

## Zur Entscheidungskonvergenz in kognitiven Systemen

Johann Kelsch\*

\*Deutsches Zentrum für Luft- und Raumfahrt (DLR), Institut für Verkehrssystemtechnik (TS), Lilienthalplatz 7, 38108 Braunschweig, Germany (Tel: +495312953433; e-mail: johann.kelsch@dlr.de)

---

**Abstract:** Das Ziel dieses Beitrags ist, eine Methode für die Analyse der Entscheidungsvorgänge innerhalb kognitiver Systeme vorzustellen. Diese Methode ist ein Teil eines Frameworks für die ganzheitliche Gestaltung dynamischer kognitiver Systeme aus dem Bereich hochautomatisierter Fahrzeugführung. Sie basiert auf der Hypothese, dass Entscheidungen in einem kognitiven System kein Ergebnis einer expliziten Berechnung in einem speziell dafür ausgewiesenen Modul sind, sondern in einem Prozess der Selbstorganisation der Systemelemente *generiert* werden können. Zur Unterstützung dieser Hypothese ist in diesem Beitrag ein skalierbares Modell der kognitiven Systeme präsentiert, das sich mithilfe der Multiagenten- und der Graphentheorie mathematisieren und analysieren lässt. Am generischen Beispiel eines Entscheidungsvorgangs beim Führen eines hochautomatisierten Fahrzeuges ist dieses Modell erläutert. Das Modell und die o.g. Methode sind auf ihre grundsätzliche Anwendbarkeit für die Gestaltung und Analyse kognitiver Systeme mittels einiger bekannter Ansätze aus der Gestaltung der geteilten Fahrzeugführung evaluiert.

---

### 1. EINLEITUNG

Wir erleben heute eine zunehmende Technisierung der Gesellschaft. Die Technik wird zu unserem ständigen Begleiter und sogar zu einem Teamplayer (Christoffersen *et al.* 2002). Computer helfen uns, komplexe Berechnungen durchzuführen sowie enorme Informationsmengen zu ordnen und zu transportieren. Unsere Maschinen, wie Flugzeuge, Schiffe, Automobile sind bereits fähig, komplexe Verkehrssituationen wahrzunehmen, sie durch Sensordatenfusion und Klassifizierung zu „begreifen“, autonome Entscheidungen zu treffen und Handlungen auszuführen. Die Maschinen werden so zunehmend kognitiv (Hollnagel *et al.* 1982). Sie werden in größere kognitive Systeme (CS) integriert, wo wir als Piloten, Steuerleute, Fahrer uns inmitten und, z.B. aufgrund der Gefährlichkeit des Verkehrs, auch in Gefahr befinden können. Dadurch entsteht ein hoher Bedarf an Modellierung und Analyse der CS, die im Mittelpunkt dieses Beitrags stehen.

Unter einem *System* versteht man eine „Menge von geordneten Elementen mit Eigenschaften, die durch Relationen verknüpft sind. Die Menge der Relationen zwischen den Elementen eines Systems ist seine Struktur. Unter Element versteht man einen Bestandteil eines Systems, der innerhalb dieser Gesamtheit nicht weiter zerlegt werden kann.“ (Springer 2013a). Die *Kognition* kann als „Fähigkeit zur Wahrnehmung und zielgerichteten Interpretation der Lebenswelt“ (Springer 2013b) und als „die von einem verhaltenssteuernden System ausgeführte Informationsumgestaltung“ (Wikipedia 2013) definiert werden. *Kognitive Systeme* sind demnach solche Systeme, die kognitive Eigenschaften, Strukturen und Verhalten aufweisen. Im Gegensatz zum kausalen Verhalten technischer dynamischer Systeme, kann das Verhalten der CS als *nichtkausal*, d.h. von außen „nicht nachvollziehbar“ bzw. „nicht vorhersehbar“ sein. Solche Systeme können ein willkürliches Verhalten zeigen. Insgesamt wird einem CS ein komplexes Verhalten unterstellt, das u.a. die

Selbstorganisation, d.h. wenn Systemelemente sich ohne eine zentralere Koordination sinnvoll ordnen, einschließt.

Diese Definitionen lassen zwei Annahmen zu:

1. CS müssen aus „kleineren“ Elementen bestehen, was sich bei einer Systemgestaltung mittels der Methode der Systemdekomposition zwecks Modellbildung vorteilhaft ausnutzen lässt.
2. Das Verhalten der CS insgesamt muss zielgerichtet und selbstorganisiert sein, was bei mehreren oder konkurrierenden Zielen einen verteilten Entscheidungsvorgang innerhalb eines CS annehmen lässt.

Mittels dieser Annahmen lassen sich zwei für diesen Beitrag relevanten Forschungsfragen nennen:

1. Wie kann ein CS zwecks Modellierung in seine Bestandteile dekomponiert werden?
2. Wie kommen innerhalb des CS die zielgerichteten Systementscheidungen zustande?

Beide Forschungsfragen werden in diesem Beitrag adressiert und diskutiert. Nach der Einleitung werden zunächst das theoretische Konzept der kognitiven Multiagenten-Systeme und das darauf bauende CS-Modell präsentiert. Anschließend wird das aufgestellte Modell anhand bekannter Ansätze aus dem Bereich der Gestaltung der hochautomatisierten Fahrzeugführung (Gasser *et al.* 2012), d.h. wenn ein Fahrer und eine Automation die Kontrolle über das Fahrzeug untereinander teilen, evaluiert.

Zur Veranschaulichung und Erläuterung der theoretischen Überlegungen ist ein praxisnahes aber dennoch generisches Beispiel verwendet. Bei dem Beispiel handelt es sich um einen Entscheidungsvorgang bei einer zwischen einem Fahrer und einer Automation geteilten Fahrzeugkontrolle (shared

control) (Sheridan *et al.* 1978). Das in dem Beispiel verwendete Fahrer-Automation-Fahrzeug-Umwelt-System und der entsprechende Entscheidungsvorgang sollen stellvertretend und allgemein die CS und die darin vorkommenden Entscheidungsvorgänge repräsentieren.

Für das Beispiel stelle man sich folgende Situation vor: Ein Fahrer fährt in einem hochautomatisierten Fahrzeug auf eine Weggabelung zu. Er und seine Automation im Fahrzeug haben unterschiedliche Absichten über die weitere Fahrtrichtung. Der Fahrer möchte nach links abbiegen und die Automation nach rechts. Der Fahrer lenkt nach links, während die Automation nach rechts lenkt (Abbildung 1).

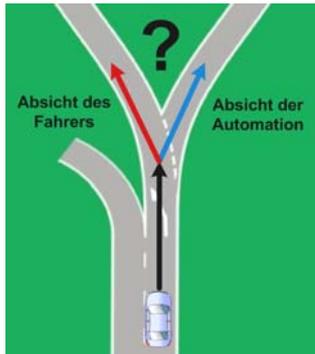


Abbildung 1: Beispiel eines Entscheidungsvorgangs in einem generischen kognitiven System

Was wird in einer solchen Situation geschehen: Fährt das Fahrzeug nach links oder nach rechts? Diese Frage lässt sich nur schwer beantworten, da das eingeführte Beispiel zwar ein relativ einfaches aber dennoch typisches CS beschreibt. Die in einem hochautomatisierten Fahrzeug vorhandenen kognitiven Systemelemente (Mensch, komplexe Automation) können unterschiedliche Motive haben, obwohl sie sich in derselben Verkehrssituation befinden. Sie befinden sich auch in einer komplexen Wechselwirkung untereinander und mit der sie umschließenden nahen Umwelt.

