Die Ausdrucksstärke der Logik erster Stufe mit eingebauten Prädikaten

Nicole Schweikardt
nisch@informatik.uni-mainz.de

Abstract: Diese Arbeit ist in der Theoretischen Informatik positioniert, und zwar in den Fachgebieten Komplexitätstheorie, mathematische Logik und Datenbanktheorie. Die Arbeit beschäftigt sich mit der Ausdrucksstärke der Logik erster Stufe auf Strukturen mit eingebauten Prädikaten wie z.B. lineare Ordnung, Addition und Multiplikation. Die Hauptergebnisse lassen sich den drei Teilbereichen „Arithmetik und Zählquantoren", „die Crane Beach-Vermutung" und „Kollaps-Resultate in der Datenbanktheorie" zuordnen. Ziel des hier vorliegenden Artikels ist, einen Einblick in die Fragestellungen und Ergebnisse zu diesen drei Themenkreisen zu geben.

1 Einleitung

In der *Komplexitätstheorie* wird die Schwierigkeit eines Problems üblicherweise durch den Zeit- oder Platzbedarf gemessen, der benötigt wird um das Problem auf einer idealisierten Rechenmaschine, z.B. der Turingmaschine, zu lösen. Fagins bahnbrechende Arbeit [Fa74] hat diese *Berechnungs*komplexität mit der *Beschreibungs*komplexität in Verbindung gebracht, d.h., mit der Komplexität bzw. der Reichhaltigkeit einer formalen Sprache, in der das jeweilige Problem beschrieben werden kann. Mittlerweile ist es gelungen, die meisten Komplexitätsklassen auf solch deskriptive Art zu charakterisieren, und zwar durch Beschreibungssprachen, die Erweiterungen der *Logik erster Stufe* sind. So hat zum Beispiel Fagin gezeigt, dass ein Problem genau dann in Polynomialzeit auf einer nichtdeterministischen Turingmaschine lösbar ist, wenn es durch eine Formel der *existentiellen Logik zweiter Stufe* beschrieben werden kann (kurz: $NP = \Sigma^1_1$). Immerman und Vardi zeigten auf, dass (auf geordneten endlichen Strukturen) die Klasse P aller auf einer deterministischen Turingmaschine in Polynomialzeit lösbaren Proleme genau die Klasse aller durch Formeln der *kleinsten Fixpunkt Logik* beschreibbaren Probleme ist (kurz: $P = FO(LFP)$). Ein weiteres Beispiel einer deskriptiven Charakterisierung einer Komplexitätsklasse ist Immermans Resultat, dass die Klasse $LOGSPACE$ aller auf logarithmischem Platz auf einer deterministischen Turingmaschine lösbaren Probleme durch Formeln der *deterministischen Transitive Closure Logik* beschrieben werden kann (kurz: $LOGSPACE = FO(DTC)$). Einen umfassenden Überblick über solche deskriptiven Charakterisierungen von Komplexitätsklassen gibt z.B. das Lehrbuch [Im99].

Eine Gemeinsamkeit all dieser deskriptiven Charakterisierungen von Komplexitätsklassen ist, dass sie bestimmte Erweiterungen der Logik erster Stufe benutzen, die ausdrucksstark

genug sind um zumindest (a) zu zählen, wie viele Elemente eine gegebene Menge hat und (b) arithmetische Operationen wie Addition und Multiplikation auszuführen. Diese Fähigkeiten des Zählens und des Ausführens von arithmetischen Operationen helfen der jeweiligen Logik dabei, Berechnungen zu beschreiben, die von Maschinen mit bestimmten Beschränkungen der Zeit- oder Platzressourcen durchgeführt werden können.
Andererseits sieht man leicht, dass die pure Logik erster Stufe für sich genommen auf geordneten endlichen Strukturen viel zu ausdrucksschwach ist um „richtige“ Berechnungen zu beschreiben. Hier stellt sich auf natürliche Weise die folgende Frage: *Was passiert wenn man die Ausdrucksstärke der Logik erster Stufe dadurch erweitert, dass man (a) die Fähigkeit des Zählens oder (b) bestimmte arithmetische Prädikate hinzufügt?*
Im vorliegenden Artikel wird untersucht, unter welchen Umständen solche Erweiterungen tatsächlich die Ausdrucksstärke der Logik erster Stufe vergrößern, und umgekehrt, für welche Art von Problemen solche Erweiterungen der Logik erster Stufe keine zusätzliche Ausdruckskraft verleihen. Die hier vorgestellten Resultate können in die folgenden drei Themenbereiche unterteilt werden:

1. Pure Arithmetik und Zählquantoren:
Hier werden rein arithmetische Strukturen, beispielsweise die natürlichen Zahlen mit linearer Ordnung und Addition, betrachtet. Die grundlegende Frage ist: *Vergrößern zusätzliche Zählquantoren tatsächlich die Ausdrucksstärke der Logik erster Stufe auf rein arithmetischen Strukturen?* Aus der in Kapitel 3 vorgestellten Antwort auf diese Frage folgt u.a. ein einfacher Beweis des Resultats von Ruhl, dass Erreichbarkeit und Zusammenhang von endlichen Graphen nicht in der Logik erster Stufe mit unären Zählquantoren und eingebauter Addition ausdrückbar sind.

