

Reflexion als sozialer Mechanismus zum strategischen Management autonomer Softwaresysteme

Ingo J. Timm und Frank Hillebrandt

Konflikt-Situationen werden in soziologischen, wirtschaftswissenschaftlichen, psychologischen sowie selbst in naturwissenschaftlichen Kontexten in ihrer Doppelrolle als Blockaden, aber auch als Attraktoren der Emergenz komplexer Systeme untersucht. Nicht zuletzt deshalb sind Konflikte auch im Bereich der Forschungen zur Künstlichen Intelligenz (KI) Ausgangspunkte für diverse Forschungsansätze. Insbesondere die Auflösung von Konflikten und das Konfliktmanagement spielen eine wichtige Rolle in Teilgebieten der KI wie Expertensystemen, Konfigurationssystemen, Planungsverfahren, Design, Wissensakquisition und -repräsentation, Verarbeitung natürlicher Sprachen, Spieltheorie, Computerunterstützte kooperative Arbeit (CSCW), Concurrent Engineering, Human-Computer-Interface (HCI) und VKI (vgl. Müller 1996: 1).

Vor allem die Theorie der Multiagentensysteme (MAS) befasst sich als ein wichtiges Teilgebiet der Verteilten Künstlichen Intelligenz (VKI) mit dem Thema Konflikt. Während bei reaktiven Systemen die Koordination durch den Entwickler explizit geplant wird, kann sich die Koordination in Systemen autonomer Agenten entweder durch Kooperation oder durch nicht kooperatives Verhalten rationaler, *egoistischer*, also ihren eigenen Zielen folgender Agenten ergeben. Die Methoden des Konfliktmanagements unterscheiden sich bei diesen Ansätzen sehr grundsätzlich voneinander. Bei reaktiven Systemen müssen die Konflikte durch den Entwickler berücksichtigt werden. Deswegen wird ihre Bearbeitung in der Regel vorab festgelegt. Das emergente Verhalten ergibt sich hier bereits aus der eigennützigen Verfolgung der jeweiligen Ziele, auch wenn sie konfliktieren. Dieser Effekt ist ein wichtiger Grund für die Relevanz der Methode der agentenbasierten Modellierung (agent-based modeling) nicht nur zur Entwicklung komplexer, neuen Anforderungen gerecht werdender MAS, sondern auch als Methode der soziologischen Forschung im Kontext von Sozialsimulationen (vgl. Axelrod, 1997; Lüde et al., 2003; Hillebrandt et al. 2004). Denn wenn Konflikte zur Laufzeit eines MAS entstehen und bewältigt werden müssen, kommt es auf die sozialen Mechanismen an, die diese Bewältigung ermöglichen. Eine soziologisch inspirierte Modellbildung zur Entwicklung entsprechender MAS kann dazu beitragen, soziale Mechanismen der Konfliktlösung mit Hilfe von Experimenten der Sozialsimulation verstehbar zu machen, um diesbezügliche soziale Prozesse zu erklären. Deshalb firmieren in dem technischen System, das wir im Folgenden beschreiben und untersuchen werden, Konflikte als expliziter Bestandteil der Systemanforderungen. Die einzelnen autonomen Agenten sind in der Lage Zielkonflikte aufzulösen, die sich durch widersprechende Ziele oder aus möglichen Widersprüchen mit Zielen der Agentenumgebung ergeben.

Eine derartige Entwicklung eines MAS zur Untersuchung von Konfliktmanagement geschieht nicht zuletzt, weil neue Entwicklungen im Internet zu sehr heterogenen artifiziellen Gesellschaften geführt haben bzw. führen werden, die nicht ausschließlich aus Agentenanwendungen bestehen, sondern zu einem großen Teil auch einfache und intelligente Dienste beinhalten, welche im Rahmen des eBusiness sowie des Semantic Web entstehen. Diese werden mit einer hohen Dynamik konfrontiert sein, da die teilnehmenden Dienste und Agentenanwendungen durch Mobilität sowohl im physischen Raum als auch im Internet ad hoc Netzwerke und Gruppen bilden. Hohe Dynamik und Heterogenität können die Autonomie der einzelnen Agenten und die Koordination im jeweiligen System erheblich beeinträchtigen. Ohne eine effiziente Koordination und einen Zielabgleich zwischen vorgegebenen und selbst entwickelten Zielen droht eine solche Agentengesellschaft außer Kontrolle zu geraten. Diesen Entwicklungen muss auf der softwaretechnischen Seite mit expliziter Betrachtung von Autonomie und dynamischer Koordination begegnet werden. Dabei spielen insbesondere aus dynamischen Kontexten entstehende Konflikte eine wichtige Rolle. Intelligente Agenten sollten so gestaltet werden, dass sie in der Lage sind, explizit und flexibel Interessens- und Zielkonflikte zu erkennen, zu bewerten und aufzulösen. Deshalb strebt die hier verfolgte Argumentation über eine Diskussion der Möglichkeiten und Grenzen des softwaretechnischen Designs von Agenten-Autonomie und sozialer Koordination zwischen autonomen Agenten die Entwicklung eines Ansatzes zur Konfliktlösung an, der mit Hilfe einer prototypischen Realisierung extensiv untersucht wird. Die dazu zu diskutierenden Eigenschaften bzw. Fertigkeiten von Agenten beziehen sich zum einen auf die Entscheidungsfindung innerhalb eines Agenten (Autonomie) und zum anderen auf die Interaktion des Agenten mit anderen Entitäten (dynamische Koordination), die sich innerhalb des Systems bewegen (vgl. u. a. Nickles und Weiß 2005; Lasarczyk und Kron 2005). Insbesondere der zuletzt genannte Aspekt ist für eine Modellierung von unterschiedlichen Formen des Konfliktmanagements zentral und lässt sich nach unserer Einschätzung mit Hilfe des Theorems des sozialen Mechanismus', das in der soziologischen Forschung breit diskutiert und etabliert ist (vgl. nur Mayntz 2002; Schimank 2002; Schmid 2005 und in sozionischer Perspektive Hillebrandt et al. 2004: 191ff.), in viel versprechender Weise bearbeiten, indem die Formen der konfliktlösenden Interaktionen zwischen autonomen Agenten komplex und multikausal modelliert werden. Dabei kommt es vor allem darauf an, soziale Mechanismen zu identifizieren und zu untersuchen, die zur effektiven Lösung von Konflikten führen. Unsere These ist, dass der soziale Mechanismus der Reflexion in diesem Zusammenhang von entscheidender Bedeutung ist. Um diese These zu belegen gehen wir wie folgt vor: In einem ersten Schritt rekonstruieren wir die Möglichkeiten und Grenzen der Software-Autonomie, um einen systematischen Rahmen für Ebenen der Autonomie in Software-Systemen (MAS) zu konstruieren. Die Anwendung solcher Systeme ermöglicht einerseits eine erhöhte Flexibilität und Robustheit. Andererseits steigt jedoch auch die Dynamik derartiger Systeme, weil Agenten widersprüchliche oder nicht abgestimmte Zielsetzungen generieren und verfolgen können (1). Die so erhöhten Anforderungen an die Gestaltung

strategischer Aspekte erfordert die interdisziplinäre Konzeptionalisierung komplexer Koordination (2). Auf dieser theoretischen Basis entwickeln wir mit Bezugnahme auf soziologische Konzepte der Reflexion ein Framework für strategisches Management (3), das wir anschließend mit Hilfe der soziologischen Systemtheorie interpretieren, um es für die Sozialsimulation des sozialen Mechanismus' der Reflexion fruchtbar zu machen (4). Am Ende steht eine kurze Schlussbemerkung zur Einordnung unserer Überlegungen in die soziologische Forschung (5).

1 Software-Autonomie und Zielkonflikte

Die Künstliche Intelligenz untersucht Systeme, deren Eigenschaften analog zu den Eigenschaften sozialer Systeme beschrieben werden können. Insbesondere werden Probleme untersucht, für deren Lösung Menschen üblicherweise eine Intelligenzleistung voraussetzen. In der Lehre hat sich der Trend herausgebildet, Methoden und Theorien der künstlichen Intelligenz entlang dem Design autonomer Systeme, so genannter Intelligenter Agenten zu diskutieren (vgl. Russel & Norvig, 2003). In diesem Rahmen hat sich die *Verteilte Künstliche Intelligenz (VKI)* etabliert (Müller, 1993). Die Forschung der VKI bemüht sich, Systeme der *Künstliche Intelligenz (KI)* in Hinsicht auf Skalierbarkeit, multiple Problemlösungsstrategien und Wiederverwendbarkeit durch autonome und kooperative Systeme zu erweitern. Nicholas Findler (1991: 23) definiert das Aufgabengebiet der VKI wie folgt: *"Distributed planning and problem solving systems handle tasks that cannot be dealt with effectively and efficiently by one single processor."* Einige Ansätze gehen bei der expliziten Definition der Aufgaben und Wirkungen der VKI weiter und sprechen in diesem Zusammenhang von einer neuen KI (New Artificial Intelligence) (vgl. Pfeiffer 1995). Der Unterschied wird bei Ferber (1999: 53) wie folgt beschrieben: *"Whereas for AI it is the individual that is intelligent, for genetics it is the organisation that displays functionalities that can be characterised as intelligent."* Ein weiterer Aspekt der neuen Künstlichen Intelligenz ist, dass die Modellierung intelligenten Verhaltens durch die Integration sensor-motorischer Fähigkeiten und intelligenter Methoden in Anlehnung an biologische kognitive Prozesse erfolgt (DC, 1996). Für die Informatik ist diese Forschungsrichtung insofern interessant, als dass die Gestaltung monolithischer Systeme für die Lösung komplexer moderner Probleme zunehmend unangemessen ist und auch eine reine Modularisierung komplexer Systeme in lose gekoppelte Subsysteme keine Vorteile bieten kann. Erst die integrierte Betrachtung von Verteilung und Interaktion ermöglicht die Gestaltung von Systemen mit inhärenter Komplexität. So stellt z.B. die zunehmend verteilte kooperative Arbeit in Netzwerken, überregionalen Unternehmensnetzwerken und virtuellen Unternehmen neue Anforderungen an die Informationstechnik. Hierbei ist gerade die Abbildung der Expertise und nicht die problemfokussierte Sicht relevant. Diese Entwicklungen führen zu einer Zunahme der Autonomie der zu Grunde liegenden Softwaresysteme. Diese ist dann gegeben, wenn ein System auf der Basis lokal autonom entscheidender Einheiten (Mikro-

Dynamik), die miteinander kooperieren, eine globale Optimierung (Makro-Dynamik) vornimmt.

Eine wesentliche Eigenschaft, die Softwareagenten zugeschrieben wird, ist Autonomie, d.h. die Agenten agieren selbständig ohne direkte Intervention von Menschen und anderen Entitäten und haben zumindest eine partielle Kontrolle über ihr Verhalten und ihren Zustand (vgl. Wooldridge & Jennings 1995; Ferber 1999; Tecucci 1998). Agenten werden insbesondere in solchen Anwendungen eingesetzt, in denen die Komplexität der Aufgabe und die Dynamik der Umgebung einer vorprogrammierten Verhaltensweise nicht adäquat sind. Daher werden Agenten so gestaltet, dass sie eine gewisse Autonomie realisieren, d.h., dass der Auftraggeber dem Agenten nicht einen Plan *mit auf den Weg gibt*, sondern dass der Agent über eigenständige Planungsalgorithmen verfügen muss (vgl. Kalenka und Jennings 1997). Im Verlauf seiner *Arbeit* wird der Agent also nicht direkt von außen gesteuert, sondern erhält durch die eigene Planung und seine Unabhängigkeit im Handeln einen gewissen Grad an Autonomie (vgl. Weiß, 1999). Russel & Norvig (2003: 37) beschreiben Autonomie wie folgt: “To the extent that an agent relies on the prior knowledge of its designer rather than on its own percepts, we say that the agent lacks autonomy.”

