Der agentenorientierte, sozionische Ansatz mit PSI

D. Dörner[1], P. Levi[2], F. Detje[1], M. Becht[2], D. Lippold[2]

[1] Institut für Theoretische Psychologie, Universität Bamberg

[2] Institut für Parallele und Verteilte Höchstleistungsrechner, Universität Stuttgart

1. Einleitung

Im Rahmen der Arbeiten zur Erklärung menschlichen Handelns in komplexen Situationen haben wir eine kognitiv-emotional-motivationale Theorie (PSI-Theorie) entwickelt, die in weiten Teilen als Computerimplementation vorliegt. Das Computersystem kann sich in (fast) beliebigen Umgebungen autonom bewegen und für seine Bedürfnisse sorgen (Dörner, 1994a, 1996, 1999). Mit der Erweiterung dieses individualpsychologischen Ansatzes um sozionische Forschungsziele soll untersucht werden, wie sich die zusätzliche Einführung eines sozialen Motivs auf die Kooperation und die Gruppenstruktur von intelligenten, multimotivierten und emotionalen Agenten auswirkt.

Bei dem sozialen Motiv handelt es sich um ein Affiliationsmotiv, operationalisiert als Bedürfnis nach „Legitimitätssignalen“ (siehe unten), die ein Agent in bestimmten Abständen benötigt. Es soll nun erforscht werden:

- wie sich aus einem solchen Bedürfnis Bindungen zwischen den Agenten entwickeln,
- wie und ob sich aus diesen Bindungen zwischen einzelnen Agenten Gruppenstrukturen entwickeln,
- welche Gruppenstrukturen spontan entstehen,
- wie und unter welchen Bedingungen Gruppen mit spezifischen Strukturen besser oder schlechter mit verschiedenen Umweltbedingungen zurechtkommen,
- welche Bedingungen (Variation von Umweltbedingungen, Agentenmerkmalen, Ressourcenknappheit) zur Stabilisierung bzw. Destabilisierung der Gruppenstruktur beitragen,
- in welchem Ausmaß sich aus der Einführung eines solchen Bedürfnisses gruppenspezifische Phänomene wie z.B. Altruismus oder soziale Konflikte ergeben.

2. Computersimulation als Mittel der Theorienbildung und Forschungsüberblick

Unseren Projektzielen liegt ein interdisziplinäres Forschungsinteresse zugrunde. Gemeinsame Überzeugung ist eine systemwissenschaftliche Vorgehensweise, die sich bei der Theorienbildung auf die formalen Mittel der Computersimulation (Modellierung) stützt. Die Funktion der systemwissenschaftlichen Modellbildung liegt ganz wesentlich in der Verknüpfung von quantitativen und qualitativen Zugängen.

Darüber hinaus lassen sich weitere Funktionen der Modellbildung angeben (vergleiche Stachowiak, 1973; Wissel, 1989; Schiepek & Strunk, 1994, S.100):

- Explikation und Formalisierung von Hypothesen

- Konsistenzprüfung von Hypothesen
- Integration von Hypothesen
- Bedeutsamkeit einzelner Hypothesen
- Verdeutlichung der Konsequenzen eines Hypothesengefüges
- Ableitung neuer Hypothesen
- Interventionen in die Dynamik eines Hypothesengefüges
- Anleitung für die Datenauswahl bei empirischen Studien
- Verbesserung des qualitativen Verständnisses
- Ableitung von qualitativen Vorhersagen
- Verbesserung (grober) quantitativer Vorhersagen
- Verbesserung der Entscheidungsgrundlage im Anwendungsbereich.

So verstanden, sind Modelle wesentliche Werkzeuge, um mit den Annahmen, die ein Forscher bezüglich des betrachteten Realitätsbereichs hat, umzugehen und die Hypothesen auf die Realität rückzukoppeln. Ein wichtiger Unterschied zwischen Modell und Theorie ist, dass eine Theorie eine symbolische Beschreibung eines bestimmten Realitätsausschnittes darstellt, ein Modell hingegen eine (teilweise) analoge Beschreibung ist. D.h. ein Modell ist entweder ein Objekt in Raum und Zeit oder zumindest ein Prozess, welcher bezüglich mancher Merkmale analog zu dem Urbild ist.

Es gibt viele unterschiedliche Auffassungen, was unter einem „Modell“ zu verstehen sei. Wir möchten mit dem „Modell“-Begriff hier vor allem eine analoge Abbildung verstehen. Ein Modell „verdinglicht“ bestimmte Ausschnitte einer Theorie. Dazu wird die formale Symbolbeschreibung partiell in ein Objekt umgesetzt, das als Gegenstand in Raum und Zeit realisiert ist (z.B. ein Flugzeugmodell) oder als Prozess Verhalten in der Zeit realisiert (z.B. eine Computersimulation).

In der Psychologie (ähnlich auch in den Kognitionswissenschaften) herrschen Kleinsttheorien eng umschriebener Phänomenbereiche vor. In den letzten zwei Jahrzehnten wurde insbesondere durch die Entwicklung sogenannter kognitiver Architekturen versucht, dieser partikularistischen Entwicklung etwas entgegen zu setzen und Rahmenmodelle und –theorien größerer Reichweite anzubieten. Ein Vorteil solcher Rahmenmodelle oder Architekturen ist, dass sie einen wissenschaftlichen Spielraum definieren, der aus bestimmten theoretischen Annahmen und praktischen Hilfsmitteln besteht.

Kognitive Architekturen sind „relatively complete proposals about the structure of human cognition“ (Anderson, 1993, S.3). In dieser Hinsicht sieht Anderson den Unterschied zu Theorien, „which address only an aspect of cognition, such as those involving the distinction between long- and short-term memory“ (Anderson, 1993, S.3).

Den Begriff der Kognitiven Architektur hat Newell in die Psychologie mit seiner Arbeit über Computerarchitekturen eingebracht (Bell & Newell, 1971). Eine kognitive Architektur stellt, in Analogie zu einem Bauplan für ein Haus, eine *vollständige Spezifikation eines Systems* dar. Allerdings bleibt diese Spezifikation *abstrakt* und erfährt ihre Konkretisierung erst beim „Bau“ des Systems.

Im wesentlichen sind zwei große Gruppen von Architekturen im Gespräch: Produktionssysteme und Neuronale Netze. Unter Produktionssystemarchitekturen werden in der Regel Mo-

delle zusammengefaßt, die semantische Regeln für symbolische Informationen formulieren. Unter dem Begriff Neuronale Netze werden üblicherweise Modelle verstanden, die Informationen auf subsymbolischer Ebene durch Aktivitätsausbreitung verarbeiten.

Haben sich bei den Produktionssystemen die Systeme Soar (Newell, 1990, 1992; Laird, Newell & Rosenbloom, 1987; Rosenbloom, Laird & Newell, 1993) und vor allem ACT (Anderson, 1983, 1993; Anderson & Lebière, 1998) etabliert, sind im Bereich der neuronalen Netze die konnektionistischen und PDP Modelle (Rumelhardt & McClelland, 1992; McClelland & Rumelhardt, 1989) in der Diskussion.

Newell & Simon (1972) haben in Produktionssystemen immer eine Möglichkeit der psychologischen Modellierung gesehen, wobei sie Produktionsregeln vor allem als theoretische Beschreibungssprache menschlicher Kognition benutzen (erste Ansätze dazu finden sich bereits bei Selz, 1913).

Die Möglichkeit der Beschreibung psychologischer, (vor allem) kognitiver Phänomene mit Hilfe von Produktionssystemen ist weitgehend unstrittig. Strittig ist hingegen die Frage, ob Produktionsregeln eine psychologische Realität haben. Halten die meisten Autoren Produktionen nur für deskriptive Elemente (Smolensky, 1987; Hinton & J.A. Anderson, 1981), sieht J.R. Anderson mehr darin, denn „the significant regularities in human behaviour emerge at the level of production rules" (Anderson, 1993, S.10), weiterhin: „Thus, a production-rule analysis is actually critical to applying the connectionist program to complex behaviour" und schließlich „ACT* is a rule-based system (the productions being the rules), but one in which we are concerned with how these rules are implemented in a psychological and neurally plausible manner" (Anderson, 1983, S.35). Nach Anderson sprechen sowohl die hohe Plausibilität als auch die guten Prognosemöglichkeiten für die psychologische Realität von Produktionsregeln. An dieser Stelle greift Anderson unseres Erachtens aber zu weit. Die menschliche Psyche mag so agieren, als ob auch sie Produktionsregeln enthalte, dies stellt aber keineswegs einen Beleg dafür dar, dass Produktionen tatsächlich als Entitäten in der Psyche realisiert sind.

In der PSI-Theorie werden auf der Grundlage der Absichts- und Handlungsregulation verschiedene psychische Phänomene erklärt. Dazu werden die molaren Phänomene einer regelbasierten Ebene komplexer psychologischer Prozesse mit der molekularen Ebene neuronaler Netzwerkelemente (theoretische Neuronen, s. bereits Dertouzos, 1965) verbunden. Konzepte z.B. zur Wahrnehmung, zu Emotionen und zu Gedächtnisprozessen sind soweit formalisiert, dass sie als Computerprogramm der menschlichen Psyche implementiert sind. Das Verhalten dieser Computerimplementation wurde bereits hinsichtlich einiger Aspekte mit dem Verhalten von Menschen (Schaub, 1993, 1997; Hille, 1997; Bartl & Dörner, 1998) und kognitiven Architekturen (Westermann & Heise, 1996; Detje, 1999) verglichen. Die sozionische Frage, die sich stellt, ist, ob diese „künstlichen Seelen" (Schaub, 1996; Hille, 1997; Dörner, 1999) auch als Grundlage zur Modellierung sozialer und gesellschaftlicher Phänomene dienen können. Es zeichnet sich ab, dass es möglich ist, menschenähnliche Artefakte zu schaffen, die in der Lage sind, über indivdualpsychologische Fragestellungen hinaus auch soziologische Prozesse und Strukturen modellieren und erklären zu können.