2. Die Crane Beach-Vermutung:
Hier werden Wortsprachen betrachtet, die einen *neutralen Buchstaben* haben, d.h., einen Buchstaben, der in jedem Wort eingefügt oder gelöscht werden kann, ohne dessen Zugehörigkeit zur Sprache zu ändern. Die grundlegende Frage ist: *Vergrößern zusätzliche arithmetische Prädikate tatsächlich die Fähigkeit der Logik erster Stufe, Sprachen mit neutralem Buchstaben zu beschreiben?* Diese Frage steht in engem Zusammenhang zu Uniformitätsmaßen für die Schaltkreiskomplexitätsklasse AC^0: Die Antwort auf obige Frage ist genau dann „ja“, wenn eine bestimmte *uniforme* Version der Schaltkreiskomplexitätsklasse AC^0 Sprachen mit neutralem Buchstaben enthält, die nicht sternfrei-regulär sind. Thériens Vermutung, dass obige Frage stets mit „nein“ beantwortet werden kann, wurde unter dem Namen *Crane Beach-Vermutung* bekannt. In Kapitel 4 werden sowohl positive als auch negative Instanzen der Crane Beach-Vermutung vorgestellt.

3. Kollaps-Resultate in der Datenbanktheorie:
Hier werden Datenbanken betrachtet, die in eine *Kontextstruktur* eingebettet sind, die aus einer unendlichen Menge von potentiellen Datenbankelementen sowie aus einer Reihe von eingebauten arithmetischen Prädikaten besteht. In der vorliegenden Arbeit werden so genannte $<$*-generische* Anfragen betrachtet, d.h. Datenbankanfragen, deren Ergebnis nur von der relativen Ordnung der einzelnen Datenbankelemente untereinander abhängt, nicht aber von deren Konstellation in Bezug auf die restlichen eingebauten Prädikate. Die grund-

legende Frage ist: *Vergrößern zusätzliche eingebaute Prädikate tatsächlich die Fähigkeit der Logik erster Stufe, $<$-generische Anfragen auszudrücken?* Man spricht von einem *Kollaps-Resultat*, wenn die Antwort auf obige Frage „nein“ ist. Die meisten bisher bekannten Kollaps-Resultate für $<$-generische Datenbankanfragen beschränkten sich auf die Klasse der *endlichen* Datenbanken. Nun kann allerdings eine Datenbankrelation, die *unendlich* viele Tupel enthält, duch einen Algorithmus repräsentiert werden, der bei Eingabe eines beliebigen Tupels ermittelt, ob das Tupel zur Datenbankrelation gehört oder nicht. In Kapitel 5 werden, mit der Methode des Ehrenfeucht-Fraïssé-Spiels, Kollaps-Resultate auch für solche *unendlichen* Datenbanken gezeigt.

Ziel des vorliegenden Artikels ist, einen Überblick über die Fragestellungen und die wichtigsten Ergebnisse der Dissertation [Sc01] zu geben. Der Rest dieses Artikels ist folgendermaßen aufgebaut: In Kapitel 2 werden einige grundlegende Begriffe erklärt. Kapitel 3 beschäftigt sich mit purer Arithmetik und Zählquantoren. Kapitel 4 handelt von der Crane Beach-Vermutung. In Kapitel 5 geht es um Kollaps-Resultate in der Datenbanktheorie.

Danksagung:
Ich danke meinem Doktorvater Clemens Lautemann für seine hervorragende Anleitung und für viele wertvolle Hinweise und Diskussionen bezüglich meines Forschungsthemas. Vielen Dank auch an Thomas Schwentick für hilfreiche Ratschläge besonders zu Beginn meiner Forschungstätigkeit; insbesondere hat er mich auf Kollaps-Resultate in der Datenbanktheorie aufmerksam gemacht. Dem Institut für Informatik und dem Fachbereich Mathematik und Informatik der Johannes Gutenberg-Universität Mainz danke ich für die guten Arbeitsbedingungen und das freundliche Arbeitsklima.

2 Grundlegende Notationen

$\mathbb{Z}$, $\mathbb{Q}$ und $\mathbb{R}$ bezeichnen die Mengen der ganzen Zahlen, der rationalen Zahlen und der reellen Zahlen. $\mathbb{N} := \{0,1,2,..\}$ ist die Menge der natürlichen Zahlen und $\mathbb{N}_{>0}$ die Menge der positiven natürlichen Zahlen. Für eine Menge A und ein festes $m \in \mathbb{N}_{>0}$ ist $A^m := \{(a_1,..,a_m) : a_1,..,a_m \in A\}$ die Menge aller m-Tupel über A. Eine m-stellige *Relation* über A ist eine Teilmenge von A^m.

Strukturen

Eine *Signatur* τ besteht aus Konstantensymbolen und Relationssymbolen; jedes Relationssymbol $R \in \tau$ hat eine feste Stelligkeit $ar(R) \in \mathbb{N}_{>0}$.
Eine τ-*Struktur* $\mathcal{A} = \langle A, \tau^{\mathcal{A}} \rangle$ besteht aus einer beliebigen Menge A, die als das *Universum* von $\mathcal{A}$ bezeichnet wird, und einer Liste $\tau^{\mathcal{A}}$, die für jedes Konstantensymbol $c \in \tau$ ein Element $c^{\mathcal{A}} \in A$ und für jedes Relationssymbol $R \in \tau$ eine Relation $R^{\mathcal{A}} \subseteq A^{ar(R)}$ enthält.

2.1 Beispiele. (a) Sei $\tau_D := \{D\}$ die Signatur, die aus einem einzigen zweistelligen Relationssymbol D besteht. Das Liniennetz der Deutschen Bahn kann man als folgende

τ_D-Struktur $\mathcal{A}_{Bahn}$ auffassen: Das Universum von $\mathcal{A}_{Bahn}$ besteht aus allen Städtenamen, und die Relation $D^{\mathcal{A}_{Bahn}}$ besteht aus allen Städtetupeln (s_1, s_2), für die es eine Direktverbindung von Stadt s_1 zu Stadt s_2 gibt.

