Quantoren zweiter Stufe zur Darstellung von Zähl- und Optimierungsfunktionen

Riko Jacob

April 1996

Studienarbeit am
Lehrstuhl für Theoretische Informatik
der Universität Würzburg,
betreut von Dr. Heribert Vollmer

Zusammenfassung

Neben der üblichen Charakterisierung von Komplexitätsklassen durch ressourcenbeschränkte (Turing-)Maschinen gibt es auch die Möglichkeit, die Komplexität eines Problems an dem Aufwand an syntaktischen Mitteln zu messen, die zu seiner Darstellung durch eine logische Formel nötig sind. Diese Betrachtungsweise geht auf ein Resultat von Fagin [Fag74] zurück, in dem er die Klasse **NP** charakterisiert.

Damit kann **NP** durch Einschränkungen der syntaktischen Mittel in Unterklassen geteilt werden. Auf diese Art werden in [PY91] **NP**-harte Optimierungsprobleme und in [SST92] Zählprobleme betrachtet.

Diese verschiedenen Ergebnisse werden in einen gemeinsamen definitorischen Rahmen gesetzt, indem sie als Quantorenanwendung in einer speziell für diesen Zweck definierten Logik charakterisiert werden.

Die einheitliche Darstellung führt zu neuen, ähnlich gebildeten Klassen. Zusammen mit einigen der Notation angepaßten Beweisen bekannter Ergebnisse illustrieren diese die Möglichkeiten, derartige Komplexitätsklassen zu bilden. Dabei wird auch die Berechnungskraft der zum Vergleich der Modelle notwendigen Ordnung des Universums betrachtet.

Inhalt

1

Einleitung

Die formale Logik ist eines der ältesten und grundlegendsten Gebiete der Mathematik. Durch das Formalisieren der Sprache mathematischer Beweise gelang es zunächst, die Grenzen des exakten Schließens auszuloten. Dabei zeigt sich auch eine enge Verbindung zur Berechenbarkeit.

Aber nicht nur für grundsätzliche Überlegungen werden Konstrukte aus der formalen Logik verwendet. Ausgehend von einem richtungweisenden Ergebnis von Fagin aus dem Jahr 1974 [Fag74], in dem er die Komplexitätsklasse **NP** charakterisiert, wird untersucht, wie die Komplexität der Darstellung eines Problems durch eine Formel der formalen Logik mit der Berechnungskomplexität zusammenhängt. Eine Richtung dieser Untersuchungen geht von Papadimitriou und Yannakakis [PY91] aus, indem sie Klassen von **NP**-harten Optimierungsproblemen durch Klassen von Formeln definieren und unterschiedliche Approximationseigenschaften zeigen können.

Diese Richtung – durch logische Formeln charakterisierte Funktionen nach der Anzahl der Quantorenwechsel in der ersten Stufe einzuteilen und ihre Struktur zu untersuchen – wird von Saluja, Subrahmanyam und Thakur in [SST92] auf Zählklassen ausgedehnt und von Thakur in seiner Dissertation zusammengefaßt [Tha92].

Die erste und entscheidende Aufgabe besteht nun darin, Begriffe und mathematische Objekte zu bilden, mit denen es möglich ist, die zunächst in unterschiedlichem Rahmen gegebenen Probleme miteinander zu vergleichen. Dazu wird hier eine Verbindung zwischen Strings, die Eingaben für Turingmaschinen darstellen, und endlichen Strukturen, für die eine logische Formel gelten kann, hergestellt.

Da die Information eines Strings im wesentlichen durch die Anordnung der Symbole gegeben ist, bietet es sich an, Strukturen über linear geordneten Grundmengen, also Teilmengen der natürlichen Zahlen, zu betrachten. Dabei ergibt sich die Schwierigkeit, daß die Darstellung dieser Ordnung in der Logik den syntaktischen Aufwand der Formel bestimmen kann und daß es Probleme z.B. auf Graphen gibt, die in der Logik auch ohne Anordnung der Elemente darstellbar sind.

Man sucht daher eine Eingabeumwandlung, die eine ausreichend schwache Reduktion darstellt, so daß die Grenzen der betrachteten Komplexitätsklassen nicht verwischt werden. Dies führt zu der Frage, welchen Einfluß eine unabhängig von der Eingabe gegebene Ordnung auf die Komplexität der logischen Formel hat.

Wie sich herausstellt, läßt sich diese Problematik nicht für alle betrachteten Situationen gleichartig lösen. Es werden daher die verschiedenen Möglichkeiten, eine Anordnung vorzunehmen und darzustellen, sorgfältig auseinandergehalten, und deren Einfluß auf die Ausdrucksstärke der Formeln untersucht.

Ein ähnliches Vorgehen, in dem eine Turingmaschine nur eine der möglichen Darstellungen des Problems akzeptieren muß, würde Probleme wie HAMILTONIAN oder MAXCLIQUE in die Klasse **P** bringen, also einen vermutlich recht starken Nichtdeterminismus verstecken. Interessanterweise fällt dieser Kollaps in der Logik nicht so gravierend aus. Es zeigt sich, daß die Hierarchie der Quantorenwechsel in der ersten Stufe alleine nicht geeignet ist, Determinismus und Nichtdeterminismus zu trennen. Dies ist in Anbetracht der vorgestellten Resultate, daß diese Hierarchien echt sind, nicht besonders überraschend.

Die andere Maßgabe für die hier vorgestellten Definitionen ist das Ziel, mehrere an verschiedenen Stellen etablierte Charakterisierungen durch logische Formeln in einen gemeinsamen definitorischen Rahmen zu fassen und sie dadurch so vergleichbar wie möglich zu machen.

Dies führt zum einen dazu, daß gleichartige Beweise in allgemeiner formulierte Sätze umgewandelt werden, und dadurch strukturelle Gemeinsamkeiten verschiedener Hierarchien klar sichtbar werden. Andererseits ergibt sich eine einheitliche Notation der bekannten Ergebnisse, die bisher nicht betrachtete Hierarchien natürlich erscheinen läßt. Die Untersuchung dieser Klassen und die in neuer Systematik angegebenen Beweise dienen dem Verständnis der zunächst ungewöhnlich erscheinenden Struktur der von Thakur in seiner Dissertation betrachteten Hierarchien.

2 Grundlagen aus der Logik

In der klassischen Logik wird definiert, unter welchen Umständen eine bestimmte Formel gültig ist. Von dieser vertrauten Situation ausgehend, wird hier eine Erweiterung vorgestellt, die es erlaubt, endlichen Strukturen natürliche Zahlen zuzuweisen. Dementsprechend wird im folgenden eine etwas ungewöhnliche, auf dieses Ziel ausgerichtete Definition der formalen Logik gegeben.

Mit ihr wird es möglich sein alle Problemklassen, die Thakur in seiner Dissertation betrachtet, als Klassen von logischen Formeln darzustellen. Durch geeignete Quantoren kann die Definition über Mächtigkeiten von Mengen in eine Quantorendarstellung in dieser erweiterten Logik umgewandelt werden.

Den untersuchten Klassen entsprechend werden die Formeln nur in pränexer Normalform definiert, was bekanntermaßen keine Einschränkung der Ausdruckskraft darstellt und hier das Zählen von Quantorenwechseln erleichtert.

Die zentrale Idee besteht darin, anstatt die Gültigkeit einer Formel zu definieren, ihr im gegebenen Zusammenhang eine natürliche Zahl zuzuweisen. Somit sondert eine Formel nicht mehr eine Teilmenge aus allen passenden Strukturen aus, sondern stellt eine Abbildung der Menge der zur Formel passenden Strukturen in die natürlichen Zahlen dar. Darin geht der Begriff der Gültigkeit auf, indem diese Abbildung als charakteristische Funktion der Menge der gültigen Strukturen interpretiert wird. Dementsprechend ist eine Formel genau dann für eine Struktur gültig, wenn sie dieser einen Wert ungleich 0 zuweist.

