anpasst, bis ein Konsens gefunden wurde.

Verwendet die Pooling-Methode Gewichte, müssen diese nicht von Anfang an unveränderbar sein. Für die Bestimmung der Gewichte sind iterative Verfahren denkbar. Dabei würden die Gruppenmitglieder beispielsweise bezüglich vergangener Leistung bewertet [Gen86b, Coo91] und [Bed01].

Zwei (LinOP und LogOP) der drei betrachteten Pooling-Methoden sind Mittelwertfunktionen. Weitere Mittel, wie das gewichtete harmonische Mittel, könnten auch noch auf ihre Brauchbarkeit in einem lokalen Kontext überprüft werden. Denkbar ist auch eine Variation der Bedingungen für die Gewichte, wie dies in [DeG91] dargestellt wird.

# Literaturverzeichnis

- [Acz80] J. Aczel und C. Wagner: „A characterization of weighted arithmetic means“. *SIAM Journal on Algebraic and Discrete Methods* **1** (3), S. 259–260, 1980.
- [Ait91] A. Aitken, C. Aitken und D. A. Stoney: *The Use of Statistics in Forensic Science*. Crc Pr Inc, November 1991.
- [Bac75] M. Bacharach: „Group Decisions in the Face of Differences of Opinion“. *Management Science* **22** (2), S. 182–191, 1975.
- [Bac79] M. Bacharach: „Normal Bayesian Dialogues“. *Journal of the American Statistical Association* **74** (368), S. 837–846, 1979.
- [Bed01] T. Bedford: *Probabilistic Risk Analysis Foundations and Methods*. Cambridge University Press, 2001.
- [Ber86] L. Berenstein, Kanal: „Consensus Rules“. *Uncertainty in Artificial Intelligence* S. 27–32, 1986.
- [Ber04] A. F. M. Bernardo, Jose M. ; Smith: *Bayesian theory*. Wiley series in probability and statistics. Wiley, Chichester [u.a.], Repr.. Aufl., 2004.
- [Bey99] Beyerer: *Verfahren zur quantitativen statistischen Bewertung von Zusatzwissen in der Messtechnik*. VDI Fortschritt-Berichte Reihe 8, Nr. 783, VDI Verlag, 1999., Universitaet Friderician Karlsruhe (TH), 1999.
- [Bey07] J. Beyerer, J. Sander und S. Werling: „Bayes’sche Methodik zur lokalen Fusion heterogener Informationsquellen“. *tm Technisches Messen* **74** (3), S. 103–111, 2007.
- [Blo98] V. Blobel und E. Lohrmann: *Statistische und numerische Methoden der Datenanalyse*. Teubner Verlag, 1. Aufl., 1998.
- [Bra95] R. Brause: *Neuronale Netze : eine Einfuehrung in die Neuroinformatik*. Leit-

- faeden der Informatik. Teubner, Stuttgart, 1995.
- [Coo91] R. Cooke: *Experts in uncertainty : opinion and subjective probability in science*. Environmental ethics and science policy series. Oxford Univ. Pr., New York [u.a.], 1991.
- [DeG74] M. H. DeGroot: „Reaching a Consensus“. *Journal of the American Statistical Association* **69**, S. 118–121, 1974.
- [DeG91] M. H. DeGroot und J. Mortera: „Optimal Linear Opinion Pools“. *Management Science* **37** (5), S. 546–558, 1991.
- [Eck06] S. Eckel: *Einsatz des Bayes-Theorems zur Erkennungsunterstuetzung bei der Satellitenbildauswertung*. Diplomarbeit am Fraunhofer IITB, 2006.
- [Fre80] S. French: „Updating of Belief in the Light of Someone Else’s Opinion“. *Journal of the Royal Statistical Society. Series A (General)* **143** (1), S. 43–48, 1980.
- [Fre81] S. French: „Consensus of opinion“. *European Journal of Operational Research* **7** (4), S. 332 – 340, 1981.
- [Fre85] S. French: „Group Consensus Probability Distributions A Critical Surcey“. *Bayesian Statistics 2* **2**, S. 183–202, 1985.
- [Gal93] S. I. Gallant: *Neural network learning and expert systems*. A Bradford book. MIT Pr., Cambridge, Mass. [u.a.], 1993.
- [Gel95] A. E. Gelfand, B. K. Mallick und D. K. Dey: „Modeling Expert Opinion Arising as a Partial Probabilistic Specification“. *Journal of the American Statistical Association* **90** (430), S. 598–604, 1995.
- [Gen84a] Genest, Weerahandi und Zidek: „Aggregating Opinions Through Logarithmic Pooling“. *Theory and Decision* **17**, S. 61–70, 1984.
- [Gen84b] C. Genest: „A Characterization Theorem for Externally Bayesian Groups“. *The Annals of Statistics* **12** (3), S. 1100–1105, 1984.
- [Gen84c] C. Genest: „A conflict between two axioms for combining subjective distributions“. *Journal of the Royal Statistical Society* 1984.
- [Gen84d] C. Genest: „Pooling Operators with the Marginalization Property“. *The Canadian Journal of Statistics / La Revue Canadienne de Statistique* **12** (2), S.

