Modellierung und Verifikation von Fähigkeiten rationaler Agenten

Nils Bulling

Institut für Informatik, Technische Universität Clausthal, Deutschland
bulling@in.tu-clausthal.de

Abstract: In dieser Dissertation wird untersucht, wie rationales Verhalten von Agenten modelliert und verifiziert werden kann. Es werden diverse Ansätze durch formale logikorientierte Methoden ausgedrückt und die Komplexität der jeweiligen Modellverifikationsprobleme (*model checking problems*) bestimmt.

1 Einleitung

Temporale Logiken haben eine lange Tradition in der Informatik. Zu den prominentesten Temporallogiken zählen die *linear time temporal logic* **LTL** und die *computation tree temporal logic* **CTL**. In diesen Formalismen lassen sich beispielsweise *Deadlock*- und Sicherheitseigenschaften ausdrücken. Beide Logiken erlangten viel Aufmerksamkeit im Zusammenhang mit der Spezifikation und Verifikation von reaktiven Systemen [CES86].

Ende des 20. Jahrhunderts wurde in der klassischen verteilten KI ein neues Paradigma eingeführt, das der *Agenten*. Dieses hat zu einem wichtigen Forschungsgebiet im Bereich autonomer und verteilter Computersysteme [Woo02, Wei99] geführt. Ein großer Teil dieser Forschung beschäftigt sich mit Aspekten von kooperativen Agenten, insbesondere ihrer kollaborativen Fähigkeiten. Die *strategische* Logik **ATL** (*alternating time temporal logic*) wurde als ein formales Mittel vorgeschlagen [AHK02], um entsprechende Eigenschaften in Multiagentensystemen (MAS) zu spezifizieren und Schlussfolgerungen über selbige anzustellen. Bis heute gehört diese Logik zu den einflussreichsten ihrer Art und erfreut sich großer Beliebtheit, weil sie viele Erweiterungen zulässt, die ihre Ausdruckskraft erhöhen und sie so in neuen Kontexten anwendbar macht.

In **ATL** wird der **CTL**-Pfadquantor E (es gibt eine Berechnung) durch sogenannte *Kooperationsmodalitäten* $\langle\langle A\rangle\rangle$ ersetzt, die es erlauben, über Fähigkeiten einer Agentengruppe A zu reden. Eine Formel $\langle\langle A\rangle\rangle\gamma$ ist wahr, genau dann, wenn die Gruppe A eine Strategie hat, die γ garantiert, unabhängig davon wie sich die übrigen Agenten außerhalb von A verhalten. In verschiedenen Artikeln wurde gezeigt, dass sich diese Logik sehr gut zur Spezifikation und Verifikation von MAS, also *offenen Systemen*, eignet. Als illustrierendes Beispiel betrachten wir Mühle. Dieses Spiel ist gelöst, d.h. der beginnende Spieler hat eine Strategie, die ihn das Spiel nicht verlieren lässt. Dies lässt sich in **ATL** als $\langle\langle 1\rangle\rangle\Box\neg\mathsf{lose}_1$ ausdrücken (Spieler 1 hat eine Strategie, so dass er für alle Gegenstrategien von Spieler 2 nie verliert). Es gilt auch $\langle\langle 1,2\rangle\rangle\Diamond\mathsf{lose}_1$ (Spieler 1 und 2 haben eine Strategie, die Spieler 1 irgendwann verlieren lässt), $\langle\langle 1,2\rangle\rangle\Diamond\mathsf{lose}_2$ und $\neg\langle\langle 1,2\rangle\rangle\Diamond(\mathsf{lose}_1 \wedge \Diamond\mathsf{lose}_2)$, wenn beide Spieler, 1 und 2, kooperieren.

ATL erlaubt es allerdings nur, pure Fähigkeiten von Agenten zu modellieren, ohne auf höhere Konzepte wie Ziele, Präferenzen, Ressourcen oder ähnliche Eigenschaften einzugehen. Obwohl $\langle\langle 1,2\rangle\rangle\Diamond\mathsf{lose}_1$ gilt, sollte ein rationaler Spieler (hier: Spieler 1) die Absicht haben, zu gewinnen. Also sollte er in diesem Fall nicht mit Spieler 2 kooperieren. Hier knüpft meine Dissertation an. Es werden Logiken erweitert, um das Verhalten *rationaler* Agenten zu untersuchen. Der Schwerpunkt liegt auf der Entwicklung von *Logiken zur Modellierung* rationaler Agenten und den Komplexitätsuntersuchungen der zugehörigen *Modellverifikationsprobleme* (*model checking problems*).

Die Untersuchung des Verhaltens rationaler Agenten ist ein Kerngebiet der KI. In den letzten Jahren haben Computersysteme viele neue Anwendungsbereiche, in denen Systeme direkt oder indirekt mit Menschen interagieren oder den Menschen auf unterschiedlichste Arten unterstützen sollen (*human-centered computing*), erschlossen.

In meiner Arbeit werden unter anderem spieltheoretische Konzepte genutzt, um die oben genannten Aspekte zu untersuchen. Die Untersuchung von rationalem Verhalten war einer der Ausgangspunkte der Spieltheorie: Bei der Analyse von Spielen wird in der Regel angenommen, dass Spieler im Sinne eines festgelegten Lösungskonzeptes (z. B. Nash-Gleichgewicht) agieren. Entsprechende Zusammenhänge zwischen **ATL** und spieltheoretischen Konzepten wurden zwar bereits von mehreren Autoren untersucht (siehe z.B. [vdHW02, vOvdHW03]). Der Schwerpunkt lag jedoch hauptsächlich auf der Charakterisierung und weniger auf der Verwendung von solchen Lösungskonzepten.