Agenten sind folglich sich selbst steuernde Software-Einheiten, die ihre internen Zustände durch (Re)Kombination von in ihnen bereits vorhandenen Elementen in Selbstorganisation herstellen. Die Definition von Autonomie wird im Bereich der VKI breit diskutiert. Ein naiver Zugang zu Autonomie nimmt die externe Sicht auf ein System ein. Dabei ist ein System dann autonom, wenn das Verhalten des Systems nicht-deterministisch ist, d.h., das System in zwei identischen Situationen unterschiedlich reagiert. Dies bedeutet dabei nicht, dass das interne Verhalten eines autonomen Systems nicht deterministisch ist. Vielmehr erreicht das System seine von außen betrachtete nicht deterministische Verhaltensweise durch die interne Speicherung von Zuständen bzw. Zustandsinformationen. Somit erhält das System die Möglichkeit, Erfahrungen im Sinne von Informationen aus der Umgebung zu speichern und in die Entscheidungsfindung zukünftiger Situationen zu integrieren. Die Betrachtung sowohl der internen als auch der externen Zustandsinformationen führt in der Regel zu einem deterministischen Systemverhalten des autonomen Softwaresystems. Daher besteht ein diffizilerer Zugang zu dem Begriff Autonomie in der Beschreibung seiner wichtigsten Eigenschaften: Pro-Aktivität, Interaktion und Emergenz. Hiermit beziehen wir in unsere Begriffsbestimmung die folgenden Aspekte zwingend ein:

- **Pro-Aktivität:** Die Entscheidung eines Subsystems (Akteurs) wird nicht auf der Basis fest verdrahteten Entscheidungsverhaltens erzeugt. Vielmehr ist ein Akteur in der Lage, in einer Situation kontextabhängig eine Aktionssequenz zu generieren und autonom auszuführen. Um ein solches System zu ermöglichen, ist die explizite Repräsentation von Zielen notwendig.

- **Interaktion:** Das einzelne Subsystem, also der Agent, der autonom agiert, ist in der Lage, mit anderen Agenten aber auch mit der Umgebung zu interagieren und so über Verhandlung den eigenen Nutzen aber auch den Nutzen des Gesamtsystems zu erhöhen. Diese Eigenschaft basiert auf der Annahme, dass ausreichend „Spielraum“ zwischen den Agenten vorhanden ist, so dass eine Koordination bzw. Interaktion von Subsystemen (Agenten) vorteilhaft sein kann.
- **Emergenz:** Die bisherigen Eigenschaften sind solche, die innerhalb der Akteure bzw. Agenten realisiert werden. Im Fokus dieses Beitrages stehen jedoch solche Systeme, die aus einer Vielzahl autonomer Einheiten bestehen, den so genannten Multiagentensystemen. Hierbei sollte das System auf seiner Makro-Ebene Eigenschaften aufweisen, die nicht Teile des individuellen Profils der Agenten darstellen. Einfach gesagt: Das gesamte System ist mehr als die Summe seiner Teile, d.h. durch die lokale Interaktion der Agenten erfolgt auf der Makro-Ebene des Systems ein Verhalten, welches im Vorfeld nicht determiniert ist, aber als vorteilhaft erscheint, was im Idealfall auch nachgewiesen werden kann.

Für die systemische Betrachtung von Autonomie in Multiagentensystemen ist der Aspekt der Emergenz besonders wichtig. Sie ist eine wichtigste Eigenschaft bei der Gestaltung von Multiagentensystemen. Dennoch ist die Definition und Bewertung von Emergenz nicht hinreichend gelöst. Der Begriff der Emergenz wird in vielen Forschungsarbeiten von den Autoren metaphorisch benutzt. Im Bereich der VKI werden im Wesentlichen drei Varianten von Emergenz unterschieden (vgl. Timm *et al.* 2001):

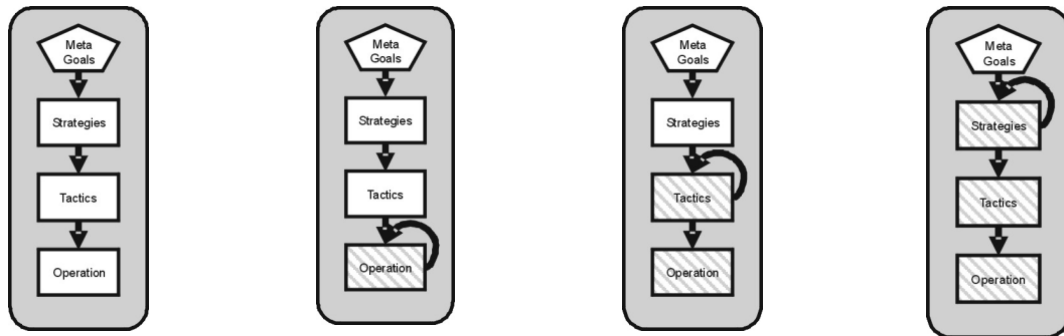
- „Emergent Properties“: Axelrod (vgl. 1997) definiert Emergenz als das Erscheinen von Eigenschaften, die als globale Effekte von lokal interagierenden Agenten auftreten: “Emergent properties are often surprising because it can be hard to anticipate the full consequences of even simple forms of interaction”. (Axelrod 1997: 4)
- „Emergent Organization“: Ferber (vgl. 1999) beschreibt die Emergenz aus einem anderen Blickwinkel, der eher auf die Struktur der Interaktion von lokal interagierenden Einheiten eingeht. Das besondere hierbei ist, dass Systeme von Agenten nicht ausschließlich als Gruppe betrachtet werden. Vielmehr koordiniert sich das System von Agenten durch lokale Interaktion und realisiert dabei eine dynamische Organisation, die Ferber mit „Emergent Organization“ bezeichnet: “They [organisational structures] can also be defined a posteriori, and we then speak of emergent organisations. These most frequently contain only reactive agents and are characterised by the absence of any predefined organisational structure, their structure being entirely the result of interactions between agents. In this case, positions and relationships are not determined in advance, but appear as the product of the

behaviours of each of the agents. More precisely, the distribution of functions and tasks follows an auto-organisation procedure, which permits an organisation to evolve and to adapt easily to modifications in the environment and to the needs of a group of agents.” (Ferber 1999: 114)

- „Emerging Intelligence“: Wooldridge und Jennings (vgl. 1995) definieren besonders in ihren frühen Arbeiten verschiedene Arten von Agentensystemen. Dabei verwenden sie die schwache Definition von Agenten (weak notion of agency) für die Minimalanforderungen, die ein Softwaresystem erfüllen muss, damit dieses als Agent definiert werden kann. Die starke Definition (strong notion of agency) verwenden sie für die Abgrenzung von Agenten und Intelligenzen. Hierbei werden viele Eigenschaften als optional angegeben. In diesem Rahmen spricht Wooldridge (vgl. 1999) aber auch von der Idee, dass intelligentes Verhalten durch die Interaktion von einfachem Verhalten realisiert werden kann, also von “the idea that intelligent behavior emerges from the interaction of various simpler behaviors.” (Wooldridge 1999: 49)

Die hier angeführten Definitionen von Emergenz sind für ihre softwaretechnische Realisation auf eine effiziente, adaptive Kommunikation von lokalen, autonomen Entscheidungssystemen, womit hier autonome Agenten gemeint sind, angewiesen.

In aktuellen Forschungsarbeiten werden unterschiedliche Grade der Autonomie von Softwareagenten betrachtet (vgl. u. a. Rovatsos & Weiss 2004; Müller, 1996). Cristiano Castelfranchi und Rosaria Conte (1992) diskutieren insbesondere in ihren frühen Arbeiten einen sehr hohen Grad der Autonomie von Softwareagenten, durch den der Einfluss von vorgegebenen Normen, Verhaltensweisen und Prozeduren als irrelevant für die Aktionsauswahl festgelegt wird. Michael Rovatsos und Gerhard Weiß (2005) nähern sich dem Begriff der Autonomie durch die Differenzierung einer externen von einer internen Sicht sowie durch die Identifikation unterschiedlicher Verhaltensperspektiven innerhalb der beiden Sichten. Eine andere Herangehensweise an die strukturierte Differenzierung von Autonomie folgt der Betrachtung der Umgebung, in der ein System eingesetzt wird bzw. werden soll. Hierbei werden unterschiedliche Ebenen der Entscheidungsfindung in Analogie zu Ansätzen in den Wirtschaftswissenschaften und der Systemtheorie mit den Begriffen Operation, Taktik und Strategie differenziert (vgl. Hentze et al. 1993; Küpper 1997). Führt man diese Ansätze zusammen, ergeben sich für die Klassifizierung von Autonomie vier Ebenen (levels of autonomy – LoA), die von streng regulierten über operativ autonomen, taktisch autonomen bis strategisch autonomen Systemen reichen (vgl. Abbildung 1).



Strenge Regulation

Operative Autonomie

Taktische Autonomie

Strategische Autonomie

Abbildung 1: Ebenen der Autonomie (LoA)

Autonome Systeme zeichnen sich durch eine von der Autonomieebene abhängige Komplexität der Entscheidungsfindung aus. Hierbei kann allgemein angenommen werden, dass die Komplexität mit der Höhe der Autonomie erheblich zunimmt. Daher sind solche Systeme nur für Anwendungen geeignet, in denen die hohe Flexibilität, die diese Systeme bieten, auch gefordert wird. Eine hieraus resultierende Grundannahme ist, dass autonome Systeme in eine variable Umgebung eingebettet sind und hierbei die Fähigkeit besitzen, diese wahrzunehmen und mit ihr zu interagieren. Von besonderem Interesse sind darüber hinaus solche Systeme, die aus mehreren autonomen Subsystemen (Agenten) bestehen wie z.B. Multiagentensysteme. Die Ausgestaltung der autonomen Systeme insbesondere bezüglich der Autonomieebenen ist dabei abhängig vom konkreten Anwendungskontext. Der Grad der Komplexität einer agenteninternen Entscheidungsfindung sollte in Abhängigkeit zu den Anforderungen der Aufgabendomäne gewählt werden. In diesem Zusammenhang hat sich in den Forschungen zur Künstlichen Intelligenz eine Klassifikation von Aufgabendomänen etabliert, die von Russell & Norvig 1994 eingeführt und 2003 verfeinert worden ist (Russell & Norvig 1994; 2003). Diese Klassifikation besteht aus sechs Kategorien (Observable, Deterministic, Episodic, Static, Discrete, Agents), die eine Unterscheidung verschiedener Problemkomplexitäten ermöglicht:

- **Observable**

Die Frage, ob ein autonomes System die gesamte Umgebung wahrnehmen kann, das System also über vollständige Informationen verfügt, die als gesichert angenommen werden können, wird durch das Kriterium „Observable“ bestimmt. Hierbei bedeutet eine vollständige Wahrnehmung (fully) die Möglichkeit des autonomen Systems auf jegliche Informationen zugreifen zu können. Gerade im Bereich der Modellierung realer Anwendungen, wie Unternehmensnetzwerke oder sozio-technische Systeme, muss jedoch davon ausgegangen werden, dass der Agent keinen vollständigen Zugriff auf alle relevanten Informationen hat (partially).

- **Deterministic**

Ein autonomes System hat prinzipiell die Möglichkeit, Aktionen von z.B. physikalischen Aktuatoren ausführen zu lassen. Für die Unterscheidung von Aufgabenkomple-

xität ist es notwendig zu wissen, in wieweit eine exakte Modellierung von Effekten der Aktionen möglich ist. Befindet sich ein autonomes System in einer deterministischen Umgebung (deterministic), so kann jeder mögliche Effekt einer Aktion exakt vorher berechnet werden. Hingegen wird bei strategischen und stochastischen Umgebungen davon ausgegangen, dass der gewünschte Effekt einer Aktion in der Umgebung nicht eintritt. Hierbei wird zudem zwischen zufälligen Ausgängen (stochastic) oder solchen Ausgängen unterschieden, bei denen ein strategisches Verhalten der übrigen Agenten (stochastic) im System angenommen werden kann.

- Episodic

Dieses Kriterium bezieht sich auf die Art und den Betrachtungszeitraum der Aufgaben des Agenten. In einer sequentiellen Umgebung liegt der Fokus auf der Fortschreibung und Anpassung des Verhaltens eines autonomen Systems über die Laufzeit des Systems. Hierbei ist es notwendig, im Verhalten des Systems aus Erfahrungen zu lernen bzw. Erfahrungen zu berücksichtigen und Erwartungen über zukünftige Veränderungen zu antizipieren. Ein autonomes System ist in einer episodischen Umgebung immer wieder mit Aufgaben konfrontiert, die in sich abgeschlossen gelöst werden können.

- Static

Ein weiteres Kriterium der Aufgabendomänen bezieht sich auf die Dynamik des Systems. Hierbei werden drei unterschiedliche Grade von Dynamik unterschieden: statisch, semi-dynamisch, dynamisch. In statischen Domänen (static) ändert sich weder das Verhalten der individuellen autonomen Systeme (Agenten) noch die Umgebung. Die Dynamik in solchen Systemen ist also auf die Aktionsausführung der Agenten beschränkt. In semi-dynamischen Umgebungen (semi) ist es möglich, dass sich das Verhalten der Agenten zusätzlich zu der Dynamik durch Aktionsausführung und der Wahrnehmung der Umgebung durch den Agenten ändert. Vollständig dynamische Umgebungen (dynamic) können darüber hinaus solche Prozesse enthalten, die die Umgebung unabhängig von den Aktionen der Agenten ändern.