- 153–163, 1984.
- [Gen86a] Genest, K. J. McConway und M. J. Schervish: „Characterization of Externally Bayesian Pooling Operators“. *The Annals of Statistics* **14** (2), S. 487–501, 1986.
- [Gen86b] C. Genest und J. V. Zidek: „Combining Probability Distributions: A Critique and an Annotated Bibliography“. *Statistical Science* **1** (1), S. 147–148, 1986.
- [Gen87] Genest und Wagner: „Further evidence against independence preservation in expert judgement synthesis“. *Aequationes Mathematicae* **32** (1), S. 74–86, Dezember 1987.
- [Gen90] C. Genest und K. J. McConnway: „Allocating the Weights in the Linear Opinion Pool“. *Journal of Forecasting* **9**, S. 53–73, 1990.
- [Giv99] G. H. Givens und P. J. Roback: „Logarithmic Pooling of Priors Linked by a Deterministic Simulation Model“. *Journal of Computational and Graphical Statistics* **8** (3), S. 452–478, 1999.
- [Gut05] A. Gut: *Probability: a graduate course*. Springer Texts in Statistics. Springer, New York, NY, 2005.
- [Hei08] M. Heizmann: *Skriptum zur Vorlesung Einfuehrung in die Informationsfusion*. Institut fuer Technische Informatik Lehrstuhl fuer Interaktive Echtzeitsysteme Fakultaeet fuer Informatik Universitaet Karlsruhe (TH), August 2008.
- [Hel90] R. Helge, M. Thomas und S. Klaus: *Neuronale Netze : eine Einfuehrung in die Neuroinformatik selbstorganisierender Netzwerke*. Addison-Wesley, Bonn, 1990. Literaturverz. S. 247 - 251.
- [Hes98] T. Heskes: *Selecting Weighting Factors in Logarithmic Opinion Pools*. In: *Advances in Neural Information Processing Systems*, Bd. 10, S. 266–272, Advances in Neural Information Processing Systems, 1998.
- [Jac95] R. A. Jacobs: „Methods for combining experts’ probability assessments“. *Neural Comput.* **7** (5), S. 867–888, 1995.
- [Kas09] C. Kascha und F. Ravazzolo: *Combining Inflation Density Forecasts*. Techn. Ber., Norges Bank, 2009.
- [Koh95] P.-A. Kohlas, Juergen; Monney: *A mathematical theory of hints : an approach to the Dempster-Shafer theory of evidence*. Lecture notes in economics and