Die Logik **ATLP** (*alternating time temporal logic with plausibilty*) [BJD09], die in meiner Dissertation eingeführt wird, verallgemeinert existierende Ansätze und eignet sich sowohl zur Charakterisierung als auch zur Nutzung von spieltheoretischen Konzepten. Dadurch wird es ermöglicht, das Verhalten von Agenten auf ein plausibles zu reduzieren und so den Strategieraum mitunter stark einzuschränken.

In einer anderen Erweiterung von **ATL** steht nicht rationales Verhalten im Sinne von verfügbaren Strategien im Vordergrund, sondern ob Koalitionen Anreize (z.B. gemeinsame Ziele) haben, zu kooperieren [BCD08]. Dieser Ansatz kombiniert **ATL** mit Argumentationstheorie.

Beide beschriebenen Szenarien unterliegen der Annahme, dass Agenten vollständige Informationen über ihre Umgebung haben. Modelltheoretisch bedeutet dies, dass je zwei verschiedene Welten im Modell unterscheidbar sind. In meiner Dissertation werden zwei Ansätze betrachtet, die es erlauben, rationales Verhalten von Agenten mit *unvollständigen* Informationen zu modellieren. Um dies zu erreichen, wird **ATLP** um klassische epistemische Konzepte erweitert. Zusammenhänge zwischen epistemischen und doxastischen Eigenschaften in Verbindung mit Rationalitätsannahmen werden untersucht [BJ09a].

In dem zweiten Ansatz zur Modellierung von unvollständigen Informationen wird die qualitative Semantik von **ATL** („gewinne oder verliere“) um wahrscheinlichkeitstheoretische Konzepte erweitert. Motiviert wird diese Erweiterung durch eine Vielzahl von Szenarien, in denen die Gegner nicht genügend Informationen haben, um die für die Proponenten ungünstigste Strategie zu identifizieren [BJ09b]. Letzteres entspricht aber genau der Annahme, die der restriktiven Semantik von **ATL** unterliegt.

Schließlich wird untersucht, wie sich Ressourcen auf das Verhalten rational agierender Agenten auswirken. Es wird gezeigt, dass viele Logiken dieser Art ein unentscheidbares Modellverifikationsproblem besitzen [BF10b].

2 Präliminarien: Strategische Logik und Modellverifikation

Die strategische Logik ATL Im Folgenden sei $\mathbb{A}\mathrm{gt} = \{1, \dots, k\}$ eine nicht leere und endliche Menge von *Agenten*, Q eine nicht leere und endliche Menge von *Zuständen* und Π eine Menge von *Propositionen*. Die strategische Logik **ATL** [AHK02] (*alternating time temporal logic*) verallgemeinert die *computation tree logic* **CTL**: Die Pfadquantoren E und A werden durch die Kooperationsmodalitäten $\langle\langle A\rangle\rangle$ ersetzt, wobei $A \subseteq \mathbb{A}\mathrm{gt}$ eine Gruppe von Agenten repräsentiert. Die Formel $\langle\langle A\rangle\rangle\gamma$ drückt aus, dass die Koalition A eine *kollektive Strategie* hat, um γ zu garantieren. Formal wird die Sprache von **ATL** durch folgende BNF-Grammatik definiert: $\varphi ::= \mathsf{p} \mid \neg\varphi \mid \varphi \wedge \varphi \mid \langle\langle A\rangle\rangle \bigcirc \varphi \mid \langle\langle A\rangle\rangle\Box\varphi \mid \langle\langle A\rangle\rangle\varphi\mathcal{U}\varphi$, wobei $A \subseteq \mathbb{A}\mathrm{gt}$ und $\mathsf{p} \in \Pi$. Für eine einelementige Gruppe $\{a\}$ wird einfachheitshalber $\langle\langle a\rangle\rangle$ geschrieben. Die temporalen Operatoren $\Box$ (immer), $\bigcirc$ (im nächsten Moment) und $\mathcal{U}$ (bis) erlauben es, Aussagen über zeitliche Abläufe entlang einer Berechnung (d.h., einer ω-Sequenz über Q) zu machen. Überdies wird $\Diamond\varphi$ (irgendwann) als $\top\mathcal{U}\varphi$ definiert. Die Formel $\langle\langle A\rangle\rangle\mathsf{r}\mathcal{U}\mathsf{s}$ drückt beispielsweise aus, dass die Gruppe A eine Strategie hat, die die Proposition s wahr macht, und dass bis dahin r gilt.

Die Semantik der Logik wird über *simultane Spielstrukturen* (*concurrent game structures*) (CGSs) definiert [AHK02]. Eine solche Struktur ist durch ein 7-Tupel $\mathfrak{M} = \langle \mathbb{A}\mathrm{gt}, Q, \Pi, \pi, Act, d, o\rangle$ gegeben. Neben den bereits bekannten beinhaltet sie folgende Elemente: (i) eine Valuationsfunktion $\pi : \Pi \to \mathcal{P}(Q)$, (ii) eine nicht leere und endliche Menge Act von Aktionen, (iii) eine Verfügbarkeitsfunktion $d : \mathbb{A}\mathrm{gt} \times Q \to \mathcal{P}(Act)$, welche angibt, welche Aktionen in welchen Zuständen von welchen Agenten ausgeführt werden können, und (iv) eine deterministische Transitionsfunktion o, die jedem Zustand $q \in Q$ und jedem Aktionsprofil $\langle\alpha_1, \dots, \alpha_k\rangle$ mit $\alpha_i \in d(i, q)$, für $1 \leq i \leq k$, einen eindeutig bestimmten Nachfolgezustand $q' = o(q, \alpha_1, \dots, \alpha_k)$ zuordnet. Anstelle von $d(a, q)$ schreiben wir $d_a(q)$. In Abbildung 1 ist eine CGS abgebildet.