- Discrete

Das Verhalten autonomer Systeme kann entweder kontinuierlich oder diskret erfolgen. Rundenbasierte Spiele, wie Schach oder Dame, sind diskrete Umgebungen (discrete). Aufgaben wie das Taxi-Fahren erfolgen hingegen in kontinuierlichen (continuous) Umgebungen. Die Unterscheidung zwischen diskret und kontinuierlich ist nicht geeignet, unterschiedliche Autonomiegrade zu unterscheiden, so dass sie im Folgenden keine weitere Berücksichtigung findet.

- Agents

Eine weitere Unterscheidung von Anwendungsdomänen erfolgt bezüglich der Anzahl der eingesetzten Agenten. Dabei ist die Frage zu beantworten, ob nur ein Agent im System eingesetzt (single) wird, oder ob es sich um ein Multiagentensystem handelt,

in dem mehr Agenten miteinander in Beziehung treten müssen, damit das System funktionieren kann (multi).

Diese Klassifikation lässt sich in Kombination mit den oben angeführten Graden der Autonomie von einzelnen Softwareagenten dazu nutzen, die Ebenen der Autonomie zu spezifizieren und bezüglich ihrer Adäquatheit in Anwendungsdomänen zu beschreiben, so dass die folgende Tabelle erstellt werden kann, die den Zusammenhang von Anwendungsdomänen und den Ebenen der Autonomie von jeweils eingesetzten Softwareagenten verdeutlicht:

Level of Autonomy	Observable	Deterministic	Episodic	Static	Agents
Strong Regulation	Fully	Deterministic	Episodic	Static	Single
Operational Autonomy	Partial	Deterministic	Episodic	Static	Multi
Tactical Autonomy	Partial	Stochastic	Episodic	Semi	Multi
Strategic Autonomy	Partial	Stochastic	Sequential	Dynamic	Multi

Tabelle 1: Anwendungsdomänen und Ebenen der Autonomie

Die Tabelle 1 zeigt die unterschiedlichen Klassifikationen der Autonomieebenen. Mit dem Begriff der strengen Regulation werden Systeme bezeichnet, die monolithisch aufgebaut sind. Hier werden alle Informationen zentral verwaltet, stehen dem Entscheidungssystem zur Verfügung und können in einer überwiegend statischen Umgebung sehr effizient eingesetzt werden. Steigt die Komplexität z.B. dadurch, dass nicht mehr alle Informationen auf einmal effizient verarbeitet werden können oder sich Teilprobleme identifizieren lassen, die mit geringem Koordinationsaufwand separat gelöst werden können, bietet sich der Einsatz von reaktiven Systemen an. Methoden der Swarm-Intelligence oder der reaktiven Agenten können beispielsweise effizient zur innerbetrieblichen Fertigungssteuerung und Tourenplanung genutzt werden (vgl. Brückner 2000), was der Ebene operativ autonomer Systeme entspricht. Diese Systeme ermöglichen eine Autonomie in der konkreten Aktionsausführung. Strategien und Taktiken sind jedoch fest implementiert und können nicht durch das Entscheidungssystem modifiziert werden. Auf der Ebene taktischer Autonomie erhalten die Systeme darüber hinaus die Möglichkeit, ihr taktisches Verhalten anzupassen. In Agentensystemen werden solche Ansätze häufig dadurch realisiert, indem den Einzel-Agenten explizite Planungsmechanismen implementiert werden, so dass sie eine Zielauswahl vornehmen können (z.B. BDI-Agenten). Eine weiterreichende Autonomie der Agenten auf der strategischen Ebene erfordert die Fähigkeit des Agenten über seine Ziele zu schlussfolgern und diese dabei zu variieren. Ein sol-

cher Ansatz findet sich beispielsweise im Algorithmus cobac, der zum Konfliktmanagement eine Kompromissbildung zwischen Zielen ermöglicht (vgl. Timm 2004).

Wenn Autonomie in beschriebener Weise als wesentliches Merkmal eines MAS realisiert wird, muss gesehen werden, dass sie eine System-Eigenschaft ist, die zu einer partiellen Unkontrollierbarkeit des Gesamtsystems führen kann, wenn etwa autonome Softwareagenten widersprüchliche oder nicht abgestimmte Zielsetzungen generieren und verfolgen. Dies wird durch die Dynamik der Vernetzung, insbesondere auch durch die Mobilität der Agenten verstärkt, die es verhindert, zur Entwurfszeit die genaue Zusammensetzung der Agentengesellschaften zu kennen und entsprechend die Zielsetzungen der Einzelagenten zu entwerfen. Eine wichtige Folgerung daraus ist: In heterogenen und dynamischen „Gesellschaften“ mit autonomen Agenten stellen sich Entscheidungen nicht so konfliktfrei dar, wie es erste Ansätze zur Agententechnologie formulieren. Ein autonomer Agent in einer solchen Gesellschaft muss deshalb über effektive Konfliktlösungsmöglichkeiten verfügen und dabei seine Ziele in Frage stellen und neue Ziele generieren. Die bereits eingangs angesprochene Unkontrollierbarkeit von Agentengesellschaften zeigt sich daran, dass die daraus emergierenden Zielkonflikte nicht in allen Situationen im Sinne des System-Designers aufgelöst werden. Die Resolution von Konflikten sollte hierbei nicht nur zu einer Entscheidung für das eine oder das andere Ziel führen, sondern auch die Generierung von neuen Zielen als Kompromisse ermöglichen. Neue Forschungsansätze in diesem Bereich müssen die (partiell) autonome Wahl und Erzeugung von Zielen und Absichten in Konfliktsituationen transparent gestalten und auf ihre Wirkungen für die Kontrollierbarkeit von dynamischen Agentengesellschaften hin untersuchen.

Spätestens an dieser Stelle wird eine soziologische Interpretation und Weiterentwicklung der Multiagentensysteme notwendig. Denn mit den beschriebenen Eigenschaften von MAS, die aus (partiell) autonomen Agenten bestehen, ergeben sich Anschlussstellen zur soziologischen Forschung. MAS entwickeln sich immer deutlicher als Formen „künstlicher“ Sozialität, in denen sich gegenseitig intransparente autonome Individuen (Agenten) begegnen und in Interaktion miteinander treten. Hier stellt sich eine der Grundfragen der Soziologie, wie Individuen (Agenten), die sich gegenseitig intransparent sind, trotz ihrer Autonomie in dauerhafter sozialer Beziehung zueinander stehen können, in analoger Weise auch für MAS. Für die Forschungen zur Verteilten Künstlichen Intelligenz stellt sich diese Frage zuerst in Bezug auf Zielkonflikte, die durch eine komplexe „Organisation“¹ von MAS bewältigt werden müssen.

Für die Gestaltung von realitätsnahen Anwendungssystemen z.B. im betrieblichen Umfeld hat die explizite Berücksichtigung des Zielsystems des Unternehmens eine besondere Bedeutung. Denn hier finden sich vielfach *inkonsistente* und sich widersprechende Ziele, die konventionell durch menschliche Entscheider behandelt bzw. ausbalanciert werden müssen. Beispiele sind „stakeholder-“ und „shareholder-value“, Reduktion der Kapitalbindung und Verkürzung der Bereitstellungszeiten. Diese Zielkonflikte spiegeln sich auf der operativen Ebene wider. Ein betriebliches Beispiel ist der Konflikt zwischen der Forderung nach einer guten Auslas-

¹ Organisation hier nicht als soziologischer Begriff verstanden, sondern als Hilfsbegriff zum System-Design.

tung der Ressourcen und einer kurzen Auftragsbearbeitung, wie er z.B. von Wiendahl und Mittendorf (2000) in der Fertigung beschrieben wird. Diese Situation ist in Abbildung 2 visualisiert (vgl. Wiendahl & Mittendorf 2000: 1f). Sie zeigt, dass eine hohe Auslastung bei großem Auftragsbestand möglich ist, der aber gleichzeitig zu längeren Durchlaufzeiten (Staus bei kritischen Ressourcen) führt. Umgekehrt erreicht man kurze Durchlaufzeiten bei kleinem Auftragsvolumen und geringer Auslastung, da dann die einzelnen Ressourcen unverzüglich den jeweiligen Auftrag abarbeiten können.

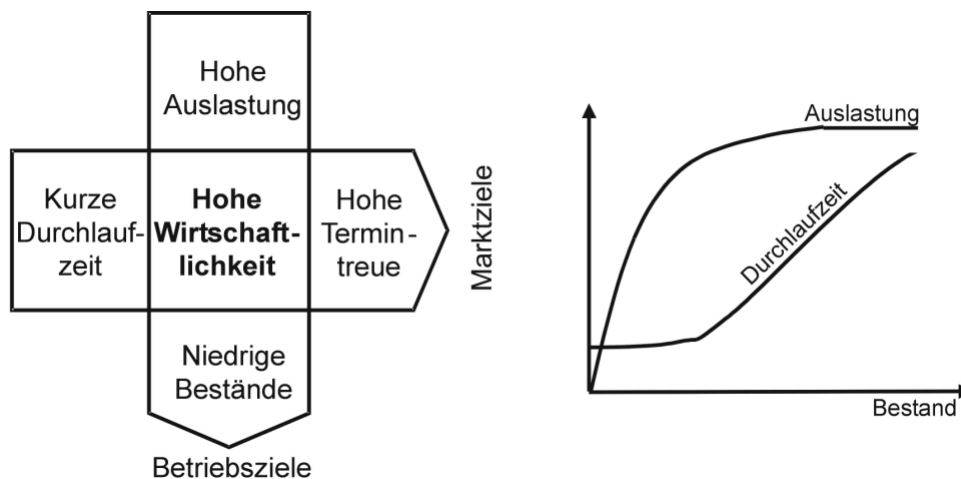


Abbildung 2: Zielkonflikte in der Fertigung

Beim Einsatz von Agententechnologie ist zu bedenken, dass für eine Akzeptanz dieser innovativen Technologie eine konstruktive Konfliktlösungsstrategie notwendig ist, die für die verantwortlichen Akteure transparent und somit nachvollziehbar ist. Wesentlich hierbei ist, dass die Zielkonflikte nicht ausschließlich innerhalb eines Agenten, sondern insbesondere auch zwischen den Agenten auftreten. Die Konflikte zwischen Agenten lassen sich in Analogie zu Konflikten in sozialen Systemen behandeln. Das Agentensystem benötigt mit anderen Worten einen sozialen Mechanismus zur Bewältigung von Konflikten, den es in Selbstorganisation anwenden kann. Nur so können die hohen Anforderungen an das MAS in einer Weise bedient werden, die eine relative Autonomie der Entscheidungsfindung innerhalb der Agenten sicherstellt. Dies ist deshalb notwendig, weil das System sich dadurch an gewandelte Umweltbedingungen anpassen kann, ohne durch den System-Designer von außen verändert werden zu müssen. Solche Anpassungen könnten dann idealerweise zur Laufzeit des System autonom initiiert werden. Zur Entwicklung eines derartigen Mechanismus der Konfliktbewältigung durch komplexere technische Koordination der Systemprozesse reichen allein technische Forschungsparadigmen, die sich auf die monokausale Gestaltung technischer Prozesse fokussieren, nicht aus. Die Konstruktion von MAS muss sich folglich immer deutlicher am Vorbild sozialer Prozesse orientieren, die von der Soziologie differenziert beschrieben und untersucht

werden.² Deshalb kann die soziologische Theoriebildung zur Weiterentwicklung von hoch komplexen, durch Selbstorganisation gekennzeichneten Systemen künstlicher Intelligenz genutzt werden, die sich dadurch auszeichnen, Konflikte in einer Weise zu lösen, die nicht nur für den Fortbestand des Systems, sondern auch für seine Reproduktion auf einem höheren Niveau der Komplexität notwendig ist. Insbesondere der differenzierte Umgang mit Zielkonflikten steigert die strategische Autonomie des technischen Systems, die sich als spezifische Form der Kommunikation, nämlich als Reflexion im System einstellt, wenn ganz bestimmte Bedingungen erfüllt sind. Um einen solchen sozialen Mechanismus zu entwickeln, ist eine soziologische Auseinandersetzung mit dem Begriff der Koordination durch Reflexion notwendig. Diese Erörterung kann dann dazu genutzt werden, den sozialen Mechanismus der Reflexion als Bestandteil des technischen Systems zu etablieren.