Unter den natürlichen Zahlen wird also die Menge $\mathbb{N} = \{0, 1, 2, \ldots\}$ verstanden.

Außerdem wird auf den Begriff der Interpretation verzichtet. Er geht in dem Begriff der Signatur, der die syntaktischen Möglichkeiten und den semantischen Rahmen einer Formel festlegt, auf. Dies entspricht der Vorstellung, daß nicht irgendwelche Modelle, die in der Logik formulierte Sätze erfüllen, gesucht werden, sondern in einem festen Mechanismus Eingaben ausgewertet werden.

Definition 1 (Signatur)
*Eine **Signatur**, auch **Ähnlichkeitstyp** (similarity type) ist ein Sechstupel $\Sigma = (\Phi, \Pi, r, V, S, \Omega)$. Dabei sind Φ und Π abzählbare disjunkte Mengen der Funktoren bzw. Prädikatssymbole. Die Funktion $r\colon \Phi \cup \Pi \to \mathbb{N}$ weist jedem dieser Symbole eine Stelligkeit zu. Die abzählbare Menge V besteht aus den Variablen erster Ordnung, die Menge S aus den endlichen Grundstrukturen (Algebren) und die Menge $\Omega \subset \Phi \cup \Pi \cup V$ aus den mit fester Semantik verbundenen Symbolen. Eine Grundstruktur besteht jeweils aus einer endlichen Menge U, dem Universum und einer Abbildung $I\colon \Omega \to U^*$, die den Standardsymbolen Funktionen und Relationen zuweist.*

Bemerkung: Diese Definition greift die Idee einer allgemeinen Interpretation oder einer Einbettung der Strukturen auf. So kann etwa als Menge S die der endlichen Teilmengen der natürlichen Zahlen mit der jeweiligen natürlichen Ordnung gewählt werden. Genauer also $\Omega = \{o\}$ und

$$S = \Big\{ \langle G, o \rangle \,\Big|\, G \subset \mathbb{N},\ |G| \in \mathbb{N},\ G \times G \supset o = \{\, (a,b) \mid a < b \,\} \Big\}.$$

Im weiteren wird der Lesbarkeit halber dieses o als $<$ geschrieben. In Abschnitt 2.4 werden einige Möglichkeiten beleuchtet, die Ordnung eines Universums darzustellen.

Bemerkung: Normalerweise werden Konstanten als nullstellige Funktionen betrachtet. Mit dieser Definition ($V \cap \Omega \neq \emptyset$) können auch „Variablen“ mit einer allgemeingültigen Bedeutung versehen werden. Dies erlaubt es, auch in Klassen, in denen eine entsprechende Relation nicht ausgewertet werden könnte, Konstanten einzuführen. Dies ist z.B. bei quantorenfreien Formeln, in denen keine Funktoren vorkommen, der Fall. Diese könnten dann mit einer entsprechend belegten „Variablen“ überprüfen, ob das kleinste Element des Universums zu einer Relation gehört, dies ist mit einem Prädikat nicht möglich.

Im weiteren wird das Universum meist ein Anfangsstück der natürlichen Zahlen sein, das mit der Gleichheitsrelation ausgestattet ist. Die Anordnung wird durch $+1$ (Nachfolger), $<$ (Ordnung) oder $+$ (Addition) dargestellt. Die Menge der Grundstrukturen legt den Rahmen fest, in dem eine Formel zu verstehen ist, und wird daher wie eine syntaktische Möglichkeit betrachtet, da sie offenbar Einfluß auf die Größe und Struktur der benötigten Formeln ausübt. Somit legt die Signatur alle syntaktischen und die elementaren semantischen Möglichkeiten der Formelbildung fest.

Da andererseits für eine gegebene Formel im allgemeinen aus dem Kontext die Signatur klar festgelegt ist, wird diese nicht unbedingt angegeben.

Außerdem ist damit auch ein Rahmen für die Eingabestrukturen gegeben, der einer Art Vorverarbeitung entspricht. So kann z.B. die Frage, ob ein gerichteter Graph der Ordnung seiner Knoten entspricht, ob also nur Pfeile von „kleineren" zu „größeren" Knoten existieren, sehr leicht beantwortet werden, wenn die Reihenfolge als Ordnung gegeben ist. Dies ist nicht mehr so einfach, wenn die Reihenfolge der Knoten durch eine Nachfolgerrelation gegeben ist.

Dementsprechend werden Klassen von logischen Formeln gebildet, in denen bestimmte Möglichkeiten zum Aufbau der Formeln zugelassen sind und deren Grundstrukturen bestimmte Symbole mit festgelegter Bedeutung umfassen. Wenn man die Möglichkeiten des Formelaufbaus der Signatur zurechnet, könnte man auch von Mengen von Signaturen sprechen, über denen Formeln gebildet werden. Ziel wird es dann sein, die dadurch definierte Menge von Problemen komplexitätstheoretisch zu betrachten.

Definition 2
Terme *über einer Signatur* $\Sigma = (\Phi, \Pi, r, V, S, \Omega)$ *genügen der folgenden induktiven Definition:*

- *Jede Variable* $x \in V$ *ist ein Term.*
- *Sind* $\varphi_1, \ldots, \varphi_r$ *Terme und* $f \in \Phi$ *ein* r*-stelliges Funktionssymbol mit* $r = r(f)$*, dann ist* $f(\varphi_1, \ldots, \varphi_r)$ *ein Term.*

Definition 3
Quantorenfreie Formeln *über einer Signatur* $\Sigma = (\Phi, \Pi, r, V, S, \Omega)$ *werden folgendermaßen definiert:*

- *Sind* $\varphi_1, \ldots, \varphi_r$ *Terme und* $R \in \Pi$ *ein* r*-stelliges Prädikatssymbol mit* $r = r(R)$*, so ist* $R(\varphi_1, \ldots, \varphi_r)$ *eine* ***atomare quantorenfreie Formel****.*
- *Sind* φ *und* ψ *quantorenfreie Formeln, so sind auch* $(\varphi \wedge \psi)$*,* $(\varphi \vee \psi)$ *und* $\neg\varphi$ *solche.*

Man schreibt $\varphi(x)$, um auszudrücken, daß x in φ frei vorkommt. In quantorenfreien Formeln sind alle Variablen aus V frei, auch wenn sie in der Formel nicht vorkommen.

Diese Formeln sind je nach vorgeschriebener Grundstruktur in den Klassen $\mathrm{QF}[=]$, $\mathrm{QF}[<]$ oder $\mathrm{QF}[=, \mathrm{succ}]$.

Definition 4 (Klassen der ersten Stufe)
Für alle $k \in \mathbb{N}$ *werden folgende Klassen von Formeln der Prädikatenlogik erster Stufe definiert:*

- *Eine quantorenfreie Formel* φ *ist in der Klasse* $\Sigma_0 = \Pi_0$*.*

- *Seien $\sigma(x) \in \Sigma_i$ und $\pi(x) \in \Pi_i$ Formeln der ersten Stufe mit freier Variable x. Dann sind*

$$\exists x\, \sigma(x) \in \Sigma_i\,,$$
$$\forall x\, \pi(x) \in \Pi_i\,,$$
$$\exists x\, \pi(x) \in \Sigma_{i+1}\,,$$
$$\forall x\, \sigma(x) \in \Pi_{i+1}\,.$$

*In diesen Formeln ist die Variable x **gebunden**, also nicht mehr frei. Die Klasse der Formeln erster Stufe ist definiert durch*

$$\mathrm{FO} = \bigcup_{i \in \mathbb{N}} \Pi_i\,.$$

Bei den Bezeichnungen der Klassen gibt der Index die Anzahl der Quantorenblöcke an, Π bzw. Σ steht für die Art des ersten vorkommenden Quantors. Ein Block von Quantoren besteht aus möglichst vielen aufeinanderfolgenden gleichartigen Quantoren. Falls bestimmte Grundstrukturen vorgesehen sind, schreibt man z.B. $\Pi_k[<]$.