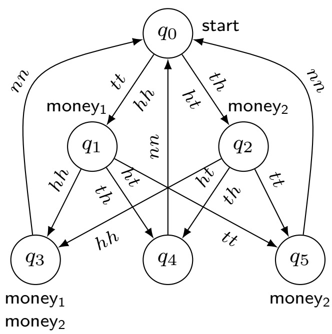

Abbildung 1: CGS $\mathfrak{M}$ mit 2 Agenten, die die Aktionen h, t, und n ausführen können. Ein Tupel th entspricht dem Aktionsprofil, bei dem Agent 1 t und Agent 2 h ausführt. start, money$_1$ und money$_2$ sind Propositionen.

Die Formel $\langle\langle A\rangle\rangle\gamma$ drückt aus, dass die Agenten in A eine sogenannte *Gewinnstrategie* haben, um γ zu verwirklichen. Es wird zwischen *speicherlosen Strategien für einen Agenten a*, $s_a : Q \to Act$ mit $s_a(q) \in d_a(q)$, die jedem

Zustand eine Aktion zuweist und *speichernutzenden Strategien* unterschieden. Letztere werden durch Funktionen $s_a : Q^+ \to Act$ mit $s_a(q_0 \dots q_n) \in d_a(q_n)$ repräsentiert. Eine *kollektive Strategie* s_A für eine Gruppe A ist ein Tupel $(s_{a_1}, \dots, s_{a_{|A|}})$, bestehend aus einer individuellen Strategie s_{a_i} für jeden Agenten $a_i \in A$. Für eine solche kollektive Strategie s_A wird $s_A|_a$ genutzt, um auf a's Strategie s_a in s_A zu referenzieren, für $a \in A$.

Führen Agenten eine Strategie s_A aus, resultiert dies in einer Menge von Pfaden in der unterliegenden CGS. Für $A = \mathbb{A}\mathrm{gt}$ wird genau ein Pfad generiert (Determinismus!). Ein *Pfad* oder eine *Berechnung* ist eine ω-Sequenz $\lambda = q_0 q_1 \dots \in Q^\omega$ von Zuständen. $\lambda[i] = q_i$ bzw. $\lambda[i,j]$ referenziert auf den i-ten Zustand bzw. auf die Sequenz $\lambda[i] \dots \lambda[k]$ für $i \leq j$. Gültig ist auch $j = \infty$. Das *Resultat einer Strategie s_A in einem Zustand q* ist die Menge $out_{\mathfrak{M}}(q, s_A)$, die alle Pfade $\lambda = q_0 q_1 \dots$ enthält, so dass $q_0 = q$ und es für jedes $i = 0, 1 \dots$ einen Aktionsvektor $\langle \alpha_1^{i-1}, \dots, \alpha_k^{i-1} \rangle$ mit $\alpha_a^{i-1} \in d_a(q_{i-1})$ für jeden Agenten $a \in \mathbb{A}\mathrm{gt}$ gibt, so dass $\alpha_a^{i-1} = s_A|_a(q_0 q_1 \dots q_{i-1})$ für jeden Agenten $a \in A$, und $o(q_{i-1}, \alpha_1^{i-1}, \dots, \alpha_k^{i-1}) = q_i$. Das Resultat einer speicherlosen Strategie wird analog definiert.

Die *speicherlose bzw. speichernutzende Semantik* von **ATL** wird wie folgt definiert (für andere Formeln ist die Semantik standard):

$\mathfrak{M}, q \models \mathsf{p}$ gdw. $\lambda[0] \in \pi(\mathsf{p})$ und $\mathsf{p} \in \Pi$;

$\mathfrak{M}, q \models \langle\langle A \rangle\rangle \gamma$ gdw. es gibt eine speicherlose bzw. speichernutzende Strategie s_A für A, so dass für alle Pfade $\lambda \in out(q, s_A)$ gilt $\mathfrak{M}, \lambda \models \gamma$ und

$\mathfrak{M}, \lambda \models \Box\gamma$ gdw. $\mathfrak{M}, \lambda[1, i] \models \gamma$ für alle $i \in \mathbb{N}_0$;

$\mathfrak{M}, \lambda \models \gamma \mathcal{U} \delta$ gdw. es gibt $i \in \mathbb{N}_0$, so dass $\mathfrak{M}, \lambda[i, \infty] \models \delta$ und $\mathfrak{M}, \lambda[j, \infty] \models \gamma$ für alle $0 \leq j < i$.

Die Logiken **ATL**$^+$ und **ATL*** erweitern **ATL**. **ATL**$^+$ erlaubt Boolesche Kombinationen von Pfadformeln, z.B. $\langle\langle A \rangle\rangle (\Diamond \mathsf{clean} \wedge \Diamond \mathsf{deliver})$. **ATL*** ist noch ausdrucksstärker: Temporaloperatoren und Kooperationsmodalitäten können beliebig kombiniert werden.

Beide Semantiken, die speicherlose und die speichernutzende, können auch um *unvollständige Informationen* erweitert werden. Dazu wird für jeden Agenten a eine *Äquivalenzrelation* $\sim_a \subseteq Q \times Q$ eingeführt. Stehen zwei Zustände in Relation, $q \sim_a q'$, bedeutet dies, dass beide Zustände für den Agenten a ununterscheidbar sind. Dies hat starke Auswirkungen auf die strategischen Fähigkeiten von Agenten.

Modellverifikation In Kombination mit diesen Logiken erlangte insbesondere das *Modellverifikationproblem* (*model checking problem*), welches zur formalen Verifikation von Systemen eingesetzt wird, große Bedeutung. Das Modellverifikationsproblem erwartet als Eingabe ein Modell $\mathfrak{M}$, einen Zustand q in diesem Modell und eine Formel φ und entspricht der Frage, ob $\mathfrak{M}, q \models \varphi$ gilt. Es wird also bestimmt, ob die Formel φ in $\mathfrak{M}$ und q wahr ist oder nicht.