(b) Sei $\tau_{\mathcal{N}} := \{0, 1, <, +\}$ die Signatur, die aus den zwei Konstantensymbolen 0 und 1, dem zweistelligen Relationssymbol $<$ und dem dreistelligen Relationssymbol $+$ besteht. Die $\tau_{\mathcal{N}}$-Struktur $\mathcal{N} := \langle \mathbb{N}, 0^{\mathcal{N}}, 1^{\mathcal{N}}, <^{\mathcal{N}}, +^{\mathcal{N}} \rangle$ wird auch als die *Presburger Arithmetik* bezeichnet, wobei $0^{\mathcal{N}}$ und $1^{\mathcal{N}}$ die natürlichen Zahlen 0 und 1 sind, $<^{\mathcal{N}}$ die lineare Ordnung auf $\mathbb{N}$ ist und $+^{\mathcal{N}}$ aus allen Tripeln $(a, b, c) \in \mathbb{N}^3$ besteht, für die $a + b = c$ gilt. □

Die Logik erster Stufe

In diesem Abschnitt wird eine sehr knappe Definition der Logik erster Stufe gegeben; eine detaillierte Einführung in die Logik erster Stufe findet sich z.B. in dem Lehrbuch [EFT96]. Sei τ eine Signatur. Wir benutzen $x_1, x_2, ..$ als Variablensymbole. *Atomare τ-Formeln* sind von der Form $y_1{=}y_2$ und $R(y_1, .., y_m)$, wobei $y_1, .., y_m$ Konstantensymbole aus τ oder Variablensymbole sind, R ein Relationssymbol aus τ und m die Stelligkeit von R ist. Eine *τ-Formel der Logik erster Stufe* (kurz: *FO*(τ)-Formel[1]) wird aus den atomaren τ-Formeln mittels der Booleschen Verknüpfungen $\neg$ (Negation), $\wedge$ (logisches „und"), $\vee$ (logisches „oder"), dem Existenzquantor $\exists$ und dem Allquantor $\forall$ aufgebaut. D.h.: Jede atomare τ-Formel ist auch eine *FO*(τ)-Formel, und sind *FO*(τ)-Formeln φ und ψ und ein Variablensymbol y gegeben, so sind auch die folgenden Formeln *FO*(τ)-Formeln: $\neg\varphi$, $\varphi \wedge \psi$, $\varphi \vee \psi$, $\exists y \varphi$ und $\forall y \varphi$.
Die *freien Variablen* einer *FO*(τ)-Formel φ sind die Variablensymbole, die in φ außerhalb des Wirkungsbereichs eines Quantors $\exists$ oder $\forall$ vorkommen. Vereinfacht schreiben wir oft $\varphi(y_1, .., y_\ell)$ um anzuzeigen, dass $y_1, .., y_\ell$ die freien Variablen von φ sind. Sind eine τ-Struktur $\mathcal{A} = \langle A, \tau^{\mathcal{A}} \rangle$ sowie Elemente $a_1, .., a_\ell$ aus dem Universum von $\mathcal{A}$ gegeben, so sagen wir „*$\mathcal{A}$ erfüllt $\varphi(a_1, .., a_\ell)$*" und schreiben „$\mathcal{A} \models \varphi(a_1, .., a_\ell)$", falls die Formel φ wahr wird, wenn man jedes Symbol aus τ mit seiner Belegung aus $\tau^{\mathcal{A}}$ und jedes freie Vorkommen der Variablen $y_1, .., y_\ell$ mit den Elementen $a_1, .., a_\ell$ interpretiert.

2.2 Beispiel. Sei $\tau_D = \{D\}$ die Signatur aus Beispiel 2.1.(a), sei $\mathcal{A}_{Bahn}$ die Struktur, die das Liniennetz der Deutschen Bahn repräsentiert, und sei $\varphi_{1\times Umsteigen}(x, y)$ die durch $D(x, y) \vee \exists z\, D(x, z) \wedge D(z, y)$ gegebene *FO*-Formel. Für beliebige Städte s_1 und s_2 gilt: $\mathcal{A}_{Bahn} \models \varphi_{1\times Umsteigen}(s_1, s_2)$ genau dann, wenn es möglich ist, von s_1 nach s_2 zu reisen und dabei höchstens einmal umzusteigen. □

Eingebaute Prädikate

In Formeln der Logik erster Stufe kann ein *eingebautes Prädikat* benutzt werden als ein k-stelliges Relationssymbol, das eine feste Interpretation hat, die unabhängig von der jeweils betrachteten konkreten Struktur ist, in der die Formel ausgewertet wird.

[1] *FO* steht für die englische Bezeichnung „*first-order logic*" der Logik erster Stufe

2.3 Definition. *Sei $k \in \mathbb{N}_{>0}$.*
Ein k-stelliges eingebautes Prädikat *P auf einem Universum $\mathbb{U}$ ist eine feste k-stellige Relation über $\mathbb{U}$, d.h., $P \subseteq \mathbb{U}^k$.*
Ein k-stelliges eingebautes Prädikat P auf Anfangsstücken von $\mathbb{N}$ *hat für jedes $N \in \mathbb{N}$ eine feste Interpretation $P^N \subseteq \{1, .., N\}^k$.* ☐

Typische Beispiele für eingebaute Prädikate auf dem Universum $\mathbb{N}$ sind das zweistellige Prädikat $<$, das alle Tupel (a, b) mit $a < b$ enthält, die dreistelligen Prädikate $+$ und $\times$, die alle Tripel (a, b, c) mit $a+b = c$ bzw. $a \cdot b = c$ enthalten und das zweistellige Prädikat *Bit*, das aus allen Tupeln (a, b) besteht, für die gilt, dass das b-te Bit der Binärdarstellung von a eine 1 ist. Auf *Anfangsstücken von* $\mathbb{N}$ werden diese Prädikate entsprechend eingeschränkt, so dass z.B. $+^N$ aus allen Tripeln $(a, b, c) \in \{1, .., N\}^3$ mit $a + b = c$ besteht.