Bemerkung: Durch diese Definitionen sind nur Formeln in pränexer Normalform zugelassen. Dies ist keine Einschränkung der Ausdrucksstärke, gelegentlich aber etwas umständlich. Daher wird im folgenden die Tatsache ausgenutzt, daß zwei Formeln mit gleichartigem Quantorenpräfix, die durch ein $\wedge$ oder ein $\vee$ verbunden sind, mit einem derartigen Quantorenpräfix geschrieben werden können. Ein Beweis, der dieses Vorgehen rechtfertigt, findet sich z.B. in [EFT92] auf Seite 149.

Bis hierher handelt es sich um nur leicht angepaßte allgemein übliche Definitionen. Die weiteren, zum Teil auch syntaktisch ungewöhnlichen Konstrukte sollen hier zunächst nur vorgestellt werden. Die Möglichkeiten, die sich aus dieser Darstellung ergeben, werden sich erst später zeigen.

Definition 5 (Zählquantor)
Sei φ eine Formel, in der die Variable x der ersten Stufe frei vorkommt, dann ist

$$\#x\, \varphi$$

eine Formel der ersten Stufe, in der x gebunden ist.

Definition 6 (Quantoren der zweiten Stufe)
Sei φ eine Formel, in der das Prädikat R bzw. der Funktor f frei vorkommen, dann sind

$$\#R\,\varphi, \quad \max_R{}^2\,\varphi, \quad \min_R{}^2\,\varphi$$

und

$$\#f\,\varphi, \quad \max_f{}^2\,\varphi, \quad \min_f{}^2\,\varphi$$

Formeln der zweiten Stufe, in denen R bzw. f gebunden sind.

Bemerkung: Die üblichen All- und Existenzquantoren der zweiten Stufe sind so darstellbar, so wird etwa für $\max^2$ abkürzend auch $\exists^2$ geschrieben. Dabei steht der Exponent lediglich als Kennzeichnung der Stufe des Quantors, dies ist vor allem beim Bilden von Formelklassen notwendig.

Definition 7 (arithmetische Modifikatoren)
Sei φ eine Formel. Dann sind

$$\neg\varphi, \quad \mathrm{mod}\,\varphi$$

Formeln.

Bemerkung: Weitere denkbare Modifikatoren dieser Art wären z.B. $(=1)$, (≤ 1) und (≥ 1), die den Wahrheitswert (0 oder 1) der angegebenen Bedingung weitergeben. Je nach Kontext könnte anstelle der 1 auch der unveränderte Wert oder der später eingeführte Bereich der Formel weitergegeben werden.

Definition 8 (prädikative Modifikatoren)
Sei φ eine Formel in der das Prädikat R bzw. der Funktor f frei vorkommen. Dann sind

$$(\text{Ordnung } R)\varphi, \quad (\text{injektiv } R)\varphi, \quad (\text{Funktion } R)\varphi$$

und

$$|R|_k\,\varphi, \quad |R|\,\varphi, \quad |R|_\infty\,\varphi, \quad \mathrm{Bin}(R)\varphi$$

Formeln der zweiten Stufe, in denen R bzw. f nach wie vor frei vorkommen.

Bemerkung: Diese Modifikatoren werden zum einen verwendet, um – ähnlich einem Orakel für eine Turingmaschine – bestimmte Relationen auszusondern, und andererseits, um bestimmte Zahlen zur Verfügung zu stellen, hier etwa die Mächtigkeit einer Relation oder eine ihr zugeordnete Binärzahl. Dies stellt offenbar kein Quantifizieren dar. Auch wenn sich hier, da der Wert einer Formel in Abhängigkeit von einer Relation verändert wird, der Begriff eines Operators anbietet, soll aus Gründen der Abtrennung von anderen so bezeichneten Konstrukten auch hier der Begriff des Modifikators verwendet werden.

Die Formeln werden den verwendeten Quantoren entsprechend in Klassen eingeteilt, deren Namen sich aus diesen Quantoren und den Symbolen aus

Ω zusammensetzen. So entsteht z.B. die Klasse $\exists^2\mathrm{FO}[<]$ aller Formeln, die aus existentiellen Quantoren der zweiten Stufe, gefolgt von einer beliebigen Formel der ersten Stufe, in der auch das Symbol $<$ verwendet werden darf, gebildet werden können.

Genauer werden die Formelklassen nach folgendem Schema bezeichnet:

$$\left\{\begin{array}{c}\#^2\\ \max^2 = \exists^2\\ \min^2\\ \cdot\end{array}\right\}\left\{\begin{array}{c}\#^1\\ \cdot\end{array}\right\}\left\{\begin{array}{c}(\text{Ordnung})\\ (\text{injektiv})\\ (\text{Funktion})\\ \cdot\end{array}\right\}\left\{\begin{array}{c}\mathrm{QF}\\ \Pi_k\\ \Sigma_k\\ \mathrm{FO}\\ \textsc{Horn}\end{array}\right\}\left\{\begin{array}{c}[<]\\ [=,\mathrm{succ}]\\ \cdot\end{array}\right\},$$

wobei $\cdot$ für die leere Alternative steht.

2.1 Semantik

Definition 9 (Struktur)
Eine ***Struktur*** $\mathcal{A}$ *über einer Signatur* $\Sigma = (\Phi, \Pi, r, V, S, \Omega)$ *ist ein Tupel* $\mathcal{A} = (A, I)$*, bestehend aus einer Menge* A*, für die es eine Grundstruktur* $(A, J) \in S$ *gibt, und der Funktion* $I\colon (\Phi \cup \Pi \cup V) \to (A^* \cup \{\bot\})$*, die jedem Prädikats- und Funktionssymbol eine* ***Belegung*** *oder das Symbol* $\bot$ *zuweist. Genauer gilt:*

- *Das Universum* A *ist eine endliche Menge; mit* $|A| = ||\mathcal{A}||$ *wird die Mächtigkeit oder Größe des Universums bzw. der Struktur bezeichnet.*
- *Die Funktion* I *stimmt auf* Ω *mit* J *überein.*
- *Jedem Prädikatssymbol* $R \in \Pi$ *der Stelligkeit* $r = r(R)$ *wird eine Teilmenge* $X \subset A^r$ *oder das Symbol* $\bot$ *zugewiesen.*
- *Jedem Funktor* $F \in \Phi$ *der Stelligkeit* $r = r(F)$ *wird eine Teilmenge* $X \subset A^{r+1}$ *oder das Symbol* $\bot$ *zugewiesen. Dieses* X *stellt eine Funktion dar, d.h. für jedes* $x \in A^r$ *gibt es genau ein* $y \in A$ *mit* $(x, y) \in X$*.*
- *Jeder Variablen* $x \in V$ *wird ein Wert aus dem Universum oder das Symbol* $\bot$ *zugewiesen.*

Definition 10
Eine Formel φ *und eine Struktur* $\mathcal{A} = (A, I)$ *über derselben Signatur* ***passen zusammen****, falls den in* φ *frei vorkommenden Symbolen durch* $\mathcal{A}$ *eine Belegung zugewiesen ist, und den gebundenen das Symbol* $\bot$*. Für nicht vorkommende Symbole wird keine Festlegung getroffen.*

Damit ergibt sich auch die Menge aller zu einer Formel passenden Strukturen. Diesen wird durch die Formel, folgenden Definitionen entsprechend, eine Zahl zugewiesen. Im Sinne einer charakteristischen Funktion ist dies mit den formalen Sprachen vergleichbar, die von Automaten und Turingmaschinen erkannt werden.

Definition 11
Sei $\mathcal{A}$ eine Struktur mit $I(x) = \perp$. Dann steht $\mathcal{A}[x \mapsto a]$ für dieselbe Struktur, lediglich mit $I(x) = a$.