Die Modellverifikationsprobleme von **ATL** bzw. **ATL*** sind $\mathbf{P}$- bzw. $\mathbf{2EXPTIME}$-vollständig bzgl. der speichernutzenden Semantik [AHK02]. Die angegebene Komplexität für

ATL lässt sich auch nicht verbessern, wenn die weniger komplexen speicherlosen Strategien verwendet werden. Der Grund dafür ist, dass sich beide Semantiken für **ATL** nicht unterscheiden: Die Sprache ist nicht ausdrucksstark genug. Ein Resultat, welches in meiner Dissertation bewiesen wird, ist, dass beide Semantiken zu unterschiedlicher Komplexität bzgl. der Sprache $\mathbf{ATL}^+$ führen. Bisher wurde angenommen, dass das Modellverifikationsproblem für $\mathbf{ATL}^+$ für beide Semantiken $\mathbf{\Delta_3^P}$-vollständig ist [Sch04]. Leider bedeutet das neue Resultat aus meiner Dissertation, dass Agenten mit Speicher schwerer (unter den üblichen Annahmen) zu verifizieren sind, als zuvor geglaubt: Das Modellverifikationsproblem für $\mathbf{ATL}^+$ unter Verwendung der speichernutzenden Semantik ist $\mathbf{PSPACE}$-vollständig [BJ10].

3 Modellierung rationaler Agenten

Analyse von rationalem Spiel In meiner Dissertation wird unter anderem untersucht, wie sich rationales Verhalten mithilfe von spieltheoretischen Mitteln analysieren lässt. **ATL** erlaubt es auszudrücken, dass eine Gruppe A eine Strategie hat, eine Formel γ zu garantieren: $\langle\langle A\rangle\rangle\gamma$. Alle möglichen Verhaltensweisen von A, als auch von den Opponenten $\mathbb{A}\mathrm{gt}\backslash A$, werden dabei betrachtet. Solch eine Aussage ist schwächer als sie auf den ersten Blick anmuten lässt. Oftmals ist es legitim anzunehmen, dass sich Agenten gemäß gewisser Rationalitätsannahmen verhalten. Sie sind nicht komplett unwissend, z. B. sollten in der Regel keine dominierten Strategien gespielt werden. Solche Beobachtungen erlauben es, den Lösungsraum einzuschränken und sich auf rationales Verhalten zu beschränken. Zu diesem Zweck stelle ich in meiner Dissertation die Logik **ATLP** (*alternating time temporal logic with plausibility*) und diverse Fragmente vor. In diesem Abstrakt wird hauptsächlich das Fragment $\mathbf{ATLP}^{\mathbf{base}}$ betrachtet. $\mathbf{ATLP}^{\mathbf{base}}$ erweitert **ATL** um die Operatoren $\mathbf{Pl}_A$ und $(\mathbf{set\text{-}pl}\ \omega)$, wobei $\omega \in \Omega$ ein *Plausibilitätsterm* aus einer nicht leeren Menge Ω ist. Der erste Operator drückt aus, dass Agenten in A rational agieren, wohingegen der zweite Operator dynamisch beschreibt, was unter Rationalität zu verstehen ist. Der Leser möge sich unter dem *Plausibilitätsterm* ω etwa eine Beschreibung von Strategien im Nash-Gleichgewicht oder dominante Strategien vorstellen. Modelle für $\mathbf{ATLP}^{\mathbf{base}}$ erweitern CGSs um folgende Elemente: Eine Menge Υ von *plausiblen* speicherlosen Strategieprofilen (*Plausibilitätsmenge*), eine Menge Ω von Plausibilitätstermen, und eine *Plausibilitätsabbildung* $[\![\cdot]\!] : Q \rightarrow (\Omega \rightarrow \mathcal{P}(\Sigma))$, welche jedem Plausibilitätsterm zustandsabhängig eine Menge von Strategien zuweist. Σ steht für die Menge aller speicherlosen Strategien. Anstelle von $[\![q]\!](\omega)$ wird $[\![\omega]\!]^q$ geschrieben. Zur Illustration wird die Formel $(\mathbf{set\text{-}pl}\ \omega_{\mathrm{NE}})\mathbf{Pl}_B\langle\langle A\rangle\rangle\gamma$ betrachtet. Sie liest sich wie folgt: Angenommen es ist rational, Strategien im Nash-Gleichgewicht zu spielen, $(\mathbf{set\text{-}pl}\ \omega_{\mathrm{NE}})$, und alle Agenten in der Gruppe B agieren gemäß dieser Rationalitätsannahme, $\mathbf{Pl}_B$, dann hat die Gruppe A eine Strategie, um γ zu garantieren.

In dieser Grundlogik ist die Denotation der Terme ω fest durch den Entwickler gegeben. Im nächsten Schritt zur vollständigen Logik **ATLP** wird $\mathbf{ATLP}^{\mathbf{base}}$ so erweitert, dass plausible Strategien durch geschachtelte **ATLP**-Formeln charakterisiert werden können. Es werden Variablen σ, deren semantische Interpretation Strategien aus Σ sind, eingeführt. Für die Interpretation von spieltheoretischen Lösungskonzepten werden *Nutzenfunktionen* (wie in der Spieltheorie üblich) benötigt. Sie geben an, wie erfolgreich Strategieprofile sind. Dazu werden temporallogische Formeln herangezogen: Erfüllt der durch ein vollständiges Strategieprofil eindeutig bestimmte Pfad die Zielformel, ist das Strategieprofil erfolgreich, sonst nicht. Daher nutzen wir auch den Begriff *Zielformel* für die Nutzenfunktion.