2 Koordination – interdisziplinär

Die steigende Mobilität von Softwareagenten und die zunehmende Verbreitung von ad hoc Netzwerken führen dazu, dass Gruppen von Softwareagenten in ihrer Zusammensetzung und somit ihren kollektiven Fertigkeiten einer hohen Dynamik ausgesetzt sein werden. Informationen und Regeln, die für die Interaktion der Agenten verfügbar sind, gelten in ad hoc Netzwerken oft nur lokal. Die Agenten müssen auf der einen Seite in der Lage sein, in heterogenen Gruppen Fertigkeiten zu spezifizieren und diese anzubieten (dynamische Teambildung). Auf der anderen Seite müssen Agenten lokale Bedingungen in ihrem Kooperationsverhalten angemessen berücksichtigen. Wird von einem dynamischen Team ein Problem gelöst, ist es notwendig, eine angemessene Kommunikation für das Feedback sowie eine Belohnung innerhalb der Gruppe zu generieren, damit die Agenten über den Erfolg ihrer Teilnahme an der Gruppe entscheiden können.

Wird in einer Gruppe von teilweise mobilen Agenten, die über ein ad hoc Netzwerk verbunden sind, eine Aufgabe zur Lösung angenommen, kann sich die Zusammensetzung der Gruppe während der Problemlösung ändern. Daher gehen wir davon aus, dass es notwendig ist, über neue Kommunikationsformen wie die „aktive“ Kommunikation nachzudenken³. Hierbei wird die Kommunikation einer Gruppe durch mediatorische Funktionalitäten organisiert, die in der Lage sind, Veränderungen der Gruppenzusammensetzungen zu erkennen. Falls eine Fertigkeit, die zur Problemlösung notwendig ist, durch Migration ausfällt, wird diese Funktionalität, die wir als „aktive“ Kommunikation bezeichnen, in föderierten Agentengesellschaften nach einem entsprechenden Agenten suchen und diesen einladen, an der Problemlösung

² Dies ist ein wichtiger Grund für die Zusammenarbeit von VKI und Soziologie, die im Forschungsfeld der Szionik (vgl. Malsch 2005: 241f.) institutionalisiert ist. An die Ergebnisse dieser Forschungen (vgl. u. a. die Beiträge in Lüde et al. 2003; Florian/Hillebrandt 2004; Fischer et al. 2005) schließen wir mit unseren Überlegungen explizit an.

³ Christoph Schlieder und Thomas Malsch haben bereits Forschungen zu Kommunikation betrieben, die über den reinen Nachrichtenaustausch hinausgehen (Malsch et al. 2006); die hier diskutierte „aktive“ Kommunikation geht über diesen Ansatz insofern hinaus, als die Kommunikation wieder an die Agenten zurückgebunden wird.

zu partizipieren. Letztendlich wird nach erfolgreicher Problemlösung die Gruppe ein Feedback erhalten. Dieses Feedback, welches in elektronischen Marktplätzen auch monetär sein kann, muss zwischen den an der Problemlösung beteiligten Agenten verteilt werden (Distribution).

Werden Konflikte und Widersprüche auf diese Weise im MAS zugelassen, müssen sie aufwändig mit technischen Mechanismen koordiniert werden, damit der Fortbestand des Systems nicht gefährdet wird. Das Zulassen von Widersprüchen und Konflikten in der technisch vermittelten Kommunikation scheint deswegen zunächst dem Paradigma der Technikentwicklung zu widersprechen, nach dem technische Systeme so berechenbar und kausal wie möglich gestaltet werden müssen, um eine effektive Operationsweise zu ermöglichen. Setzt man jedoch die Autonomie der an MAS beteiligten Agenten als Bedingung für die hohe Lösungskapazität des Systems für komplexe Probleme, stellt sich dieser Sachverhalt in neuer Weise dar. Um dies genauer zu sehen, lohnt sich ein Blick auf einige grundlegende Einsichten der soziologischen Systemtheorie in die Operationsweise und Reproduktionsform sozialer Systeme. Niklas Luhmann sieht Konflikte nicht als Blockaden der Reproduktion des Systems, sondern ganz im Gegenteil als notwendige Bedingung einer dynamischen Reproduktion komplexer sozialer Systeme. Komplex sind soziale Systeme nach ihm dann, wenn die Elemente, aus denen das soziale System besteht, für Luhmann Kommunikationseinheiten⁴, nicht mehr zu jedem anderen Element in Beziehung gesetzt werden können (vgl. Luhmann 1984: 46), wenn also ereignishaft Kommunikationseinheiten als Anschluss-Elemente des sozialen Systems nicht mehr auf alle zuvor gezeitigten Elemente referieren können, wenn also die Systemelemente des sozialen Systems vielschichtig und multidimensional aufeinander bezogen sind.⁵ Derartig strukturierte Systeme „benötigen ... ein recht hohes Maß an Instabilität, um laufend auf sich selbst und ihre Umwelt reagieren zu können, und sie müssen diese Instabilitäten laufend reproduzieren“ (Luhmann 1984: 501). Widersprüche und Konflikte erscheinen nun als Garanten dieser Instabilitäten. Die Funktion von Widersprüchen besteht darin, „die Formeinheit eines Sinnzusammenhangs zu wahren, ja herauszuheben, aber die Erwartungssicherheit, die damit normalerweise verbunden ist, nicht zu stärken, sondern im Gegenteil: aufzulösen“ (Luhmann 1984: 501). Widersprüche dienen folglich „der Reproduktion des Systems durch die dazu rötige Reproduktion von Instabilitäten“ (Luhmann 1984: 504). Instabilität ist dann gegeben, wenn eine Unsicherheit des Anschlusswertes von Ereignissen gegeben ist, wenn also

4 An dieser Stelle setzen wir bestimmte Grundannahmen der Luhmannschen Theorie sozialer Systeme voraus, etwa die, dass Kommunikation als dreifacher Selektionsprozess verstanden werden muss, der erst abgeschlossen ist, wenn Information (erster Selektionsprozess), Mitteilung (zweiter Selektionsprozess) und Verstehen (dritter Selektionsprozess, verstanden als Anschluss einer Kommunikationseinheit im sozialen System), zusammengeführt worden sind. Vgl. hierzu ausführlich Luhmann 1984: 193-201 und insbesondere 194.

5 Hier ist die grundlegende Unterscheidung von sozialen Systemen in Interaktion, Organisation und Gesellschaft durch Luhmann (vgl. 1984: 16) wichtig, weil Interaktionen, verstanden als Kommunikation unter Anwesenden, die sozialen Systeme sind, die als einzige ohne die Entstehung von Komplexität auskommen können, um sich zu reproduzieren, während Organisationen und vor allem die Gesellschaft als umfassendes soziales System, in dem alle anderen sozialen Systeme zwangsläufig vorkommen müssen, sich konstitutiv durch Komplexität auszeichnen. Wir werden uns zur Interpretation des hier zu diskutierenden technischen Systems vor allem auf das Organisationssystem als Analogiebegriff zum technischen System konzentrieren.

Anschlusselemente, also Kommunikationseinheiten in kontingenter Weise auf die bereits gezeitigten Elemente des sozialen Systems bezogen sind. Diese Instabilität ist notwendig, damit die Reproduktion des Systems nicht durch Variabilität gefährdet wird, die sich aus wandelnden Umweltanforderungen für soziale Systeme zwangsläufig ergibt. Denn die Umwelt des Systems ist nicht statisch und dies können komplexe soziale Systeme in ihrer Operationsweise nicht ignorieren, wenn sie sich auf dem erreichten Komplexitätsniveau reproduzieren wollen. Daraus folgt: „Das System immunisiert sich nicht gegen das Nein, sondern mit Hilfe des Nein; es schützt sich nicht gegen Änderungen, sondern mit Hilfe von Änderungen gegen Erstarrung in eingefahrenen, aber nicht mehr umweltadäquaten Verhaltensmustern.“ (Luhmann 1984: 507; Hervorhebungen weggelassen)

Konflikte haben gerade in Bezug auf diese Eigenschaft sozialer Systeme die wichtige Funktion, Instabilitäten operativ herzustellen, weil sie als „operationalisierte, Kommunikation gewordene Widersprüche“ (Luhmann 1984: 537) gefasst werden können. Gleichsam muss das System eine konstruktive Operationalisierung von Instabilitäten erzeugen, damit die Instabilität des Systems nicht zu einem exklusiven Dauerzustand wird, was die Reproduktion des Systems wiederum gefährden würde. Das heißt: Eine Immunisierung sozialer Systeme durch das konstruktive Operationalisieren von Widersprüchen und Konflikten ist eine Existenzbedingung komplexer sozialer Systeme. Wie wir weiter unten genauer zeigen werden, gelingt diese konstruktive Operationalisierung von Widersprüchen und Konflikten insbesondere in Organisationssystemen häufig durch die Institutionalisierung von Reflexion, die eine Form von Systemrationalität erzeugt, die auch für die Entwicklung und dynamische Reproduktion von komplexen MAS von großer Bedeutung ist. Denn eine dynamische Koordination von Widersprüchen und Konflikten ist nur möglich, wenn das System zur Reflexion fähig ist. Institutionalisierte, also auf Dauer gestellte Reflexion ist ein sozialer Mechanismus für das Konfliktmanagement in dynamischen Systemen.

Diese Argumentation ist auch für die Entwicklung von im hier eingeführten Sinne komplexen MAS von nicht zu unterschätzender Bedeutung. Wenn nämlich MAS Widersprüche und Konflikte enthalten, die nicht durch Negation, sondern durch konstruktive Operationalisierung bewältigt werden müssen, damit das MAS auf dynamische Umweltbedingungen flexibel reagieren kann, ist es notwendig, soziale Mechanismen in technischen Systemen zu erzeugen, die diese Funktion bedienen können. Die Institutionalisierung von Reflexion ist, so unsere These, ein sozialer Mechanismus zur operativen Bewältigung von durch Widersprüche und Konflikte entstehender Instabilität, die dann auch in technischen Systemen nicht mehr als zu vermeidender, sondern als zu operationalisierender Zustand begriffen werden kann, indem das technische System mit durch Reflexion gerichteter Sensibilität ausgestattet wird. Die Analogiebildung sehen wir hier im Systemtyp der Organisation. Als Referenztheorie zur Weiterentwicklung von MAS im hier dargelegten Sinne soll uns die Systemtheorie der Organisation dienen. Um dies zu begründen, gehen wir zunächst auf die bisher in den Forschungen zur Konstruktion von MAS entwickelten Formen des strategischen Managements ein.

3 Strategisches Management von MAS

Die Wirtschaftsinformatik, in der die Gestaltung und das Management von betrieblichen Anwendungssystemen realisiert werden, verwendet dazu Modellierungs- und Analyseansätze wie Spieltheorie, Entscheidungstheorie oder Prozessmodellierungsmethoden. Dies geschieht, um die Entscheidungsfindung in einer sich schnell ändernden, also dynamischen Welt zu verbessern. Zusätzlich besteht über diese Änderungen und die neuen Randbedingungen oft Unsicherheit (Risikomanagement). Bei der Anwendung von Agententechnologie fehlen jedoch Methoden zur Entwicklung von Agentensteuerungsmechanismen, die eine Adaptation einer mittelfristigen Strategie (z.B. durch Modifikation der Zielaktivierungsfunktionen) ermöglichen und damit eine Optimierung zwischen langfristigen unternehmerischen Zielen und kurzfristigen Änderungen der Umwelt (z.B. Marktschwankungen) realisieren. In den Termini der Agententechnologie handelt es sich dabei um eine Balance zwischen deliberativem und reaktivem Verhalten. Hierbei müssen eingesetzte Agenten fähig sein, angemessen und in *Echtzeit* in den zu Grunde liegenden partiell emergenten Plänen mit kritischen Pfaden, Engpässen und Ausfallrisiken (z.B. Unfälle) umzugehen. Ein besonderes Problem stellt die zentrale Rolle der Kommunikation dar. Im Rahmen von Standardisierungen, z.B. FIPA (1996) oder XML (W3C 2002), werden hierfür verschiedene Aspekte festgelegt, die das Austauschformat, die Struktur sowie die Abfolge von Nachrichten festlegen. Während Struktur und Format notwendige Bestandteile einer Standardisierung sind, bedeutet die Festlegung der Abfolge von Nachrichten in Kommunikationsprotokollen eine zusätzliche Einschränkung. Diese Schwierigkeiten haben ein besonderes Gewicht bei der verteilten Planung in Multiagentensystemen (vgl. Pechoucek 2003: 2). Unterschiedliche Partner können unterschiedliche Kommunikationsgewohnheiten haben, was zum einen aus den heterogenen Umfeldern entsteht. Andererseits wird dieses explizit zur Differenzierung von Angeboten im Bereich elektronischer Märkte gefordert (vgl. Lee et al. 2003). Als Beispiel soll ein Einkaufsagent herangezogen werden, der in internationalen, ihm unbekanntem Märkten operiert und sich auf die im jeweiligen lokalen und zeitlich veränderlichen Markt üblichen Verhandlungsstrategien einstellen muss. Derartige Agenten sollten in der Lage sein, sich auf solche Unterschiede dynamisch einzulassen und ihr Kommunikationsverhalten entsprechend zu verändern. Allgemeiner formuliert handelt es sich hierbei um ein Koordinationsproblem in Unternehmensnetzwerken. Eine Grundannahme für den Einsatz autonomer Systeme in solchen Umgebungen ist, dass durch das autonome Verhalten Flexibilität ermöglicht wird, die ein robusteres Verhalten ermöglicht (vgl. Kirn 2006). Reale Anwendungen besitzen eine Komplexität, die häufig auf die Nutzung von autonomen Systemen der höchsten Ebene (LoA 4, strategisch autonome Systeme) hindeutet. Hierbei betrachten die Agenten die individuelle Zielerreichung. Solche Systeme sind aber Teil von Organisationen, deren Grenzen sich über das autonome System hinaus erstrecken. In betriebswirtschaftlichen Anwendungen sind für diese übergeordneten Systeme