Für das unten definierte $\mathrm{val}(\mathcal{A}[x \mapsto a] \models \varphi)$ *wird kurz auch* $\mathrm{val}(\varphi(x \mapsto a)$ *oder nur* $\mathrm{val}(\varphi(a))$ *geschrieben.*

Definition 12 (Wert von Termen)
Der Wert eines Terms ergibt sich induktiv. Sei e ein Term und $\mathcal{A} = (A, I)$ eine passende Struktur. Dann gilt:

- *Ist $e = v \in V$ eine Variable, so hat e den Wert $I(v) \in A$.*
- *Ist $e = f(\varphi_1, \ldots, \varphi_n)$ und haben die Terme $\varphi_1, \ldots, \varphi_n$ die Werte $x_1 \in A, \ldots, x_n \in A$, dann hat e den Wert $I(f)(x_1, \ldots, x_n) \in A$.*

Da Term und Struktur zueinander passen, wird nie der Wert $\perp$ zugewiesen. Die so definierten Werte liegen also innerhalb des Universums und haben mit den Werten von Formeln, die im folgenden definiert werden, nichts zu tun.

Definition 13 (Wert einer Formel der ersten Stufe)
*Einer Formel τ wird für eine passende Struktur $\mathcal{A} = (A, I)$ induktiv der Wert $m = \mathrm{val}(\mathcal{A} \models \tau) \in \mathbb{N}$ und der **Bereich** $N \in \mathbb{N}$ zugewiesen:*

- *Falls $\tau = R(e_1, \ldots, e_r)$, so hat τ den Bereich $N = 1$.
Haben die Terme $e_1, \ldots, e_r$ die Werte $x_1, \ldots, x_n$, dann gilt $m = 1$ falls $(x_1, \ldots, x_n) \in I(R)$, sonst $m = 0$.*
- *Falls $\tau = \varphi \vee \psi$ mit Formeln φ und ψ vom Bereich 1, so gilt $m := \max\{\mathrm{val}(\mathcal{A} \models \varphi), \mathrm{val}(\mathcal{A} \models \psi)\}$. τ hat Bereich $N = 1$.*
- *Falls $\tau = \varphi \wedge \psi$ mit Formeln φ und ψ vom Bereich 1, so gilt $m := \min\{\mathrm{val}(\mathcal{A} \models \varphi), \mathrm{val}(\mathcal{A} \models \psi)\}$. τ hat Bereich $N = 1$.*
- *Falls $\tau = \exists x\, \varphi(x)$ mit einer Formel φ vom Bereich 1, so gilt $m = \max\{\, \mathrm{val}(\mathcal{A}[x \mapsto a] \models \varphi) \mid x \in |\mathcal{A}|\,\}$. Der Bereich ist der von φ.*
- *Falls $\tau = \forall x\, \varphi(x)$ mit einer Formel φ vom Bereich 1, so gilt $m = \min\{\, \mathrm{val}(\mathcal{A}[x \mapsto a] \models \varphi) \mid x \in |\mathcal{A}|\,\}$. Der Bereich ist der von φ.*

- *Falls $\tau = \neg\varphi$ mit φ vom Bereich M ist, so hat τ den Bereich $N = M$, und es gilt $m := N - \mathrm{val}(\mathcal{A} \models \varphi)$.*

Der Anschauung entsprechend, bedeutet $\mathrm{val}(\mathcal{A} \models \varphi) \neq 0$, daß φ für $\mathcal{A}$ gültig ist, oder auch, daß $\mathcal{A}$ die Formel φ erfüllt. Dementsprechend steht $\mathcal{A} \models \varphi$ als Abkürzung für $\mathrm{val}(\mathcal{A} \models \varphi) \neq 0$. So haben auch All- und Existenzquantor die gewohnte Bedeutung und können der Definition entsprechend auch als $\min^2$ und $\max^2$ geschrieben werden. Die Tatsache, daß zwei Formeln φ und ψ über derselben Signatur für alle passenden Strukturen den gleichen Wert ergeben, die dargestellten Abbildungen also identisch sind, wird als $\mathrm{val}(\varphi) \equiv \mathrm{val}(\psi)$ geschrieben. Im folgenden steht auch weiterhin der Begriff einer gültigen Formel für eine, die nicht zu 0 auswertet. Der Bereich einer Formel gibt an, zu welchen Zahlen sie auswerten könnte. Die zugewiesenen Werte liegen also zwischen 0 und dem Bereich einschließlich.

Man überprüft leicht, daß die gewohnten, und manchmal zur Definition genutzten, Gleichheiten $\varphi \wedge \psi := \neg(\neg\varphi \vee \neg\psi)$ und $\forall x\varphi := \neg\exists x\neg\varphi$ gelten.

Die Definition der Semantik der Quantoren zweiter Stufe wird zurückgestellt, um zunächst einige einfache Eigenschaften von Formeln und Strukturen aufzuzeigen.

Definition 14 (Substruktur)
Für eine Signatur $\Sigma = (\Phi, \Pi, r, V, S, \Omega)$ seien $\mathcal{A} = (A, I)$ und $\mathcal{B} = (B, J)$ Strukturen über Σ. Sei $B \subseteq A$, und für alle $X \in \Phi \cup \Pi$ gelte

$$I(X) \cap (B^{r(X)} \cup \{\bot\}) = J(X)$$

sowie für alle $X \in V$

$$I(X) = J(X)\,.$$

Dann ist $\mathcal{B}$ eine ***Substruktur*** *von $\mathcal{A}$. Dabei bleiben Funktionen solche, sie sind also auch im neuen Universum total. Das bedeutet, daß der Bildbereich $f(B) \subset B$ ist. Außerdem ist die so entstehende Strukturgrundmenge wieder ein Element von S.*

Für eine Teilmenge $B \subset A$ heißt die Struktur $\mathcal{B} = (B, J)$ mit

$$J(x) := I(x) \cap (B^{r(x)} \cup \{\bot\})$$

für $x \notin V$ bzw. für $x \in V$

$$J(x) := I(x)$$

die ***durch B induzierte Substruktur*** *von $\mathcal{A}$.*

Bemerkung: Die Forderung, daß eine Substruktur mit denselben Basisrelationen und Funktionen wieder eine zulässige Grundmenge sein muß, ist recht stark. Die einzige hier vorgestellte nichtleere Basisrelation, bei der überhaupt nichttriviale Substrukturen möglich sind, ist eine totale Ordnung.

Die folgenden Überlegungen werden in späteren Beweisen benötigt, und daher hier kurz begründet. Eine ausführlichere Darstellung dieser Sachverhalte findet sich z.B. in [EFT92] ab Seite 51.

Lemma 15
Für eine quantorenfreie Formel φ mit Bereich 1 und zwei passende Strukturen $\mathcal{A} = (A, I) \subseteq \mathcal{B} = (B, J)$ über einer Signatur $\Sigma = (\Phi, \Pi, r, V, S, \Omega)$ gilt:

$$\mathcal{A} \models \varphi \;\Leftrightarrow\; \mathcal{B} \models \varphi.$$

Beweis: Da es sich um Substrukturen handelt, gilt $I|_V = J|_V$, den Variablen werden dieselben Werte zugewiesen. Dies überträgt sich – durch die induktive Definition der Werte, und weil $J(f)|_A = I(f)$ für alle $f \in \Phi$ gilt – auf die Werte der Terme, die alle in A liegen. Dort werden auch den Prädikaten durch I und J dieselben Wahrheitswerte zugeordnet. So ergeben sich beim induktiven Auswerten von $\vee$, $\wedge$ und $\neg$ auch dieselben Wahrheitswerte für alle Teilformeln, also auch für φ. ■

Lemma 16 (Einschränkung rein all-quantifizierter Formeln)
Für eine quantorenfreie Formel φ mit Bereich 1 und eine Struktur $\mathcal{I}$ über einer passenden Signatur Σ gelte

$$\mathcal{I} \models \forall x_1 \cdots \forall x_n \, \varphi \,.$$

Für alle Substrukturen $\mathcal{J} \subseteq \mathcal{I}$ gilt dann

$$\mathcal{J} \models \forall x_1 \cdots \forall x_n \, \varphi \,.$$

Beweis: Offenbar paßt auch $\mathcal{J}$ zu φ. Aus der Voraussetzung ergibt sich, daß für beliebige $a_i \in |\mathcal{J}|$ die Struktur $\mathcal{I}[x_1 \mapsto a_1, \ldots, x_n \mapsto a_n]$ die Formel φ erfüllt.