- mathematical systems ; 425. Springer, Berlin, 1995. (Berlin ...) : DM 106.00 ; (New York ...).
- [Lad77] R. Laddaga: „Lehrer and the consensus proposal“. *Syntese* 1977.
- [Lin82] D. V. Lindley: „Scoring Rules and the Inevitability of Probability“. *International Statistical Review / Revue Internationale de Statistique* **50** (1), S. 1–11, 1982.
- [Lin83] D. V. Lindley: „Theory and Practice of Bayesian Statistics“. *Journal of the Royal Statistical Society. Series D (The Statistician)* **32** (1/2), S. 1–11, 1983.
- [Lin87] D. V. Lindley: „The Probability Approach to the Treatment of Uncertainty in Artificial Intelligence and Expert Systems“. *Statistical Science* **2** (1), S. 17–24, 1987.
- [Mad64] A. Madansky: *Externally Bayesian Groups*. Technical Report RM-4141-PR, RAND Corporation, 1964.
- [McC78] K. J. McConway: *The combination of expert' opinions in probability assessment: Some theoretical considerations..* Dissertation, University College London, 1978.
- [McC81] K. J. McConway: „Marginalization and Linear Opinion Pools“. *Journal of the American Statistical Association* **76** (374), S. 410–414, 1981.
- [Mei05] S. Meintrup, David ; Schaeffler: *Stochastik : Theorie und Anwendungen*. Statistik und ihre Anwendungen. Springer, Berlin, 2005. Literaturverz. S. 589 - 593; kart. : EUR 34.95, sfr 59.50.
- [Mor77] P. A. Morris: „Combining Expert Judgments: A Bayesian Approach“. *Management Science* **23** (7), S. 679–693, 1977.
- [Pen05] D. M. Pennock und M. P. Wellman: „Graphical Models for Groups: Belief Aggregation and Risk Sharing“. *Decision Analysis* **2** (3), S. 148–164, 2005.
- [Rai70] H. Raiffa: *Decision analysis : Introductory lectures on choices under uncertainty*. Addison-Wesley series in behavioral science: Quantitative methods. Addison-Wesley, Reading, Mass. [u.a.], 2. print.. Aufl., 1970.
- [Roj93] R. Rojas: *Theorie der neuronalen Netze : eine systematische Einfuehrung*. Springer-Lehrbuch. Springer, Berlin, 1993.

- [San06] J. Sander und J. Beyerer: *Fusion agents - realizing Bayesian fusion via a local approach*. In: *Multisensor Fusion and Integration for Intelligent Systems, 2006 IEEE International Conference on*, S. 249–254, Sept. 2006.
- [San08] J. Sander und J. Beyerer: „Decreased complexity and increased problem specificity of Bayesian fusion by local approaches“. *Information Fusion, 2008 11th International Conference on* S. 1–8, 30 2008-July 3 2008.
- [San09a] J. Sander: *Further Investigation of Focussed Bayesian Fusion*. Techn. Ber., Lehrstuhl fuer Interaktive Echtzeitsysteme Institut for Anthropomatics, 2009.
- [San09b] J. Sander und J. Beyerer: „A Local Approach for Bayesian Fusion: Mathematical Analysis and Agent Based Conception“. *Robotics and Autonomous Systems* **57** (3), S. 259–267, März 2009.
- [San10] J. Sander, M. Heizmann, I. Goussev und J. Beyerer: *Global evaluation of focussed Bayesian fusion*. In: *Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications, Proceedings of SPIE Vol. 7710*, J. J. Braun (Hrsg.), Mai 2010.
- [Sei09] T. Seidenfeld: *Decisions and Games CMU Formal Philosophy Summer School 2009*, 6 2009.
- [Sha76] G. Shafer: *A mathematical theory of evidence*. Princeton university press, 1976.
- [Sto61] M. Stone: „The Opinion Pool“. *The Annals of Mathematical Statistics* **32** (4), S. 1339–1342, 1961.
- [Swa92] B. Swain: „Consensus Theoretical Classification Methodes“. *IEEE Transactions on Systems, man and cybernetics* **22**, S. 688 – 704, 1992.
- [Wag82] C. Wagner: „Allocation, Lehrer models, and the consensus of probabilities“. *Theory and Decision* **14** (2), S. 207–220, Juni 1982.
- [Wag84] C. Wagner: „Aggregating subjective probabilities: some limitative theorems“. *Notre Dame J. Formal Logic* **25** (3), S. 233–240, 1984.
- [War07] E. B. J. S. Wark: *Concepts, Models, and Tools for Information Fusion*. Artech House, 2007.
- [Win68] R. L. Winkler: „The Consensus of Subjective Probability Distributions“. *Management Science* **15** (2), S. B61–B75, 1968.