3 Arithmetik und Zählquantoren

In diesem Kapitel werden die rein arithmetischen Strukturen $\langle \mathbb{N}, < \rangle$, $\langle \mathbb{N}, <, + \rangle$ und $\langle \mathbb{N}, <, +, \times \rangle$ betrachtet. Wir erweitern die Logik erster Stufe um so genannte *Zählquantoren*:
Sei τ eine Signatur mit $\tau \subseteq \{<, +, \times\}$. Die τ-Formeln der *Logik erster Stufe mit unären Zählquantoren* (kurz: *FOunC*(τ)-Formeln) erhält man genauso wie die *FO*(τ)-Formeln, mit dem Zusatz, dass außer den Quantoren $\exists$ und $\forall$ auch noch unäre Zählquantoren der Form $\exists^{=x} y$ zugelassen sind. D.h.: Ist φ eine *FOunC*(τ)-Formel mit freien Variablen $y, z_1, .., z_\ell$, so ist $\exists^{=x} y\, \varphi$ eine *FOunC*(τ)-Formel mit den freien Variablen $x, z_1, .., z_\ell$. Für natürliche Zahlen $a, d_1, .., d_\ell$ gilt: $\langle \mathbb{N}, \tau^{\mathbb{N}} \rangle \models \big(\exists^{=x} y\, \varphi\big)(a, d_1, .., d_\ell)$ genau dann, wenn es genau a verschiedene natürliche Zahlen b gibt, so dass $\langle \mathbb{N}, \tau^{\mathbb{N}} \rangle \models \varphi(b, d_1, .., d_\ell)$.
Analog kann man für jede Zahl $k \in \mathbb{N}_{>0}$ die Klasse der *FOk-aryC*(τ)-Formeln definieren, indem man Zählquantoren der Form $\exists^{=x} y_1, .., y_k\, \varphi$ zulässt, die erlauben, die Anzahl der k-Tupel zu zählen, für die die Formel φ wahr wird.
Hinsichtlich der Ausdrucksstärke dieser Logiken ergibt sich in [Sc01] folgendes Bild:

$$
\begin{array}{ccccccc}
\mathit{FO}(<,+,\times) & = & \mathit{FOunC}(<,+,\times) & = & \mathit{FOk\text{-}aryC}(<,+) & & \forall\, k \geqslant 2 \\
| & & & & & & \\
\mathit{FO}(<,+) & = & \mathit{FOunC}(<,+) & = & \mathit{FOunC}(<) & & \\
| & & & & & & \\
\mathit{FO}(<) & & & & & &
\end{array}
$$

Abbildung 1: Die Ausdrucksstärke der Logik erster Stufe mit und ohne Zählquantoren auf rein arithmetischen Strukturen mit Universum $\mathbb{N}$. Eine senkrechte Linie bedeutet, dass die obere Formelklasse eine größere Ausdrucksstärke besitzt als die untere Formelklasse.

Das darunter wohl interessanteste neue Resultat ist, dass die Logik erster Stufe über $\langle \mathbb{N}, <, + \rangle$ ausdrucksstark genug ist, um sämtliche unären Zählquantoren zu simulieren, d.h.:

3.1 Theorem (*FOunC*$(<,+) =$ *FO*$(<,+)$ auf $\mathbb{N}$).
Zu jeder FOunC$(<,+)$*-Formel* φ *mit freien Variablen* $y_1,..,y_\ell$ *gibt es eine FO*$(<,+)$*-Formel* ψ *mit freien Variablen* $y_1,..,y_\ell$*, so dass für alle natürlichen Zahlen* $a_1,..,a_\ell$ *gilt:* $\langle\mathbb{N},<,+\rangle \models \varphi(a_1,..,a_\ell)$ *genau dann, wenn* $\langle\mathbb{N},<,+\rangle \models \psi(a_1,..,a_\ell)$. □

Dieses Resultat lässt sich auch auf Anfangsstücke von $\mathbb{N}$ übertragen, das heißt: Zu jeder *FOunC*$(<,+)$-Formel $\varphi(y_1,..,y_\ell)$ gibt es eine *FO*$(<,+)$-Formel $\psi(y_1,..,y_\ell)$, so dass für alle natürlichen Zahlen N und alle Zahlen $a_1,..,a_\ell \in \{1,..,N\}$ gilt:
$\langle\{1,..,N\},<,+\rangle \models \varphi(a_1,..,a_\ell)$ genau dann, wenn $\langle\{1,..,N\},<,+\rangle \models \psi(a_1,..,a_\ell)$.

Unter Anwendung von Theorem 3.1 erhält man (a) einen einfachen Beweis des Resultats von Ruhl [Ru99], dass Erreichbarkeit und Zusammenhang von endlichen Graphen nicht in der Logik erster Stufe mit unären Zählquantoren und eingebauter Addition ausdrückbar sind, und (b) einen Beweis dafür, dass die *Crane Beach-Vermutung* u.a. für *FOunC*$(<,+,\times)$ falsch ist.