explizite Kennlinien oder Ziele bekannt, die teilweise auch formalisiert, z.B. in Form von Kennzahlen, vorliegen. Die Flexibilität, die erreicht werden soll, bedarf einer effizienten Verarbeitung des Entscheidungsverhaltens in den verteilten Systemen. Da die Entscheidungskomplexität in autonomen Systemen stark von den zur Verfügung stehenden Aktionen eines Agenten sowie von der Anzahl seiner Ziele abhängt, ist eine Berücksichtigung der Organisationsziele in der individuellen Entscheidungsfindung in jedem Entscheidungsschritt weder effizient noch gewünscht.⁶ Die Autonomie ist eine erwünschte Eigenschaft in Bezug auf Flexibilität, birgt jedoch inhärent das Risiko, dass durch die Interaktion lokal optimierender Einheiten keine globale Optimierung, z.B. bezogen auf die Ziele der Organisation, erfolgt. Die Unabhängigkeit der Entscheidungsfindung kann also zu unerwarteten und ungewünschten Systemzuständen führen, da konfligierende oder inkonsistente Zielmengen von den beteiligten Einzelagenten nicht erkannt und explizit berücksichtigt werden können.

Der konventionelle Ansatz zur Berücksichtigung strategischer Aspekte innerhalb eines Systems führt über die Manipulation bzw. Anpassung des Systems durch menschliche Nutzer bzw. Eigner, so dass die Rahmenbedingungen wieder eingehalten werden. Eine erste Möglichkeit der Flexibilisierung ist die Nutzung von Systemkomponenten, die die einzelnen Subsysteme überwachen und bei Zielverfehlung oder ungewünschten Systemzuständen in die Autonomie der Subsysteme eingreifen. In diesen Ansätzen wird die Autonomie der Teilsysteme (Agenten) durch einen externen Eingriff eingeschränkt. Im Rahmen unserer bisherigen Ausführungen gehen wir davon aus, dass die Autonomie eine positive und vom Systemdesigner bewusst gewählte Eigenschaft der Agenten ist. Wird die Agententechnologie z.B. im Rahmen von Unternehmensnetzwerken zur Verhandlung der heterogenen Interessen genutzt, so wäre ein Eingriff von außen in ein autonomes Subsystem, in diesem Fall in die Entscheidungsstruktur eines Unternehmensrepräsentanten, nur schwer vorstellbar. Daher schlagen wir einen neuen Ansatz zur Implementierung eines strategischen Managements vor, der auf der Hypothese beruht, dass ein übergeordnetes und formalisiertes Zielsystem vorliegt und dieses durch eine weitgehend autonome Interaktion der Agenten eingehalten werden kann. Die Vorstellung dieses Mechanismus, den wir als „Strategisches Management“ bezeichnen, orientiert sich an der Beispieldomäne einer betrieblichen Werkstattfertigung, in der die Agenten Ressourcen wie Werkzeugmaschinen repräsentieren, sowie technologisch an der BDI-Agentenarchitektur (vgl. Rao & Georgeff, 1996). Die BDI-Architektur mit den Erweiterungen zum Konfliktmanagement, die von Ingo Timm (vgl. 2004) eingeführt sind, ermöglicht es, die unterschiedlichen Ebenen der Autonomie abzubilden. Im Rahmen des strategischen Managements werden in Abhängigkeit des Grades der Zielverfehlung in der Architektur bestimmte autonome Verhaltensweisen eingeschränkt, so dass das Gesamtsystem wieder in einen stabilen Zustand, also einen Zustand ohne Zielverfehlung, zurückkehren kann.

⁶ Zur Veranschaulichung sei hier angemerkt, dass eine Entscheidungsfindung z.B. auf strategischer Ebene mit einer exponentiellen Komplexität in Abhängigkeit der Anzahl der Ziele erfolgt. Nun würde das wiederum bedeuten, dass z.B. die Hinzunahme von fünf Organisationszielen zu einer Komplexitätssteigerung von 5 im Exponenten verursacht.

Die grundlegende Annahme dieses Ansatzes ist, dass die Ziele formalisiert zur Verfügung stehen und automatisch ausgewertet werden kann, in wie weit ein Ziel erreicht oder nicht erreicht ist. In der Anwendungsdomäne einer Werkstattfertigung sind solche Zielsysteme durch die Spezifikation von Kennzahlen sowie deren gewünschter Ausprägung vorgegeben. So wird beispielsweise für eine Werkstatt häufig das Ziel „hohe Auslastung“ definiert, dass durch die Kennzahl „Auslastung“ gemessen werden kann (vgl. Nyhuis & Wiendahl 1999). Hierbei wird angestrebt, diese Kennzahl zu maximieren. Die Zielspezifikation sollte für die Nutzung im strategischen Management eine Funktion zur Messung der Zielerreichung enthalten. Diese sollte zumindest drei Ausprägungen von Zielerreichung unterscheiden können: Ziel erreicht (grüner Bereich), Ziel leicht verfehlt (gelber Bereich) und Ziel stark verfehlt (roter Bereich). In der im Technologie-Zentrum Informatik (TZI) der Universität Bremen durchgeführten prototypischen Implementierung dieses Konzepts sind sowohl die Definition von kennzahlenbasierten Zielen sowie die Definition der unterschiedlichen Zielverfehlungen leicht konfigurierbar berücksichtigt worden.

Die individuellen Akteure in dem System sollen durch autonome Subsysteme implementiert werden. Hierfür werden Intelligente Agenten auf der Basis der BDI/Diskursagenten vorgeschlagen. Das strategische Management soll durch einen 4-stufigen Prozess unterstützt werden, der in Abbildung 3 visualisiert ist.

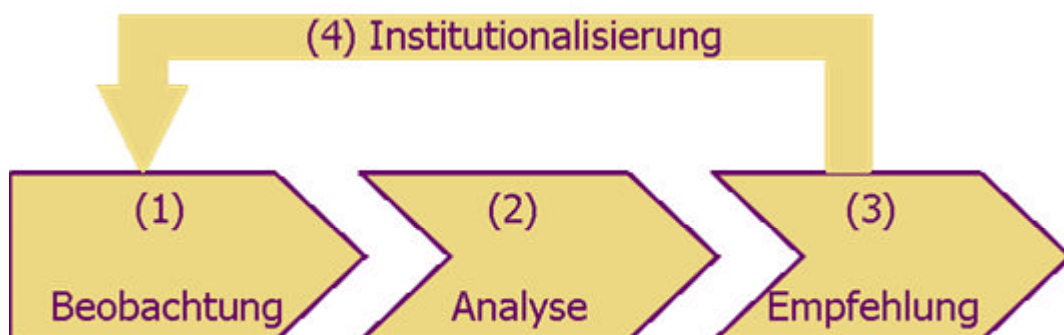


Abbildung 3: Phasen des Strategischen Managements in Multiagentensystemen (eigene Darstellung)

Im Normalbetrieb befindet sich das System in der ersten Phase (Beobachtung). In dieser Phase melden alle Agenten ihre individuellen Erfolgsparameter an eine spezielle Instanz, die die Systemziele sowie deren Erfüllungsgrad auf der Basis dieser individuellen Erfolgsparameter berechnet. Die Zielerfüllung sowie die aktuellen Ziele werden allen Teilnehmern des Systems zur Verfügung gestellt. Die Realisierung einer solchen Instanz kann entweder über eine Blackboard-Architektur erfolgen, in der das Blackboard zusätzlich die Fähigkeit erhält, neben der Kommunikationsvermittlung zwischen den Agenten auch die Berechnung der Ziele vorzunehmen. Im Rahmen des in Bremen entwickelten Prototyps wird diese Instanz durch einen Agenten realisiert, der eine Verarbeitung der individuellen Erfolgsparameter vornimmt und die Systemziele sowie deren Erfüllungsgrad regelmäßig kommuniziert. Hierbei wird auch die

Tendenz der Zielerfüllung mit abgeschätzt. In dieser Phase sollen die individuellen Agenten, ihr operatives Verhalten im Hinblick auf den Zielerfüllungsgrad von Systemzielen optimieren. Auf der operativen Ebene hat ein Agent die Möglichkeit, Pläne bzw. Aktionssequenzen zu generieren, um kurzfristige Ziele (Intention bzw. Absicht), die durch die taktischen und strategischen Ebenen vorgegeben sind, zu verfolgen. Wenn ein Agent bei dieser Planung alternative Aktionen oder Aktionssequenzen identifizieren kann, die für die Zielerreichung der individuellen Ziele gleichwertig sind, so wählt er die Aktionssequenzen, die unter Berücksichtigung der Zielerreichungsgrade des Systems am besten geeignet sind. In einer Realisierung kann diese Verfahrensweise so implementiert werden, dass ein Agent für die Erreichung des individuellen kurzfristigen Ziels unterschiedliche Aktionssequenzen simuliert und die Aktionssequenz auswählt, bei der die Systemziele am besten bewertet werden. Diese Adaption des Systemverhalten durch individuelle Berücksichtigung der aktuellen Zielerfüllungsgrade wird besonders in solchen Situationen gut funktionieren, in denen die Ziele nicht stark abweichen, d.h. wir propagieren die Nutzung dieses Ansatzes für Zielereichungsgrade grün und gelb. Wird über längere Zeit ein Systemziel leicht verfehlt (gelb) und die individuelle Adjustierung der operativen Autonomie war nicht in der Lage, die Zielabweichung zu reduzieren, so wechselt das System in Phase 2 (Analyse). In dieser Phase sollen die Agenten ihre individuellen Zielsetzungen im Rahmen der globalen Zielverfehlung interpretieren. Idealerweise sollen die Agenten jene kurzfristigen Ziele (Intentionen) vermeiden, die in Konflikt oder in Widerspruch zu solchen Zielen stehen, die auf systemischer Ebene verfehlt werden. Hierbei handelt es sich wiederum um Prozesse, die in der individuellen Autonomie behandelt und nicht von außen gesteuert werden. Ein Agent untersucht seine alternativen kurzfristigen Ziele, die er zur Verfügung hat. Hierbei werden die Interdependenzen zwischen den alternativen Zielen und den Systemzielen berechnet und es werden die individuellen Ziele als Alternativen in den nächsten Aktionen weiter verfolgt, die minimale Konflikte oder Widersprüche zu den Systemzielen aufweisen. Ist etwa in der Fertigung die Auslastung zu gering, so wird ein Ressourcenagent versuchen, neue Aufträge zu akquirieren, um das Systemziel der Auslastung zu unterstützen. In Hinblick auf die Autonomieebenen bedeutet diese Art der Adjustierung des Agentenverhaltens, dass der einzelne Agent seine taktische Autonomie einschränkt.