Zur Illustration wird erneut die CGS $\mathfrak{M}$ aus Abbildung 1 betrachtet. Es wird angenommen, dass die beiden Agenten die Zielformeln η_1 bzw. η_2 aus Abbildung 2 besitzen. Die Abbildung zeigt das Normalformspiel $\mathcal{S}(\mathfrak{M}, (\eta_1, \eta_2), q_0)$, welches aus $\mathfrak{M}$, η_1, η_2 und dem Startzustand q_0 gewonnen wird. Beispielsweise erfüllt der Pfad $(q_0 q_1 q_3)^\omega$, der zur Strategie (s_{hhh}, s_{hhh}) gehört, beide Zielformeln, und entsprechend ist der jeweilige Erfolg dieser Strategie für beide Agenten 1. Dieses Vorgehen erlaubt es, CGSs mit strategischen Spielen zu verknüpfen, und so Formeln mit Lösungskonzepten in Relation zu setzen. Beispielsweise charakterisiert die Formel

$$BA_a^{\eta_a}(\sigma) \equiv (\textbf{set-pl}\ \sigma[\mathbb{A}\mathrm{gt} \setminus \{a\}])\mathbf{Pl}\Big((\langle\!\langle a \rangle\!\rangle \eta_a) \rightarrow (\textbf{set-pl}\ \sigma)\langle\!\langle \emptyset \rangle\!\rangle \eta_a\Big)$$

eine Beste-Antwort-Strategie (BA) σ_a von Agent a für ein gegebenes Strategieprofil $\sigma_{\mathbb{A}\mathrm{gt}\setminus\{a\}}$ der Gegner, wobei η_a die Zielformel von a ist. Die Formel liest sich wie folgt: Angenommen alle Spieler in $\mathbb{A}\mathrm{gt}\setminus\{a\}$ spielen gemäß σ ((**set-pl** $\sigma[\mathbb{A}\mathrm{gt} \setminus \{a\}]$)) und a hat eine Strategie, um seine Zielformel η_a zu erfüllen ($\langle\!\langle a \rangle\!\rangle \eta_a$), dann wird diese auch dann erfüllt, wenn a die Strategie σ_a spielt ((**set-pl** $\sigma)\langle\!\langle \emptyset \rangle\!\rangle \eta_a$). Es gibt also keine bessere Strategie für a. In meiner Dissertation wird gezeigt, wie man durch Formeln von **ATLP** viele Standardlösungskonzepte charakterisieren kann [BJD09].

Die Logik **ATLP** erweitert die Grundlogik außerdem um die Möglichkeit, über Strategien zu quantifizieren und Plausibilitätsterme beliebig zu schachteln. Es wird gezeigt, dass **ATLP**$^{\textbf{base}}$ $\mathbf{\Delta_3^P}$-vollständig und die allgemeine Logik **PSPACE**-vollständig ist. Weiterhin werden die oberen Komplexitäten von allen dazwischenliegenden Fragmenten bestimmt.

$\eta_1\backslash\eta_2$	s_{hhh}	s_{hht}	s_{hth}	s_{htt}	s_{thh}	s_{tht}	s_{tth}	s_{ttt}
s_{hhh}	**1,1**	**1,1**	0,0	0,0	0,1	0,1	0,1	0,1
s_{hht}	**1,1**	**1,1**	0,0	0,0	0,1	0,1	0,1	0,1
s_{hth}	0,0	0,0	0,1	0,1	0,1	0,1	0,1	0,1
s_{htt}	0,0	0,0	0,1	0,1	0,1	0,1	0,1	0,1
s_{thh}	0,1	0,1	0,1	0,1	**1,1**	**1,1**	0,0	0,0
s_{tht}	0,1	0,1	0,1	0,1	**1,1**	**1,1**	0,0	0,0
s_{tth}	0,1	0,1	0,1	0,1	0,0	0,0	0,1	0,1
s_{ttt}	0,1	0,1	0,1	0,1	0,0	0,0	0,1	0,1

Abbildung 2: Strategien und die dazugehörigen Pfade für die CGS aus Abbildung 1. Als Zielformeln werden $\eta_1 \equiv \Box(\neg\mathsf{start} \rightarrow \mathsf{money}_1)$ und $\eta_2 \equiv \Diamond\mathsf{money}_2$ gewählt. Pareto-optimale Profile sind fett und Nash Gleichgewichte grau hinterlegt dargestellt. Dabei ist $s_{x_0x_1x_2}$ die Strategie, in der der Spieler Aktion $x_i \in \{h, t\}$ in Zustand q_i spielt, für $i = 0, 1, 2$ und n in q_j für $j = 3, 4, 5$.

In meiner Dissertation werden auch syntaktische Fragmente und eingeschränkte Klassen von Modellen identifiziert, die ein handhabbares Modellverifikationsproblem besitzen. Es wird insbesondere gezeigt, dass Modellverifikation in polynomieller, deterministischer Zeit möglich ist, wenn die Menge der plausiblen Strategien eine gewisse Struktur hat, nämlich wenn diese rechteckig ist. Letzteres bedeutet, dass die Plausibilitätsmenge gegen Kombinationen von Teilstrategien abgeschlossen ist.

Unvollständige Informationen und Ungewissheit Die Logik **ATLP** erlaubt es, Aussagen über rationales Verhalten von Agenten mit perfekten Informationen zu treffen. In der

Dissertation wird **ATLP** daher erweitert, um auch Agenten mit unvollständigen Informationen modellieren zu können. Die resultierende Logik **CSLP** [BJ09a] beinhaltet nun auch epistemische Operatoren. Außerdem erlaubt das Plausibilitätskonzept eine nicht klassische Unterscheidung von epistemischen und doxastischen Eigenschaften. Das Zusammenspiel dieser Konzepte wird analysiert und die **KDT45**-Axiome werden betrachtet.