4 Die Crane Beach-Vermutung

Die Crane Beach-Vermutung ist nach dem Ort benannt, an dem sie erstmals formuliert (und teilweise auch bewiesen) wurde, und zwar Crane Beach, St. Philip, Barbados. Um die Vermutung präzise wiedergeben zu können, benötigen wir einige weitere Notationen: Ein *Wort* über einem endlichen Alphabet A ist eine endliche Zeichenkette, die aus Buchstaben aus A gebildet ist. Wir schreiben A^* um die Menge aller solchen Worte zu bezeichnen. Eine *Wortsprache* (oder kurz *Sprache*) ist eine Teilmenge von A^*. Um Worte durch Strukturen zu repräsentieren benutzen wir die Signatur $\{<\}\cup\tau_{\mathsf{A}}$, wobei τ_{A} für jeden Buchstaben $\mathsf{a} \in \mathsf{A}$ ein unäres Relationssymbol Q_{a} enthält. Jedes Wort $w = w_1 \cdots w_N \in \mathsf{A}^*$ der Länge N wird repräsentiert durch die $(\{<\}\cup\tau_{\mathsf{A}})$-Struktur $\langle w,<\rangle := \langle\{1,..,N\},<,\tau_{\mathsf{A}}^w\rangle$, wobei τ_{A}^w die Liste der Relationen $Q_{\mathsf{a}}^w := \{i \in \{1,..,N\} : w_i = \mathsf{a}\}$ ist, für alle $\mathsf{a} \in \mathsf{A}$. Das heißt, das Universum von $\langle w,<\rangle$ besteht aus allen Positionen in w, und die Aussage $Q_{\mathsf{a}}(i)$ ist für eine Position i genau dann wahr, wenn an i-ter Position in w der Buchstabe a steht. Man kann nun *FO*$(<,\tau_{\mathsf{A}})$-Formeln benutzen um Aussagen über Worte aus A^* zu machen. Einer Formel φ, die keine freien Variablen hat, ordnen wir die Wortsprache $L(\varphi)$ zu, die aus denjenigen Worten $w \in \mathsf{A}^*$ besteht, für die gilt $\langle w,<\rangle \models \varphi$; hier sagen wir auch „die Formel φ *beschreibt* die Wortsprache $L(\varphi)$“. Um mehr Wortsprachen durch Formeln beschreiben zu können, kann man zusätzliche eingebaute Prädikate zulassen.

4.1 Beispiel. Die Wortsprache $L_1 := \{\mathsf{a}^n\mathsf{b}^m : n \in \mathbb{N}_{>0}, m \in \mathbb{N}\}$ lässt sich durch die *FO*$(<,\tau_{\mathsf{A}})$-Formel $\varphi_1 := \exists x\, \psi_1(x)$ beschreiben, wobei $\psi_1(x) := \forall y\, (y{<}x \wedge Q_{\mathsf{a}}(y)) \vee (y{=}x \wedge Q_{\mathsf{a}}(y)) \vee (x{<}y \wedge Q_{\mathsf{b}}(y))$.
Die Sprache $L_2 := \{\mathsf{a}^n\mathsf{b}^n : n \in \mathbb{N}_{>0}\}$ lässt sich durch die *FO*$(<,+,\tau_{\mathsf{A}})$-Formel $\varphi_2 := \exists x \Big(\psi_1(x) \wedge \exists z\, (x{+}x{=}z \wedge \forall y\, (y{<}z \vee y{=}z))\Big)$ beschreiben. □

Es ist seit langem bekannt, dass man mit *FO*$(<,\tau_{\mathsf{A}})$-Formeln genau die sternfrei regulären Wortsprachen beschreiben kann und dass die Sprache L_2 aus Beispiel 4.1 *nicht* regulär ist. Beispiel 4.1 zeigt also, dass zusätzliche eingebaute Prädikate die Ausdrucksstärke der Logik erster Stufe wirklich vergrößern können. — Man weiß sogar, dass für jede beliebige Menge $\mathfrak{Präd}$ von eingebauten Prädikaten die $FO(<, \mathfrak{Präd}, \tau_{\mathsf{A}})$-Formeln genau diejenigen Wortsprachen beschreiben können, die zur „$\mathfrak{Präd}$-uniformen“ Version der Schaltkreiskomplexitätsklasse AC^0 gehören. Andererseits ist bekannt, dass die sehr einfache, unter dem Namen PARITY bekannte Sprache $\{w \in \{\mathsf{a}, \mathsf{b}\}^* : \text{die Anzahl der } \mathsf{a}\text{s in } w \text{ ist gerade}\}$ noch nicht einmal zur nicht-uniformen Version von AC^0 gehört und daher auch nicht durch $FO(<, \mathfrak{Präd}, \tau_{\mathsf{A}})$-Formeln beschreibbar ist. Eine Ursache hierfür ist, dass der Buchstabe b ein für die Sprache PARITY *neutraler* Buchstabe ist, d.h., dass man durch Einfügen oder Löschen von bs in einem Wort aus $\{\mathsf{a}, \mathsf{b}\}^*$ nicht dessen Zugehörigkeit bzw. Nichtzugehörigkeit zur Sprache ändern kann. So kann zum Beispiel auch das Leerzeichen als neutraler Buchstabe für LaTeX-Code sowie für die meisten Programmiersprachen aufgefasst werden. In Analogie zur Nicht-Beschreibbarkeit von PARITY formulierte Thérien die folgende Vermutung:

4.2 Definition (Crane Beach-Vermutung).
Sei $\mathfrak{Präd}$ eine Menge von eingebauten Prädikaten und sei F eine Formelklasse (z.B. FO oder FOunC). Die Crane Beach-Vermutung *für $F(<, \mathfrak{Präd})$ ist genau dann wahr, wenn für jedes Alphabet A und jede Sprache $L \subseteq \mathsf{A}^*$, die einen neutralen Buchstaben hat, gilt: Falls L durch eine $F(<, \mathfrak{Präd}, \tau_{\mathsf{A}})$-Formel beschreibbar ist, so kann L auch durch eine $F(<, \tau_{\mathsf{A}})$-Formel beschrieben werden.* □

Übersetzt in die Terminologie der Schaltkreis-Komplexitätstheorie besagt die Crane Beach-Vermutung für $FO(<, \mathfrak{Präd})$, dass alle zur $\mathfrak{Präd}$-uniformen Version von AC^0 gehörigen Sprachen mit neutralem Buchstaben sogar sternfrei regulär sind.