Die Forschung zur Verteilten Künstlichen Intelligenz (VKI) beschäftigt sich von Haus aus mit Gruppen bzw. Gesellschaften von Agenten und muss sich daher auch mit "sozialen" Fragestellungen auseinandersetzen. Müller (1993) identifiziert das Gesellschaftskonstruktionsproblem als eine der zentralen Fragestellungen der VKI. Die bisherigen Arbeiten aus diesem Bereich lassen sich unterscheiden nach solchen, in denen "nur" soziale Beziehungen zwischen künstlichen Agenten untersucht werden und solchen, in denen die künstlichen Agenten vorwiegend mit Menschen kommunizieren.

Eine starke Strömung innerhalb der "künstlichen" Multi-Agenten-System (MAS) Forschung beschäftigt sich mit sogenannten "BDI"-Architekturen (BDI = "belief", "desire", "intention"; Rao, Georgeff & Sonenberg, 1992; Singh, 1994; Haddadi, 1996; Castelfranchi & Conte, 1995). Die Mehrzahl dieser Arbeiten versucht mit Hilfe verschiedener erweiterter Modallogiken grundlegende Zusammenhänge zu formulieren. Wenn der Begriff "sozial" verwendet wird, handelt es sich meist um die Übertragung von mentalen Konzepten eines Individuums auf eine Menge von Agenten. So werden z.B. bei Rao, Georgeff & Sonenberg (1992) "soziale" Pläne mit Hilfe modaler und temporaler Logik formalisiert. Soziale Pläne sind dort Pläne, in denen andere Agenten oder Gruppen von anderen Agenten und deren Wissen (Beliefs) und Absichten (Intentions) vorkommen. Dies ist sicher ein wesentlicher Aspekt für Agentengesellschaften. Da BDI und verwandte Architekturen jedoch von einem grundsätzlich rationalen Agentenbild ausgehen, sind darüberhinausgehende soziologisch oder psychologisch motivierte Theorien selten die Grundlage für die erstellten Modelle.

Viele der eher praktisch orientierten Arbeiten zu MAS (BDI basiert oder nicht) (Gasser, 1998; Fischer, Müller & Pischel, 1995) werden auf Probleme aus dem Produktionsumfeld (Produktionsautomatisierung) angewendet, wo Lösungen auf rein rationaler Ebene gefunden werden können und z.B. Emotionen, die integraler Bestandteil menschlichen Handelns sind, nicht hilfreich scheinen.

Im Oz Projekt an der Carnegie Mellon University (Bates, Loyall & Reilly, 1998) entstehen Agenten mit einem möglichst breiten Spektrum integrierter Fähigkeiten und Verhaltensweisen. Diese Fähigkeiten beinhalten zielgerichtetes reaktives Verhalten, emotionale interne Zustände und emotionales Verhalten, soziales Wissen und Verhalten und eingeschränkte Fähigkeiten natürliche Sprache zu verstehen und zu erzeugen. Dem emotionalen Verhalten liegt jedoch ein anderes Modell zugrunde als in den hier angestrebten Forschungen (die Modelle werden von den Forschern selbst als "shallow" bezeichnet). Das Modell ist nicht wie das von PSI bedürfnisorientiert aufgebaut und enthält somit auch kein Affiliationsbedürfnis. Obwohl auch Agentengesellschaften simuliert werden, liegt der Schwerpunkt der Arbeiten auf der Modellierung, Simulation und Untersuchung von Individualverhalten.

Doran (1995)untersuchte die Entstehung von Hierarchien in frühgeschichtlichen Kulturen. Dabei wurde mit Hilfe einer Multi-Agenten Simulation eine Theorie des Archäologen Mellers (1985) abgebildet, die besagt, dass für die Entstehung von Hierarchien Ressourcenknappheit der wesentliche Einflussfaktor ist. Die Theorie konnte anhand der Simulation nachvollzogen werden. Weitere Untersuchungen dieser Arbeitsgruppe beschäftigen sich mit der Frage,

inwieweit die Fähigkeit, die Zukunft vorhersagen zu können, für Agentengesellschaften von Vor- bzw Nachteil ist. Weiterhin wurde der Einfluss von kollektiven Irrtümern auf Agentengruppen untersucht. Ferner ist geplant, Emotionen in die Agenten einzubauen. Die Agentensimulationen laufen über mehrere Generationen (Schönherr, 1998; Doran, 1996).

Als Beispiel für die Arbeiten die sich mit der Agent-Mensch Interaktion beschäftigen soll hier Julia (Foner, 1993) erwähnt werden. Julia lebt in einer Welt mit verschiedenen Räumen, in denen sie sich bewegen kann, und die von anderen Agenten oder Menschen bevölkert wird. Die Teilnehmer in einer solchen Welt kommunizieren ausschließlich über textbasierte Dialoge. Die Fähigkeiten von Julia basieren jedoch nicht auf Modellen menschlichen Verhaltens. Julia reagiert lediglich auf verschiedene Stichworte und wählt dann aus einer Fülle von vorgefertigten Antworten regelbasiert aus. Die in einem Turing-Test Szenario erzielten Ergebnisse sind zwar recht beeindruckend; der Erkenntnisgewinn jedoch eher gering.

Die Untersuchungen innerhalb der Artificial Life Forschung (Adami, 1998) beschäftigt sich überwiegend mit der Simulation reaktiver MAS (Ferber, 1994) und die Leitbilder und Untersuchungsgegenstände kommen häufig aus der Insektenwelt. Es werden meist sehr große Population untersucht und der Aspekt der Vermehrung / Fortpflanzung spielt eine wesentliche Rolle. Grand & Cliff (1998) untersuchten relativ komplexe Agenten mit genetisch spezifizierten neuronalen Netzen, genetisch spezifiziertem biochemischen System und der Fähigkeit zur geschlechtlichen Vermehrung. Die einzelnen Gene repräsentieren Eigenschaften wie z.B. die “biochemische” Struktur, Aufbau des neuronalen Netzes und nicht-äußerliche Eigenschaften wie z.B. Neugier.

3. Grundlagen der PSI-Implementation und der Umwelt von PSI

Die PSI-Theorie stiftet einen Rahmen für die Abbildung verschiedener psychischer Phänomene. Dabei wird über rein kognitive Aspekte hinausgegangen: Emotionen, Bewusstsein und Wahrnehmung werden als Prozesse eines informationsverarbeitenden Systems beschrieben.

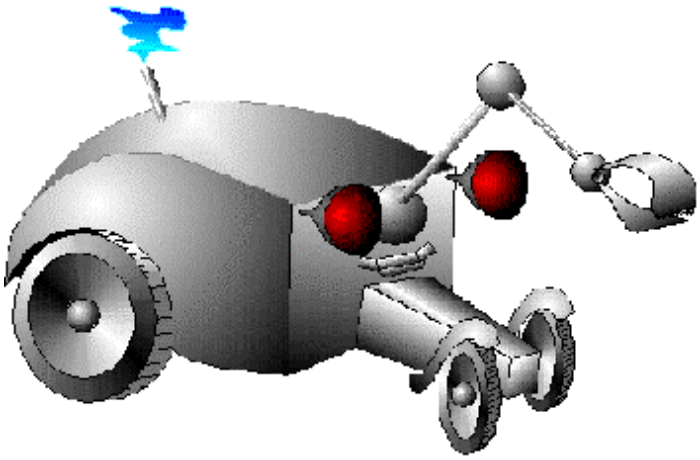

Abbildung 1: „James“, eine PSI-Inkarnation.

Abbildung 1 zeigt eine „Inkarnation“ des allgemeinen PSI-Modells als ‚James‘, der die Aufgabe hat, auf einer Insel (Abbildung 2) nach sogenannten ‚Nukleotiden‘ zu suchen (‚Nukleotide‘ sind Klumpen eines spontan strahlungsarmen und deshalb ungefährlichen Brennmateri-

als für Kernkraftwerke; die Sammlung von Nukleotiden ist nur eine Beispielaufgabe für James; man kann ihm andere Aufgaben geben).

James ist ein computersimulierter „Agent“, in dessem Inneren sich ein Dampfmaschinenmodell als Kraftzentrum befindet. Diese Dampfmaschine muss mit Wasser versorgt werden und ausserdem mit Brennstoff. Wasser kann in Teichen oder Bächen vorkommen; Brennstoff kann erworben werden, indem James nach Sonnenblumen bzw. Haselnüssen sucht, deren Öl für den Brenner der Dampfmaschine verwendet werden kann.