Die Kenntnis der genannten Systemelemente und der noch grob formulierten Beziehung dazwischen lässt sich präzisieren. Dies kann mittels des Konzeptes der kognitiven Multiagenten-Systeme zwecks einer Systemdekomposition und -modellierung gemacht werden.

## 2. KOGNITIVE MULTIAGENTEN-SYSTEME

Auf das in der Einleitung eingeführte Beispiel sowie auf ähnliche CS lässt sich das Konzept der Multiagenten-Systeme anwenden (Kelsch *et al.* 2013). Damit lassen sich die CS strukturell zerlegen. Dabei können die im CS vorhandenen Elemente als sogenannte *Agenten* modelliert werden.

Bei der Definition eines Agenten existiert in der einschlägigen Literatur noch keine allgemeingültige Meinung. Ein Agent kann „von außen betrachtet“ definiert werden. Eine der gängigsten Definitionen dafür lautet: „an agent is any entity able to act“ (Castelfranchi 1998). Man kann einen Agenten auch „von innen betrachtet“ definieren: „An agent is a highly organized complex system which has intrinsic goals“ (Vakarelov 2009). Insgesamt kann ein Agent als eine abge-

grenzte und zielgerichtet agierende Einheit mit verschiedenen möglichen Informationszuständen innerhalb seiner komplexen Strukturen verstanden werden. Agenten haben eine eingeschränkte Sicht auf die globalen Systemzustände und das Systemverhalten sowie einen hohen Grad an Autonomie. Sie besitzen bestimmte Eigenschaften, wie eine Interaktionsbereitschaft innerhalb lokaler geometrischer und sozialer Umgebungen, eine Robustheit gegenüber offenen Umgebungen etc. (Luck *et al.* 2005). Das interaktive Agentenverhalten eines Multiagenten-Systems kann emergente Systemeigenschaften, z.B. mittels einer Selbstorganisation, generieren. Genau diese Eigenschaften, die denen einer Kognition sehr nahe stehen, wie die Zielgerichtetheit, Autonomie, Emergenz, Selbstorganisation und Robustheit gegenüber offenen Umgebungen, macht Multiagenten-Systeme nützlich für die Modellierung von CS.

Unter Verwendung des Multiagenten-Konzeptes kann eine Dekomposition für das in der Abbildung 1 eingeführte CS durchgeführt werden (Abbildung 2). Zwei Typen der Agenten können beim Modellieren unterschieden werden: Kognitive und nichtkognitive Agenten (Kelsch *et al.* 2013). Die *kognitiven Agenten* (CA, CAs für Mehrzahl) sind solche, die komplexe Modelle über sich selbst und über ihre Umwelt aufbauen können. Dadurch kann ihr Verhalten z.B. adaptiv und adaptierbar sein. Menschen (bzw. Fahrer) und komplexe Automation gehören zu diesem Agententyp. Die *nichtkognitiven Agenten* (NA, NAs für Mehrzahl) sind reaktive Systemelemente, wie natürliche und künstliche Umweltobjekte. Die Automation eines hochautomatisierten Fahrzeuges wäre ein CA. Das Fahrzeug als ein reaktives Artefakt wäre ein NA.

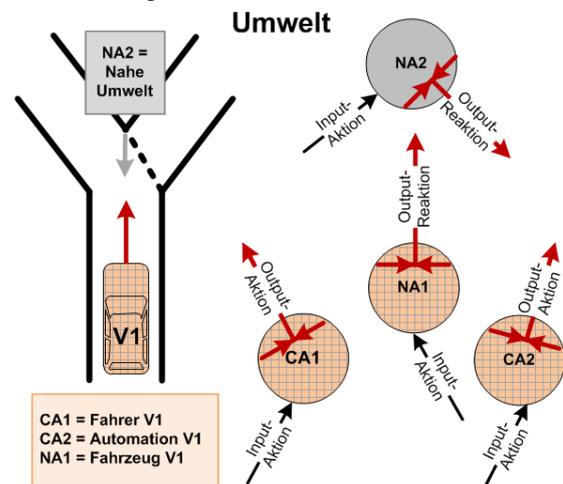


Abbildung 2: Agentenbasierte Dekomposition eines hochautomatisierten Fahrzeuges und seiner nahen Umwelt

Die eingeführte Beispielsituation lässt sich folgendermaßen zerlegen: V1 ist ein hochautomatisiertes Fahrzeug, das sich auf einer Straße bewegt. Es enthält drei Agenten: CA1, CA2 und NA1. CA1 ist der Fahrer, CA2 ist die Fahrzeugautomation und NA1 ist das Fahrzeug. Der Fahrer und die Automation sind CAs. Das Fahrzeug, auf dieser Ebene betrachtet, ist ein NA, weil es nur auf die Aktionen des CA1, CA2 und NA2 reagieren kann. Würde man eine solche Dekomposition aus einer abstrakteren Systemsicht anstreben, um z.B. Verkehrsflüsse zu zerlegen, so könnte das Fahrzeug V1 als ein einzelner CA erachtet werden. Somit wäre V1 eine *kognitive*

*Population* im Sinne des Multiagenten-Konzeptes. Die relevante nahe Umwelt, wie die sich verzweigende Straße, eventuelle weitere Fahrzeuge, Hindernisse etc., kann im NA2 aggregiert werden. Dies zeigt u.a. die Skalierbarkeit der Dekompositionsmethode. Alle CAs und NAs bilden ein CS, das sich innerhalb der offenen Umwelt befindet.

In Bezug auf ihre Grenzen besitzen Agenten ein *äußeres Verhalten*, das auf der Systemebene *wirkt*. Sie können durch Ereignisse angestoßene Aktionen und Reaktionen senden und Aktionen anderer Agenten empfangen. Dies ist in der Abbildung 2 durch Pfeile und Pfeilspitzen dargestellt. Die Pfeile sind die Aktionen und die Pfeilspitzen die Ereignisse, die diese Aktionen verursachen oder hervorrufen. Durch die Gesamtwirkung aller Aktionen entsteht auf der Systemebene eine komplexe *Wechselwirkung*, die emergente Systemeigenschaften hervorbringen kann. In der einschlägigen Literatur spricht man in solchen Fällen von der möglichen *negativen* oder *positiven* Wechselwirkung, die jeweils als *Konkurrenz* oder *Kooperation* bezeichnet werden können (Castelfranchi 1998, Kelsch *et al.* 2013). Die CA-Wechselwirkung bestimmt auf der Systemebene das *Verhalten des CS*.

**Beispiel:** Der Fahrer in einem hochautomatisierten Fahrzeug aus der Abbildung 1 und die Fahrzeugautomation sind CAs. Sie können sich jederzeit in unterschiedlichen kognitiven Zuständen befinden und willkürliche Handlungen ausführen, obwohl sie objektiv in derselben Verkehrssituation sind. Dies kann zu einer negativen CA-Wechselwirkung auf der Ebene des CS, also einer Konkurrenz führen. Genauso kann es in dem Beispiel zu abgestimmten Handlungen und einer intakten Kooperation und damit insgesamt zu einer positiven CA-Wechselwirkung zwischen dem Fahrer und der Automation kommen. Dafür müssen die kognitiven Zustände und Zustandsübergänge der beiden Agenten, z.B. mittels eines zusätzlichen HMI-Agenten bzw. Arbiters, durch eine geeignete Mensch-Maschine Interaktion bzw. Arbitrierung (Kelsch 2012a) koordiniert und untereinander abgestimmt sein.