Ein Hauptresultat in [Sc01] ist, dass die Crane Beach-Vermutung für $FO(<, +, Q, \mathfrak{Teil}_Q)$ wahr ist, wobei Q eine bestimmte, unendlich große Menge von natürlichen Zahlen und $\mathfrak{Teil}_Q$ die Liste aller Teilmengen von Q ist. Als weiteres *positives* Resultat wird gezeigt, dass für *jede* Menge $\mathfrak{Präd}$ von eingebauten Prädikaten die Crane Beach-Vermutung für $BC(EFO)(<, \mathfrak{Präd})$ wahr ist, wobei $BC(EFO)$ die Einschränkung der Logik erster Stufe auf Boolesche Kombinationen von rein existentiellen Formeln ist. Der Beweis dieser beiden Resultate wird mit der Methode des *Ehrenfeucht-Fraïssé Spiels* (siehe z.B. [Im99]) geführt, und zwar indem eine Gewinnstrategie für einen der beiden Spieler in einem „einfachen“ Spiel (d.h. in einem Spiel, in dem man die Prädikate aus $\mathfrak{Präd}$ ignorieren kann) übersetzt wird zu einer Gewinnstrategie in einem „schwierigen“ Spiel (in dem auch die Prädikate aus $\mathfrak{Präd}$ berücksichtigt werden müssen).

Andererseits sind auch eine Reihe von *negativen* Instanzen der Crane Beach-Vermutung bekannnt. So konnte Immerman beispielsweise zeigen, dass die Vermutung für $FO(<, +, \times)$ *falsch* ist, d.h., es gibt eine Sprache L mit neutralem Buchstaben, die durch eine $FO(<, +, \times)$-Formel beschrieben werden kann, nicht jedoch durch eine $FO(<)$-Formel. Unter Benutzung von Theorem 3.1 wird in [Sc01] gezeigt, dass die Crane Beach-Vermutung für $FOunC(<, \mathfrak{Präd})$ genau dann *falsch* ist, wenn man mit Hilfe der Prädikate in $\mathfrak{Präd}$ eine Menge von natürlichen Zahlen beschreiben kann, die *nicht* semi-linear ist. Dies hat z.B. zur Folge, dass die Vermutung für $FOunC(<, +, \times)$ und für $FOunC(<, Q)$ falsch ist (wobei Q die oben erwähnte Menge ist, mit der die Vermutung für $FO(<, +, Q, \mathfrak{Teil}_Q)$

wahr ist).

In ihrer ursprünglichen, uneingeschränkten und von Immerman falsifizierten Version besagte die Crane Beach-Vermutung, dass bereits $FO(<)$-Formeln ausreichen um alle Sprachen mit neutralem Buchstaben zu beschreiben, die mit *beliebigen* eingebauten Prädikaten in der Logik erster Stufe beschreibbar sind. Ein Resultat aus [Sc01] zeigt das Ausmaß, in dem diese Version der Crane Beach-Vermutung falsch ist: Es gibt keine *abzählbare* Menge $\mathfrak{Präd}$ von eingebauten Prädikaten und keine abzählbare Erweiterung F der Logik erster Stufe, so dass man durch $F(<, \mathfrak{Präd})$-Formeln alle Sprachen mit neutralem Buchstaben beschreiben kann, die mit *beliebigen* eingebauten Prädikaten in der Logik erster Stufe beschreibbar sind.

5 Kollaps-Resultate in der Datenbanktheorie

Relationale Datenbanken werden in der Datenbanktheorie oft als relationale Strukturen über einem festen, möglicherweise unendlichen Universum $\mathbb{U}$ modelliert (siehe z.B. das Lehrbuch [AHV95]). Eine Datenbank über $\mathbb{U}$ kann also als eine τ-Struktur $\mathcal{A} = \langle \mathbb{U}, \tau^{\mathcal{A}} \rangle$ aufgefasst werden, wobei τ eine Signatur ist, die aus endlich vielen Relationssymbolen besteht. Der *aktive Domain* von $\mathcal{A}$, kurz: $adom(\mathcal{A})$, ist die Menge aller Elemente aus $\mathbb{U}$, die zu mindestens einem Tupel gehören, das in einer der Relationen aus $\tau^{\mathcal{A}}$ liegt. Somit ist $\mathbb{U}$ die Menge aller *potentiellen* Datenbankelemente, während $adom(\mathcal{A})$ die Menge der Elemente ist, die *tatsächlich* in der Datenbank $\mathcal{A}$ vorkommen. Eine Datenbank wird als *endlich* bezeichnet, falls ihr aktiver Domain endlich ist.

Eine *Boolesche Anfrage* Q ist eine Anfrage, die jeder Datenbank $\mathcal{A}$ eine „Antwort“ $Q(\mathcal{A}) \in \{ja, \ nein\}$ zuordnet. Ein Beispiel für eine Boolesche Anfrage ist die Anfrage Q_{PARITY}: „Ist die Anzahl der Elemente im aktiven Domain *gerade*?“. Analog kann man auch k-stellige Anfragen Q betrachten, die jeder Datenbank $\mathcal{A}$ eine k-stellige Relation $Q(\mathcal{A}) \subseteq \mathbb{U}^k$ zuordnen. Der Einfachheit halber werden im Folgenden ausschließlich Boolesche Anfragen betrachtet, alle erwähnten Ergebnisse gelten aber analog auch für k-stellige Anfragen. Als *Anfragesprache* werden wir hier die Logik erster Stufe betrachten, die allgemein als der relationale Kern der Datenbankanfragesprache SQL angesehen wird.