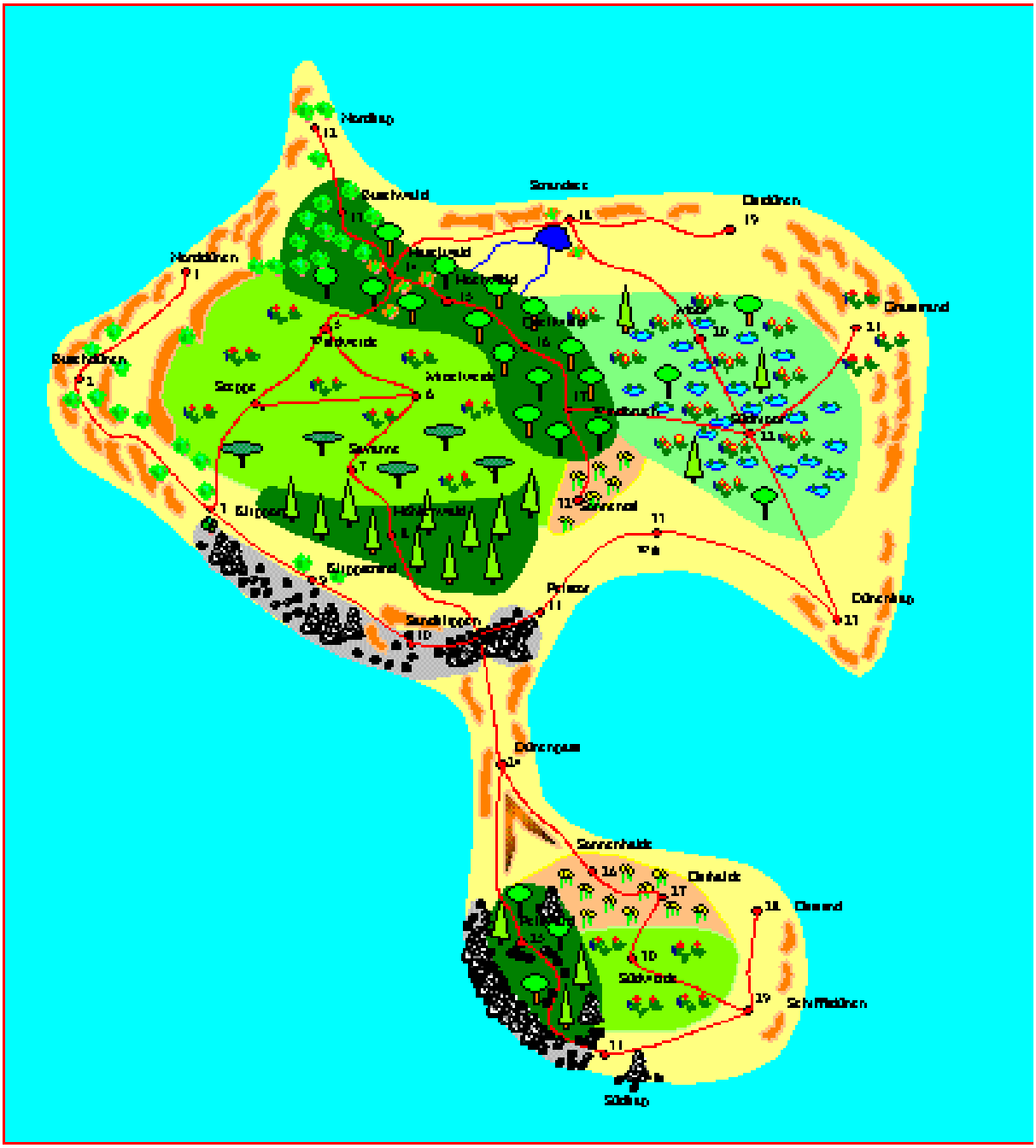

Abbildung 2: Eine „Insel“ als Beispiel für eine konkrete Lebensumwelt.

James muss sich ausserdem vor Gefahren schützen: Steinschläge können seine Aussenhaut beschädigen, Schwefelgase können die Rohrleitungen im Inneren des Roboters beschädigen. James spürt solche Beschädigungen als „Schmerzen“ und lernt in Zukunft die entsprechenden

Orte auf der Insel zu meiden. Das PSI-System James ist also lernfähig; ausser dass es lernt, welche Orte man vermeiden sollte, lernt es, an welchen Orten Wasser, Nüsse oder Sonnenblumen gefunden werden können; es lernt also Ziele. Zudem lernt es, wie man zu den entsprechenden Orten kommen und auf welche Weise man die Objekte manipulieren kann. Es erwirbt also Pläne, Wissen um die Art und Weise, wie Dinge zu manipulieren sind und ausserdem Ortswissen, Wissen um die Wege, die man einschlagen muss, um bestimmte Ziele zu erreichen. Aufgrund solchen Wissens kann James planen.

Die PSI-Modellstruktur ermöglicht nicht nur, dass James Ziele und Methoden erlernen kann, vielmehr verfügt James auch über eine perzeptive Lernfähigkeit. Für Objekte, die er nicht kennt, legt James neue Schemata in seinem Gedächtnis an, die ihm die Wiedererkennung ermöglichen. Objekte kann James in beliebiger Form erkennen und wiedererkennen. Die Fähigkeit zum perzeptiven Lernen bedeutet, dass man konkrete Realisationen von PSI, wie z.B. James auf der Insel, in beliebigen Umwelten agieren lassen kann. Wenn nur die Objekte der jeweiligen Realität als Strichmuster erkenn- und unterscheidbar sind, kann PSI mit ihnen operieren. Abbildung 3 zeigt ein Situationsbild aus der Insellandschaft, das PSI erkennen und dessen Objekte es wiedererkennen kann.

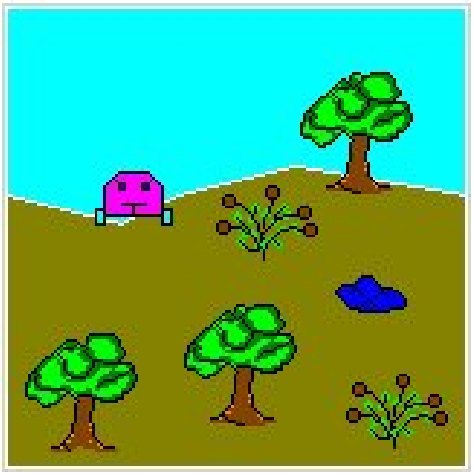

Abbildung 3: Blickfeld von James auf der Insel.

James verfügt über eine Reihe von Bedürfnissen. Neben den basalen Motiven, die der Existenzerhaltung dienen, also neben „Durst“, „Hunger“ (Brennstoffbedürfnis), „Schmerzvermeidung“ verfügt James noch über „informationelle“ Bedürfnisse. Eines davon ist das Bedürfnis, die Bestimmtheit seiner Umgebung möglichst zu maximieren, das andere ist das Bedürfnis, die Kompetenz, grob gesagt: die Handlungsfähigkeit, zu erhöhen. Diese Bedürfnisse entwickeln sich bei James durch das Erlernen von Zielen zu Motiven. D.h. im Laufe der Zeit lernt James, welche Situationen geeignet sind, bestimmte Bedürfnisse zu befriedigen.

Maximale (subjektive) Bestimmtheit ist dann vorhanden, wenn James in der Lage ist, die Effekte seiner eigenen Handlungen und den Ablauf von Geschehnissen fehlerfrei zu prognostizieren. Eine korrekte Prognose (z.B. des Effektes einer Handlung) ist ein „Bestimmheitsereignis“ (B-Ereignis) und erhöht den „Bestimmtheitspegel“, eine inkorrekte Prognose ist ein Unbestimmheitsereignis (UB-Ereignis) und senkt den Bestimmtheitspegel. James leidet unter einem niedrigen Bestimmtheitspegel und bemüht sich, diesen zu heben, indem er versucht, herauszufinden, welchen Gesetzmäßigkeiten die Geschehnisse in seiner Umgebung gehor-

chen, indem er also spezifisch exploriert. Das Bedürfnis nach Bestimmtheitssignalen und das Bedürfnis, Unbestimmtheit zu vermeiden, ist ein „informationelles“ Bedürfnis; PSI (bzw. James) strebt hier nicht nach bestimmten Formen von Materie, sondern nach Information (Dörner, 1994a, 1999).

Ein anderes informationelles Bedürfnis ist das Bedürfnis nach Kompetenz. Dieses besteht darin, dass James versucht, die Frequenz von Kompetenzereignissen (K-Ereignissen) zu maximieren. Ein K-Ereignis besteht für James darin, dass eine bestimmte seiner Aktionen einen Effekt hat. Ein besonders starkes K-Ereignis ergibt sich dann, wenn Aktionen zu Bedürfnisbefriedigungen führen. Dabei ist das K-Ereignis um so stärker, je größer das Bedürfnis war, bevor es befriedigt wurde, je größer also die Reduktion des Bedürfnisses ist.

Ein Mangel an K-Ereignissen führt zu einem niedrigen Kompetenzpegel; dieser wiederum kann dazu führen, dass James versucht, die Anzahl von K-Ereignissen zu steigern. Dies kann z.B. dadurch erfolgen, dass James „auf Abenteuer ausgeht“, d.h. Situationen hoher Unbestimmtheit anstrebt, wenn sein Bestimmtheitspegel hoch ist. James setzt sich also „mutwillig“ riskanten Situationen aus, um so neue Kompetenzen in neuen Situationen zu erwerben. Die Reduktion von Unbestimmtheit ist ein starkes K-Ereignis.

Ausser über die so eben dargestellten motivationalen „Antriebe“ verfügt die PSI-Modellstruktur noch über emotionale Steuerungen (Dörner, 1993, 1994b). Diese bestehen darin, dass in Abhängigkeit vom Gesamtbedürfnisdruck und in Abhängigkeit vom Bestimmtheits- und Kompetenzpegels bestimmte Verhaltenstendenzen auftreten bzw. bestimmte innere „Modulationsparameter“ gesetzt werden. Bei den Modulationsparametern handelt es sich um die Aktiviertheit, die Selektionsschwellle und den Auflösungsgrad.

Wenn z. B. der Bestimmtheitspegel niedrig ist und die Kompetenz hoch, so wird einmal die Aktiviertheit hoch sein und es gibt bei James eine Tendenz zu eher aggressiven Explorationen. James Verhalten weist unter diesen Umständen Kühnheit und Wagemut auf (s. das Grimmsche Märchen ‘Von einem der auszog, das Fürchten zu lernen’). Ist unter solchen Umständen aber der Kompetenzpegel niedrig, so wird James eher zum Rückzug in bekannte, gut beherrschte Bereiche neigen und wird nicht dazu tendieren, sich erhöhter Unbestimmtheit auszusetzen. Es zeigt in einer solchen Situation eher Fluchttendenzen, also „Angst“.

Die Aktiviertheit steuert ausserdem die Modulationsparameter ‘Auflösungsgrad’ und ‘Selektionsschwelle’. Der Auflösungsgrad ist die Genauigkeit von Vergleichen (in der Wahrnehmung oder beim „Denken“). Je höher der Auflösungsgrad, desto mehr Vergleiche gibt es zwischen den wahrgenommenen Reizen und internen Schemata; je geringer der Auflösungsgrad, desto weniger Vergleiche. Da Wahrnehmen (die „Identifizierung von Etwas als Etwas“) auf Vergleichen der Reize mit Gedächtnisschemata basiert, ergibt sich bei einem geringen Auflösungsgrad „Überinklusivität“. James wird unter solchen Umständen dazu tendieren, Dinge die einander nur ähnlich sind, miteinander zu verwechseln. Die Überinklusivität ist der Preis, der für die höhere Geschwindigkeit von Vergleichen bei einem geringeren Auflösungsgrad gezahlt werden muss. Überinklusivität bedeutet auf der einen Seite schnelle Reaktionsfähigkeit (Bedingungen werden nur grob in Rechnung gestellt, deshalb werden Aktionen für machbar

gehalten, die tatsächlich nicht ausführbar sind), auf der anderen Seite sind genau aus diesem Grund die Aktionen sehr riskant.