# Anhang

# A. Mathematische Grundlagen

## A.1. Stochastik Grundlagen

Stochastik als Überbegriff für Wahrscheinlichkeitstheorie und Statistik ist ein Teilgebiet der Mathematik. In der mathematischen Stochastik werden Zufallsexperimente untersucht und beschrieben. Alle möglichen Ereignisse, die auftreten können, sind in der Menge  $Z$  enthalten. Eine Wahrscheinlichkeit  $p$  ist ein Maß und ordnet somit einer Ereignismenge  $A \subseteq Z$  einen Wert zu. Dieser schwankt zwischen „unmöglich“ und „sicher“. Eine Wahrscheinlichkeit muss folgende Kolmogorov'schen Gesetze [Gut05] (Seite 10) erfüllen:

- $\sum_{A \subseteq Z} p(A) = 1$
- $0 \leq p(A) \leq 1, \forall A \subseteq Z$
- $p(A \cup B) = p(A) + p(B), \forall A, B \subseteq Z, A \cap B = \emptyset$

Durch diese Axiome lassen sich verschiedene Kombinationen von Wahrscheinlichkeiten  $A$  und  $B$  bilden. Dabei ist das Nichteintreten von  $A$  mit  $\bar{A}$  bezeichnet.

- $p(A \text{ oder } B) = p(A) + p(B) - p(A \text{ und } B)$
- $p(A \text{ und } B) = p(A) \cdot p(B|A)$

Dabei ist  $p(B|A)$  die *bedingte* Wahrscheinlichkeit. Dies ist die Wahrscheinlichkeit für  $B$  unter der Voraussetzung, dass Ereignis  $A$  eingetreten ist. Ein *Likelihood* ist eine bedingte Wahrscheinlichkeit  $p(B|A)$ , welche als Funktion von  $A$  aufgefasst wird.

Zwei Ereignisse heißen unabhängig, wenn eine der folgenden Bedingungen erfüllt sind:

- $p(A|B) = p(A)$
- $p(B|A) = p(B)$
- $p(A \text{ und } B) = p(A) \cdot p(B)$



Diese Bedingungen bedeuten, dass das Wissen über das Eintreten eines Ereignisses das Eintreten des anderen nicht beeinflusst.

Zwei Ereignisse  $A$  und  $B$  heißen bedingt unabhängig bezüglich eines Ereignisses  $C$ , wenn eine der folgenden Bedingungen erfüllt ist:

- $p(A|B \text{ und } C) = p(A|C)$
- $p(B|A \text{ und } C) = p(B|C)$
- $p(A \text{ und } B|C) = p(A|C) \cdot p(B|C)$

Dies bedeutet: Ist  $C$  gegeben, dann sind  $A$  und  $B$  unabhängig.

## A.2. Unsicherheit

Unsicherheit wird in vielen Bereichen wie der Philosophie, Physik, Wirtschaftswissenschaft und Statistik unterschiedlich definiert. Sie beschreibt das Fehlen von Sicherheit bezüglich Messgrößen oder zukünftiger Ereignisse. Werden beispielsweise Messungen durchgeführt, so kann es durch Messfehler zu unterschiedlichen Messergebnissen kommen. Unsicherheit wird durch eine Wahrscheinlichkeit repräsentiert. Lindley sagt in [Lin87, Lin83]: „Die einzige befriedigende Beschreibungsform von Unsicherheit ist eine durch Wahrscheinlichkeit.“

## A.3. sigma-Algebren

Eine  $\sigma$ -Algebra ist ein Grundbegriff der Maßtheorie und der Stochastik. Es handelt sich dabei um ein Mengensystem  $\mathcal{Z} \subseteq \mathcal{P}(Z)$  über einer nicht leeren Grundmenge  $Z$ . Dabei wird dieses Mengensystem  $\sigma$ -Algebra genannt, wenn folgendes gilt:

- $Z \in \mathcal{Z}$ : Die Grundmenge  $Z$  ist in  $\mathcal{Z}$  enthalten.
- $A \in \mathcal{Z} \Rightarrow \bar{A} \in \mathcal{Z}$ :  $\mathcal{Z}$  ist abgeschlossen gegenüber der Komplementbildung.
- $\forall A^i \in \mathcal{Z} \Rightarrow \bigcup_{i \in \mathbb{N}} A^i \in \mathcal{Z}$ :  $\mathcal{Z}$  ist abgeschlossen gegenüber der endlichen Vereinigung.