Wichtig zu betonen ist, dass hier die Interaktion ungleich der Wechselwirkung ist. Eine Interaktion sind zwei gegenseitige Aktionen der Agenten, ein Interaktionsprozess ist ein über die Zeit verteilter Aktionsaustausch zwischen den CAs. Eine Wechselwirkung ist dagegen ein emergenter Systemprozess, der auf der übergeordneten CS-Ebene zusätzliche Systemeigenschaften generieren kann, die keinem CA eigen sind.

Agenten besitzen eine *innere Struktur*, die durch Ereignisse und Aktionen in bestimmte Zustände versetzt werden kann. Dies ist durch die zusammenlaufenden Pfeile im Inneren der die Agenten kennzeichnenden Kreise der Abbildung 2 angedeutet. Bei den einfacheren NAs sind diese Aktionen und Zustandsänderungen einer eher kausalen, reaktiven Natur, wohingegen bei den komplexen CAs die Aktionen und Zustandsänderungen nichtkausal, intentional sein können. Die innere Struktur der Agenten kann nach dem Konzept der *Aktionstension* modelliert werden (Kelsch *et al.* 2012b).

Aktionstension selbst ist ein operationalisiertes Modell der in jede Kognition inkludierten *Motivation*. Sie beschreibt die Art und die Reihenfolge verschiedener sinnvoller und möglicher kognitiver Zustände, Zustandsübergänge und dazugehö-

riger Aktionen die denselben *optimalen Systemzustand* zum Ziel haben. Eine Aktionstension beschreibt in einer integrierten Weise, wie dringend eine bestimmte Aktion ist. Sie fasst zwei sich ausschließende (bzw. dichotome) Aktionen oder Aktionsabfolgen zusammen, die immer in Richtung des optimalen Systemzustandes zeigen. Je weiter der aktuelle kognitive Zustand vom optimalen entfernt ist, desto dringender wird die jeweilige Aktion.

**Beispiel:** Für die Situation aus der Abbildung 1 lässt sich eine Aktionstension modellieren, die die dichotomen Aktionen „Lenken-nach-links“ und „Lenken-nach-rechts“ zusammenfasst. Sowohl der Fahrer als auch die Automation werden durch diese Tension „motiviert“ und lenken entsprechend nach links oder nach rechts in Abhängigkeit von ihrem intern als optimal definierten Systemzustand. Wenn der Fahrer den Zustand „Nach-links-Fahren“ als optimal für sich definiert, dann wird er nach links lenken. Je weiter das Fahrzeug vom Zustand „Nach-links-Fahren“ entfernt ist, desto stärker wird der Fahrer nach links lenken oder auch andere Strategien wählen, um das Fahrzeug in seiner Bewegung nach links zu manipulieren. Ein Beispiel einer solchen Strategie könnte die vollständige Kontrollübernahme durch das Abschalten der Automation sein. Die Automation wird entsprechend nach rechts lenken, wenn sie den optimalen Systemzustand als „Nach-rechts-Fahren“ definiert. Ungeachtet der Einfachheit dieses Beispiels sei angemerkt, dass das Konzept der Aktionstension eine Schlüsselrolle für die dargestellte Art der Modellierung der kognitiven Multiagenten-Systeme hat.

Wenn man Systeme im Kontext einer Kognition analysiert, müssen die Begriffe „Aktion“ und „Wirkung“ etwas genauer angeschaut werden. Laut der verwendeten Definition für Agenten führen sie ihre Aktionen aufgrund zielgerichteter bzw. intentionaler Zustände und Zustandsübergänge innerhalb ihrer Strukturen aus. Für CAs, aufgrund der zielgerichteten Natur der Kognition selbst, lässt sich ein intentionales Handeln, also eine beabsichtigte *Manipulation* der Umwelt statt nur einer gerichteten Aktion in die Umwelt hinein, umso mehr annehmen. Auch eine unbeabsichtigte Aktion, wie eine physikalische oder physiologische Reaktion eines Agenten, kann auf der Ebene des CS eine komplexe Wirkung entfalten (z.B. „man kann nicht nicht kommunizieren“, Watzlawick *et al.* 1967). Damit lässt sich annehmen, dass CAs grundsätzlich im neutralen Sinn manipulativ und manipulierbar sind.

Die Wechselwirkung der benachbarten CAs und damit das CS-Verhalten entstehen eben durch diese Manipulationen und Manipulierbarkeit. Dabei muss berücksichtigt werden, dass im Kontext einer Kognition eine Manipulation nur dann wirkt, wenn sie mit genügend Nachdruck, Aufforderung bzw. *Salienz* auf der Senderseite präsentiert wird. Ebenso wichtig ist die Empfängerseite, die eine hereinkommende Manipulation mit einer Hinwendung, Abwendung bzw. generell einer *Aufmerksamkeit* belegt. Dieser Zusammenhang ist wichtig zum Verständnis und zum Modellieren der *Struktur der CS*. Er äußert sich in der *vollständigen Verbundenheit* aller benachbarten Agenten unter der Berücksichtigung der Salienz- und der Aufmerksamkeitskomponente einer Manipulation.

Über *multimodale Kommunikationskanäle* können die Agenten Output-Manipulationen (OM, OMs für Mehrzahl) senden

und Input-Manipulationen (IM, IMs für Mehrzahl) empfangen. Damit können Agenten einander manipulieren und eine emergente negative oder positive Wechselwirkung (Konkurrenz oder Kooperation) auf der CS-Ebene generieren.

Eigenmanipulationen (EM, EMs für Mehrzahl) sind aufgrund interner kognitiver Prozesse bei den CAs grundsätzlich immer vorhanden. Sie können als eine persistente Berücksichtigung der eigenen Meinung, kognitiver Modelle bzw. als die „kognitive Hysterese“ (Norman *et al.* 1986) interpretiert werden. Wenn ein CA frei von Manipulationen der anderen Agenten ist, sollte er nach der EM handeln. NAs besitzen keine oder nur rudimentäre physikalisch bedingte EMs, z.B. durch die Masseträgheit generierte Hysterese oder Zustandspersistenz. Wenn NAs nicht von Außen manipuliert werden, handeln sie grundsätzlich auch nicht.

Über die auf Kommunikationskanälen geführten Manipulationen können Agenten miteinander gekoppelt oder voneinander entkoppelt werden. Dies kann mit einem entsprechenden *Kopplungsgrad* angegeben werden (Flemisch *et al.* 2010). CAs können die Kopplungsgrade ihrer IMs und OMs jederzeit selbst kontrollieren. Je größer die Aufmerksamkeit eines CA gegenüber einer IM oder die Saliens einer OM sind, desto größer ist die Wirkung der Manipulation. Bei fehlender Aufmerksamkeit und/oder Saliens fehlt auch die Kopplung. Bei gezielter Abwendung (negative Aufmerksamkeit, wie Trotz, Abneigung) oder unangenehm wirkenden Manipulationen (negative Saliens, wie Ekelwirkung) kann es zu einer Gegenkopplung kommen. Durch die beschriebenen Kopplungsvorgänge können CS spontan entstehen, zerfallen oder modifiziert werden.