Bei erhöhter Aktivierung, also bei erhöhtem Bedürfnisdruck, steigt ausserdem die 'Selektionsschwelle'. Die Selektionsschwelle bestimmt die „Konfliktsensibilität" von James. Da die allgemeine PSI-Modellstruktur ja multimotiviert ist, streiten gewöhnlich Motive um den Zugang zur Handlungsregulation. Bei einer niedrigen Selektionsschwelle finden leicht „Absichtswechsel" statt; James geht als z.B. von der Absicht, Brennstoff zu suchen über zur Absicht, Wasser aufzutanken, wenn sich gerade die Gelegenheit dafür ergibt. Oder er bricht Explorationen ab. Eine hohe Selektionsschwelle erschwert solche Absichtswechsel, macht James also gewissermaßen konzentrierter, aber auch sturer und rigider.

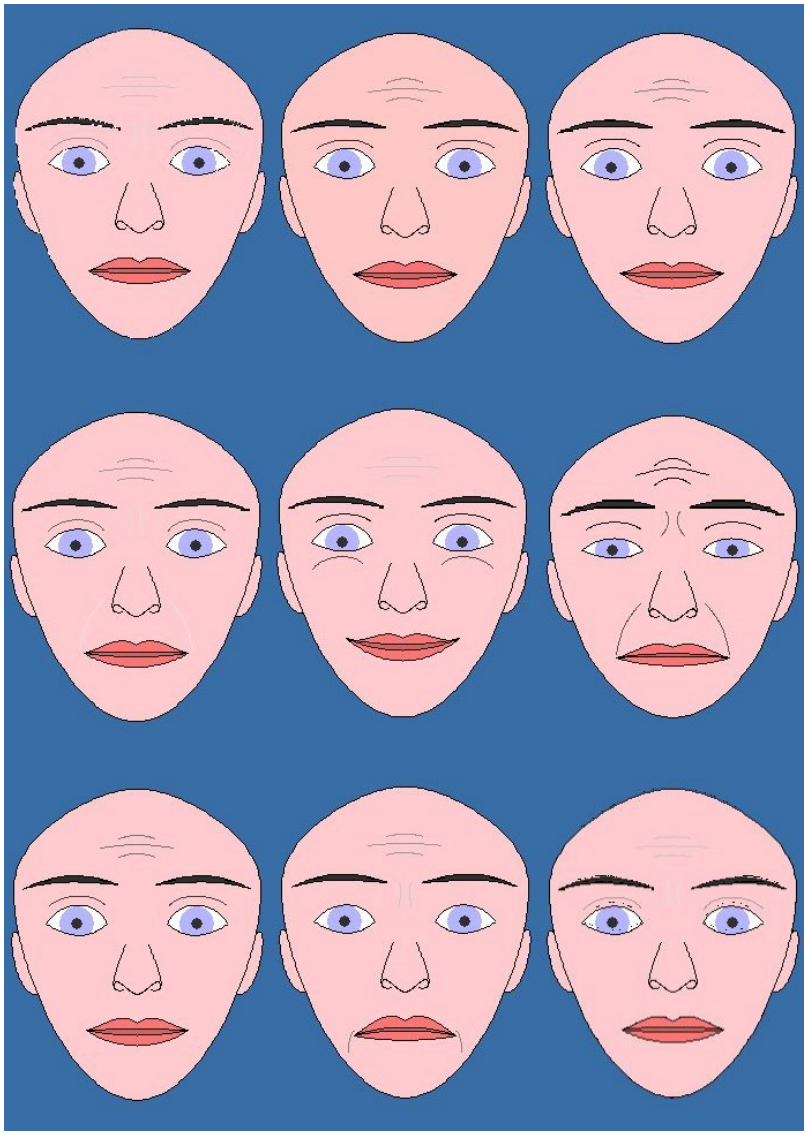

Abbildung 4: Gesichtsausdruck von James. In der Mitte: erfreut, in der Mitte unten: hilflos, in der Mitte rechts: schmerzerfüllt, oben links: entpannt.

Das Zusammenwirken der verhaltenssteuernden Wirkung des Kompetenzpegels und des Bestimmtheitspegels zusammen mit den Schwankungen der Modulationsparameter „Aktiviertheit", „Selektionsschwelle", „Auflösungsgrad" erzeugt bei James unterschiedliche Verhaltensweisen, die als verschiedene Emotionen identifizierbar sind (und die man auch dem Ge-

sichtsausdruck von James ablesen kann). Wir haben das PSI-Modell zum Zwecke der besseren Beobachtbarkeit mit einem Schemagesicht ausgestattet, dessen „Muskeln“ von den internen Zuständen von PSI, respektive James, entsprechend den Theorien über den Gesichtsausdruck und Gefühl angesteuert werden (Ekman, 1988; Hara, 1997). Abbildung 4 zeigt einige Beispiele für den Gesichtsausdruck von PSI (bzw. James) in verschiedenen emotionalen Lagen.

Das Handeln von James wird initiiert durch die Auswahl eines bestimmten Motivs oder Bedürfnisses (z.B. Wasserbeschaffung, Exploration, Kompetenzerwerb oder Schadensbehebung). Je nach Wissen um Methoden und Ziele aktiviert James dann einen Automatismus zur Zielerreichung, plant bestimmte Aktionen, wenn genügend Wissen über die jeweilige Realität in seinem Gedächtnis vorhanden ist oder geht – das ist die ultima ratio – zu einem Verhalten nach dem Prinzip „Versuch und Irrtum“ über.

Die Arbeiten an der PSI-Theorie haben auch zur Konstruktion verschiedener Simulationsprogramme geführt, die sowohl als Umwelten für PSI als auch als Untersuchungsgegenstand für menschliche Versuchspersonen dienen (u.a. „Garten“, „Labyrinth“, „BioLab“, „Landschaft“, „Insel“). Das Programm, dass wir für erste Untersuchungen im Rahmen des sozionischen Projekts herangezogen haben, ist das Umweltszenario „Insel“. Das Simulationsprogramm „Insel“ (gleiches gilt für „Landschaft“) ist durch seinen Programmaufbau eher als ein Konstruktionstool zur Konkretisierung verschiedener Umwelten (was die Agenten angeht) bzw. verschiedener Spiele (was die Versuchspersonen angeht) anzusehen. Die Variation der Umwelt ist in einem einfachen Texteditor möglich (die zu beachtenden syntaktischen Randbedingungen sind minimal), die Protokollierung ist vollständig und die integrierten Elemente sind auf die Fähigkeiten der Agenten und die Anforderungen an das Projekt abgestimmt.

Die „Insel“ (Abbildung 2) besteht aus einer beliebigen Anzahl von „Orten“ die auf beliebige Weise miteinander verbunden sein können („Wege“). Ein jeder Ort der Insel besteht aus einer beliebigen Anzahl von Objekten, die von den Agenten auf beliebige Weise manipuliert werden können und diese Manipulationen können beliebige Auswirkung auf den Agenten und / oder die Umwelt haben. Das Programm „Insel“ liest jeweils die Spezifikationen für eine konkrete Insel ein (ein erneutes Compilieren des Programms ist unnötig), lediglich semantische Kriterien beschränken die Möglichkeiten (so sollten Gegebenheiten, die dem „gesunden Menschenverstand“ widersprechen, vermieden werden).

Der „autonome Agent“ PSI, hier als „James“ (Abbildung 1) hat die Aufgabe, seine existentiellen Bedürfnisse nach Nahrung, Wasser und Schadensvermeidung zu befriedigen, gleichzeitig aber auch nach „Nukleotiden“ zu suchen und diese aufzusammeln. Zur Befriedigung der existentiellen Bedürfnisse gibt es auf der Insel Orte, an denen für seinen Betrieb notwendiges Wasser und Brennstoff vorhanden ist. Nukleotide sind z.B. in Dünen oder Felsen „versteckt“. Diese Objekte geben Nukleotide erst nach einer entsprechenden Manipulation frei. Die Wege zwischen zwei Orten zeichnen sich durch unterschiedlich gute Bodenverhältnisse aus (schlechter Boden verursacht Schaden) und durch ein gewisses (von Weg zu Weg variierendes) Risiko von Steinschlag aus, der ebenfalls zu Schaden führt.

James lernt die Objekte der Umwelt zu manipulieren und die Orte zur Befriedigung seiner Bedürfnisse anzustreben bzw. jene, die Gefahr bedeuten, zu vermeiden.

Um in seiner Umwelt tätig zu werden, stehen „James“ verschiedene Operatoren zur Verfügung (Abbildung 5).

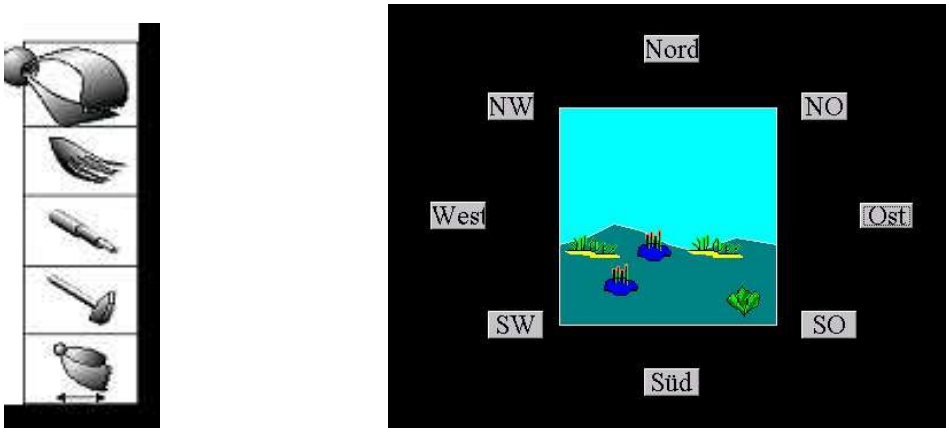

Abbildung 5: Die Manipulationsmöglichkeiten (Multi-Funktionsarm und Richtungsoperatoren).