Im Bereich der *Constraint Datenbanken* (siehe [KLP00]) betrachtet man Datenbanken, die in eine *Kontextstruktur* eingebettet sind, die aus dem Universum $\mathbb{U}$ aller potentiellen Datenbankelemente sowie zusätzlichen eingebauten Prädikaten auf $\mathbb{U}$ besteht. Als solche Kontextstrukturen werden z.B. $\langle \mathbb{N}, <, + \rangle$ oder auch $\langle \mathbb{R}, <, +, \times \rangle$ genutzt. Datenbankanfragen, die in der Logik erster Stufe formuliert sind, können nun außer auf die Datenbankrelationen aus τ auch auf die eingebauten Prädikate der Kontextstruktur zugreifen. Die Nutzung solcher eingebauten Prädikate kann die Ausdrucksstärke der Logik erster Stufe enorm erhöhen. Beispielsweise ist bekannt, dass obige Anfrage Q_{PARITY} in der Logik erster Stufe ausgedrückt werden kann, falls die Datenbank in die Kontextstruktur $\langle \mathbb{N}, <, +, \times \rangle$ eingebettet ist, nicht aber falls ausschließlich auf die Datenbankrelationen in τ zugegriffen werden kann.

Zumeist verlangt man, dass eine Anfrage *generisch* ist, d.h., dass sie auf „identischen“

Datenbanken, unabhängig von deren konkreter Repräsentation, auch identische Ergebnisse liefert. Wann genau zwei Datenbanken als identisch angesehen werden, hängt vom jeweiligen Anwendungskontext ab. Im Bereich der *räumlichen Datenbanken*, in denen die Datenbanken in eine linear geordnete Kontextstruktur eingebettet sind, werden zwei Datenbanken $\mathcal{A} = \langle \mathbb{U}, \tau^{\mathcal{A}} \rangle$ und $\mathcal{B} = \langle \mathbb{U}, \tau^{\mathcal{B}} \rangle$ z.B. genau dann als „identisch" angesehen, wenn sie $<$-*isomorph* sind. Falls das Kontextuniversum $\mathbb{U}$ aus der Menge der natürlichen Zahlen besteht, bedeutet dies, dass es eine ordnungserhaltende, bijektive Abbildung von $adom(\mathcal{A})$ nach $adom(\mathcal{B})$ gibt, die die Relationen von $\mathcal{A}$ auf die Relationen von $\mathcal{B}$ abbildet. Falls das Kontextuniversum aus den reellen Zahlen besteht, bedeutet es, dass es eine ordnungserhaltende, bijektive Abbildung von $\mathbb{R}$ nach $\mathbb{R}$ gibt, die die Relationen von $\mathcal{A}$ auf die Relationen von $\mathcal{B}$ abbildet.

Eine Boolesche Anfrage wird nun als $<$-*generisch* bezeichnet, falls sie auf $<$-isomorphen Datenbanken identische Ergebnisse liefert. Beispielsweise ist die oben erwähnte Anfrage Q_{PARITY} $<$-generisch.

Analog zur Crane Beach-Vermutung stellt sich nun folgende Frage: Gegeben eine Kontextstruktur $\langle \mathbb{U}, <, \mathfrak{Präd} \rangle$ und eine Klasse $\mathscr{K}$ von Datenbanken über $\mathbb{U}$, kann $FO(<, \mathfrak{Präd})$ mehr $<$-generische Anfragen beschreiben als $FO(<)$ — oder gibt es umgekehrt zu jeder $FO(<, \mathfrak{Präd}, \tau)$-Formel φ eine $FO(<, \tau)$-Formel ψ, die auf allen Datenbanken aus $\mathscr{K}$ äquivalent zu φ ist? Falls letzteres der Fall ist, so spricht man von einem *Kollaps-Resultat*; genauer sagt man „*die Kontextstruktur* $\langle \mathbb{U}, <, \mathfrak{Präd} \rangle$ *erlaubt den natürlich-generischen Kollaps für FO auf der Datenbankklasse* $\mathscr{K}$". Im Falle der Gültigkeit eines solchen Kollapses könnte man beispielsweise dem Nutzer erlauben, zur einfacheren Formulierung seiner Datenbankanfrage die Prädikate in $\mathfrak{Präd}$ zu Hilfe zu nehmen und dann automatisch diese $FO(<, \mathfrak{Präd})$-Anfrage in eine äquivalente und effizient auszuwertende $FO(<)$-Anfrage umwandeln.