Es gilt also insbesondere $a_i \in |\mathcal{J}| \subseteq |\mathcal{I}|$.

Mit Lemma 15 gilt also φ für beliebige Belegungen der x_i durch $a_i \in |\mathcal{J}|$, und somit die Behauptung. ■

Lemma 17 (Erweiterung rein existentiell quantifizierter Formeln)
Für eine quantorenfreie Formel φ mit Bereich 1 und eine Struktur $\mathcal{I}$ über einer passenden Signatur Σ gelte

$$\mathcal{I} \models \exists x_1 \cdots \exists x_n \varphi \,.$$

Für alle passenden Strukturen mit $\mathcal{I} \subseteq \mathcal{J}$ gilt dann

$$\mathcal{J} \models \exists x_1 \cdots \exists x_n \varphi \,.$$

Beweis: Aus der Voraussetzung ergibt sich, daß es $a_1, \ldots, a_n \in |\mathcal{J}|$ gibt, so daß die Struktur $\mathcal{I}[x_1 \mapsto a_1, \ldots, x_n \mapsto a_n]$ die Formel φ erfüllt. Da diese a_i insbesondere Elemente von $|\mathcal{I}|$ sind, gilt mit Lemma 15 die Behauptung. ∎

Definition 18
Zwei Strukturen $\mathcal{I}$ und $\mathcal{J}$ über derselben Signatur $\Sigma = (\Phi, \Pi, r, V, S, \Omega)$ sind ***isomorph*** *zueinander, wenn es eine bijektive Abbildung $\tau\colon |\mathcal{I}| \to |\mathcal{J}|$ der Universen ineinander gibt, für die im Bezug auf alle Relationssymbole P und Funktionssymbole f (auch solche aus Ω) gilt:*

$$P_{\mathcal{I}}(\mathbf{x}) \;\Leftrightarrow\; P_{\mathcal{J}}(\tau(\mathbf{x}))$$

bzw.

$$\tau(f_{\mathcal{I}}(\mathbf{x})) = f_{\mathcal{J}}(\tau(\mathbf{x}))$$

Satz 19
Für zwei isomorphe Strukturen $\mathcal{I}$, $\mathcal{J}$ und eine passende Formel ψ über einer Signatur Σ gilt:

$$\mathrm{val}(\mathcal{I} \models \psi) \;=\; \mathrm{val}(\mathcal{J} \models \psi)$$

Beweis: Dies folgt aus der Definition von Isomorphie und der induktiven Definition der Semantik. ∎

2.2 Relationen und Funktionen

Es zeigt sich, daß Relationen und Funktionen sehr ähnlich sind und in vielen Situationen gegenseitig ersetzt werden können. Mit bestimmten syntaktischen Möglichkeiten kann man in einem Satz (einer Formel) Funktionen durch Prädikate austauschen und umgekehrt.

Dabei kann auf Prädikate nicht vollständig verzichtet werden, da sonst nach der Definition keine Formeln entstehen können. Man kann sich aber sicherlich auf die Gleichheitsrelation beschränken und alle anderen Prädikate durch Funktionen nach $\{0, 1\}$ ersetzen. Dafür ist lediglich ein zweielementiges Universum nötig. Es zeigt sich in Abschnitt 3.1, daß dies keine wirkliche Einschränkung darstellt.

Die Umwandlung von Funktionen in Prädikate ist sicher interessanter. Daß sie grundsätzlich möglich ist, ist in Anbetracht der Darstellung als Mengen bzw. Relationen nicht überraschend. Es zeigt sich jedoch, daß die Verwendung von Funktionssymbolen eine syntaktisch stärkere Möglichkeit darstellt. Daher soll hier nicht nur die Art der Äquivalenz beider Logiken sondern auch der syntaktische Umbauaufwand genauer dargestellt werden.

Da die Logik ohne Funktoren in einem bestimmten Sinn schwächer ist, und damit andere syntaktische Mittel mehr Klassen von Formeln trennen können (Satz 45), wird in Abschnitt 4.2 eine solche Logik verwendet.

Lemma 20
Seien φ eine Formel über einer Signatur $\Sigma = (\Phi, \Pi, r, V, S, \Omega)$ und $\mathcal{J}$ eine Struktur. Dann gibt es über der Signatur $\Sigma = (\emptyset, \Pi', r', V', S', \Omega')$ eine Formel φ' und eine zu $\mathcal{J}$ semantisch äquivalente Struktur $\mathcal{J}'$, so daß φ und φ' äquivalent sind, also

$$\mathrm{val}(\mathcal{J} \models \varphi) \quad = \quad \mathrm{val}(\mathcal{J}' \models \varphi')\ .$$

Semantisch äquivalent bedeutet hier, daß Funktoren f durch Prädikate P_f ersetzt werden. Falls dem Funktor eine Funktion zugewiesen ist, wird diese durch die entsprechende Relation ersetzt.

Beweis:

Die Signatur Σ' wird so gewählt, daß die im folgenden beschriebene Formel φ' zu ihr paßt.

Für jeden in φ vorkommenden Funktor $f(x_1, \ldots, x_n)$ wird eine neue Variable y zu V hinzugenommen. Sei φ' die Formel, die entsteht, wenn das vorkommende $f(x_1, \ldots, x_n)$ durch y ersetzt wird. Dann ist φ äquivalent zu $\exists y P_f(x_1, \ldots, x_n, y) \wedge \varphi'$ und auch zu $\forall y P_f(x_1, \ldots, x_n, y) \rightarrow \varphi'$. Durch iteriertes Anwenden hiervon und durch die leicht mögliche äquivalente Umformung entsteht eine funktorenfreie Formel in Pränexform.

Falls φ nicht quantorenfrei war, ist φ' in derselben Σ_k- bzw. Π_k-Klasse von Formeln, da durch die oben angegebene Alternative ein beliebiger bestehender Quantorenblock vergrößert werden kann. Ansonsten gibt es ein φ' aus Σ_1 und ein anderes äquivalentes in Π_1.

Falls die Eigenschaft des Prädikates, eine Funktion darzustellen, nicht gesichert ist, kann dies durch die Teilformel

$$\begin{array}{rl} \forall \mathbf{x}\, \exists \mathbf{y} & P_f(\mathbf{x}, \mathbf{y}) \\ \forall \mathbf{x} \forall \mathbf{y} \forall \mathbf{z} & P_f(\mathbf{x}, \mathbf{y}) \wedge P_f(\mathbf{x}, \mathbf{z}) \rightarrow \mathbf{y} = \mathbf{z} \end{array}$$

überprüft werden.

Dementsprechend wird φ' dann mindestens zu einer Π_2-Formel. ∎

An dieser Stelle kann erstmals ein sogenannter Modifikator sinnvoll eingesetzt werden.

Der Modifikator $(\text{Funktion } R)$ überprüft die Eigenschaft, daß R eine Funktion darstellt. Daher hat $(\text{Funktion } R)\varphi$ denselben Wert wie φ, falls R funktional ist, sonst den Wert 0. Das Konstrukt $\exists f$ kann also im wesentlichen durch $\exists P_f(\text{Funktion } P_f)$ ersetzt werden. Dies ist tatsächlich so etwas wie ein Orakel. Dieses Vorgehen führt allerdings nicht in allen vorstellbaren Situationen zum gewünschten Ergebnis, so etwa in der Situation, daß ein All- oder Minimierungsquantor vorausgeht. Dann muß in dem Fall, daß f keine Funktion ist, ein anderer sinnvoll definierter Wert weitergegeben werden. Da dies hier nicht auftritt, wird von einer entsprechenden Definition abgesehen.