Er kann sich durch Betätigung der Richtungsoperatoren von Ort zu Ort bewegen, sofern ein Weg vorhanden ist. Er kann sich den Objekten eines Ortes nähern (dies ist notwendig, um sie z.B. zu manipulieren) und sich von diesen Objekten auch wieder entfernen. Ferner ist James mit einem Multi-Funktions-Arm ausgerüstet, der es ihm erlaubt, die Objekte auf verschiedene Weise zu manipulieren. So kann er sieben, greifen, trinken, hämmern, schütteln. Die spezifische Wirkung (z.B. ob bedürfnisbefriedigend, bedürfniserzeugend, erfolglos oder umweltändernd) einer Manipulation eines bestimmten Objekts ist frei vorgebbar und wird aus einer entsprechenden Konfigurationsdatei durch das „Insel-Programm“ eingelesen (für Details siehe Detje, 2000)[1].

4. Ein agentenorientierter, sozionischer Forschungsansatz

In unserem Forschungsprojekt sollen die Entstehung und die Formen sozialer Strukturen in einer Gesellschaft von „autonomen Agenten“ untersucht werden, die sich ergeben, wenn man die autonomen Agenten mit einem „Affiliationsmotiv“ versieht. Es werden in diesem Projekt Agenten untersucht, die ausser künstlicher Intelligenz eine emotional-motivationale Handlungsregulation besitzen. Die Entscheidungsfindung der Agenten findet nicht hinsichtlich eines oder mehrerer vorgegebener Ziele statt, sondern die Ziele lassen sich auf grundlegende Bedürfnisse wie Hunger, Durst, Kompetenz und Bestimmtheit zurückführen. Insbesondere sollen die Agenten als Voraussetzung für soziale Bindungen ein Affiliationsbedürfnis erhalten. Dieser Ansatz begründet sich zum einen aus einer soziologischen Fragestellung zum anderen aber auch aus den z.B. in (Castelfranchi & Conte, 1998) dargestellten Grenzen der durch die Spieltheorie motivierten ökonomischen und strategischen Rationalität für Multiagentensysteme.

---

[1] Das PSI-Programm und die „Insel“-Umgebung stehen über URL: http://www.uni-bamberg.de/ppp/insttheopsy zum Download bereit.

Die Agenten handeln in einer simulierten Welt, die als „Insel" konzipiert ist, aber zu einer dreidimensionalen Welt ausgebaut wird und den Multi-Agenten-Betrieb zulässt. Die Agenten sollen zunächst das Ziel haben, ihr Leben zu erhalten und darüber hinaus das Nebenziel, sich die Insel „untertan" zu machen. Ziel des Projektes ist es, zu untersuchen, unter welchen Bedingungen sich aus einer Ansammlung individueller Agenten interessante makroskopische Strukturen herausbilden, welcher Art sie sind und wie diese zu erklären sind. Dazu muss eine Umwelt für die Agenten geschaffen werden, die sehr unterschiedliche Interaktionsmöglichkeiten bietet und die Ausbildung sehr vielfältiger Verhaltensmuster ermöglicht.

Es soll ein Simulationssystem für Gruppen intelligenter, multimotivierter, emotionaler Agenten entstehen. Mithilfe dieses Simulationssystems sollen bezüglich eines homogenen (d.h. nur aus Agenten bestehenden) Systems folgende Fragen untersucht werden:

- Welche makroskopischen Muster/Strukturen können in solchen Gruppen durch Selbstorganisation entstehen?
- Welche Rückwirkungen der makroskopischen Muster auf die Mikroebene der Agenten lassen sich beobachten?
- Wie robust sind die makroskopischen Muster gegenüber Veränderungen der Umwelt, bzw. der Agenten?

Weiterhin sollen Hybridgesellschaften aus Agenten und Menschen untersucht werden:

- Wie kann der Mensch im Sinne eines makroskopischen Ordnungsbegriffs integriert werden?
- Wie wirkt sich die Emotionaltät der Agentengesellschaft auf den Menschen aus?
- Wie lassen die sich simulativ gewonnenen Ergebnisse soziologisch interpretieren?

Die PSIs leben auf ihren Inseln bislang als 'Robinsons' ohne 'Freitag'. Sie versuchen, ihre jeweilige Lebensumwelt zu erkunden und sorgen sich um ihren Lebensunterhalt, um ihre Existenz. Sie versuchen ihre Kompetenz zu erweitern, indem sie explorieren. Aber sie leben einsam und haben auch keinerlei Bedürfnis nach der Gesellschaft anderer PSIs. Genau das wollen wir ändern. Wir wollen die PSIs mit einem Affiliationsbedürfnis versehen. Dies soll folgendermaßen beschaffen sein: PSI soll bei anderen PSIs Legitimitätssignale (L-Signale) (s. Boulding, 1978, S. 196 ff) wahrnehmen können und ausserdem Anti-Legitimitätssignale (AL-Signale). PSI bekommt – strukturell ganz gleichartig mit dem Bedürfnis nach einem bestimmten Wasserpegel oder einem „Bestimmtheitspegel" – ein Bedürfnis nach einem L-Signal bzw. ein Affiliationspegel. Jedes eintreffende L-Signal wird den Pegel ansteigen lassen, jedes AL-Signal aber wird den Pegel absenken.

Was sind L-Signale? Bei uns Menschen z.B. das Lächeln, viele Formen des Körperkontaktes, z.B. Streicheln, Umarmungen, darüber hinaus aber auch erlernte Verhaltensformen oder bestimmte Merkmale. Die Gleichartigkeit der „Uniformierung" wird also als L-Signal gewertet (genauer gesagt: als Ankündigung für mögliche L-Signale und mögliche Hilfeleistungen); das mag ein Grund dafür sein, dass Institutionen, die auf einen festen Zusammenhalt Wert legen, für irgendeine Art von Uniformierung – und sei es nur in der Form von Anstecknadeln am Revers – Wert legen. Ein L-Signal signalisiert implizit, dass man innerhalb seiner Gruppe willkommen ist, dass man Hilfe und Beistand erwarten kann.

Bei den PSIs wirken L-Signale normstiftend. Dies geschieht dadurch, dass die PSIs L-Signale nur dann absondern, wenn sie Bedürfnisbefriedigungen erfahren. Die Auffüllung also des Wasser- oder Brennstofftanks, die Verminderung von Unbestimmtheit, die Vermehrung der Kompetenz führt zur Absonderung von L-Signalen, die von anderen PSIs aufgenommen werden können. Die PSIs brauchen solche L-Signale, um das entsprechende Affiliationsmotiv zu befriedigen. Ein Mittel, um andere PSIs zu veranlassen, L-Signale abzusondern, ist folglich, diesen anderen PSIs Bedürfnisbefriedigungen zu beschaffen, umgangssprachlich gesagt: ihnen zu helfen. Hilfe (in beliebiger Art) ist also ein Instrument zu Erwerb von L-Signalen. Die PSIs werden erlernen, dass man durch bestimmte „altruistische" Aktivitäten L-Signale erwerben, bzw. Anti-L -Signale vermeiden kann; auf diese Weise etablieren sich soziale Normen.

Wichtig ist weiterhin, dass man die PSIs dazu bringt, ggf. supplikative Signale auszusenden. Formal gesehen ist ein supplikatives Signal (umgangssprachlich kann man das als Bitten oder Betteln bezeichnen) ein Ankündigungssignal für ein potentielles L-Signal. Im Klartext heisst ein supplikatives Signal: „Wenn Du mir hilfst, werde ich mich dankbar zeigen!" (In der Übersetzung: „Wenn du meine Bedürfnisse befriedigst, werde ich dir L-Signale zusenden!").

Weiterhin ist es wichtig, dass, wie bei allen anderen Bedürfnissen, auch im Zusammenhang mit L-Signalen Ziele gelernt werden können. Die PSIs lernen, wo sie Haselnussbüsche zu suchen haben oder wo es Wasser gibt. Und wenn sie „Hunger" oder „Durst" haben, streben sie diese Orte an. Das sind gewissermaßen die Eingangspforten für konsummatorische Endhandlungen und die darauf folgenden Bedürfnisbefriedigungen. Gleichartiges kann mit den L-Signalen geschehen. Wenn ein PSI mit z.B. grüner Aussenhaut (damit es unterscheidbar ist) „vereinsamt" ist (d.h. wenn sein L-Pegel niedrig ist) auf ein anderes PSI trifft, dessen Aussenhaut rot bemalt ist, und das ihm eine Fülle von L-Signalen verschafft, dann wird dieses rot bemalte PSI für unser einsames PSI eine Quelle für L-Signale werden, so wie ein Platz, an dem Haselnussbüsche wachsen, ein Ort ist, den man aufsucht, um seinen Hunger zu befriedigen. Etwas légère ausgedrückt: das einsame PSI wird sich in das rote PSI „verlieben". Es wird seine Nähe anstreben, wann immer der Pegel im L-Tank absinkt. Wenn sich aber nun das rote PSI einem anderen zuwendet, und seinen ja nicht unendlichen Vorrat an L-Signalen einem blauen PSI zukommen lässt, so mag es geschehen, dass das ursprünglich einsame PSI sich nunmehr frustriert sieht, in seinem Bemühen, von dem roten PSI L-Signale zu bekommen. Es ist ja auf dieses gewissermaßen „geprägt". Mit ein bisschen Analysefähigkeit könnte es herausbekommen, woran das liegt; „blau" zieht alle L-Signale von „rot" ab, also könnte es für „grün" ein Ziel werden, „blau" zu beseitigen. — Keineswegs muss also notwendigerweise die Einführung des Bedürfnisses nach L-Signalen zu einer harmonischen Gesellschaft von PSIs führen; vielmehr könnte es Krach und Streit und Eifersucht geben.