Einen anders motivierten Zugang zu unvollständigen Informationen verfolgt die Logik **pATL** (*alternating time temporal logic with probabilistic success*) [BJ09b]. Die Idee lässt sich am besten anhand einer anderen aber äquivalenten Interpretation der **ATL** Formel $\langle\langle A\rangle\rangle\gamma$ verdeutlichen: A hat eine Gewinnstrategie, die gegen die für A *ungünstigste* Strategie von A's Opponenten erfolgreich ist. Dies bedeutet implizit, dass die Opponenten die Fähigkeit haben müssen, diese ungünstigste Strategie herauszufinden. Dafür müssen aber z. B. entsprechende Kommunikationswege vorhanden sein. Diese Annahme ist nicht immer gerechtfertigt. Darüber hinaus ist die Semantik von **ATL** qualitativ: Besteht nur eine Möglichkeit, die Formel falsch zu machen, egal wie unwahrscheinlich es ist, diese Möglichkeit zu finden, dann ist die Formel falsch. In vielen Szenarien (z.B. in diversen nicht sicherheitskritischen Systemen) ist es jedoch wünschenswert, Erfolg zu einer gewissen Wahrscheinlichkeit zu garantieren. Betrachten wir etwa den Fall, in dem die Opponenten n Strategien haben, von denen nur eine schlecht für A ist, also die betrachtete Formel γ falsch machen würde. Es gilt $\neg\langle\langle A\rangle\rangle\gamma$. Weiterhin sei es der Fall, das jeder Opponent seine (individuelle) Strategie unabhängig von den anderen Opponenten wählt, z. B. aufgrund fehlender Kommunikationsmöglichkeiten. In diesem Fall ist γ jedoch mit einer Wahrscheinlichkeit von $1-\frac{1}{n}$ wahr.

Dieses Beispiel dient als Motivation für **pATL**. Kooperationsmodalitäten werden zu $\langle\langle A\rangle\rangle^{p}_{\omega}\gamma$ erweitert. Sie drücken aus, dass die Gruppe A eine Strategie hat, γ mit einer Wahrscheinlichkeit von $p \in [0,1]$ zu garantieren, wenn das erwartete Verhalten der Gegner durch eine Wahrscheinlichkeitsverteilung ω modelliert wird.

Die generische Logik $\mathbf{pATL}_{\mathcal{BH}}$ erlaubt eine beliebige probabilistische Voraussage über das Verhalten der Gegner. Meine Dissertation befasst sich genauer mit zwei Instanzierungen: gemischte Strategien (hierzu wird die Logik $\mathbf{pATL}_{MMS}$ eingeführt) und Verhaltensstrategien (*behavioural strategies*) (dies führt zur Logik $\mathbf{pATL}_{BS}$). Es wird unter anderem gezeigt, dass $\mathbf{pATL}_{MMS}$ **ATL** verallgemeinert: $\mathfrak{M}, q \models_{\mathbf{ATL}} \langle\langle A\rangle\rangle\gamma$ gdw. $\mathfrak{M}, q \models_{\mathbf{pATL}_{MMS}} \langle\langle A\rangle\rangle^{1}_{\omega}\gamma$, dies jedoch im Allgemeinen nicht für $\mathbf{pATL}_{BS}$ gilt. Überraschend sind insbesondere die Komplexitäten der Modellverifikationsprobleme beider Logiken. Obwohl die Theorie hinter $\mathbf{pATL}_{BS}$ sehr viel komplexer als die hinter $\mathbf{pATL}_{MMS}$ ist – die Semantik von der ersten Logik wird über kontinuierlichen Maßintegralen definiert und die der zweiten „nur" über einer endlichen Summe – verhält es sich mit der Schwierigkeit der Verifikationsprobleme entgegengesetzt. Das Modellverifikationsproblem für $\mathbf{pATL}_{MMS}$ ist $\mathbf{PP}$-hart und in $\mathbf{PSPACE}$ und für $\mathbf{pATL}_{BS}$ ist das Problem nur $\mathbf{P}$-vollständig [BJ09b].

Agenten mit beschränkten Ressourcen Ein anderer Aspekt rationalen Verhaltens, der in meiner Dissertation betrachtet wird, ist der Effekt von Ressourcen auf die Fähigkeiten von Agenten. In **ATL** wird angenommen, dass Agenten alle ihre zur Verfügung stehenden Aktionen (immer) ausführen können. Es werden Erweiterungen von **ATL** betrachtet, in denen Aktionen Ressourcen konsumieren und produzieren können. Insbesondere die Produktion von Ressourcen stellt sich als kostenintensiv heraus. Es wird bewiesen, dass das Modellverifikationsproblem oftmals unentscheidbar ist, wenn der Verbrauch als auch die Produktion von Ressourcen möglich ist.

Zunächst wird die Auswirkung der Hinzunahme von Ressourcen zu **CTL**, also dem 1-Agenten-Fall, untersucht. Es wird gezeigt, wie ressourcenbeschränkte Systeme modelliert und durch die Logik **RTL** analysiert werden können. Das Gegenstück der **CTL**-Formel $\mathsf{E}\gamma$ (es gibt eine Berechnung, entlang welcher γ gilt) ist die Formel $\langle\rho\rangle\gamma$: Für eine gegebene Menge ρ an verfügbaren Ressourcen gibt es eine Berechnung, die γ erfüllt und die mit den gegebenen Ressourcen realisierbar ist. Überdies werden verschiedene Sprachen unterschiedlicher Ausdrucksstärke vorgestellt und untersucht. Das Hauptresultat, welches Ergebnisse über (entscheidbare) Petri-Netz-Probleme nutzt, zeigt, dass das Modellverifikationsproblem für **RTL** entscheidbar ist [BF10a].