Es genügt also, sich auf eine formale Logik mit Prädikaten zu beschränken, da die syntaktische Kraft dieses Formalismus durch den Modifikator ausgedrückt werden kann. Diese Art der Vereinheitlichung der syntaktischen Möglichkeiten ist Ziel der getroffenen Definitionen und führt zu weiteren Modifikatoren und Quantoren.

2.3 Hornformeln

Definition 21 (Hornformel)
*Eine **Hornformel** ist eine Formel $\varphi = \exists^2 \mathbf{R}\, \forall x_1 \cdots \forall x_n\, \vartheta$ über einer Signatur $\Sigma = (\emptyset, \Pi, r, V, S, \Omega)$. Dabei ist der Rumpf der Formel ϑ eine quantorenfreie Formel erster Stufe in konjunktiver Normalform wobei in jeder Disjunktion (Hornklausel) nur ein positives Literal enthalten ist. Positive Literale, die Eingaberelationen oder Relationen aus Ω darstellen, also die freien Symbole zweiter Stufe von φ, werden dabei nicht gezählt.*

Bemerkung: Die besondere Behandlung der fest vorgegebenen Prädikate erweist sich als sinnvoll, da für die algorithmische Auswertung nur die quantifizierten Prädikate relevant sind. Daher werden nur die Prädikate gezählt, die durch Quantoren zweiter Stufe gebunden sind. Die Eigenschaft, eine Hornformel zu sein, kann durch Hinzunehmen von Quantoren verlorengehen.

Der besonderen Struktur von Hornformeln entsprechend, kann ϑ auch als Menge von Implikationen, den Hornklauseln, interpretiert und geschrieben werden. Die Einschränkung, daß es nur ein positives Literal pro Klausel geben darf, führt zu folgenden Möglichkeiten:

$P_1(\mathbf{x}), \ldots, P_n(\mathbf{y}) \Rightarrow Q(\mathbf{z})$**:** echte Implikationen;
aus mehreren positiven Prädikaten folgt ein positives Prädikat.

$P(\mathbf{x})$**:** positive Grundwahrheit;
das Prädikat an der angegebenen Stelle gilt auf jeden Fall.

$P_1(\mathbf{x}), \ldots, P_n(\mathbf{y}) \Rightarrow \Box$**:** Fehlerklausel;
diese Prädikate dürfen in dieser Kombination nicht wahr sein.

Die freien Prädikate von φ werden immer in den Bedingungsteil der Implikationen aufgenommen, da sie bei der Entscheidung, ob es erfüllende Relationen $\mathbf{R}$ gibt, als fest angesehen werden müssen.

Die Klasse der Hornformeln, die eine Nachfolgerrelation verwenden, wird als $\exists^2\text{HORN}[=, \mathrm{succ}]$ bezeichnet.

2.4 Totale Ordnungen

Wie schon bei der Definition der Signaturen angedeutet, stehen die Ordnung des Universums und das Prädikat, das sie darstellt, hier im Mittelpunkt des Interesses. Dies liegt nicht zuletzt daran, daß ohne sie die Repräsentation der zeitlichen Abfolge der Berechnung einer Turingmaschine in der Logik kaum vorstellbar ist. Dementsprechend sollen hier zunächst der Begriff der Ordnung genau geklärt und einige Eigenschaften aufgezeigt werden.

Definition 22 (Ordnungsrelation)
*Eine zweistellige Relation φ heißt eine **totale Ordnung**, wenn sie folgende Formel erfüllt:*

$$\begin{aligned} \forall x \forall y \forall z \quad & (\varphi(x,y) \vee \varphi(y,x) \vee (x = y)) && (2.1) \\ \wedge \quad & (\neg\varphi(x,y) \vee \neg\varphi(y,x)) && (2.2) \\ \wedge \quad & \neg\varphi(x,x) && (2.3) \\ \wedge \quad & ((\varphi(x,y) \wedge \varphi(y,z)) \to \varphi(x,z)) && (2.4) \end{aligned}$$

Bemerkung: Diese Definition durch eine logische Formel gibt zugleich eine Formel erster Stufe an, die überprüft, ob ein Prädikat eine Ordnungsrelation darstellt. Für sie steht auch der Modifikator $(\mathrm{Ordnung}\ o)$, der überprüft, ob es sich bei der Relation o um eine Ordnung handelt. Dies ist ein interessantes Strukturmittel für Zählklassen und für Klassen, die in der ersten Ordnung keinen Allquantor erlauben.

Der folgende Satz zeigt, daß diese Definition eine Ordnungsrelation im wesentlichen eindeutig festlegt.

Satz 23
Die Automorphismengruppe einer totalen Ordnung über einem endlichen Universum besteht nur aus der Identität.

Beweis: Induktion über die Mächtigkeit $n = ||\mathcal{A}||$ des Universums.

Induktionsanfang: Da über der leeren Menge nur die leere Abbildung besteht, und diese als Identität zu verstehen ist, gilt die Aussage.

Induktionsschritt: Die Aussage gelte für Universen der Mächtigkeit n. Sei $\mathcal{A} = (A, \varphi)$ eine Ordnungsrelation mit $|A| = n + 1$. Dann gibt es ein $\alpha \in A$ für das $\forall x((x = \alpha) \vee \varphi(\alpha, x))$ gilt, da sonst $\forall x\, \exists y\, ((x \neq y) \wedge \varphi(x, y))$ gelten würde. Von einem beliebigen Element x ausgehend, gibt es dann eine Folge $(x_i)_i$ mit $x = x_0$, wobei für alle i gilt $\varphi(x_i, x_{i+1})$. Da das Universum endlich ist, gibt es dem Schubfachprinzip gemäß $i \neq j$ mit $x_i = x_j$. Mit der Transitivität von φ folgt $\varphi(x_i, x_j)$ im Widerspruch zu 2.3.

Sei $\mu\colon A \to A$ ein beliebiger Automorphismus. Dann folgt $\mu(\alpha) = \alpha$ aus $\forall x(x = \mu(\alpha)) \vee \varphi(\mu(\alpha), x)$. Die Widerspruchsannahme $\mu(\alpha) \neq \alpha$ führt zu $\varphi(\mu(\alpha), \alpha)$ und $\varphi(\alpha, \mu(\alpha))$, was mit 2.3 nicht zu vereinbaren ist.

Mit der Induktionsvoraussetzung und Lemma 16 folgt, daß $\mu|_{A\setminus\{\alpha\}}$ nur die Identität sein kann. Somit gilt die Induktionsvoraussetzung für $n + 1$, und damit die Aussage des Satzes. ■

Definition 24
Ein Tripel $(\mathrm{succ}, \alpha, \omega)$ *mit einer binären Relation* succ *und zwei unären Relationen* α *und* ω *heißt eine* ***Nachfolgerrelation****, wenn sie die folgende Formel erfüllt:*

$$\begin{array}{cll} & \forall x \forall a \forall o & \alpha(a) \wedge \omega(o) \to (\neg\sigma(x, a) \wedge \neg\sigma(o, x)) \\ \wedge & \forall x\, \exists y & \omega(x) \vee \sigma(x, y) \\ \wedge & \forall x \forall y \forall z & \sigma(x, y) \wedge \sigma(x, z) \to y = z \end{array}$$