**Beispiel:** In einem hochautomatisierten Fahrzeug, ist es möglich, über eine Steer-by-Wire-Lenkung den Fahrer von der Fahrzeug-Kontrolle (bzw. -Manipulation) abzukoppeln. Der Fahrer kann sich auch selbst ab- und einkoppeln, indem er die Stellteile für die Fahrzeugkontrolle (Lenkrad, Pedale etc.) loslässt oder stärker anfasst. Damit kann er jederzeit jegliche IMs ignorieren und seinen Willen in Bezug auf die Fahrzeugkontrolle durchsetzen. Dies gilt für jeden CA. Dadurch kann das Fahrer-Fahrzeug System modifiziert werden.

Das beschriebene Konzept der kognitiven Multiagenten-Systeme lässt sich folgendermaßen kurz zusammenfassen:

*Struktur der Agenten:* Agenten besitzen komplexe Strukturen, die sich in unterschiedlichen (kognitiven) Zuständen befinden können. Diese Strukturen lassen sich mittels der Aktionstensionen operationalisieren und modellieren.

*Verhalten der Agenten:* Agenten streben intern gegen die von ihnen als optimal definierten Zustände der Aktionstensionen mittels (kognitiver) Zustandsübergänge. Die Zustandsübergänge innerhalb der Agentenstrukturen generieren die Aktionen (bzw. Manipulationen) der Agenten, die über Kommunikationskanäle im CS verteilt werden und auf der CS-Ebene zur CA-Wechselwirkung beitragen.

*Struktur der CS:* Alle strukturell abgrenzbaren Elemente eines CS, die selbständige Aktionen durchführen können, sind Agenten, wobei man CAs und NAs unterscheiden kann. Alle

benachbarten Agenten sind durch gegenseitige Manipulationen vollständig miteinander verbunden.

*Verhalten der CS:* Alle Manipulationen der Agenten wirken auf der CS-Ebene und können dort eine positive oder eine negative Wechselwirkung (Konkurrenz oder Kooperation) generieren, die das Verhalten des CS bestimmen.

### 3. MATHEMATISIERUNG DES MODELS

Kognitive Multiagenten-Systeme lassen sich mathematisch beschreiben, was zum formalen Aufstellen der CS-Modelle verwendet werden kann. Mittels einer Zeitreihenentwicklung des jeweiligen Modells können Entscheidungskonvergenzen innerhalb des modellierten CS untersucht werden.

Eine Manipulation  $m$  kann durch die jeweilige gerichtete Aktion bzw. Handlung  $h$  und sie umschließende Ereignisse  $e$  definiert werden. Eine Tension  $T$  kann durch die jeweiligen Tensionszustände  $z_m$  mit dem optimalen Zustand  $z_0$ , die dazugehörigen Zustandsübergänge  $u$  und dadurch generierte Manipulationen  $m$  definiert werden (1).

$$\bar{m} := \{e_{vor}, \bar{h}, e_{nach}\} \quad (1)$$

$$T := \{z_{-m} \dots z_0 \dots z_m, u_1, u_2, \dots, -\bar{m} \dots 0 \dots \bar{m}\}$$

Ein Agent  $A$  kann als

$$A := \{S_A, V_A\}, \quad \text{mit} \quad (2)$$

$$S_A := \{T_1, T_2, \dots\}, V_A := f(S_A, t)$$

definiert werden, wobei  $S_A$  die Struktur des Agenten und  $V_A$  das Verhalten des Agenten als Funktion der zeitlichen Entwicklung seiner Struktur ist (2). Ein CS kann dann als

$$CS := \{S_{CS}, V_{CS}\}, \quad \text{mit} \quad (3)$$

$$S_{CS} := \{A_1 \dots A_n, m_{A_1 A_1} \dots m_{A_1 A_n}, m_{A_2 A_1} \dots m_{A_2 A_n}, \dots\},$$

$$V_{CS} := f(S_{CS}, t)$$

definiert werden. Dabei ist  $S_{CS}$  die Struktur des CS,  $m$  sind die Agenten verbindenden Manipulationen und  $V_{CS}$  ist das Verhalten des CS als Funktion seiner Struktur und ihrer zeitlichen Entwicklung (3).

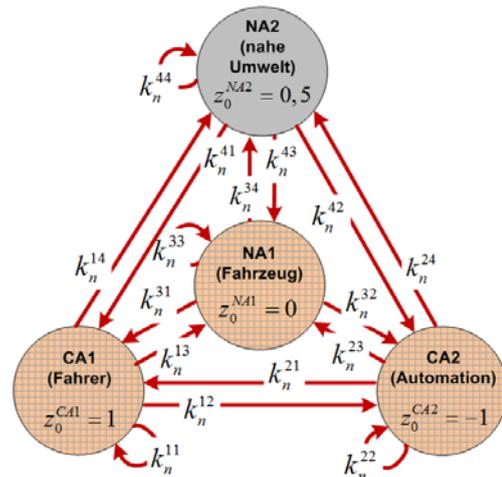


Abbildung 3: Dekomposition eines hochautomatisierten Fahrzeuges und seiner nahen Umwelt als vollständiger Graph

Unter Berücksichtigung von (1), (2) und (3) lässt sich das in den Abbildungen 1 und 2 eingeführte CS als ein vollständiger Graph darstellen. Die Abbildung 3 zeigt diesbezüglich ein agentenorientiertes Modell eines hochautomatisierten Fahrzeuges und seiner nahen Umwelt. Agenten sind als Knoten und Kommunikationskanäle dazwischen sind als Kanten modelliert. CA1 ist der Fahrer, CA2 ist die Fahrzeugautomation, NA1 ist das Fahrzeug und NA2 ist die nahe Umwelt. Die  $z_0$  der Agenten sind die jeweiligen optimalen Tensionszustände auf derselben Aktionstension (lenken nach links...0...lenken nach rechts) zum Zeitpunkt  $t_0$  (Ausgangslage). Der Fahrer möchte nach links (1), die Automation möchte nach rechts (-1) fahren, das Fahrzeug fährt noch gerade aus (0) und die Straße hat wegen der entsprechenden Spurmarkierung eine etwas stärkere Tendenz nach links (0,5).

Jeder Agent hat einen Kommunikationskanal zu jedem anderen Agenten inklusive sich selbst. Bei mehreren, z.B. multimodalen, Kanälen können diese als zusätzliche Kanten oder als weitere Agenten (Knoten) modelliert werden. Kopplungsgrade auf den Kommunikationskanälen zwischen den Agenten sind durch entsprechende Gewichte  $k_n$  an den Kanten des Graphs angegeben, wodurch der Kopplungsgrad zum Zeitpunkt  $t_n$  angegeben werden kann. Kopplungsgrade können einen beliebigen Wert zwischen  $-1$  und  $1$  annehmen. Ein Kopplungsgrad kann als Produkt des Hinwendungs- bzw. Aufmerksamkeitskoeffizienten  $a_n$  und des Nachdruck- bzw. Salienz-Koeffizienten  $s_n$  weiter zerlegt werden (4).

$$k_n = a_n \cdot s_n \quad (4)$$

Das Produkt aus einem Kopplungsgrad und einer auf dem entsprechenden Kanal geführten Manipulation kann als eine *Ignoranz*, *Anregung* oder *Hemmung* interpretiert werden.