Aufbauend auf dem 1-Agenten-Fall werden ressourcenbeschränkte *Multi*agentensysteme betrachtet. Obwohl die Grundidee dieselbe ist, stellt sich dieses Szenario als weitaus komplexer dar. In meiner Arbeit wird analysiert, *was* in ressourcenbeschränkten Systemen prinzipiell verifizierbar ist (also, was entscheidbar ist). Handhabbare Teilsysteme sollen in zukünftiger Forschung identifiziert werden. Im Folgenden wird kurz auf die verschiedenen Logiken und deren Eigenschaften eingegangen. Zunächst werden die drei Logiken **RAL**, **RAL**$^+$ und **RAL***, die mit den drei bekannten Logiken **ATL**, **ATL**$^+$ und **ATL*** korrespondieren, definiert. **RAL** ist die eingeschränkteste Logik und **RAL*** die ausdrucksstärkste. Es werden wieder speicherlose (durch ein r gekennzeichnet) und speichernutzende Strategien (durch ein R gekennzeichnet) betrachtet. Das erste Ergebnis ist ein negatives: Das Modellverfikationsproblem ist sogar schon für die ausdrucksschwächste Sprache mit speicherlosen Strategien (d.h. für **RAL**$_r$) und einem einzigen Agenten unentscheidbar [BF10b]!

	$\mathcal{L}_{RAL^*}$	$\mathcal{L}_{RAL^+}$	$\mathcal{L}_{RAL}$	$pr\text{-}\mathcal{L}_{RAL^*}$	$pr\text{-}\mathcal{L}_{RAL^+}$	$pr\text{-}\mathcal{L}_{RAL}$
$\models_R$	U^1	U^1	U^1	U^1	U^1	U^1
$\models_r$	U^1	U^1	U^1	U^2	U^2	U^2
rf+$\models_R$ / $\models_R^\infty$	U^2	U^2	U^2	U^2/ U^2_∞	**?** / U^2_∞	**?** / U^2_∞
rf+$\models_r$	**?**	**?**	**?**	**?**	**?**	**?**
$\models_R^k$, $\models_r^k$	**D**	**D**	**D**	**D**	**D**	**D**

Abbildung 3: Übersicht über die Komplexitäten der Modellverifikationsprobleme. Jede Zelle entspricht der Logik, die sich aus Sprache und Semantik ergibt. Der Inhalt einer Zelle gibt an, ob das Problem entscheidbar (D) oder unentscheidbar (U^x) ist. x gibt die Anzahl der notwendigen Agenten an.

Zum Beweis der Unentscheidbarkeit wird das Halteproblem von 2-Zählerautomaten auf das entsprechende Modellverifikationsproblem reduziert. Dabei simulieren zwei verschiedene Ressourcentypen die Zählerstände vom Automaten, und ein Agent (der als Gegner fungiert) simuliert die Berechnungen des Automaten.

Ist also nun alles andere auch unentscheidbar? Ohne weitere restriktive Annahmen ist dies offensichtlich der Fall. Aus diesem Grund werden zwei weitere einschränkende Annahmen gemacht. In den *ressourcenflachen* Varianten (kurz *rf*) können geschachtelte Kooperationsoperatoren nicht dynamisch auf die noch vorhanden Ressourcen zugreifen. Bei jedem Vorkommen eines Operators $\langle\langle A\rangle\rangle$ muss die Ressourcenverteilung explizit angegeben werden. In der zweiten Einschränkung, der *Proponenten-Restriktivität* (kurz *pr*), verbrauchen nur die Proponenten Ressourcen (d.h., die Agenten, die in den Kooperationsmodalitäten auftauchen). Die Opponenten sind diesbezüglich nicht eingeschränkt.

Weiterhin wird gezeigt, dass das Problem selbst im *pr*-Kontext für einen einzigen Agenten unentscheidbar bleibt, wenn speichernutzende Strategien verwendet werden. Ähnlich verhält es sich mit den *rf*-Varianten, in der Reduktion werden jedoch zwei Agenten genutzt. Auch für die Kombination von beiden Einschränkungen (d.h. im Fall von rf-pr-$\mathbf{RAL}^*_R$) bleibt das Problem im Allgemeinen unentscheidbar. Es werden allerdings zwei Agenten und eine ausdrucksstärkere Sprache benötigt (oder eine etwas stärkere Semantik, auf die hier aber nicht weiter eingegangen werden kann).

Eine Übersicht über alle Resultate ist in Abbildung 3 dargestellt. Wie die Tabelle zeigt, sind noch nicht alle Kombinationen vollständig untersucht. In meiner Arbeit wurden auch *entscheidbare* Szenarien betrachtet, in dem die Menge der Ressourcen beschränkt oder eine beschränkte Semantik genutzt worden ist.

4 Schlussfolgerungen

In meiner Dissertation wurden formale Ansätze zur Modellierung von rationalen Agenten vorgestellt. Die meisten von ihnen basieren auf der strategischen Logik **ATL** [AHK02].

Ein großer Teil meiner Arbeit befasste sich mit rationalem Verhalten aus einer spieltheoretischen Perspektive. Dazu wurden zunächst die bereits vorhandenen Verbindungen zwischen strategischen Logiken und Spielen vorgestellt und erweitert. Daraus resultierte die flexible und ausdrucksstarke Logik **ATLP**, welche vorhandene Logiken verallgemeinert und es erlaubt, spieltheoretische Lösungskonzepte zu charakterisieren und diese auch zu nutzen, um die Effekte von Plausibilitätsannahmen über Multiagentensystemen zu analysieren. Dadurch ist es möglich, die Menge der Strategien einzuschränken und spezifischere Aussagen über Agenten zu machen. Im Einklang mit den Zielen der Arbeit – nämlich der Modellierung und der Verifikation von MAS – wird die Komplexität des Modellverifikationsproblems untersucht.