Da  $Z$  in der  $\sigma$ -Algebra  $\mathcal{Z}$  enthalten ist und nach der zweiten Regel auch das Komplement von  $Z$   $\bar{Z} = \emptyset$  in  $\mathcal{Z}$  enthalten ist, ist die leere Menge immer in  $\mathcal{Z}$  enthalten [Mei05]. A

erzeugt eine  $\sigma$ -Algebra:

$$\mathcal{A} = \sigma(A) = \bigcap_{\mathcal{B} \in F(A)} \mathcal{B},$$

mit  $F(A) = \{B \subseteq \mathcal{P}(Z) \mid A \subseteq B, B \text{ ist } \sigma\text{-Algebra}\}$ .  $\sigma(A)$  ist dann der  $\sigma$ -Operator [Mei05].

Eine  $\sigma$ -Algebra heißt *tertiär*, wenn sie mehr als drei nicht leere und paarweise unabhängige Elemente enthält, die ungleich dem trivialen  $\sigma(A) = \{\emptyset, A, \bar{A}, Z\}$  mit  $A \subset Z$ , ist [Wag82].

## A.4. Scoring rules

Eine Bewertungsfunktion bildet eine Schätzung oder eine mit Unsicherheit behaftete Meinung und einen Ist-Zustand auf eine Skala ab. Damit kann die Meinung bewertet werden. Je besser die Schätzung auf den Ist-Zustand passt, um so höher sollte die Bewertungsfunktion sein. Erreicht ein Schätzer eine gute Bewertung, nennt man ihn „gut kalibriert“. Es gibt verschiedene Arten von Bewertungsfunktionen. *Strictly proper scoring rules* sind dann maximal, wenn der Schätzer ehrlich schätzt. Eine verfälschte Schätzung, die die Bewertung erhöht, ist damit unnötig. Beispiele für proper scoring rules sind:

$$n_1(p(\cdot), z^i) = 2p(z^i) - \sum_{j=1}^n f(z^j)^2,$$

$$n_2(p(\cdot), z^i) = \frac{p(z^i)}{\sqrt{\sum_{j=1}^n f(z^j)^2}},$$

$$n_3(p(\cdot), z^i) = \log p(z^i).$$

Diese heißen quadratische, sphärische und logarithmische Bewertungsfunktion. Werden Aussagen über ein Ereignis (z.B. den Ausgang eines Spiels) gewertet, kann anhand des nach n Durchgängen erhaltenen Gewinns eine Bewertungsfunktion definiert werden.

## B. Daten der Beispiele

Hier werden zur besseren Nachvollziehbarkeit die Daten für das Beispiel aus Abschnitt 5.2 nachgetragen: Die Beobachtungen, welche den Likelihood  $p(d|z)$  erzeugen, werden in Tabelle B.1 und B.2 dargestellt.

Die Wahrscheinlichkeit der Gruppenmitglieder für die gemeinsame Referenz ( $z^1$ ) wurde in Abbildung 5.2 dargestellt. Tabelle B.3 stellt die gerundeten Zahlenwerte zu dieser Abbildung dar. In Tabelle B.5 wird das gerundete Ergebnis eines LinOP mit gleichen Gewichten, vom Beispiel in Abschnitt 5.2 dargestellt. Die Abbildung 5.1 zeigt diese Ergebnisse graphisch.

Die Tabelle B.4 stellt das Ergebnis des LinOP mit den Gewichten die aus der Bewertung der fünf Agenten der gemeinsamen Referenz erzeugt wurden, dar. Die Abbildung 5.3 zeigt diese Ergebnisse graphisch.