**Beispiel:** Bei  $k_n = 0$  ist der entsprechende Kanal vollständig geschlossen bzw. die Agenten sind entkoppelt (Ignoranz). Der Fahrer lässt z.B. die Stellteile los und steuert das Fahrzeug nicht mehr. Bei  $k_n = 1$  ist der Kommunikationskanal vollständig offen bzw. die Agenten sind gekoppelt (Anregung). Der Fahrer lenkt z.B. aktiv und regt sein Fahrzeug damit ebenfalls zum Lenken an. Bei  $k_n = -1$  existiert eine negative Rückkopplung bzw. die Agenten sind gegengekoppelt (Hemmung). Der Fahrer lenkt z.B. nach links, die Automation auch. Der Fahrer empfindet das zusätzliche Lenken der Automation aber als unangenehm und steuert dagegen, womit das nach links Lenken des Fahrers gehemmt wird.

Die gesamten Kopplungsgrade eines CS zum Zeitpunkt  $t_n$  können in einer Adjazenz- bzw. Kopplungsmatrix  $K_n$  zusammengefasst und analysiert werden (5).

$$K_n = \begin{pmatrix} k_n^{11} & k_n^{21} & k_n^{31} & k_n^{41} \\ k_n^{12} & k_n^{22} & k_n^{32} & k_n^{42} \\ k_n^{13} & k_n^{23} & k_n^{33} & k_n^{43} \\ k_n^{14} & k_n^{24} & k_n^{34} & k_n^{44} \end{pmatrix} \quad (5)$$

Die Multiplikation der Kopplungsmatrix mit dem Vektor der Output-Manipulationen der Agenten beschreibt in allgemeiner und vereinfachten Form den eingangsseitigen Wahrnehmungsschritt, den kognitionsinternen Bewertungsschritt und den ausgangsseitigen Manipulationsschritt der Agenten.

**Beispiel:** Der Fahrer und die Automation produzieren zwecks geteilter Fahrzeugkontrolle zum Zeitpunkt  $t_{n-1}$  normierte OMs  $o^{CA1}$  und  $o^{CA2}$  (nach links lenken und nach rechts lenken). Am Fahrzeug kommt die IM  $i^{NA1}$ , am Fahrer und an der Automation kommen jeweils die multimodalen Feedbacks (IMs)  $i^{CA1}$  und  $i^{CA2}$  zum Zeitpunkt  $t_n$  an. Der resultierende Vektor  $I_n$  der IMs am Fahrer, an der Automation, am Fahrzeug und an der Umwelt zum Zeitpunkt  $t_n$  kann mit (6) angegeben werden.

$$I_n = \begin{pmatrix} i_n^{CA1} \\ i_n^{CA2} \\ i_n^{NA1} \\ i_n^{NA2} \end{pmatrix} = K_n \cdot \begin{pmatrix} o_{n-1}^{CA1} \\ o_{n-1}^{CA2} \\ o_{n-1}^{NA1} \\ o_{n-1}^{NA2} \end{pmatrix} = \begin{pmatrix} k_n^{11} o_{n-1}^{CA1} + k_n^{21} o_{n-1}^{CA2} + k_n^{31} o_{n-1}^{NA1} + k_n^{41} o_{n-1}^{NA2} \\ k_n^{12} o_{n-1}^{CA1} + k_n^{22} o_{n-1}^{CA2} + k_n^{32} o_{n-1}^{NA1} + k_n^{42} o_{n-1}^{NA2} \\ k_n^{13} o_{n-1}^{CA1} + k_n^{23} o_{n-1}^{CA2} + k_n^{33} o_{n-1}^{NA1} + k_n^{43} o_{n-1}^{NA2} \\ k_n^{14} o_{n-1}^{CA1} + k_n^{24} o_{n-1}^{CA2} + k_n^{34} o_{n-1}^{NA1} + k_n^{44} o_{n-1}^{NA2} \end{pmatrix} \quad (6)$$

Solche Notation gestattet folgende drei Anmerkungen:

1.) Bereits bei einem sehr einfachen CS, wie es in der Abbildung 1 gezeigt ist, liegt eine relativ hohe Komplexität einer möglichen Ausprägung der CA-Wechselwirkung vor. Es wird außerdem deutlich, dass die IMs sich als Zeitreihen mit multiplen Abhängigkeiten von den EMs und OMs jeweils anderer Agenten entwickeln. Die Auswirkungen der Interaktionen zwischen den CAs können durch komplexe Rückkopplungen zur CA-Wechselwirkung beitragen.

2.) Die Operatoren (\*) und (+) können folgendermaßen interpretiert werden:

(\*) - **Operator** stellt durch die Multiplikation der Manipulation mit ihrem Kopplungsgrad eine anregende, hemmende oder ignorierende Beziehung zwischen dem Sender und dem Empfänger der Manipulation her. Je größer der Betrag dieses Produktes ist, desto „wirkungsvoller“ ist die jeweilige Manipulation in ihrem anregenden oder hemmenden Charakter.

(+) - **Operator** bildet den Prozess des „gewichteten Bewertens“ ab. Das Ergebnis der Summe kann als eine insgesamt wahrgenommene IM des jeweiligen CA interpretiert werden.

3.) Die jeweiligen Produkte der normierten OMs mit ihren Kopplungsgraden wachsen im Betrag nie über  $1$  an. Das macht insofern Sinn, dass der jeweilige Agent einen anderen nur zu einer maximal möglichen Handlung bzw. Gegenhandlung anregen bzw. hemmen kann. Der Betrag einer solchen maximal möglichen Handlung kann höchstens  $1$  werden.

**Beispiel:** Wenn der Fahrer nach links lenkt und von der Automation zum Rechtslenken angeregt wird, dann könnte er maximal nach rechts lenken. Danach wird seine Handlung eine Sättigung erreichen (mehr nach rechts lenken als maximal kann er nicht). Damit wäre der Agent durch den Bewertungsschritt vollständig „überzeugt“ bzw. „umgestimmt“.

Weil nach (6) die IMs durch die Summenbildung in ihrem Betrag größer  $1$  sein können, müssen sie anschließend mit der Funktion (7) auf den Wertebereich  $[-1; 1]$  begrenzt werden.

$$b: R^n \mapsto R^n, \quad (7)$$

$$x_j \in R \mapsto \begin{cases} -1; & \text{für } x_j \leq -1 \\ 1; & \text{für } x_j \geq 1 \\ x_j; & \text{sonst} \end{cases}, \quad \text{für } j \in N_0$$

**Anmerkung:** Die vorgeschlagene Begrenzung stellt eine deutliche Vereinfachung des Modells dar. Dies kann unkritisch sein, wenn, wie hier, nur mittels einer Tension modelliert wird. Näher an der Wirklichkeit ist aber eine agenteninterne Wechselwirkung in ihren Strukturen (Tensionen). Die möglichen „Manipulationsüberläufe“, die hier begrenzt werden, können z.B. zum Modellieren der Lernprozesse oder der Aufmerksamkeitssteuerung beitragen.