Einen umfassenden Überblick über solche natürlich-generischen Kollaps-Resultate für die Klasse aller *endlichen* Datenbanken findet man z.B. in [KLP00]. Nun kann allerdings eine Datenbankrelation, die *unendlich* viele Tupel enthält, duch einen Algorithmus repräsentiert werden, der bei Eingabe eines beliebigen Tupels ermittelt, ob das Tupel zur Datenbankrelation gehört oder nicht. Daher wäre es wünschenswert, Kollaps-Resultate auch für *unendliche* Datenbanken zu erzielen. Die meisten bisher bekannten Methoden zum Nachweis von Kollaps-Resultaten nutzen jedoch ganz wesentlich die Endlichkeit der betrachteten Datenbanken. In [Sc01] wird die Methode der *Übersetzung von Gewinnstrategien im Ehrenfeucht-Fraïssé Spiel* vorgestellt und genutzt, um Kollaps-Resultate auch für unendliche Datenbanken zu erzielen. Es stellt sich sogar heraus, dass der natürlich-generische Kollaps *genau dann* gilt, wenn die Übersetzung von Gewinnstrategien im Ehrenfeucht-Fraïssé Spiel möglich ist. Mittels dieser Methode wird beispielsweise gezeigt, dass die Kontextstrukturen $\langle \mathbb{N}, <, \mathfrak{Teil} \rangle$ und $\langle \mathbb{N}, <, +, Q, \mathfrak{Teil}_Q \rangle$ den natürlich-generischen Kollaps für *FO* auf der Klasse *aller beliebigen* Datenbanken erlauben. Hierbei besteht $\mathfrak{Teil}$ aus allen Teilmengen von $\mathbb{N}$, Q ist die bereits in Kapitel 4 erwähnte unendlich große Menge von natürlichen Zahlen, und $\mathfrak{Teil}_Q$ besteht aus allen Teilmengen von Q.

Diese Kollaps-Resultate lassen sich vom Kontextuniversum $\mathbb{N}$ auch auf das Kontextuniversum $\mathbb{R}$ übertragen, so dass man folgendes Ergebnis erhält: Die Kontextstrukturen $\langle \mathbb{R}, <, \mathfrak{Teil} \rangle$ und $\langle \mathbb{R}, <, +, Q, \mathfrak{Teil}_Q, \mathbb{Q}, \mathbb{Z}, \mathfrak{Gruppen} \rangle$ erlauben den natürlich-generischen Kollaps für *FO* auf der Klasse aller $\mathbb{N}$-*einbettbaren* Datenbanken. Hierbei besteht $\mathfrak{Teil}$ aus allen Teilmengen von $\mathbb{R}$ und $\mathfrak{Gruppen}$ aus allen Teilmengen von $\mathbb{R}$, die bezüglich $+$ eine Gruppe sind und die Zahl 1 enthalten. Eine Datenbank heißt $\mathbb{N}$-*einbettbar*, falls ihr aktiver Domain ordnungserhaltend in die natürlichen Zahlen eingebettet werden kann. Eine Datenbank heißt $\mathbb{N}$-*repräsentierbar*, falls jede ihrer Relationen aus höchstens abzählbar vielen mehrdimensionalen „Rechtecken“ besteht, die durch eine $\mathbb{N}$-einbettbare Menge repräsentiert werden können. Mittels eines *Lifting Theorems* können Kollaps-Resultate für die Klasse der $\mathbb{N}$-einbettbaren Datenbanken sogar auf die (größere) Klasse der $\mathbb{N}$-repräsentierbaren Datenbanken übertragen werden.

Literatur

[AHV95] Abiteboul, S., Hull, R., und Vianu, V.: *Foundations of databases.* Addison-Wesley. 1995.

[EFT96] Ebbinghaus, H.-D., Flum, J., und Thomas, W.: *Einführung in die mathematische Logik*. Spektrum Akademischer Verlag. 4. Auflage. 1996.

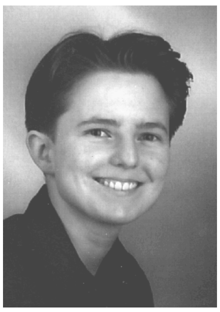
[Fa74] Fagin, R.: Generalized first-order spectra and polynomial-time recognizable sets. In: Karp, R. M. (Hrsg.), *Complexity of Computation.* Volume 7 of *SIAM-AMS Proceedings.* S. 43–73. 1974.

[Im99] Immerman, N.: *Descriptive complexity*. Springer-Verlag. 1999.

[KLP00] Kuper, G. M., Libkin, L., und Paredaens, J. (Hrsg.): *Constraint databases.* Springer-Verlag. 2000.

[Ru99] Ruhl, M.: Counting and addition cannot express deterministic transitive closure. In: *LICS'99: 14th Annual IEEE Symposium on Logic in Computer Science.* IEEE Computer Society. S. 326–334. Trento, Italy. July 1999.

[Sc01] Schweikardt, N.: *On the Expressive Power of First-Order Logic with Built-In Predicates*. Dissertation. Johannes Gutenberg-Universität Mainz, Fachbereich Mathematik und Informatik. Dezember 2001. Veröffentlicht beim Logos-Verlag, Berlin, 2002, ISBN 3-8325-0017-0.



Nicole Schweikardt, geboren 1973, absolvierte von 1992 bis 1998 ein Studium in Mathematik und Informatik an der Johannes Gutenberg-Universität Mainz, das sie im Februar 1998 als Diplom-Mathematikerin abschloss. Ihr anschließendes Promotionsstudium in Theoretischer Informatik mit den Nachbarfächern Praktische Informatik und Reine Mathematik beendete sie im Juni 2002 mit der Note Summa Cum Laude. Von März 1998 bis September 2002 war sie als wissenschaftliche Mitarbeiterin am Institut für Informatik der Johannes Gutenberg-Universität Mainz tätig. Ihre Dissertation entstand im Rahmen des DFG-Forschungsprojekts „Untersuchung der Ausdrucksstärke verschiedener eingebauter Relationen im Kontext von MonadicNP mit dem Ziel des Nachweises unterer Schranken“. Seit Oktober 2002 wird sie durch ein einjähriges Post-Doc Stipendium des DAAD gefördert um an der University of Edinburgh an der Erforschung der Ausdrucksstärke und Komplexität monadischer Fixpunktsprachen zu arbeiten.