	d1	d2	d3	d4	d5
$z^1$	0.0374556592707371	0.05643766339691114	0.0460166193610843	0.063391154786102	0.0365266637750653
$z^2$	0.0366636318663375	0.0465059000558207	0.0332783974473212	0.0451848064826909	0.0530308313701784
$z^3$	0.0654546897153217	0.0505088304712526	0.0360939943857164	0.0328707103170299	0.056030259266063
$z^4$	0.0524797986762285	0.0421022298200988	0.0640445031024476	0.0465768674526497	0.0342093566484005
$z^5$	0.0528022138438268	0.0457238338879921	0.0578373751205543	0.0634287848607374	0.061706254835181
$z^6$	0.0583804890259396	0.0418174578773792	0.0648622918410834	0.0487733830551668	0.0464318254024655
$z^7$	0.0500355306876588	0.0451200601916529	0.0662933391499896	0.048142854034478	0.0227674787518739
$z^8$	0.045623080562185	0.055869909762151	0.0417379148941134	0.064506898419671	0.0724403504284605
$z^9$	0.0455025449852504	0.0485358984797991	0.0401071182742954	0.0380397119840512	0.0582056314200435
$z^{10}$	0.0511479140829731	0.0566833854831034	0.0648903127959861	0.0364744895862742	0.0546969428080013
$z^{11}$	0.0299876508081206	0.0382315596442192	0.0567146501079415	0.0695650442430724	0.0729545263398244
$z^{12}$	0.0511133008033844	0.0533122678127127	0.0515339628661338	0.0521345003237732	0.0432006155510922
$z^{13}$	0.0365344375826519	0.0497543138599508	0.0605075822875282	0.0228122695848914	0.0543989998197122
$z^{14}$	0.0561365231466708	0.0499063335094145	0.0440153712713669	0.0493096273245691	0.0443034805319228
$z^{15}$	0.0453644740069213	0.0500728742751365	0.0540903574072725	0.0584299481724048	0.0542280093511733
$z^{16}$	0.0510133672412667	0.0625733134412652	0.0499255208179272	0.0458422030179398	0.0526579425085034
$z^{17}$	0.0499069746527072	0.0633017730198185	0.0446290702621314	0.0481160522343965	0.0621273064235853
$z^{18}$	0.0667510563208139	0.0408035121282541	0.0423850318506777	0.0515655158459117	0.0524896849024111
$z^{19}$	0.0517163552254851	0.0458206956641131	0.0525063995381657	0.0635166113339975	0.0200615907397968
$z^{20}$	0.0659303074955198	0.0569181866467539	0.0285301872182633	0.0513185669401921	0.0475322491262456

**Tabelle B.1.:** Wahrscheinlichkeitsverteilung der Beobachtungen  $z^1, \dots, z^{20}$  der zehn Informationsquellen  $d_1, \dots, d_5$  für das Beispiel in Abschnitt 5.2.

	d6	d7	d8	d9	d10
$z^1$	0.0442454832067784	0.0313039481526371	0.0356453936288481	0.0260303334600258	0.0492781074678681
$z^2$	0.0859257639829498	0.0693819042912448	0.0509354305604001	0.0426700463046824	0.042990725169804
$z^3$	0.0418029205038702	0.0514237193521013	0.0577201499508727	0.0617404732003015	0.0447700490819479
$z^4$	0.059151448636116	0.0856129658794883	0.0396191924744763	0.0710679678968012	0.0530781715389066
$z^5$	0.0534839504603976	0.0640622432694782	0.0592480250436426	0.0552645092010326	0.0547854409803378
$z^6$	0.0454238702323324	0.052213270367289	0.026421198495246	0.0480984846109494	0.0631493654659439
$z^7$	0.0605836877120743	0.0426568094810384	0.0472146700311433	0.0532219497114916	0.0424479096027718
$z^8$	0.0422282002628138	0.0521238187302096	0.0248758674725424	0.0370197776013825	0.0325832306859837
$z^9$	0.0482766012243187	0.0521692107106329	0.057799091803151	0.0550179123691526	0.05674877904601
$z^{10}$	0.068807456891548	0.022354445283703	0.040250969717407	0.0352376814497027	0.06806854598639
$z^{11}$	0.0443368518977667	0.0655741197146622	0.0504736394703167	0.056013931296298	0.0353505090361774
$z^{12}$	0.0266650466224615	0.0436549808824555	0.0403036134807127	0.0723281840365573	0.049434576952263
$z^{13}$	0.0597425541493754	0.0332547607494589	0.0638507871938074	0.0582387836776235	0.0412264182724214
$z^{14}$	0.0438491753255662	0.0385968805181564	0.0441391524801791	0.0435184292074813	0.0468289821636162
$z^{15}$	0.0559964508011265	0.0486939581542261	0.0753596646617504	0.0440350907833036	0.0540705667997011
$z^{16}$	0.0546202164408901	0.0656815236089145	0.0667143676183343	0.0545186227666535	0.0560720787246062
$z^{17}$	0.0346820819937937	0.0566307378035725	0.0720200785201994	0.0549742921153176	0.0376005761715907
$z^{18}$	0.0542361502364276	0.0490626038276614	0.0565042352500143	0.0378864626169245	0.0692988116693989
$z^{19}$	0.0375278716346847	0.0370101366374024	0.0437590130459697	0.0432482066738505	0.0548698428360987
$z^{20}$	0.0384142177847082	0.0385379625856676	0.0471454591009863	0.0498688610204677	0.0473473123481626

**Tabelle B.2.:** Wahrscheinlichkeitsverteilung der Beobachtungen  $z^1, \dots, z^{20}$  der Zehn Informationsquellen  $d6, \dots, d10$  für das Beispiel in Abschnitt 5.2.