Definition 25 (passende Nachfolgerrelation)
Eine totale Ordnung φ *und eine Nachfolgerrelation* (σ, α, ω) ***passen zusammen****, falls gilt*

$$\forall x \forall y \quad \sigma(x, y) \to \varphi(x, y).$$

Lemma 26
Für jede totale Ordnungsrelation gibt es genau eine passende Nachfolgerrelation und umgekehrt. Die zu einer Ordnung passende Nachfolgerrelation kann mit Hilfe einer $\exists^2\Pi_2[<]$*-Formel zur Verfügung gestellt werden.*

Beweis: Die (nicht reflexive) transitive Hülle der Nachfolgerrelation ist eine totale Ordnung. Der eindeutige irreduzible Kern der Ordnungsrelation ist eine Nachfolgerrelation. Die gewünschte Formel ergibt sich direkt aus den letzten beiden Definitionen. ■

Lemma 27 (Successor über k Variablen)
Seien $\mathcal{A} = (A, \mathrm{succ}, \alpha, \omega)$ eine Struktur und $(\mathrm{succ}, \alpha, \omega)$ eine Nachfolgerrelation. Dann liefert folgende induktive Definition für alle $k \in \mathbb{N}$, $k > 1$ eine Nachfolgerrelation S_k über A^k ([Pap94], S. 174):

$$\begin{aligned}
\forall x \forall y\, S_1(x,y) &\leftrightarrow \mathrm{succ}(x,y) \\
\forall x_1 \cdots \forall x_k \forall y_1 \cdots \forall y_k\, S_k(x_1,\ldots,x_k,y_1,\ldots,y_k) &\leftrightarrow \\
&\Big(\mathrm{succ}(x_k,y_k) \wedge (x_1,\ldots,x_{k-1}) = (y_1,\ldots,y_{k-1})\Big) \vee \\
&\vee \Big(\omega(x_k) \wedge \alpha(y_k) \wedge S_{k-1}(x_1,\ldots x_{k-1},y_1,\ldots y_{k-1})\Big)
\end{aligned}$$

Beweis: Die Eigenschaften der Definition vererben sich im Fall der ersten Zeile direkt von succ, in der zweiten Zeile wird ein Übertrag in die bereits definierten S_{k-1} realisiert, so daß auch hier der Nachfolger eindeutig festgelegt wird. ■

Satz 28 (Horn-Successor über k Variablen)
Sei $\mathcal{A} = (A, \mathrm{succ}, \alpha, \omega)$ eine Struktur, wobei $(\mathrm{succ}, \alpha, \omega)$ eine Nachfolgerrelation ist. Dann gibt es für alle $k \in \mathbb{N}$ eine Menge von Hornklauseln, die genau dann erfüllt ist, wenn für alle $i \leq n$ die S_i die nach obiger Konvention festgelegten Nachfolgerrelationen für A^i sind.

Beweis: Es genügt, Hornklauseln anzugeben, die die Formeln in Lemma 27 erzwingen, die also genau dann erfüllt werden, wenn die S_k dieser Definition entsprechen.

Dies gelingt mit den folgenden, über alle vorkommenden Variablen allquantifizierten Implikationen:

$$\begin{aligned}
\mathrm{succ}(x,y) &\Rightarrow S_1(x,y) \\
\mathrm{succ}(x_k,y_k) \wedge (x_1 = y_1) \wedge \cdots \wedge (x_{k-1} = y_{k-1}) &\Rightarrow S_k(x_1,\ldots,x_k,y_1,\ldots,y_k) \\
\omega(x_k) \wedge \alpha(y_k) \wedge S_{k-1}(x_1,\ldots x_{k-1},y_1,\ldots y_{k-1}) &\Rightarrow S_k(x_1,\ldots,x_k,y_1,\ldots,y_k) \\
y_i \neq z_i \wedge S_k(\mathbf{x},\mathbf{y}) \wedge S_k(\mathbf{x},\mathbf{z}) &\Rightarrow \Box
\end{aligned}$$

Die letzte Zeile entspricht also k Klauseln, für jedes $1 \leq i \leq k$ eine. Alle diese zusammen entsprechen $\mathbf{y} \neq \mathbf{z} \wedge S_k(\mathbf{x},\mathbf{y}) \wedge S_k(\mathbf{x},\mathbf{z}) \Rightarrow \Box$.

Dem funktionalen Charakter der Nachfolgerrelation entsprechend genügt es sicherzustellen, daß alle benötigten Elemente vorkommen (Zeilen 1–3) und alle Funktionswerte eindeutig festgelegt sind (letzte Zeile). ■

2.5 Allgemeine Quantoren zweiter Stufe

Eine Motivation für diese Arbeit war, die in Thakurs Dissertation vorgestellten Ergebnisse in einen einheitlichen definitorischen Rahmen zu bringen. Eine solche Möglichkeit ist ein erweitertes Quantorenkonzept, in dem diese Ergebnisse in einem etwas allgemeineren Zusammenhang darstellbar sind, und Gemeinsamkeiten der verschiedenen Resultate sichtbar werden.

Da diese Quantoren direkt auf die Art der Probleme zugeschnitten sind, werden sie erst in den jeweiligen Kapiteln angegeben, hier wird zunächst nur der gemeinsame definitorische Rahmen gegeben.

Hier treten erstmals nichttriviale Bereiche der Form $|A|^k \cdot 2^{|A|^l}$ auf.

Definition 29 (Zählquantor)
Sei $\mathcal{A} = (A, I)$ eine Struktur, φ eine passende Formel. Die Formel $\#x\,\varphi(x)$ mit einer Variablen x der ersten Stufe hat den Wert

$$\mathrm{val}\,(\#x\,\varphi(x)) = \sum_{a \in A} \mathrm{val}\,\varphi(a)\ ,$$

die Formel $\#R\,\varphi(R)$ mit einer Variablen R der zweiten Stufe mit Stelligkeit $l = r(R)$ hat den Wert

$$\mathrm{val}\,(\#R\,\varphi(R)) = \sum_{R \subseteq A^l} \mathrm{val}\,\varphi(R)\ .$$

Dabei vergrößert sich der Bereich der Formel um den Faktor $|A|$ im ersten Fall, im zweiten Fall um den Faktor $2^{|A|^l}$.

Dieser Quantor stellt für Formeln φ mit Bereich 1 tatsächlich ein Zählen dar: Durch die Summation wird genau die Anzahl der erfüllenden Belegungen für R gezählt.

Lemma 30
Zählquantoren, auch verschiedener Stufen, können miteinander vertauscht werden:

$$\mathrm{val}(\#X\#Y\,\varphi) \quad \equiv \quad \mathrm{val}(\#Y\#X\,\varphi)$$

Beweis: Dies ergibt sich direkt daraus, daß endliche Summationen vertauschbar sind. ∎

Dadurch ist eine Blockbildung zu Vektoren von Prädikaten und von Variablen erster Stufe gerechtfertigt.

Definition 31
Quantoren $\min^2_R \varphi(R)$ *und* $\max^2_R \varphi(R)$ *bilden aus einer Funktion über* R *den entsprechenden Wert.*

$$\mathrm{val}(\min_R{}^2 \varphi(R)) := \min\{\mathrm{val}\,\varphi(R) \mid R \subseteq A^k\}$$

$$\mathrm{val}(\max_R{}^2 \varphi(R)) := \max\{\mathrm{val}\,\varphi(R) \mid R \subseteq A^k\}$$

Der Bereich dieser Formeln bleibt unverändert.