Insgesamt ist uns nicht bekannt, welche verschiedenen Formen von sozialen Beziehungen sich aus der Einführung des Bedürfnisses nach L-Signalen bei den PSIs ergibt. Es interessiert uns nun herauszufinden, was unter diesen Umständen geschieht. In welcher Weise wird sich die Gesellschaft der PSIs strukturieren? Wird sich so etwas wie Führerschaft entwickeln? Werden sich feste soziale Beziehungen entwickeln? Inwieweit werden die PSIs kooperieren? Was ergibt sich, wenn das Bedürfnis nach L-Signalen mit dem Bedürfnis nach Haselnüssen konkurriert?

Es erscheint uns, dass die Einführung eines Bedürfnisses nach Legitimitätssignalen die einfachste Form ist und zugleich die am eleganteste (von den uns bekannten), um soziale Beziehungen in einer Gruppe autonomer Agenten zu entwickeln. Es ist uns aber zum großen Teil unklar, welche Folgen sich daraus für PSI-Gruppen und auch für den Überlebenskampf der Einzelnen ergeben. Wir wollen daher in einer Reihe von Studien untersuchen, in welcher Weise sich soziale Beziehungen in einer PSI-Gesellschaft entwickeln, wie solche Beziehungsgefüge in Abhängigkeit von verschiedenen Bedingungen wachsen, sich ändern und auch wieder verfallen. Im einzelnen sollen folgende Bedingungen variiert werden:

Variation der Umwelt: Man kann 'schwierige' und 'leichte' Umwelten konstruieren, also Inseln, die wenig Gefahren und viele Gelegenheiten enthalten und solche, die viele Gefahren und nur schwer zugängliche und auch knappe Ressourcen enthalten. In den „schwierigen" Umwelten kommt natürlich der Hilfsbereitschaft eine besondere Rolle zu. Werden sich unter solchen Bedingungen die sozialen Beziehungen anders entwickeln als unter mehr „paradiesischen" Bedingungen mit vielen, leicht zugänglichen Ressourcen und wenig Gefahr? Werden unter allzu schwierigen Verhältnissen die sozialen Beziehungen zerfallen und einem „rette-sich-wer-kann!" - Verhalten Platz machen? Das kann sich durchaus ergeben, da nach der Konstruktion der PSIs sich die sozialen Bedürfnisse ja gegen die anderen Bedürfnisse durchsetzen müssen. Unter schwierigen Bedingungen kann sich durchaus eine „Mackie-Messer" - Gesellschaft entwickeln: „Erst kommt das Fressen, dann kommt die Moral". — Vielleicht sind hier auch die Übergänge bedeutsam; was geschieht mit den sozialen Strukturen, wenn man die Lebensumstände sich langsam verbessern oder verschlechtern lässt?

Variation der Personenparameter: Es soll untersucht werden, wie sich verschiedene Fähigkeitsprofile und „Persönlichkeitsmerkmale" der PSIs auf die sozialen Strukturen auswirken. Wir können die PSIs mit jeweils verschiedenen körperlichen, geistigen und Temperamentseigenschaften ausstatten. Wir können manche PSIs so gestalten, dass es ihnen z.B. leicht fällt, Felsbrocken zu zerschlagen; anderen fällt es schwer oder es ist ihnen gar nicht möglich. Manche können sich z.B. schnell bewegen, andere nur langsam. — Wir können manche PSIs mit einem guten, andere mit einem schlechten Gedächtnis ausstatten oder mit einer guten oder schlechteren Planungsfähigkeit. Wie wirken sich solche Unterschiede auf die sozialen Beziehungen aus? Wir können manche PSIs so gestalten, dass sie eher resignativ-depressiv sind (das kann man erreichen, indem man den Kompetenzpegel nur langsam ansteigen, aber schnell verfallen lässt) oder „kühn" und „wagemutig" (langsam absinkender und schnell ansteigender Kompetenzpegel). — Wir möchten homogene oder hochgradig gemischte PSI-Populationen untersuchen und überprüfen, in welcher Weise „Gleichheit" oder „Ungleichheit" sich auf die Entwicklung sozialer Gefüge auswirkt. Auch hier haben wir im Augenblick keine konkreten Hypothesen, in welcher Weise sich in verschiedenen PSI-Gesellschaften die sozialen Beziehungen entwickeln, wenn man Persönlichkeitsunterschiede einführt.

Man kann die Zusammenhänge zwischen den PSI-Agenten aus drei verschiedenen Blickwinkeln betrachten, nämlich im Hinblick auf die Eigenarten der Umwelt, im Hinblick auf die Eigenschaften der PSI-Agenten und im Hinblick auf die Vorgaben für Gruppenstrukturen.

Im Hinblick auf diese Parameter des Gesamtsystems lassen sich folgende Gruppen von Fragestellungen untersuchen:

1. In welche Kombinationen aus Umwelt-, PSI-Agenten und Gruppeneigenschaften stellen sich stabile Verhältnisse ein, welche Kombinationen „gehen unter", welche Kombinationen sind entwicklungsfähig?

2. Welche Phänomene aus den Bereichen Soziologie, Sozialpsychologie und Allgemeiner Psychologie lassen sich im Hinblick auf die stabilen, labilen und „untergehenden" Kombinationen beschreiben und analysieren?

3. Welche Theorien aus den Bereichen Soziologie, Sozialpsychologie und Allgemeine Psychologie sind mit den Ergebnissen vereinbar?

Diese Untersuchungen bauen schrittweise aufeinander auf. Zunächst sollte man versuchen, zu ermitteln, welche Motivausstattungen, Temperamentseigenschaften, kognitive Fähigkeiten mit schwierigen, einfachen, oder sich verändernden Umwelten kompatibel sind und welche nicht.

Sodann soll untersucht werden, welche sozialpsychologischen oder soziologischen Phänomene in den Gruppen auftreten. Hier ist z.B. zu untersuchen:

- In welchem Ausmaß und in welcher Weise tritt Führerschaft und Gefolgschaft, überhaupt Hierarchie, auf?
- Inwieweit führt die „Rücksichtstendenz" zu mittelmäßigen Leistungen der Gruppenmitglieder und hemmt oder verhindert extrem gute, vermindert aber auch extrem schlechte Leistungen (Badke-Schaub, 1993).
- Tritt so etwas wie *social loafing* (Harkins, Latané & Williams, 1980) auf? Dies ist die Tendenz zum „sozialen Faulenzen"; einige Agenten ruhen sich auf Kosten der anderen aus.
- Als *Groupthink* hat Janis (1972) eine mächtige Tendenz von Gruppen beschrieben, suboptimal zu entscheiden. Janis versteht unter Groupthink ein extremes Streben nach Harmonie und Einigkeit in einer Gruppe. Dieses Harmoniestreben schwächt die Fähigkeit der Gruppe zum flexiblen Handeln. Bei den PSI-Agenten könnte so etwas wie „Groupthink" in Ansätzen dadurch auftreten, dass die eigene Kompetenz aufgrund befriedigender Affiliation als hoch eingeschätzt wird. Kompetenz ist bei den PSI-Agenten ja abhängig von der Anzahl und der Stärke der Bedürfnisbefriedigungen. Bei einem starken Affiliationsbedürfnis und starken Bedürfnisbefriedigungen stellt sich ein starkes Kompetenzempfinden ein, welches dann die einzelnen Gruppenmitglieder verleitet, sich auf riskante Unternehmungen einzulassen, denen ihre Fähigkeiten nicht gewachsen sind.

Damit sind einige sozialpsychologisch / soziologische Phänomene angerissen, die im Rahmen von PSI-Agenten-Szenarien untersucht werden sollen.

Neben der Möglichkeit bestimmte allgemeinpsychologische, sozialpsycholgische und soziologische Phänomene abzubilden und zu analysieren, können im Rahmen der PSI-Agenten-Szenarios auch theoretische Konzepte der Psychologie und Soziologie abgebildet und analysiert werden (vergleiche Hannemann, 1988).

- Allgemeinpsychologisch ist dabei vor allem das gruppenspezifische *Affiliationsmotiv* zu denken: Welche Rolle spielt es? Ist ein gesondertes Affiliationsmotiv überhaupt notwendig, etc. Aber auch der wesentliche Aspekte der funktionalen Einbettung des Emotionskonzeptes und dessen Rolle für soziale Strukturen ist darstellbar und analysierbar.

- Sozialpsychologische Fragen der Rollenaufteilung und der *Verteilung von verschiedenen Kompetenzen* auf verschiedene Gruppenmitglieder sind darstellbar. Welche Funktion, welchen Nutzen haben unterschiedliche Gruppenstrukturen in Abhängigkeit von Variationen in den Eigenschaften der Agenten und/oder der Umwelt.

- Soziologische Fragen und Konzepte lassen sich in Hinblick auf die Beziehung zwischen verschiedenen Teilgruppen des Szenarios darstellen und deren Beitrag zu Funktion und Handlungsweise des gesamten Agentenkollektivs. Welche *Organisationsform der Agenten* jenseits der Gruppe hat in Abhängigkeit von welchen Eigenschaften der Agenten und Teilgruppen welche Effekte?

- Generell dürften Konzepte aus dem Umfeld der verschiedenen Systemtheorien die interessantesten Einblicke eröffnen. Aspekte der *Selbstorganisation* und der *Synergie* werden durch das konzipierte PSI-Agenten-Szenario im Rahmen von sozialpsychologischen und soziologischen Blickwinkeln *experimentell* untersuchbar.