	<i>Agent</i> <sub>1</sub>	$\S$ <i>Agent</i> <sub>2</sub>	<i>Agent</i> <sub>3</sub>	<i>Agent</i> <sub>4</sub>	<i>Agent</i> <sub>5</sub>	<i>Bayes</i>
	0.0212	0.0658	0.0651	0.0496	0.0244	0.0072

**Tabelle B.3.:** Dargestellt sind die bedingten Wahrscheinlichkeiten  $p(d_i|z^j)$

	<i>Agent</i> <sub>1</sub>	<i>Agent</i> <sub>2</sub>	<i>Agent</i> <sub>3</sub>	<i>Agent</i> <sub>4</sub>	<i>Agent</i> <sub>5</sub>	<i>Bayes</i>
$z^1$	0	0	0	0	0	0.0072
$z^2$	0.0397	0	0	0	0	0.0380
$z^3$	0.0395	0	0	0	0	0.0378
$z^4$	0.0894	0	0	0	0	0.0856
$z^5$	0.1760	0	0	0	0	0.1685
$z^6$	0	0.0358	0	0	0	0.0342
$z^7$	0	0.0237	0	0	0	0.0227
$z^8$	0	0	0	0	0	0.0165
$z^9$	0	0.0467	0	0	0	0.0447
$z^{10}$	0	0	0.0292	0	0	0.0280
$z^{11}$	0	0	0.0502	0	0	0.0481
$z^{12}$	0	0	0.0278	0	0	0.0266
$z^{13}$	0	0	0	0.0218	0	0.0209
$z^{14}$	0	0	0	0.0215	0	0.0206
$z^{15}$	0	0	0	0.0998	0	0.0956
$z^{16}$	0	0	0	0	0.1474	0.1412
$z^{17}$	0	0	0	0	0.0646	0.0618
$z^{18}$	0	0	0	0	0.0646	0.0619
$z^{19}$	0	0	0	0	0	0.0115
$z^{20}$	0	0	0	0	0.0225	0.0216

**Tabelle B.4.:** Dargestellt sind Ergebnisse der mit der gemeinsamen Referenz gewichteten Wahrscheinlichkeiten.

---

	<i>Agent</i> <sub>1</sub>	<i>Agent</i> <sub>2</sub>	<i>Agent</i> <sub>3</sub>	<i>Agent</i> <sub>4</sub>	<i>Agent</i> <sub>5</sub>	<i>Bayes</i>
$z^1$	0	0	0	0	0	0.0072
$z^2$	0.0225	0	0	0	0	0.0380
$z^3$	0.0224	0	0	0	0	0.0378
$z^4$	0.0508	0	0	0	0	0.0856
$z^5$	0.1000	0	0	0	0	0.1685
$z^6$	0	0.0629	0	0	0	0.0342
$z^7$	0	0.0417	0	0	0	0.0227
$z^8$	0	0	0	0	0	0.0165
$z^9$	0	0.0822	0	0	0	0.0447
$z^{10}$	0	0	0.0509	0	0	0.0280
$z^{11}$	0	0	0.0876	0	0	0.0481
$z^{12}$	0	0	0.0485	0	0	0.0266
$z^{13}$	0	0	0	0.0289	0	0.0209
$z^{14}$	0	0	0	0.0285	0	0.0206
$z^{15}$	0	0	0	0.1326	0	0.0956
$z^{16}$	0	0	0	0	0.0962	0.1412
$z^{17}$	0	0	0	0	0.0421	0.0618
$z^{18}$	0	0	0	0	0.0421	0.0619
$z^{19}$	0	0	0	0	0	0.0115
$z^{20}$	0	0	0	0	0.0147	0.0216

**Tabelle B.5.:** Dargestellt sind Ergebnisse des LinOP mit gleichen Gewichten für alle Gruppenmitglieder