In vielen denkbaren und realen Szenarien wird jedoch eine unkoordinierte individuelle Handlung bei einer globalen Verfehlung nicht zu zufrieden stellenden Ergebnissen führen. Problematisch wird diese Situation dann, wenn viele Agenten versuchen, auf die gleiche Weise ein Globalziel zu erreichen. Dann kann das Globalziel übererreicht werden, was wiederum dazu führen kann, dass ein anderes Ziel verfehlt wird. Angenommen, die Fertigung hat sehr kurze Durchlaufzeiten und eine geringe Auslastung. Die Auslastung kann dann durch Annahme neuer Aufträge erhöht werden. Auf der Basis theoretischer und praktischer Erkenntnisse in der Fertigungstechnik ist aber bekannt, dass maximale Auslastung und minimale Durchlaufzeit nicht gleichzeitig erreicht werden können (die Ziele sind gegenläufig, vgl. Wiendahl & Mittendorf 2000: S. 1). Versuchen viele Agenten gleichzeitig die Auslastung zu erhöhen, so

kann es zu einer neuen Zielverfehlung bei der Durchlaufzeit kommen. Eine solche Wechselwirkung lässt sich nur durch koordinierte Aktionen verhindern. Daher wird in einem Fall andauernder starker Zielverfehlung (roter Bereich) die Phase 3 des strategischen Managements aktiviert. Diese Phase benötigt nicht nur eine passive Darstellung der Ziele und deren Zielerreichung, sondern darüber hinaus auch eine Koordinationsinstanz, die als Mediator eine Verhandlung über die Anpassung des individuellen Verhaltens ermöglichen soll. Es wird davon ausgegangen, dass durch die Schwere der Zielverfehlung die Agenten ein Interesse daran haben, an einer gemeinsamen Lösung zu partizipieren, damit eine robuste und verlässliche Systemausführung weiterhin gewährleistet werden kann. Daher adjustieren die Agenten in dieser Phase nicht die operative oder taktische Autonomie, sondern adaptieren ihre strategische Komponente, so dass ein Eingriff in die strategische Autonomie erfolgt, soweit diese implementiert ist. In dieser Phase werden nicht nur die zuletzt genutzten kurzfristigen Ziele und deren Interdependenzen mit den globalen Zielen untersucht; vielmehr werden in diese Analyse auch die persistenten Ziele (Desire) eingeschlossen. Der Mediator initiiert in dieser Situation einen Prozess, in dem die Agenten ihre unterschiedlichen Interessen verhandeln und sich über eine Anpassung ihrer Ziele einigen. Im Anschluss wählen die Agenten ihre Ziele, Pläne und Aktionen auf der Basis des Verhandlungsergebnisses aus. Der Mediator bleibt solange aktiv, bis die Zielerreichung des Systems sich wieder im grünen oder zumindest im gelben Bereich befindet. Im Rahmen des Agentensystems wird dieser Prozess durch eine vermittelte Rekonfiguration der individuellen Zielsysteme der Agenten erreicht. In unserer prototypischen Realisierung wird jedem Ziel eines Agenten eine Gewichtung gegeben. Durch diesen Mediationsprozess werden die Gewichte der Ziele ähnlich einem statistischen Lernverfahren angepasst, um eine für das System optimale Konfiguration zu ermöglichen. Nach erfolgreichem Abschluss der Phase 3 wird in Phase 4 (Institutionalisierung) das Verhandlungsergebnis, welches zum Erfolg geführt hat, als Regel für das System abgespeichert. So kann in einer vergleichbaren Situation auf die rechen-, kommunikations- und zeitkomplexe Berechnung verzichtet und direkt wieder ein stabiler Systemzustand (grüner Zielerfüllungsgrad) erreicht werden. Für unsere Implementierung werden in dieser Phase die kurzzeitig veränderten Zielgewichte in den Normalbetrieb übernommen.

4 Künstliche Sozialität in einer dynamischen Welt

Die hier vorgestellte Heuristik, also der technische Prozess Beobachtung, Analyse, Empfehlung und Institutionalisierung erinnert in vielen Punkten an die grundlegende soziologische Theorie der Reflexion sozialer Systeme, wie sie Luhmann formuliert hat. Luhmann hat in seiner allgemeinen Theorie die Frage nach der Reflexion von sozialen Systemen gestellt und in Teilen beantwortet. Interessant ist, dass er seine diesbezüglichen Überlegungen auf „Automaten“ (ebd.: 619) bezieht, wie sie im Anschluss an von Neumann entstanden sind, also auf Maschinen, die in sich selbst wieder vorkommen und deshalb als Maschinen mit der Fähigkeit

zur Reflexion verstanden werden können, die ihre Prozesse autonom, Luhmann würde sagen, selbstreferentiell reproduzieren und sich als System dadurch erhalten. Heinz von Förster (1993: 135) hatte in diesem Zusammenhang von nicht trivialen Maschinen gesprochen. Eine „triviale“ Maschine ist nach ihm „eine Anordnung von Regeln und Gesetzen, durch die gewisse Tatbestände in andere überführt werden“. Eine „nicht-triviale Maschine“ hat für ihn (ebd.: 136) „Regeln, die die ... tatbestandsändernden Regeln selbst wieder ändern: eine Maschine in einer Maschine, sozusagen eine ‚Maschine zweiter Ordnung‘“. Das hier entwickelte und diskutierte technische System des multikausalen Konfliktmanagements muss als eine Maschine zweiter Ordnung im Sinne Försters verstanden werden, weil hier zur Laufzeit neue „Kommunikationsregeln“ emergieren, die vom Systemdesigner nicht vorab festgelegt werden können und dennoch die Systemzustände nachhaltig verändern.

Die so möglich werdenden technischen Prozesse der Koordination von Kommunikation lassen sich mit Hilfe der Luhmannschen Theorie zum Zusammenwirken von Selbstreferenz, Reflexivität, Reflexion und Systemrationalität, die analog zum in Abbildung 3 veranschaulichten technischen Prozess Beobachtung, Analyse, Empfehlung und Institutionalisierung betrachtet werden kann, soziologisch interpretieren und verstehbar machen. Wird nämlich die Idee des multikausalen Konfliktmanagements technischer Systeme, also der komplexen Koordination von Kommunikationselementen auf den von Luhmann für die Soziologie etablierten Begriff der Selbstreferenz als basale und konstitutive Operationsweise von sozialen, aus Kommunikationselementen emergierenden Systemen bezogen, erscheinen Reflexivität und Reflexion als Sonderformen der Selbstreferenz (vgl. Luhmann 1984: 601f.) sozialer Systeme, also quasi als ihre nicht trivialen Formen. Reflexivität ist als prozessuale Selbstreferenz definiert, während Reflexion dann vorliegt, wenn die Selbstbeschreibung des Systems sich an der Unterscheidung von System und Umwelt orientiert, wenn sich also das System selbst als Ganzes in Differenz zu seiner Umwelt thematisiert. Basale Selbstreferenz bezieht sich auf die Ebene des einzelnen Kommunikationselements, also auf die elementare Ebene sozialer Systeme. Ohne diese dauerhafte Referenz auf die Elemente des Systems ist die Aufrechterhaltung der System-Umwelt-Differenz nicht möglich, so dass basale Selbstreferenz den Fortbestand, also die dauerhafte Reproduktion eines sozialen Systems erst ermöglicht. Prozessuale Selbstreferenz, also Reflexivität, orientiert sich an der Differenz von vorher/nachher und ist folglich auf die Varianz der möglichen Anchlusselemente im Prozess der Systemreproduktion bezogen. Sie bezieht sich auf die Ebene des Prozesses, also auf die prozessuale Ebene des Systems. Hier wird der Prozess der Verkettung von Elementen reflexiv, wenn sich ein neues Element etwa darauf bezieht, ob es anschlussfähig ist. Fragen wie, „Habe ich Sie richtig verstanden?“, oder „Ist dies so gemeint?“, sind typische Kommunikationseinheiten, die auf prozessuale Selbstreferenz hinweisen. Reflexion bezieht sich auf das System selbst, also auf die Ebene der Einheit der Differenz zwischen System und Umwelt. Hier geht es also um die Reflexion der Einheit des Systems.

Dieser Begriff der Reflexion kann auf Planung (Management) und Rationalität bezogen werden (vgl. Luhmann 1984: 635ff.). Systemrationalität wird innerhalb der Theorie selbstreferentieller Systeme als spezifischer, nämlich als anspruchsvollster Operationsmodus selbstreferentieller Systeme bestimmt (vgl. Kneer 1992: 95).⁷ Sie ist eine besonders unwahrscheinliche Form der Reflexion: „Rationalität ist erst gegeben, wenn der Begriff der Differenz selbstreferentiell benutzt wird, das heißt, *wenn auf die Einheit der Differenz reflektiert wird.*“ (Luhmann 1984: 640) Bei Planung und Rationalität kann es demnach nicht mehr um Kausalitäten gehen, die eindeutig rekonstruiert bzw. präkonstruiert werden können. Ein anspruchsvoller Begriff von Rationalität, der dem hier nachgezeichneten Begriff der Reflexion gerecht zu werden vermag, muss mindestens multikausal konstruiert werden. Denn nach Luhmann formuliert der Rationalitätsbegriff „die anspruchsvollste Perspektive der Selbstreflexion eines Systems. [...] Er bezeichnet ... den Schlusspunkt der Logik selbstreferentieller Systeme“ (Luhmann 1984: 645f.) Rationalität bedeutet die geplante Erzeugung von Selektivitätsverstärkung mittels Reflexion. Die Bedingungen für Rationalität sind im Anschluss an diese Definition sehr vielschichtig anzusetzen, denn bei Selektivitätsverstärkung wird die System-Umwelt-Differenz als Schema der Informationsgewinnung verwendet, indem Systeme „sich selbst durch ihre Differenz zur Umwelt bestimmen und dieser Differenz in sich selbst operative Bedeutung, Informationswert, Anschlusswert verleihen“ (Luhmann 1984: 641). Während Reflexion, also die Selbstreferenz des Systems in Differenz zu seiner Umwelt, per se Selektivitätsverstärkung bewirkt, wird bei Systemrationalität die Selektivitätsverstärkung durch den geplanten Wiedereintritt der Differenz von System und Umwelt in das System erzeugt. Der Übergang zwischen Reflexion und Systemrationalität ist deshalb fließend. Die Frage ist dann, „wie ein real operierendes System die Differenz von System und Umwelt, die es selbst erzeugt und deren Reproduktion es sich selbst verdankt, in das System wiedereinführen kann“ (Luhmann 2000: 462). Geschieht dies, erzeugt das System eine „Temporalisierung von Komplexität“ (Luhmann 2000: 464). Rationalität erscheint dann als institutionalisierte, geplante Reflexion, die im System eine geordnete Irritation erzeugt und dadurch Selektivitätsverstärkung bewirkt. Institutionalisierte Reflexion, die das Reflexionspotenzial des Systems dauerhaft abrufbar macht, erzeugt die Möglichkeit, „dass das System seine Einwirkungen auf die Umwelt an den Rückwirkungen auf es selbst kontrollieren muss, wenn es sich rational verhalten will“ (Luhmann 1984: 642).

Dieser Prozess ist hoch komplex und bedarf einer genauen Analyse. Hier kommt der Begriff des sozialen Mechanismus ins Spiel, weil gefragt werden muss, welche Mechanismen eine Institutionalisierung von Reflexion ermöglichen. Es gibt in sozialen Systemen Planungs- und Managementinstanzen, die sich als institutionalisierte Reflexion des Systems fassen lassen. Das von uns entwickelte System zeichnet diesen Institutionalisierungsprozess für technische Systeme nach und ermöglicht dadurch die experimentelle Untersuchung des sozialen Mecha-

7 „Aus Selbstreferenz kann man nicht auf Rationalität schließen. Selbstreferenz ist Bedingung für Steigerung, für Steigerung der Einschränkung, für Aufbau von Ordnung durch Reduktion von Komplexität.“ (Luhmann 1984: 638)

nismus' der Institutionalisierung von Reflexion. Es soll mit anderen Worten geklärt werden, in wiefern sich Reflexion institutionalisieren kann und welche sozialen Mechanismen dazu wirksam werden. Wenn basale und prozessuale Selbstreferenz nicht mehr ausreichen, um eine Systemreproduktion zu ermöglichen, wird Reflexion notwendig. Wird im System erkannt und kommuniziert, dass eine Reproduktion des Systems auf Dauer gefährdet ist, wenn Reflexion nicht dauerhaft geschieht, institutionalisiert sich Reflexion zur Herstellung von Systemrationalität. Dieser Institutionalisierungsprozess ist mit sozialen Mechanismen verbunden, die nachgezeichnet werden müssen, um Erklärungen für die Institutionalisierung von Reflexion zu gewinnen.