5. Zusammenfassung der ersten Forschungsergebnisse

In den bisherigen Arbeiten mit dem PSI-System standen (individual-)psychologische Fragestellungen im Vordergrund. Da wir in diesem Projekt die Entstehung und die Formen sozialer Strukturen in einer Gesellschaft von PSI-basierten, autonomen Agenten untersuchen werden, haben wir in einem ersten Schritt den PSI-Agenten mit einem sozialen Motiv versehen.

Vor dem Hintergrund des Implementationsstands von PSI haben wir schrittweise Änderungen an PSI und der Umwelt vorgenommen, die sich sozionischen Forschungszielen nähern, um die durch diese Änderungen bewirkten Konsequenzen aufzuzeigen und zu analysieren. Es bleibt zu bedenken, dass in diesem ersten Schritt PSI / James weiterhin allein auf seiner Insel leben musste, denn sowohl die Implementation der PSI-Theorie als auch die Umwelt waren zum Zeitpunkt dieser Experimente noch nicht auf einen Multi-Agenten-Betrieb umgestellt. Was wir hingegen tun konnten, war PSI mit einem weiteren Bedürfnis, weiteren Handlungsmöglichkeiten und die Insel mit weiteren Objekten auszustatten und die Konsequenzen dieser Änderungen zu beobachten.

Für diese Versuchsreihe variierten wir die folgenden Parameter:

- Wir führten zuerst ein Affiliationsmotiv ein (A+), beließen aber die Handlungsmöglichkeiten (M-) und die Umwelt, wie sie waren, d.h. PSI hat keine Möglichkeit, das neue Bedürfnis zu befriedigen (Gruppe A+M-I).

- Dann führten wir einen Operator ein (M+), der es PSI ermöglicht, an einem adäquaten Objekt das Affiliationsbedürfnis zu befriedigen, änderten aber die Umwelt nach wie vor nicht, d.h. es gibt kein entsprechendes Objekt (Gruppe A+M+I).

- Als nächstes sahen wir an einem Ort bzw. an mehreren Orten auch die entsprechenden Objekte vor (IxT), so dass PSI sein Affiliationsbedürfnis durch Aufsuchen dieser Orte

und die Anwendung des entsprechenden Operators auf die dafür vorgesehenen Objekte befriedigen kann (Gruppe A+M+I1T bzw. A+M+I8T).

- Als Kontrollgruppe dienen uns PSIs, die weder ein Affiliationsbedürfnis, noch erweiterte Handlungsmöglichkeiten, noch neue Objekte in der Umwelt besitzen (Gruppe A-M-I).

Wir erhielten folgende Ergebnisse: Statten wir PSI mit einem Affiliationsbedürfnis aus, das aufgrund gleichbleibender Umwelteigenschaften unbefriedigt bleiben muss, so wechselt PSI öfter die Orte, exploriert die Objekte eines Ortes aber seltener. Auch die internen Prozesse und Zustände verändern sich. PSI nimmt weniger Motivwechsel vor und es fallen ihm weniger zielführende Assoziationen ein. Kann PSI sein Affiliationsbedürfnis befriedigen, weil es Objekte in der Umwelt gibt, die dies ermöglichen, so erlischt diese Tendenz zum „Herumstreunen“. Die durchschnittliche Aktivierung steigt an und das Verhalten wird insgesamt wieder erfolgreicher und stärker geprägt von Versuchen, die Objekte der Umwelt zu erkunden. Dies in zunehmend stärkerem Maße, wenn sich die Objektanzahl, die das Bedürfnis befriedigen kann, erhöht (für Details siehe Detje, 2001).

6. Die nächsten Schritte

Für die Untersuchung verschiedener PSI-Gemeinschaften wird ein Multi-Agenten-Simulator entwickelt, der das Handeln einer Menge von Agenten in einer dynamischen Umwelt ermöglicht. Die in dieser Umwelt handelnden Gruppen von Agenten können aus künstlichen oder menschlichen Individuen bestehen, oder hybrider, d.h. gemischter, Natur sein. Parallel zur Entwicklung einer geeigneten Umgebung für die Agenten, wird ein formales Modell (im Sinne der Informatik) der zugrunde liegenden PSI-Architektur entwickelt (formuliert in der "Unified Modeling Language"), um die Funktionen der Modellbildung (siehe oben) optimal zu unterstützen.

Um im Rahmen des sozionischen Forschungsansatzes das Verhalten von Mengen unterschiedlicher Agenten in Gesellschaften und das Verhalten dieser Gesellschaften insgesamt zu untersuchen, stellen sich die folgenden Aufgaben in Bezug auf die PSI-Theorie:

1. Prüfung der Kommunikationsmöglichkeiten von PSI, damit beliebige Agenten Informationen über sich, die Umwelt und die anderen Agenten austauschen können.

2. Prüfung der Fähigkeiten von PSI zur Kooperation und zur Konkurrenz mit anderen Agenten. Unterschiedliche Verfahren zur Kooperation sollen von den künstlichen Agenten selbst erlernt werden.

3. Erweiterung der kognitiven Fähigkeiten, insbesondere der Sprache, der Lernstrategien und der Planungsheurismen, damit PSI-Gruppen in dynamischen Umgebungen agieren können und dort komplexe, interaktive Verhaltensmuster und gesellschaftliche Strukturen entstehen können.

Für eine Kommunikation zwischen Agenten müssen diese in der Lage sein,

- andere Agenten als solche zu erkennen,
- Information zu verarbeiten, die andere Agenten erzeugen und

- auch selbst in der Lage sein, Information weiterzugeben.

Die Fähigkeit zur Erkennung anderer Agenten unterscheidet sich nicht prinzipiell von der Fähigkeit zur Erkennung der Objekte in der Umgebung, so dass PSI dazu bereits in der Lage ist. Insofern kann dieses Problem als gelöst gelten, sofern die individuellen Agenten, individuelle (z.B. optische) Merkmale besitzen. Die Fähigkeit zum Empfang und zur Verarbeitung von Signalen (oder Nachrichten) ist PSI konzeptuell gegeben, da PSI nicht prinzipiell auf die Wahrnehmung optischer Reize beschränkt ist. Dasselbe gilt für das Aussenden von Signalen (Nachrichten), da PSI nicht prinzipiell auf das Manipulieren von Objekten beschränkt ist.

Die Kommunikation zwischen zwei (oder mehreren) Agenten kann z.B. über die Lage bestimmter Objekte auf der Insel stattfinden (Agent A gibt Agent B die Information, dass es an der Küste Hasselnusssträucher gibt). Es können aber auch Verfahrensweisen (wie kommt man zur Küste) oder innere Zustände (z.B. über Energiemangel) weitergegeben werden (Agent A könnte Agent B die Information geben, dass er Nahrung braucht).

Die Sprache der Agenten müsste Begriffe definieren („Semantik"; was das Schema-Gedächtnis von PSI „automatisch" leistet) und über Regeln verfügen, diese zu Begriffsfolgen zu verknüpfen („Syntax"; was über PSIs „Aktionsschemata" prinzipiell bereits möglich ist). Schließlich müssten kommunizierende PSI-Agenten Regeln zur Verwendung der Begriffe bzw. Begriffsfolgen festlegen (Pragmatik), was sie, wie andere Ziele auch, erlernen sollten. Als „sprachlicher Ausdruck" könnten dabei in erster Linie das schon vorhandene Gesicht von PSI dienen. Ein Lächeln könnte hier z.B. bedeuten „Danke schön!"

Aufbauend auf die Fähigkeit zur Kommunikation werden die praktischen Verfahren zur Kooperation entwickelt. Abstrakte Beschreibungsverfahren für Kooperation in Multiagentensysteme wurden von Muscholl & Levi (1996) vorgestellt. Die Übertragbarkeit auf PSI-basierte Multiagentensysteme wird derzeit geprüft. Muscholl & Levi (1996) sowie Becht, Muscholl & Levi (1997) haben auf der Modellierungsebene Verfahren zur Erfüllung von kooperativen Aufgaben mit Hilfe einer Spezifkationssprache (siehe Becht, Muscholl & Levi, 1998) unabhängig von konkreten Agenten beschrieben. Die Möglichkeit zum Einsatz in PSI-basierten Gruppen wird derzeit überprüft. Die dort entwickelten Konzepte sind geeignet, um kooperative Lernverfahren mit unterschiedlichen und wechselnden Agentengruppen zu spezifizieren und zu realisieren.

Unseres Erachtens können nur dann komplexe und soziologisch relevante Verhaltensweisen und gesellschaftliche Strukturen entstehen, wenn PSI neben motivationalen Antrieben, einer emotionalen Verhaltensmodulation auch über menschenähnliche, kognitive Fähigkeiten verfügt. Zur Erweiterung der kognitiven Fähigkeiten sind als Voraussetzung die derzeit in PSI realisierten Verfahren des Lernens und Planens sowie die Anforderungen an das Lernen in Bezug auf die Kooperation von Agenten zu untersuchen, formal zu beschreiben und weiterzuentwickeln. Dies betrifft vor allem die folgenden Aspekte:

- Welche Informationen werden mit den anderen Agenten ausgetauscht?
- Wird diese Information zum individuellen Planen genutzt?

- Auf welche Weise gewinnt ein PSI-Agent seine Erfahrungen durch reine Beobachtung (Modelllernen)?
- Welches Format haben diese Erfahrungen?
- Welches Vorwissen der Agenten ist nötig, damit sie situationsadäquat lernen?

Welchen Einfluss die Fähigkeiten der künstlichen Agenten zur Kommunikation, Kooperation und die erweiterten Fähigkeiten zum Lernen und Planen für die Bildung sozialer Strukturen haben, bleibt noch eine offene Forschungsfrage.

7. Literatur

Adami, C. (1998). *Introduction to Artificial Live*. New York u.a.: Springer.

Anderson, J. R. (1983). *The Architecture of Cognition*. Cambridge, MA: Harvard University Press.

Anderson, J. R. (1993). *Rules of the Mind*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Anderson, J. R. & Lebière, C. (1998). *The Atomic Components of Thought*. Mahwah, NJ: Lawrence Erlbaum Associates.