Die CA-Wechselwirkung und somit das Verhalten eines CS lassen sich mittels einer Zeitreihenanalyse untersuchen. Um Zeitreihen entwickeln zu können, muss berücksichtigt werden, dass die IMs der einen Agenten gleichzeitig OMs der anderen sind. Des Weiteren werden die IMs durch die Multiplikation mit der Kopplungsmatrix zu OMs des nächsten Rechenschrittes transformiert. Dies lässt sich mithilfe von (5), (6) und (7) als (8) angeben.

$$\begin{aligned} I_n &= O_n \\ O_{n+1} &= b(K_n \cdot I_n) \end{aligned} \quad (8)$$

**Anmerkung:** Die (8) zeigt insgesamt eine Rekursion, die entweder als ein digitales Schaltwerk oder als ein zeitdiskreter Regelkreis interpretiert werden kann. Dies kann ggf. Ansätze der Stabilitäts- und Komplexitätsanalyse, z.B. mittels der Eigenwertanalyse der Kopplungsmatrix, Lyapunov-Exponenten (Schumacher *et al.* 2001) etc., erlauben.

In diesem Abschnitt wurde die Methode der Mathematisierung der Modelle der kognitiven Multiagenten-Systeme dargestellt. Mittels einer Zeitreihenentwicklung des jeweiligen Modells kann die Entscheidungskonvergenz innerhalb des modellierten CS untersucht werden. Dies soll im nächsten Abschnitt mittels des in der Abbildung 1 eingeführten Beispiels sowie einiger bekannter Ansätze aus der Gestaltung der geteilten Fahrzeugführung zum Zweck der Evaluation der vorgestellten Methode vertieft werden.

#### 4. EVALUATION

Das erste Ziel der Evaluation war die Überprüfung, ob bekannte Ansätze der Gestaltung hochautomatisierter Fahrzeugführung sich mit der beschriebenen Methode der Systemdekomposition und -modellierung abbilden lassen. Als konkrete Anforderung für das erste Ziel der Evaluation sollten bei bekannten Systemkonfigurationen erwartete Systementscheidungen im Modell zustande kommen. Das positive Ergebnis würde die Dekompositionsmethode und die damit erstellten Modelle zwar nicht vollständig validieren, dennoch erste Hinweise geben, dass sie einen wertvollen Beitrag zur Analyse und zur Gestaltung von CS liefern können.

Das zweite Ziel der Evaluation war zu überprüfen, dass Systementscheidungen in dem auf die vorgeschlagene Weise erstellten Modell, tatsächlich ohne sie explizit zu treffen zu-

stande kommen können. Als konkrete Anforderung dafür sollte ein konvergenter und stabiler Verlauf der entscheidungsrelevanten Output-Variablen herauskommen. Das positive Ergebnis würde zwar nicht nachweisen, dass alle Entscheidungsprozesse in allen CS durch Konvergenz zustande kommen, dennoch wäre es ein Hinweis darauf, dass dies eine weiter zu verfolgende Arbeitshypothese ist.

Für das Adressieren der beiden Ziele der Evaluation wurde das in der Abbildung 1 eingeführte Beispiel-CS herangezogen. Ebenfalls vorgestellt ist die Systemdekomposition dafür nach dem Konzept der kognitiven Multiagenten-Systeme (Abbildung 2). In der Abbildung 3 wurde bereits der Graph der CS-Dekomposition dargestellt.

Eine Grundkonfiguration der Kopplungsmatrix (5) des CS wurde so angenommen, dass  $k^{11}=k^{22}=1$  und  $k^{33}=k^{44}=0$  mit der Begründung, dass die ersten beiden Agenten (Fahrer und Automation) CAs mit EMs und die beiden letzten Agenten (Fahrzeug und nahe Umwelt) NAs ohne EMs sind. Der Fahrer und die Automation können in die Fahrzeugkontrolle eingekoppelt sein ( $k^{13}, k^{23}$ ) und einander manipulieren ( $k^{12}, k^{21}$ ). Beide nehmen das Feedback des Fahrzeuges wahr ( $k^{31}=k^{32}=1$ ). Beide nehmen die nahe Umwelt wahr ( $k^{41}=k^{42}=1$ ), ohne sie direkt zu beeinflussen ( $k^{14}=k^{24}=0$ ). Außerdem wurde angenommen, dass weder das Fahrzeug die nahe Umwelt, noch die Umwelt das Fahrzeug beeinflussen kann ( $k^{34}=k^{43}=0$ ). Dies ist eine Vereinfachung, da bei noch genauer Betrachtung das Fahrzeug die wahrgenommene nahe Umwelt beeinflussen sollte, weil es sich darin bewegt, worauf hin seine Position in der nahen Umwelt und somit seine Beziehung dazu verändert wird. Die Grundkonfiguration der Kopplungsmatrix (5) und des Vektors der Output-Manipulationen zum Zeitpunkt  $t_0$  ergeben sich bei allen Annahmen zu (9):

$$K = \begin{pmatrix} 1 & k_n^{21} & 1 & 1 \\ k_n^{12} & 1 & 1 & 1 \\ k^{13} & k^{23} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad O_0 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0,5 \end{pmatrix} \quad (9)$$

Mit Hilfe von (7), (8) und (9) lassen sich Zeitreihen für unterschiedliche Systemkonfigurationen entwickeln. Es wurden Systemkonfigurationen mit folgenden erwarteten Systementscheidungsergebnissen, Ergebnissen der Zeitreihenentwicklung bis zum 11. Schritt und der Interpretation der Ergebnisse ausgewählt (Tab. 1):

Zeitreihenentwicklung 1	
Grundkonfiguration	Erwartetes Ergebnis
Fahrer lenkt nach links, Assistenz lenkt nach rechts, das Fahrzeug fährt gerade aus, und der Verlauf der Straße hat eine Tendenz nach links. Fahrer und Assistenz sind in die Kontrolle eingekoppelt und beeinflussen sich nicht gegenseitig.	Fahrer und Assistenz sollten sich entscheiden, gemeinsam nach links zu lenken und das Fahrzeug sollte nach links fahren.
$K = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad O_0 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0,5 \end{pmatrix}$	