In der soziologischen Forschung zu sozialen Mechanismen wird allgemein anerkannt, „dass es in der Welt des Sozialen kaum universelle, deterministische Zusammenhänge zwischen zwei in ihren möglichen Ausprägungen konstant bleibenden Variablen gibt“ (Mayntz 2002: 21). Angesichts dessen bietet es sich an, den Begriff des sozialen Mechanismus für die soziologische Modellbildung der Institutionalisierung von Reflexion fruchtbar zu machen. Mit dem Begriff des sozialen Mechanismus' werden soziale Wirkungszusammenhänge beschrieben, die sich kausal rekonstruieren lassen. Soziale Mechanismen sind zunächst nicht mehr als ein Set relativ einfacher Ideen mit empirischem Bezug, die der Hypothesenbildung zur Erforschung sozialer Phänomene dienen. Sie stellen nach Renate Mayntz (2002: 25) „verallgemeinerte Wirkungszusammenhänge“ dar und sind als Unterkategorie der Kategorie Prozess anzusehen, weil sie als Wirkungszusammenhänge das „Wie“ des Zustandekommens sozialer Prozesse thematisieren, indem sie die „Mechanik“ sozialer Prozesse nachzeichnen (vgl. ebd.). Der Begriff des sozialen Mechanismus beschreibt empirische Regelmäßigkeiten, die als Prozesse der Stabilisierung oder Bildung sozialer Strukturen identifiziert werden können. Dennoch können soziale Mechanismen nicht ausschließlich mechanistisch im ingenieurwissenschaftlichen Sinne begriffen werden, indem ein monokausaler Wirkungszusammenhang zwischen feststehenden Variablen konstruiert wird. Die Ergebnisse sozialer Mechanismen, also ihre Folgen, die sich als Strukturbildungen darstellen, sind variabel. Sie können vorher nicht festgelegt werden. Genau diese Ergebnisoffenheit sozialer Mechanismen muss im Simulationsexperiment berücksichtigt werden, damit diese Experimente nicht nur theoretische Annahmen nachvollziehen, sondern auch Lösungsansätze für Theorieprobleme aufzeigen. Deshalb ist es in der sozionischen, auf Sozialsimulation zielenden Modellbildung, die im Spannungsfeld zwischen Reduktion und Adäquanz (vgl. Schimank 2002) angesiedelt ist, notwendig, bestimmte soziale Mechanismen als Instrumente der Erklärung von Strukturierungsprozessen auszuwählen und ihre Erklärungskraft zu plausibilisieren.

Ein weiterer Vorteil des Begriffs sozialer Mechanismus für die sozionische Modellbildung besteht darin, dass er sich gehaltvoll auf technische Mechanismen in Multiagentensystemen beziehen lässt. Zu beachten ist nämlich, dass in der sozionischen Simulationsforschung zusätzlich zu den soziologischen Kriterien auch technische Kriterien der formalisierenden Modellbildung berücksichtigt werden müssen. Die Auswahl der zu untersuchenden sozialen Zu-

sammenhänge ist also in nicht geringem Maße auf den Stand der entwickelten Technik bezogen. Im Kontext von Multiagentensystemen lassen sich soziale Mechanismen in iterativer Weiterentwicklung von hier bereits vorhandenen Mechanismen wie Aufgabendelegation und Ressourcenallokation modellieren und implementieren.⁸ Diese Weiterentwicklung von Multiagentensystemen, wie sie hier für das technische Konfliktmanagement durchgeführt wird, erzeugt nützliche Instrumente für die Sozialsimulation und kann selbstverständlich nur in Bezug auf eine soziologische Referenztheorie, wir haben zu diesem Zweck die Theorie der Reflexion Luhmanns gewählt, vollzogen werden. Dazu muss die Qualität der Theorie, die als Ausgangspunkt der Modellbildung fungiert, geprüft werden. Ebenso wenig wie sich eine Theorie mit einem inhärenten Widerspruch zur Modellbildung eignet, weil aus dem Theorie-widerspruch beliebig viele Aussagen ableitbar sind, eignet sich eine Theorie, deren Grundannahmen so schwach sind, dass nur wenige oder keine Aussagenzusammenhänge aus ihnen abgeleitet werden können. Die verwendete Theorie muss theoretische Annahmen enthalten, die sich mathematisch modellieren lassen, so dass die Konsequenzen der so formulierten Annahmen formal abgeleitet werden können. Nur so stellen sich die modelltheoretischen Vorteile der genaueren Spezifizierung der Theorie, der Vermeidung von Redundanzen und der Erhöhung der deduktiven Kraft einer Theorie ein (vgl. Ziegler 1972: 13ff; Müller-Benedict 2003: 26).

Wenn unstrittig ist, dass Mechanismen Kausalgesetze zu Grunde liegen (vgl. Schmid 2005: 51), muss geklärt werden, wie Kausalität in Bezug auf soziale Mechanismen gefasst werden kann. Es kann hier nicht um Monokausalität gehen, diese Sichtweise der Sozialität, die sich auf einfache Gesetzmäßigkeiten stützt, wird nicht nur von Luhmann mit guten Gründen, die sich in der Komplexität der sozialen Reproduktion schnell finden lassen, abgelehnt. Dennoch wird man im Kontext eines Erklärungsanspruchs soziologischer Forschung den Kausalitätsbegriff nicht vollständig ablehnen können, wie es in neueren Texten Luhmanns zuweilen geschieht. Benötigt wird im Kontext von Forschungen zur Sozialsimulation sozialer Mechanismen vielmehr eine der Komplexität sozialer Wirkungszusammenhänge angemessene Neudefinition des Kausalitätsbegriffs, wie es auch Luhmann in einem seiner frühen Arbeiten empfiehlt:

„Zunehmende Differenzierung bedingt ... wirksamere Mechanismen der Absorption von Komplexität durch Selektivitätsverstärkung. Diese Formel müssen wir an die Stelle der alten Antithese von Arbeitsteilung und Koordination bzw. von Differenzierung und Integration setzen.“ (Luhmann 1970: 103)

Es geht also bei der Betrachtung des sozialen Mechanismus' der Reflexion um die kausale Rekonstruktion von „Selektivitätsverstärkung“, also um die Klärung der Frage, *wie mit Hilfe einer institutionalisierten Reflexion neue Anschlussmöglichkeiten für das soziale System entstehen und wie diese dann genutzt werden, um Systemkomplexität handhabbar zu machen.*

⁸ Ein Beispiel hierfür ist die Weiterentwicklung des Äquivalententausches zum sozialen Mechanismus des Gagentausches mit Struktur bildenden Effekten (vgl. Alam et al. 2005).

Monokausale Erklärungen müssen, kurz gesagt, durch multikausale Erklärungen ersetzt werden, wenn soziologische Erklärungen der Komplexität der Sozialität, wie sie mit der soziologischen Theorie selbstreferentieller Systeme konstruiert wird, gerecht werden wollen. Dies gilt besonders für die soziologische Erklärung der Institutionalisierung von systemischer Reflexion, die einen der anspruchsvollsten Mechanismen der Reproduktion sozialer Systeme darstellt.

An dieser Stelle lässt sich mit gutem Recht fragen, ob es nötig ist, die komplexen Sachverhalte eines soziologisch fundierten Reflexions- und Rationalitätsverständnisses für technische Systeme zu reformulieren, oder ob technische Systeme mit einem kausalen Begriff der Rationalität hinreichend beschrieben sind, ob also eine Semantik der kausalen Mechanik von Prozessen für die Beschreibung technischer Zusammenhänge ausreicht. Diese Frage ist davon zu unterscheiden, ob technische Systeme die Anforderungen, die an sie gestellt werden, mit einfachen Mechanismen bewältigen können. Wie wir gleich eingangs deutlich gemacht haben, führt eine Steigerung des Grades der Autonomie der Entscheidungskompetenzen von Softwareagenten in einem Multiagentensystem zur Steigerung der Komplexität der zwischen derartigen Agenten ablaufenden Prozesse. Setzt man nun die Autonomie von Agenten als Bedingung der Möglichkeit zur Bewältigung von bestimmten Anforderungen an das technische System, wie wir es skizziert haben, handelt man sich zwangsläufig das Problem der Entstehung und Handhabung von komplexen Prozessen innerhalb von Multiagentensystemen ein. Technische Systeme ähneln dadurch immer mehr sozialen Systemen. Eine soziologische Theorie, die den Rationalitätsbegriff unreflektiert in monokausaler Weise benutzt, reicht zur Analyse der hier entstehenden technischen Prozesse nicht mehr aus, sowie sie der Analyse von gesellschaftlichen und anderen sozialen Prozessen nicht gewachsen ist, wenn es um das Aufdecken von komplexen Zusammenhängen der Sozialität gehen soll. Einfache, unreflektierte Theorien der Sozialität, die mit monokausalen Denkfiguren arbeiten, wie etwa die Rational Choice Theorie, eignen sich lediglich zur Beschreibung eines sehr kleinen Ausschnittes der Sozialität, etwa wenn sich Entscheidungen, eine sehr spezifische und unwahrscheinliche Form sozialer Praxis, kausal auf ihre Wirkungen hin rekonstruieren lassen. Dies ist unter sehr komplexen Umweltbedingungen ein höchst unwahrscheinlicher Fall. Ein eindimensionaler Begriff der Rationalität hat nur für derartig unwahrscheinliche Prozesse eine gewisse Plausibilität. Für das technische System ist dies genau umgekehrt. Kausale Prozesse sind die Regel. Multikausale Prozesse sind unwahrscheinlich und es wurde bisher innerhalb der Technikentwicklung daran gearbeitet, Technik so kausal wie möglich zu konstruieren. Nicht zuletzt deshalb ist Technik das Paradigma für Kausalität. Werden technische Systeme jedoch aufgrund der an sie gestellten Anforderungen im Paradigma der Autonomie von Agenten konstruiert, ändert sich dieser Sachverhalt grundlegend. Plötzlich entstehen Prozesse, die sich nicht mehr in monokausaler Weise rekonstruieren lassen, weil sie in ihren Wirkungen variabel sind. Sie zielen nicht mehr auf die Aussonderung von überraschenden, nicht in das Kausalitätsschema passenden Ereignissen, sondern haben das Ziel, aus dem in Beziehung setzen von autonomen

Agenten Ereignisse entstehen zu lassen, die zur Designzeit des technischen Systems nicht bekannt sind. Diese Flexibilisierung des technischen Systems, die man mit Luhmann als geplante Selektivitätsverstärkung innerhalb des Systems interpretieren kann, wird mit dem Ziel durchgeführt, dem System einen variablen Umgang mit potenziell unbekanntem Umweltereignissen zu ermöglichen.

Dies erinnert sehr an die Formen der Reflexion in organisierten Sozialsystemen, die dort sehr häufig im Management institutionalisiert sind. Denn mit dem Management verfügt das Unternehmen als Organisation über ein Reflexionszentrum (vgl. Paetow 2004: 52). Wird Reflexion im Management institutionalisiert, erhöht dies unter bestimmten Bedingungen die Irritabilität des Systems, indem Selektivitätsverstärkungen erzeugt werden. Die Systemtheorie der Organisation nach Luhmann (vgl. 1992; 2000) betrachtet diese Prozesse in erster Linie deskriptiv: Ihr geht es, wie Kai Paetow (vgl. 2004: 47ff.) nachweist, vorrangig um die Beschreibung der Autokatalyse des Organisationssystems, die mit Hilfe des Begriffs der soziokulturellen Evolution rekonstruiert wird. Die Frage nach der reflexiven Planung der Organisation, die „bestimmte künftige Merkmale des Systems [fixiert] und sie zu verwirklichen [sucht]“ (Luhmann 1984: 635), stellt sich in diesem Zusammenhang nur am Rande der Argumentation, wenn etwa gefragt wird, wie kommunizierte Planung zur Irritation des Systems führen kann und wie die Autokatalyse des Systems durch eine Verstetigung, also Institutionalisierung der Reflexion dieser Planung beeinflusst wird.

Aus der Beantwortung dieser Fragen mit Hilfe einer Theorie selbstreferentieller Systeme entwickelt die systemtheoretische Organisationssoziologie Gründe für einen generellen Planungs- und Steuerungspessimismus.⁹ Dennoch kann mit den hier vorgestellten Begriffen Widerspruch, Konflikt, Selbstreferenz, Reflexion, Rationalität und Planung eine Hypothesenbildung über die sozialen Mechanismen der Institutionalisierung von Reflexion und Konflikt-Management in Unternehmen entwickelt werden. Gerade die Umsetzung dieser Begriffe in einem technischen System erlaubt es, Fragen zu klären, die sich auf die Folgen bestimmter Reflexions- und Planungsschritte im Management einer Organisation beziehen. Denn unbestritten ist auch in der systemtheoretischen Organisationssoziologie, dass in Organisationen regelmäßig planende und reflektierende Entscheidungen getroffen werden, die nicht ohne Folgen für die Art der Reproduktion des Systems bleiben. Die Systemtheorie gibt jedoch mit Recht zu bedenken, dass die mit der Planung und der Reflexion kommunizierten Ziele im Reflexions- und Planungsprozess in der Regel nicht erreicht werden, dass die Zielerreichung also insgesamt durch Unsicherheit gekennzeichnet ist. Gerade dies macht es aber nötig, Reflexions- und Planungsprozesse der Organisation sehr genau zu erforschen. Die Simulationsforschung kann dazu einen Beitrag leisten, indem sie bestimmte Grundannahmen der Systemtheorie, die als analytische Modelle der Organisation verstanden werden können, zur konstruktiven Modellbildung nutzt. Dies ermöglicht eine empirische Sozialsimulation von Pro-

⁹ Typisch ist hierfür die folgende Aussage: „Es bleibt der Evolution überlassen, welche Entwicklungsschwerpunkte, welche Funktionssysteme, welche Strukturen sich ... mehr bewähren als andere.“ (Luhmann 1997: 770f.)

dabei darin, die Ausgestaltung des Mechanismus' zur Adjustierung von Autonomie in Abhängigkeit des Systemzustands und der Zielsetzung der Agenten zu implementieren. Hierbei liefert das soziologische Konzept der Reflexion, wie es von Luhmann formuliert wird, die entscheidenden Impulse zur Weiterentwicklung und Strukturierung des strategischen Managements von MAS. Denn die Variation der Systementwicklung beruht auf der Analyse und Interpretation des konkreten Systemzustandes durch das System selbst. Dieser Prozess geschieht in Hinblick auf eine potentielle oder bereits erfolgte Zielverfehlung des MAS. Durch diesen Mechanismus ist das System zur Selbstorganisation der eigenen Systemprozesse fähig und kann analog zu sozialen Systemen, wie sie Luhmann definiert, untersucht werden, indem die systemtheoretische Steigerungsformel basale Selbstreflex, Reflexivität, Reflexion und Systemrationalität auf den technischen Prozess der Institutionalisierung des strategischen Managements im MAS angewendet wird.