Badke-Schaub, P. (1993). *Gruppen und komplexe Probleme: Strategien von Kleingruppen bei der Bearbeitung einer simulierten AIDS-Ausbreitung*. Frankfurt a. M.: Peter Lang.

Bartl, C. & Dörner, D. (1998). *Comparing the Behaviour of PSI with Human Behaviour in the BioLab Game*. Bamberg: Lehrstuhl Psychologie II, Memorandum Nr. 32.

Bates, J., Loyall, A. B. & Reilly, W. S. (1998). An Architecture for Action, Emotion, and Social Behavior. In: *Readings in Agents* (eds.: M. Huhns & P. Singh), San Francisco: Morgan Kaufmann, pp. 225-231.

Becht, M., Muscholl, M. & Levi, P. (1997). Ein Framework für Kooperationsverfahren zwischen Roboteragenten. In: *Autonome Mobile Systeme* (eds: P. Levi, Th. Bräunl & N. Oswald), Berlin: Springer, pp. 166-177.

Becht, M., Muscholl, M. & Levi, P. (1998). Transformable Multi-Agent Systems: A specification language for cooperation processes. In: *Proceedings of the World Automation Congress* (ed.: M. Jamshidi*)*. Albuquerque, TSI Press.

Bell, C. G. & Newell, A. (1971). *Computer Structures: Readings and Examples*. New York: McGraw-Hill.

Boulding, K.E. (1978). *Ecodynamics*. Beverly Hills, CA: Sage.

Castelfranchi, C. & Conte, R. (1995). *Cognitive and Social Action*. London: UCL Press.

Castelfranchi, C. & Conte, R. (1998). Limits of economic and strategic rationality for agents and MA systems. In: *Robotics and Autonomous Systems*, Elsevier 1998.

Dertouzos, M. L. (1965). *Treshold Logic. A Synthesis Approach*. Research Monographs No. 32. Cambridge, Mass.: MIT Press.

Detje, F. (1999). *Handeln erklären.* Wiesbaden: DUV.

Detje, F. (2000): *"Insel". Dokumentation – Versuche – Ergebnisse*. Bamberg: Institut für Theoretische Psychologie, Memoramdum Nr. 39.

Detje, F. (2001): *PSI - Erste Schritte in Richtung sozialen Verhaltens*. Bamberg: Institut für Theoretische Psychologie, Memoramdum Nr. 41.

Doran, J. (1995). Modeling early Human Society. *ASIB Quarterly*, pp. 31-33.

Doran, J. (1996). Simulating Societies using Distributed Artificial Intelligence. In: *Social Science Microsimulation* (eds.: K. G. Troitzsch, U. Mueller, G. N. Gilbert & J. E. Doran), Berlin: Springer, pp. 381-393.

Dörner, D. (1993). Wissen, Emotionen und Handlungsregulation oder die Vernunft der Gefühle. *Zeitschrift für Psychologie, 2,* 167 - 202.

Dörner, D. (1994a). Über die Mechanisierbarkeit der Gefühle. In: Krämer, S. (Hrsg.), *Geist - Gehirn - künstliche Intelligenz: zeitgenössische Modelle des Denkens*. Berlin: de Gruyter, 130 - 161.

Dörner, D. (1994b). Emotionen, kognitive Prozesse und der Gebrauch von Wissen. In: Klix, F. & Spada, H. (Hrsg.): *Kognition und Emotion*. Göttingen: Hogrefe.

Dörner, D. (1996). Eine Systemtheorie der Motivation. In: Kuhl, J. & Heckhausen, H. (Hrsg.), *Enzyklopädie der Psychologie: Motivationspsychologie*. Göttingen: Hogrefe, C/IV/4, 329-357.

Dörner, D. (1999). *Bauplan für eine Seele*. Reinbek: Rowohlt.

Ekman, P. (1988). *Gesichtsausdruck und Gefühl*. Paderborn: Junfermann.

Ferber, J. (1994). Simulating with reactive agents. In: *Many-Agent Simulation and Artificial Life* (eds.: Hillebrand, E. & Stender, J.), IOS Press, pp. 8-30.

Fischer, K., Müller, J. P. & Pischel, M. (1995). A pragmatic BDI Architecture. In: *Intelligent Agents II. Agent Theories, Architectures, and Languages* (eds.: Wooldridge, Müller & Tambe). ATAL 95, pp. 203-218.

Gasser, L. (1998). Social Conceptions of Knwledge and Action: DAI Foundations and Open Systems Semantics. In: *Readings in Agents* (eds.: M. Huhns & M. Singh), San Francisco, CA: Morgan Kaufmann, pp. 389-404.

Grand, S. & Cliff, D. (1998). Creatures: Entertainment Software Agents with Artificial Life. In: *Autonomous Agents and Multi-Agent Systems* (eds.: N. Jennings, K. Sycara, Georgeff & M. Kluwer), pp. 39-57.

Haddadi, A. (1996). *Communication and Cooperation in Agent Systems: A Pragmatic Theory*. Number 1056 in Lecture Notes in Artificial Intelligence. Berlin: Springer.

Hanneman, R. A. (1988). *Computer-assisted Theory Building*. *Modeling Dynamic Social Systems.* Newbury Park: Sage.

Hara, F. (1997). Expression for Human-Artifacts Communication. *Proceedings of the International Symposium on System Life*. Tokyo, 21-22. Juli 1997.

Harkins, S. G., Latané, B. & Williams, K. (1980). Social loafing: Allocating effort or taking it easy? *Journal of Experimental Social Psychology, 16,* 457-465.

Hille, K. (1997). *Die „künstliche Seele". Analyse einer Theorie*. Wiesbaden: Deutscher Universitäts-Verlag.

Hinton, G. E. & Anderson, J. A.(Hrsg.) (1981). *Parallel Models of Associative Memory*. Hilldale, NJ: Lawrence Erlbaum.

Janis, I. J. (1972). *Victoms of Groupthink*. Boston: Houghton Mifflin.

Laird, J., Newell, A. & Rosenbloom, P. (1987). SOAR: An Architecture for General Intelligence. *Artificial Intelligence, 33,* 1-64.

McClelland, J. L. & Rumelhart, D. E. (1989[9]). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 2 Psychological and Biological Models*. Cambridge, MA: MIT Press.

Mellars, P. A. (1985). The Ecological Basis of Social Complexity in the Upper Palaeolithic of Southwestern France. In: *Prehistoric Hunter-Gatherers: the emergence of cultural complexity* (eds.: Douglas Price & James A Brown), New York: Academic, pp. 271-297.

Müller, J. (1993). *Verteilte Künstliche Intelligenz: Methoden und Anwendung*. Mannheim: BI Wissenschaftsverlag.

Muscholl, M. & Levi, P. (1996). Entscheidungsnetzwerke für selbstorganisierende Roboterarchitekturen. In: *Autonome Mobile Systeme* (eds.: Schmidt, G. & Freyberger, F), Berlin: Springer, pp. 236-245.

Newell, A. & Simon, H. A. (1972[.3]). *Human Problem Solving*. Englewood Cliffs, NJ: Prentice-Hall.

Newell, A. (1990). *Unified Theories of Cognition. (The William James Lectures, 1987)*. Cambridge, MA: Havard University Press.

Newell, A. (1992). Unified Theories of Cognition and the Role of Soar. In: Michon, J. A. & Akyürek, A. (Hrsg.), *Soar: A Cognitive Architecture in Perspective. A Tribute to Allen Newell*. Dordrecht: Kluwer Academic Publishers, 25-79.

Rao, A. S., Georgeff, M. P. & Sonenberg, E. A. (1992). Social Plans: A Preliminary Report. In: E. Werner and Y. Demazeau (Hrsg.): *Third European Workshop on Modelling Autonomous Agents in a Multi-Agent World (DAI-3)*. North-Holland, Amsterdam: Elsevier, pp. 57-76

Rosenbloom, P. S., Laird, J. E. & Newell, A. (Hrsg.) (1993): *The Soar papers: Research on integrated intelligence*. 2 Bände. Cambridge, MA: MIT Press.

Rumelhart, D. E. & McClelland, J. L. (1992[.10]). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Vol. 1 Foundations*. Cambridge, MA: MIT Press.

Schaub, H. (1993). *Modellierung der Handlungsorganisation*. Bern: Huber.

Schaub, H. (1996). Künstliche Seelen - Die Modellierung psychischer Prozesse. *Widerspruch, 29,* 56-82.

Schaub, H. (1997). Modelling Action Regulation. In: Brezinski, J., Krause, B. & Maruszewski, T. (Hrsg.), *Idealization VIII: Modelling in Psychology*. Amsterdam: Rodopi, 97-136.

Schiepek, G. & Strunk, G. (1994). *Dynamische Systeme. Grundlagen und Analysemethoden für Psychologen und Psychiater*. Heidelberg: Asanger.

Schönherr, M. (1998). Irren ist Okay - Drei Provokationen zur Agententheorie. *c't, 23,* S.110.

Selz, O. (1913). *Über die Gesetze des geordneten Denkverlaufs*. Stuttgart: Spemann.

Singh, M. P. (1994). *Multiagent Systems: A theoretical Framework for Intentions, Know-How, and Communication*. Number 799 in Lecture Notes in Artificial Intelligence. Berlin: Springer.

Smolensky, P. (1987). Connectionist AI, Symbolic AI, and the Brain. *Artificial Intelligence Review, 1,* 95-109.

Stachowiak, H. (1973). *Allgemeine Modelltheorie*. Wien: Springer.

Steels, L. (1996). *The Spontaneous Self-organization of an Adaptive Language.* Oxford: Oxford University Press.

Westermann, R. & Heise, E. (1996). Motivations- und Kognitionspsychologie: Einige intertheoretische Verbindungen. In: Kuhl, J. & Heckhausen, H. (Hrsg.): *Enzyklopädie der Psychologie, Band C/IV/4 (Motivation, Volition und Handlung).* Göttingen: Hogrefe, 275-327.

Wissel, C. (1989). *Theoretische Ökologie*. Berlin: Springer.