Ergebnis der Zeitreihenentwicklung 1	
	<p>Fahrer opt. Zust.</p> <p>Automation opt. Zust.</p> <p>Fahrzeug Zust.</p> <p>Umwelt Zust.</p>
Interpretation der Zeitreihenentwicklung 1	
<p>Nach der Wahrnehmung der Tendenz des Straßenverlaufs und einer kurzen Verzögerung entscheidet sich die Assistenz, der Intention des Fahrers zu folgen. Das Fahrzeug fährt nach links, der Endzustand ist stabil. Die Systementscheidung ist erwartungskonform konvergiert.</p>	
Zeitreihenentwicklung 2	
Grunkonfiguration	Erwartetes Ergebnis
<p>Fahrer lenkt nach links, Assistenz lenkt nach rechts, das Fahrzeug fährt gerade aus, und der Verlauf der Straße spielt keine Rolle. Fahrer und Assistenz sind in die Kontrolle eingekoppelt und beeinflussen sich nicht gegenseitig.</p>	<p>Keine Systementscheidung sollte zustande kommen. Das Fahrzeug fährt in den Fehlerzustand "Fahren gerade aus".</p>
$K = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad O_0 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}$	
Ergebnis der Zeitreihenentwicklung 2	
	<p>Fahrer opt. Zust.</p> <p>Automation opt. Zust.</p> <p>Fahrzeug Zust.</p> <p>Umwelt Zust.</p>
Interpretation der Zeitreihenentwicklung 2	
<p>Systementscheidung konvergiert nicht. Das Fahrzeug fährt gerade aus. Der Zustand ist instabil, weil stark konfliktbehaftet.</p>	
Zeitreihenentwicklung 3	
Grunkonfiguration	Erwartetes Ergebnis
<p>Fahrer lenkt nach links, Assistenz lenkt nach rechts, das Fahrzeug fährt gerade aus, und der Verlauf der Straße hat eine Tendenz nach links. Fahrer und Assistenz beeinflussen sich nicht gegenseitig. Die Assistenz ist in die Kontrolle eingekoppelt, Fahrer wird entkoppelt.</p>	<p>Das System sollte insgesamt nach rechts fahren und dort einen stabilen Zustand erreichen.</p>
$K = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad O_0 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0,5 \end{pmatrix}$	
Ergebnis der Zeitreihenentwicklung 3	
	<p>Fahrer opt. Zust.</p> <p>Automation opt. Zust.</p> <p>Fahrzeug Zust.</p> <p>Umwelt Zust.</p>
Interpretation der Zeitreihenentwicklung 3	
<p>Systementscheidung konvergiert in Richtung der Assistenzintention. Das Fahrzeug fährt nach rechts nach einem kurzen Schwinger, bleibt aber stabil. Der Fahrer wird mit einer kurzen Verzögerung umgestimmt. Dies ist eine bekannte Strategie aus der Gestaltung der Fahrer-Fahrzeug-Interaktion in Notfallsituationen (Heesen et al. 2014), wenn der Fahrer für kurze Zeit komplett oder teilweise von der Kontrolle entkoppelt wird. Die Entkopplung der Assistenz statt des Fahrers würde eine entgegengesetzte Entscheidungskonvergenz vorweisen. Dies entspricht der Strategie der Kontrollübernahme durch den Fahrer z.B. mittels des Wechsels des Automationslevels mittels einer Zustandsübergang (Schieben et al. 2008) oder durch das Abschalten/Übersteuern der Assistenz (Heesen et al. 2014).</p>	

Zeitreihenentwicklung 4	
Grunkonfiguration	Erwartetes Ergebnis
<p>Fahrer lenkt nach links, Assistenz lenkt nach rechts, das Fahrzeug fährt gerade aus, und der Verlauf der Straße hat eine Tendenz nach links. Fahrer und Assistenz sind in die Kontrolle eingekoppelt. Assistenz beeinflusst den Fahrer, der Fahrer beeinflusst nicht die Assistenz.</p>	<p>Das System sollte insgesamt nach rechts fahren und dort einen stabilen Zustand erreichen.</p>
$K = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad O_0 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0,5 \end{pmatrix}$	
Ergebnis der Zeitreihenentwicklung 4	
	<p>Fahrer opt. Zust.</p> <p>Automation opt. Zust.</p> <p>Fahrzeug Zust.</p> <p>Umwelt Zust.</p>
Interpretation der Zeitreihenentwicklung 4	
<p>Systementscheidung konvergiert in Richtung der Assistenzintention. Das Fahrzeug fährt nach rechts nach einer kurzen Verzögerung, bleibt aber stabil. Der Fahrer wird mit einer kurzen Verzögerung umgestimmt. Dies ist eine bekannte Standardstrategie aus der Gestaltung der Fahrer-Fahrzeug-Interaktion mittels einer indirekten Beeinflussung des Fahrers durch z.B. ein passives multimodales Feedback (Kelsch 2012a). Für den umgekehrten Fall, dass der Fahrer die Assistenz beeinflusst, würde ein entgegengesetztes Ergebnis herauskommen, d.h. das System würde nach links fahren.</p>	

Tab.1: Vier (plus zwei) getestete Fahrer-Automation-Fahrzeug-Umwelt-Systemkonfigurationen mit deren erwarteten Ergebnissen, Zeitreihenentwicklungen und Interpretation

## 5. DISKUSSION

Die Evaluationsergebnisse legen nahe, dass das nach dem Konzept der kognitiven Multiagenten-Systeme erstellte Fahrer-Automation-Fahrzeug-Umwelt Modell ein erwartetes und adäquates Verhalten liefert. Dies ist insofern mit einer Einschränkung zu sehen, dass es hierbei um nur sechs (vier evaluierte plus zwei spiegelverkehrte zu Zeitreihenentwicklung 3 und 4) Konfigurationen handelt. Nichtsdestotrotz decken diese Konfigurationen bereits mehrere bekannte Ansätze aus dem Bereich der Gestaltung hochautomatisierter Fahrzeugführung. Es sind z.B. die Auskopplungen des Fahrers und der Automation abbildbar, was dem Konzept der Automationslevel (Parasuraman et al. 2000) und den Transitionen (Schieben et al. 2008) dazwischen nahe kommt. Des Weiteren ist das Konzept der indirekten Beeinflussung des Fahrers durch das passive Automationsfeedback abbildbar, das einen der grundlegendsten Ansätze der Mensch-Maschine Interaktionsgestaltung darstellt. Die indirekte Beeinflussung der Automation ist ebenfalls abbildbar, die einer indirekten Fahrzeugkontrolle entspricht. Angesichts der Vielfalt der Konzepte, die in einem Modell zusammenhängend (vor-)untersucht werden können, lässt sich der Nutzen der vorgeschlagenen Modellierungsmethode der CS erkennen. In der Entscheidungskonvergenz innerhalb der zeitlichen Entwicklung des aufgestellten Modells konnten Konzepte der impliziten Arbitrierung (Kelsch 2012a) und des adaptiven Automationsverhaltens (Griesche et al. 2012) generiert werden. Damit konnte die Arbeitshypothese, dass Entscheidungen innerhalb kognitiver Systeme konvergieren, mit der Einschränkung der wenigen Testläufe bestätigt werden.

Insgesamt lässt sich die vorgeschlagene Methode der Modellierung der CS den aktionsorientierten Methoden zuordnen, wie der hierarchischen Methode von (Brooks 1996). Wegen des Konzeptes der Aktionstension steht sie aber auch nicht im direkten Konflikt zu den BDI-Architekturen (belief, desire, intention) (Rao *et al.* 1991) aufgrund der Berücksichtigung von Systemzielen. Diese sind prinzipiell im optimalen Zustand der Aktionstensionen vorhanden. Eine der wichtigsten Annahmen der Methode, die der Entscheidungskonvergenz, entspricht dem allgemeinen Verständnis der zellulären Automaten und deren generativem Gestaltungsansatz (Wolfram 2002). Alles zusammen legt nahe, dass die vorgeschlagene Methode der Dekomposition und der Modellierung der CS zielführend sein kann und weiter zu verfolgen ist.