Ausgehend von **ATLP** wurde eine entsprechende Variante für unvollständige Informationen eingeführt. Die Logik **CSLP** hat ähnliche Eigenschaften und erlaubt Schlussfolgerungen über rationale Agenten mit eingeschränktem Wissen.

In diesem Kontext wurden zwei weitere Logiken untersucht. **CoalATL** kombiniert **ATL** mit Argumentationstheorie und erlaubt die Untersuchung von rationalen Koalitionen und ihrer Stabilität. **pATL** unterliegt der Annahme, dass es den Gegnern nicht immer möglich ist, uneingeschränkt zu kommunizieren, um so die für die Proponenten ungünstigste Gegenstrategie zu identifizieren. In der Logik wird das Verhalten der Gegner durch wahrscheinlichkeitstheoretische Vorhersagen modelliert.

Schließen wurden Ressourcen und deren Auswirkungen auf rationales Verhalten betrachtet. Dies geschah in einem temporalen und in einem Multiagentenkontext. Es wurde gezeigt, dass die Verifikation von ressourcenbeschränkten Agenten im Allgemeinen sehr viel schwerer (größtenteils unentscheidbar) ist.

Literatur

[AHK02] R. Alur, T. A. Henzinger und O. Kupferman. Alternating-Time Temporal Logic. *Journal of the ACM*, 49:672–713, 2002.

[BCD08] Nils Bulling, Carlos Chesnevar und Jürgen Dix. Modelling Coalitions: ATL + Argumentation. In Lin Padgham und David Parkes, Hrsg., *Proceedings of the Seventh International Joint Conference on Autonomous Agents and Multi-Agent Systems*, Seiten 681–688, Estoril, Portugal, May 2008. ACM Press.

[BF10a] Nils Bulling und Berndt Farwer. Expressing Properties of Resource-Bounded Systems: The Logics RBTL and RBTL*. In J. Dix, M. Fisher und P. Novak, Hrsg., *Post-Proceedings of CLIMA '09*, number 6214 in LNCS 6214, Seiten 22–45, Hamburg, Germany, September 2010.

[BF10b] Nils Bulling und Berndt Farwer. On the (Un-)Decidability of Model-Checking Resource-Bounded Agents. In *Proceedings of the 19th European Conference on AI (ECAI 2010)*, Seiten 567–572, Lisbon, Portugal, August 16-20 2010.

[BJ09a] Nils Bulling und Wojciech Jamroga. Rational Play and Rational Beliefs under Uncertainty. In *Proc. of AAMAS'09*, p. 257–264, Hungary, May 2009. ACM Press.

[BJ09b] Nils Bulling und Wojtek Jamroga. What Agents Can Probably Enforce. *Fundamenta Informaticae*, 93:81–96, 2009.

[BJ10] Nils Bulling und Wojciech Jamroga. Verifying Agents with Memory is Harder than It Seemed. *AI Communications*, 23(4):389–403, December 2010.

[BJD09] Nils Bulling, Wojtek Jamroga und Jürgen Dix. Reasoning about Temporal Properties of Rational Play. *Annals of Mathematics and AI*, 53(1-4):51–114, 2009.

[CES86] E.M. Clarke, E.A. Emerson und A.P. Sistla. Automatic Verification of Finite-State Concurrent Systems Using Temporal Logic Specifications. *ACM Transactions on Programming Languages and Systems*, 8(2):244–263, 1986.

[Sch04] P. Y. Schobbens. Alternating-Time Logic with Imperfect Recall. *Electronic Notes in Theoretical Computer Science*, 85(2):82–93, 2004.

[vdHW02] W. van der Hoek und M. Wooldridge. Tractable Multiagent Planning for Epistemic Goals. In C. Castelfranchi und W.L. Johnson, Hrsg., *Proceedings of the First International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS-02)*, Seiten 1167–1174. ACM Press, New York, 2002.

[vOvdHW03] S. van Otterloo, W. van der Hoek und M. Wooldridge. Knowledge as Strategic Ability. *Electronic Lecture Notes in Theoretical Computer Science*, 85(2), 2003.

[Wei99] G. Weiss, Hrsg. *Multiagent Systems. A Modern Approach to Distributed Artificial Intelligence*. MIT Press: Cambridge, Mass, 1999.

[Woo02] M. Wooldridge. *An Introduction to Multi Agent Systems*. John Wiley & Sons, 2002.

Dr. Nils Bulling wurde 1981 in Hamburg geboren. Nach seinem Informatikstudium an der TU Clausthal war er dort als wissenschaftlicher Mitarbeiter beschäftigt und promovierte in 2010 bei Prof. Dr. J. Dix. Im Januar 2011 wurde er zum Akademischen Rat a. Z. an der TU Clausthal ernannt. Seine Forschungsgebiete sind in Bereichen der Logik, Multiagentensysteme und Verifikation angesiedelt. In den letzten Jahren publizierte er zu diesen Themen regelmäßig Artikel in Journalen, auf Konferenzen und Workshops. Für viele Zeitschriften, internationale Konferenzen und Workshops fungierte er als Gutachter und war Mitglied bei diversen Programmkomitees von internationalen Veranstaltungen. Er hielt regelmäßig Vorträge auf Konferenzen, wurde zu diversen Dagstuhlseminaren eingeladen und gab Kurse auf internationalen Summerschools.