Die dadurch ermöglichte interdisziplinäre Konzeptionalisierung von Formen der technischen und sozialen Koordination von Kommunikationsprozessen kann als Ausgangspunkt einer umfassenden Untersuchung des Reflexionsmechanismus genutzt werden, um zum einen technische Systeme zu optimieren und um zum anderen den sozialen Mechanismus der Reflexion in seiner Operationsweise besser zu verstehen, als dies in Luhmanns Theorievorgabe geschieht. Folglich ist der Ertrag unserer Überlegungen nicht nur die Aufarbeitung und Klassifikation von technischer Autonomie, sondern auch die Fixierung eines interdisziplinären Konzeptes der dynamischen Koordination von komplexen Systemprozessen, die wir an der technischen Herausforderung der Koordination von Zielkonflikten in MAS, die aus relativ autonomen Agenten bestehen, durchgeführt haben. Dies führt zu einem neuen Modell des strategischen Managements von MAS, das sich nicht mehr primär auf die Aussonderung, sondern vor allem auf die kreative Absorption von Widersprüchen und Konflikten konzentriert. Diese technische Neuentwicklung erfüllt letztlich zum anderen die Bedingungen zur interdisziplinären Weiterentwicklung des sozialen Mechanismus' der Reflexion, wie er von Luhmann in die soziologische Forschung eingeführt wird, weil sie eine Kalibrierung von komplexen Modellen der Sozialsimulation ermöglicht. Dadurch kann die Operationsweise des sozialen Mechanismus der Reflexion umfassend experimentell untersucht werden, was zu einer Weiterentwicklung der soziologischen Theoriebildung genutzt werden kann.

6 Literatur

- Alam, Shah Jamal, Frank Hillebrandt, and Michael Schillo 2005: Sociological Implications of Gift Exchange in Multiagent Systems, in: Journal of Artificial Societies and Social Simulation (JASSS), 8(3) <http://jasss.soc.surrey.ac.uk/8/3/5.html>.
- Axelrod, Robert M. 1997: The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration. (Princeton Studies in Complexity) Princeton, NJ : Princeton University Press, 1997.

- Brückner, Sven 2000: Return from the Ant – Synthetic Ecosystems for Manufacturing Control. Dissertation. Mathematisch-Naturwissenschaftliche Fakultät II, Humboldt-Universität zu Berlin.
- Caltelfranci, Christiano and Rosaria Conte 1992: Emergent functionality among intelligent systems: Cooperation within and without minds. *Journal on Artificial Intelligence and Society*, 6(1):78–87.
- Ferber, Jacques 1999: *Multi-Agent Systems - An Introduction to Distributed Artificial Intelligence*. Harlow, UK: Addison-Wesley.
- Findler, Nicholas 1991: Uttam sengupta: An overview of some recent and current research in the ai lab at arizona state university. *AI Magazine*, 12(3):22–29.
- Fischer, Klaus, Michael Florian und Thomas Malsch (Hg.) 2005: *Socionics: Contributions to the Scalability of Complex Social Systems*, Berlin et al., Springer.
- Florian, Michael und Frank Hillebrandt (Hg.) 2004: *Adaption und Lernen von und in Organisationen*. Beiträge aus der Sozionik, Wiesbaden: VS-Verlag,
- Förster, Heinz von 1993: *KybernEthik*, Berlin: Merve.
- Hentze, Joachim, Peter Brose und Andreas Kammel 1993: *Unternehmensplanung - Eine Einführung*. 2. Auflage. Bern: UTB für Wissenschaft Uni-Taschenbücher GmbH
- Hillebrandt, Frank, Daniela Spresny und Matthias Hamsch 2004: *Sozialsimulation, Gabentausch und soziales Lernen. Konzeptionelle Überlegungen aus der Sozionik*, in: Michael Florian und Frank Hillebrandt (Hg.): *Adaption und Lernen von und in Organisationen*. Beiträge aus der Sozionik, Wiesbaden: VS-Verlag, S. 187-228.
- Kalenka, Susanne und Nicholas R. Jennings 1997: Socially responsible decision making by autonomous agents. In Kepa Korta, Ernest Sosa und Xavier Arrazola (Hrsg.): *Cognition, Agency and Rationality*, Dordrecht, The Netherlands: Kluwer, S. 135–149.
- Kneer, Georg 1992: Bestandserhaltung und Reflexion. Zur kritischen Reformulierung gesellschaftlicher Rationalität, in: Werner Krawietz und Michael Welker (Hg.): *Kritik der Theorie sozialer Systeme*. Auseinandersetzungen mit Luhmanns Hauptwerk, Frankfurt/M.: Suhrkamp, S. 86-112.
- Küpper, Hans-Ulrich, 1997: *Controlling - Konzepte, Aufgaben und Instrumente*. 2. Auflage. Stuttgart: Verlag Schäffer-Poeschel.
- Lasarczyk, Christian and Thomas Kron 2005: *Coordination in Scaling Actor Constellations*, in: Fischer, Klaus, Michael Florian und Thomas Malsch (Hg.) 2005: *Socionics: Contributions to the Scalability of Complex Social Systems*, Berlin et al., Springer, pp 199-217.
- Lüde, Rolf von, Daniel Moldt, Rüdiger Valk 2003: *Sozionik – Modellierung soziologischer Theorie*, Münster und Hamburg, Lit-Verlag.
- Luhmann, Niklas 1970: *Reflexive Mechanismen*, in: ders.: *Soziologische Aufklärung*, Bd. 1: Aufsätze zur Theorie sozialer Systeme, Opladen: Westdeutscher.
- Luhmann, Niklas 1984: *Soziale Systeme. Grundriss einer allgemeinen Theorie*, Frankfurt/M., Suhrkamp.

- Luhmann, Niklas 1992: Organisation, in: Küpper, Willi und Günther Ortman (Hg.): Mikropolitik. Rationalität, Macht und Spiele in Organisationen, 2., durchgesehene Auflage, Opladen: Westdeutscher, S. 165-185.
- Luhmann, Niklas 1997: Die Gesellschaft der Gesellschaft, Frankfurt/M., Suhrkamp.
- Luhmann, Niklas 2000: Organisation und Entscheidung, Wiesbaden, Westdeutscher.
- Malsch, Thomas 2004: Kommunikationsanschlüsse. Zur soziologischen Differenz zwischen realer und künstlicher Sozialität, Wiesbaden: VS-Verlag.
- Malsch, Thomas, Christoph Schlieder, Peter Kiefer, Maren Lübcke, Rasco Perschke, Marco Schmitt und Klaus Stein, Klaus 2006: Communication between Process and Structure: Modelling and Simulating Message Reference Networks with COM/TE. In JASS, to appear.
- Mayntz, Renate 2002: Zur Theoriefähigkeit makro-sozialer Analysen, in: dies. (Hg.): Akteure – Mechanismen – Modelle. Zur Theoriefähigkeit makro-sozialer Analysen, Frankfurt/M., New York, Campus, S. 7-43.
- Müller, Heinz-Jürgen 1993: Verteilte Künstliche Intelligenz - Methoden und Anwendungen. Mannheim: BI Wissenschaftsverlag
- Müller, Jörg P. 1996: The Design of Intelligent Agents: A Layered Approach. Lecture Notes in Artificial Intelligence: Bd. 1177. Berlin: Springer.
- Müller-Benedict, Volker 2003: Modellierung in der Soziologie. Heutige Fragestellungen und Perspektiven, in: Soziologie, 1/2003, S. 21-36.
- Nikles, Matthias and Gerhard Weiß 2005: Multiagent Systems without Agents – Mirror-Holons from the Compilation and Enactment of Communication Structures, in: Fischer, Klaus, Michael Florian und Thomas Malsch (Hg.) 2005: Socionics: Contributions to the Scalability of Complex Social Systems, Berlin et al., Springer, pp. 263-288.
- Nyhuis, Peter und Hans-Peter Wiendahl (1999): Logistische Kennlinien – Grundlagen, Werkzeuge und Anwendungen, Berlin: Springer.
- Paetow, Kai 2004: Systemevolution und Lernen in der Organisation, in: Michael Florian und Frank Hillebrandt (Hg.): Adaption und Lernen von und in Organisationen. Beiträge aus der Sozionik, Wiesbaden: VS-Verlag, S. 41-67.
- Pfeiffer, R. 1995: Cognition - Perspectives from Autonomous Agents. Robotics and Autonomous Systems 1995 (1995), Nr. 15, S. 25-46
- Rao, Anand S. und Michael P. Georgeff (1995): BDI-Agents: From Theory to Practice. In: Proceedings of the First International Conference on Multiagent Systems. San Francisco, California : AAAI-Press, S. 312-319.
- Rovatsos, Michael und Gerhard Weiss 2005: Autonomous Software. In Chang, S. K. (Hg.): Handbook of Software Engineering and Knowledge Engineering, volume 3: Recent Advances, River Edge, New Jersey. World Scientific Publishing.
- Russel, Stuart J. und Peter Norvig 1994: Artificial Intelligence: A Modern Approach. Englewood Cliffs, New Jersey: Prentice Hall.

- Russel Stuart J. und Peter Norvig 2003: *Artificial Intelligence: A Modern Approach*. 2nd Edition, Englewood Cliffs, New Jersey: Prentice Hall.
- Schimank, Uwe 2002: Theoretische Modelle sozialer Strukturdynamiken: Ein Gefüge von Generalisierungsniveaus, in: Mayntz, Renate (Hg.): *Akteure – Mechanismen – Modelle. Zur Theoriefähigkeit makro-sozialer Analysen*, Frankfurt/M., New York: Campus, S. 151-178.
- Schmid, Michael 2005: Soziale Mechanismen und soziologische Erklärungen, in: Hans-Jürgen Aretz und Christian Lahusen (Hg.): *Die Ordnung der Gesellschaft*, Frankfurt/M. et al.: Peter Lang, S. 35-82.
- Timm, Ingo J. 2004: *Dynamisches Konfliktmanagement als Verhaltenssteuerung Intelligenter Agenten*, Dissertationen in der Künstlichen Intelligenz (DISKI). infix - AKA Verlagsgruppe, Köln.
- Timm, Ingo J., Hans Kurt Tönshoff, Otthein Herzog and Peer-Oliver Woelk 2001 : *Synthesis and Adaption of Multiagent Communication Protocols in the Production Engineering Domain*. In: Peter Butala und Kanji Ueda (Hrsg.): *Proceedings of the 3rd International Workshop on Emergent Synthesis (IWES '01)*. Bled, Slovenia: LAKOS, Fakulteta za strojninstvo, S. 73–82.
- Weiß, Gerhard (Hrsg.) 1999: *Multiagent Systems - A Modern Approach to Distributed Artificial Intelligence*. Cambridge, Massachusetts: The MIT Press.
- Wiendahl, Hans Peter und Martin Mittendorf 2000: Kapitel Anmerkungen zum Gegenstand des Wörterbuchs In: *Wörterbuch der PPS: Deutsch-Englisch, Englisch-Deutsch*. Berlin: Springer-Verlag, S. 1–8.
- Wooldridge, Michael 1999: *Intelligent Agents*. In: Gerhard Weiss (Hrsg.) 1999 (siehe oben), S. 27–78.
- Wooldridge, Michael und Nicholas R. Jennings 1995: *Intelligent Agents: Theory and Practice*. In: *The Knowledge Engineering Review* 10 (1995), Nr. 2, S. 115–152.
- Ziegler 1972: *Theorie und Modell. Der Beitrag der Formalisierung zur soziologischen Theoriebildung*, München, Wien, Oldenbourg.