## 6. AUSBLICK

Entscheidungskonvergenzen sollten in Zukunft für komplexere Systeme evaluiert werden. Die vorgestellte CS-Modellierungsmethode kann dafür auf mehrschichtige Modelle erweitert werden. Im Erfolgsfall können damit z.B. komplexere kognitive Modelle, wie das der drei Level der Kontrolle (Rasmussen 1996) oder das der Situation Awareness (Endsley 1995) formal nachgebildet und evaluiert werden. Interferenzeffekte zwischen den einzelnen Modalitäten der Wahrnehmung, der kognitiven Informationsverarbeitung und der anschließenden Handlung der Agenten könnten damit untersucht werden. Interessant erscheint auch die Erweiterung der Methode in Richtung des Modellierens des Lernverhaltens, Aufmerksamkeitseffekte, mentaler Modelle etc. Ebenfalls interessant erscheint die Anbindung der vorgestellten Methode an die Komplexitätstheorie mit ihrer mathematischen Methodensammlung, wie die Analyse der Hyperstrukturen, der Stabilität, der Emergenz etc.

Durch die Skalierbarkeit der vorgestellten Methode und der damit erstellten Modelle könnten auch komplexere soziotechnische Systeme mit den entsprechenden verteilten sozialpsychologischen Effekten der Kooperation, der Konkurrenz etc. nachgebildet und untersucht werden. Durch den im Kern der Methode liegenden *generativen Ansatz* könnten vielleicht sogar Effekte in beliebiger sozial-, kognitions- oder neuropsychologischer Tiefe modelliert und untersucht werden. Damit könnten solche grundsätzlichen Forschungsfragen angegangen werden, wie: Gibt es in kognitiven Systemen ein Kleinstelement, so etwas wie ein „Kognitron“?

## LITERATUR

- Brooks, R. (1986): *A robust layered control system for a mobile robot*, IEEE Journal of Robotics and Automation 2 (1): pp. 14-23
- Castelfranchi, C. (1998). *Modeling Social Action for AI Agents*. Journal Artificial Intelligence - Special issue: artificial intelligence 40 years later archive Volume 103, Aug. 1-2, 1998
- Christoffersen, K.; Woods, D. D. (2002): *How to make automated systems team players*. In: Salas, E. (Hrsg.): *Advances in Human Performance and Cognitive Engineering*. Research Bd. 2. Elsevier Science Ltd., 2002, S. 1-12
- Endsley, M. R. (1995): *Toward a theory of situation awareness in dynamic systems*. Human Factors 37, S. 32-64
- Flemisch, F.; Kelsch, J.; Heesen, M.; Löper, C. (2010): *Wie erkunden wir die Lebenswelt kooperativer Technik? Skizze des Gestaltungsraums haptisch-multimodaler Kopplung Mensch, Co-Automation und Regelstrecke als Teil einer kooperativen Bewegungsbeeinflussung*. GfA Frühjahrskonferenz 2010, Darmstadt
- Gasser, T.M.; Arzt, C.; Ayoubi, M.; Bartels, A.; Bürkle, L.; Eier, J.; Flemisch, F.; Häcker, D.; Hesse, T.; Huber, W.; Lotz, C.; Maurer, M.; Ruth-Schumacher, S.; Schwarz, J.; Vogt, W. (2012): *Rechtsfolgen zunehmender Fahrzeugautomatisierung*. *Berichte der Bundesanstalt für Straßenwesen*, Fahrzeugtechnik, F83, S. 9
- Griesche, S.; Kelsch, J.; Heesen, M.; Martirosjan, A. (2012): *Adaptive Automation als ein Mittel der Arbitrierung zwischen Fahrer und Fahrzeugautomation*. AAET 2012, Braunschweig
- Heesen, M.; Dziennus, M.; Hesse, T.; Schieben, A.; Brunken, C.; Löper, C.; Kelsch, J.; Baumann, M. (2014): *Interaction design of automatic steering for collision avoidance: challenges and potentials of driver decoupling*. IET Intelligent Transport Systems 01/2014
- Hollnagel, E.; Woods, D. D. (1982): *Cognitive Systems Engineering: New wine in new bottles*. International Journal of Human-Computer Studies, Volume 51, Nr. 2, Aug. 1999, S. 339-356
- Kelsch, J. (2012a): *Arbitration between Driver and Automation: why Overriding is just the Tip of the Iceberg*. InteractIVe Summer School, 04.-06.06.12, Corfu Island, Greece
- Kelsch, J., Heesen M., Hesse T., Baumann M. (2012b): *Using human-compatible reference values in design of cooperative dynamic human-machine systems*. EAM 2012, 11-12.09.2012, Braunschweig, Germany
- Kelsch, J.; Temme, G.; Schindler, J. (2013): *Arbitration based framework for design of holistic multimodal human-machine interaction*. AAET 2013, 6.-7. Feb. 2013, Braunschweig, Germany
- Luck, M., McBurney, P., Shehory, O.; Willmott, S.; AgentLink Community (2005): *Agent Technology: Computing as Interaction. A Roadmap for Agent Based Computing*. AgentLink III, Sept. 2005, ISBN: 0854328459
- Norman, D. A.; Draper, S. (1986): *User centered system design: New perspectives on human-computer interaction*. Lawrence Erlbaum Associates, Mahwah, S. 419
- Parasuraman, R.; Sheridan, T. B.; Wickens, C. D. (2000): *A Model for Types and Levels of Human Interaction with Automation*. IEEE Transactions On Systems, Man, And Cybernetics – Part A: Systems And Humans, Vol. 30, No. 3, S. 286-297
- Rao, A. S.; Georgeff, M. P. (1991): *Modeling rational agents within a BDI-architecture*. Proceedings of the International Conference on Knowledge, Representation and Reasoning, S. 473-484
- Rasmussen, J. (1986): *Information processing and human-machine interaction: An approach to cognitive engineering*. New York, North-Holland, S. 101-115
- Schieben, A.; Flemisch, F.; (2008): *Who is in control? Exploration of transitions of control between driver and an eLane vehicle automation*. In: VDI-Berichte 2048. VDI Verlag. Wolfsburg, ISBN 9783180920481, S. 455-469
- Schumacher, W.; Leonhard, W. (2001): *Grundlagen der Regelungstechnik*. Vorlesungsskript Stand: 17.10.2001. Technische Universität Braunschweig
- Sheridan, T. B.; Verplank, W. (1978): *Human and Computer Control of Undersea Teleoperators*. Cambridge, MA: Man-Machine Systems Laboratory, Department of Mechanical Engineering, MIT
- Springer Gabler Verlag (Hrsg.) (2013a): *Gabler Wirtschaftslexikon*, Stichwort: System, <http://wirtschaftslexikon.gabler.de/Archiv/3210/system-v11.html>, Zugriff am 22.07.2013
- Springer Gabler Verlag (Hrsg.) (2013b): *Gabler Wirtschaftslexikon*, Stichwort: Kognition, <http://wirtschaftslexikon.gabler.de/Archiv/9957/kognition-v8.html>, Zugriff am 23.07.2013
- Vakarelov, O. (2009): *The Cognitive Agent*. Department of Philosophy, University of Arizona, Tucson, Arizona, Arizona. Online: <http://www.u.arizona.edu/~okv/Papers/CogAgent.pdf>.
- Watzlawick, P.; Helmick Beavin, J.; Jackson, D.D. (1967): *Pragmatics of Human Communication: A Study of Interactional Patterns, Pathologies, and Paradoxes*. Norton, ISBN: 0393010090
- Wikipedia (2013): *Kognition*. <http://de.wikipedia.org/wiki/Kognition>. Zugriff am 23.07.2013
- Wolfram, S. (2002): *A New Kind of Science*. Wolfram Media. ISBN 9781579550080