Definition 32
Der Modifikator $|R|_M\, \varphi(R)$ *hat den Wert* M*, falls* $\varphi(R)$ *den Wert 0 hat, ansonsten hat die Formel den Wert „Mächtigkeit der Relation* R *als Menge".*

$$\mathrm{val}\,(|R|_M \varphi(R)) := \begin{cases} M & \textit{falls}\ \mathrm{val}\,\varphi(R) = 0 \\ |R| & \textit{falls}\ \mathrm{val}\,\varphi(R) \neq 0 \end{cases}$$

Der Bereich ergibt sich zu n^k*, wobei* k *für die Stelligkeit von* R *und* n *für die Größe des Universums steht.* $|\cdot|$ *steht für* $|\cdot|_0$*, und* $|\cdot|_\infty$ *steht für* $|\cdot|_N$*, wobei* N *der aktuelle Bereich ist.*

Ein existentieller Quantor $\exists X$ ist darstellbar als $\max^2_X$, ein Allquantor $\forall X$ ergibt sich als $\min^2_X$, ein Quantor eindeutiger Existenz $\exists! X$ als das hier nur angedeutete $(=1)\#X$. Die entsprechenden bekannten Darstellungen sind daher als Abkürzungen für Formeln zu verstehen, die nur aus $\max^2$, $\min^2$ und den anderen hier vorgestellten Quantoren und Modifikatoren gebildet sind.

3

Logik und Berechnung

In der klassischen Logik wird normalerweise die Frage gestellt „Ist der Satz φ gültig?“. Die andere Frage „Gibt es ein Modell für einen Satz φ?“ ist einer Berechnung schon etwas näher, wie sich im folgenden zeigt.

3.1 Formeln zum Entscheiden von Problemen

Definition 33
Unter dem ***generalisierten Spektrum eines Satzes*** φ *wird die Menge der passenden Strukturen verstanden, die ein gültiges Modell für den Ausdruck bilden.*

$$\mathcal{S}(\varphi) = \{\, \mathcal{A} \text{ paßt zu } \varphi \ \mid \mathcal{A} \models \varphi \,\}$$

Je nachdem, welche syntaktischen Möglichkeiten für φ zugelassen werden, können unterschiedlich schwierige Probleme charakterisiert werden. So ergeben sich syntaktisch definierte Klassen von generalisierten Spektren.

Die hier vorgestellten Klassen haben, ähnlich den durch Turingmaschinen definierten, die Eigenschaft, daß endlich viele Ausnahmen in jedem Fall gesondert behandelt werden können.

Satz 34
Sei φ der Rumpf einer Hornformel mit Nachfolgerrelation, für die (mit endlich vielen Ausnahmen) für eine Menge L von Strukturen über einer Signatur gilt:

$$\mathcal{I} \in L \ \Leftrightarrow_{\text{fast überall}} \mathcal{I} \models \exists \mathbf{R}\, \forall \mathbf{x}\, \varphi$$

Dann gibt es ein ϑ, das Rumpf einer Horn Formel ist, für das ohne Ausnahmen gilt:

$$\mathcal{I} \in L \ \Leftrightarrow \ \mathcal{I} \models \exists \mathbf{R}\, \exists \mathbf{A}\, \forall \mathbf{x}\, \vartheta$$

Dabei ist, falls die Formel erfüllt wird, $\mathbf{A}$ eindeutig festgelegt.

Beweis: Da φ nur für endlich viele Ausnahmen falsch ist, gibt es ein $M \in \mathbb{N}$, so daß für alle Ausnahmen $\mathcal{I}$ die Größe des Universums $||\mathcal{I}|| < M$ ist. Es werden M einstellige Prädikate A_i eingeführt, die, falls das Universum nicht zu klein ist, die ersten M Elemente des Universums numerieren, indem A_i genau für das i-te Element gilt. Dementsprechend wird A_1 mit α identifiziert. Folgende Klauseln garantieren, daß diese Prädikate eindeutig so festgelegt werden: ($1 \leq i \leq M$)

$$\begin{aligned} A_i(x) \wedge \mathrm{succ}(x,y) &\Rightarrow A_{i+1}(y) \\ x \neq y \wedge A_i(x) \wedge A_i(y) &\Rightarrow \square \end{aligned}$$

Für jede Struktur der Größe $k < M$, die nicht in L ist, wird die Klausel

$$A_k(x_k) \wedge \omega(x_k) \wedge_{i \leq k} A_i(x_i) \wedge \neg\mathcal{I} \quad \Rightarrow \quad \square$$

hinzugenommen, wobei $\neg\mathcal{I}$ für die Beschreibung der Struktur steht. Darin sind alle Prädikate – mit allen Kombinationen der x_i eingesetzt – enthalten. Diese Literale werden genau dann negiert, wenn sie in der entsprechenden Struktur nicht gelten. Dies ist auch in Hornklauseln erlaubt, da es sich um Eingaben handelt.

Die Klauseln von φ der Form

$$\gamma \quad \Rightarrow \quad \delta$$

werden in Klauseln der Form

$$A_M(x_M) \wedge \gamma \quad \Rightarrow \quad \delta$$

umgewandelt.

Falls $k = ||\mathcal{I}|| < M$ und $\mathcal{I} \in L$ gilt, ist $A_M(x)$ für alle x falsch, die Klauseln, die aus φ entstandenen sind, sind automatisch richtig, sowie alle Klauseln, die $A_k(x_k) \wedge \omega(x_k)$ nicht enthalten. Da die anderen, die diese Bedingung enthalten, nur andere Strukturen ausschließen, gilt die Behauptung. Falls $\mathcal{I} \notin L$ gilt, gibt es eine entsprechende Klausel, die dafür sorgt, daß die Behauptung gilt.

Ist $k = ||\mathcal{I}|| \geq M$, dann gilt $A_M(x_M)$ für ein x_M. Da alle anderen Variablen unabhängig davon all-quantifiziert sind, erzwingen die Klauseln dieselben Belegungen der Prädikate wie in φ. Dies ist genau dann konsistent möglich, wenn dies in φ möglich war. ∎

Um durch Turingmaschinen definierte Komplexitätsklassen mit derartigen generalisierten Spektren vergleichbar zu machen, muß die Eingabe einer Struktur in eine solche Maschine festgelegt werden.

Definition 35 (Standardkodierung)
Sei $\mathcal{A} = (A, \mathbf{P}) = (A, P_1, \ldots P_k)$ eine Eingabestruktur mit $A = \{1, \ldots, n\}$, das Universum ist also geordnet. Die P_i haben alle die Stelligkeit d. Falls

dies nicht so ist, werden sie durch ein kartesisches Produkt mit dem Universum A an den hinteren Stellen ersetzt.

$$\forall \mathbf{z}\, P(x_1, \ldots, x_i) \leftrightarrow P(x_1, \ldots, x_i, \mathbf{z})$$

Dann versteht man unter ***Standardkodierung*** *den Bitstring $w = e(\mathcal{A})$, der folgende Eigenschaften hat:*

- *Der String w ist in Blöcke der Länge n^d aufgeteilt. Jeder solche Block kann ein P_i darstellen.*
- *Der erste Block besteht aus $(1^n 0^{n^2-n})^{n^{d-2}}$, also im wesentlichen einer unären Darstellung der Größe des Universums.*
- *Dann steht für jede der Relationen ein Block. Dieser ist an der Stelle $\sum_j x_j \cdot n^{j-1}$ genau dann 1, wenn $P_i'(x_1, \ldots, x_d)$ also $P_i(x_1, \ldots, x_{r(i)})$ gilt. Dies entspricht der Interpretation als Zahl im n-System mit niederwertiger Stelle links.*

Die Länge der Standardkodierung ist also durch ein Polynom in der Größe des Universums beschränkt, so daß in den hier betrachteten Klassen beides als Eingabelänge äquivalent ist. Es gibt also ein Polynom p, für das $|e(\mathcal{A})| \leq p(|\mathcal{A}|)$ gilt.

Außerdem können an eine solche Standardkodierung beliebige Zeichen angehängt werden, ohne die Information zu verändern, da die Länge der relevanten Information sich bereits aus dem ersten Wechsel von 1 auf 0 errechnen läßt. Dies ermöglicht es insbesondere diese auf das unendlich lange Arbeitsband einer Turingmaschine über $\{\sqcup, 1\}$ zu kodieren. Bei dieser Konvention handelt es sich nur um ein technisches Detail, man könnte mit beinahe denselben Beweisen auch ein Eingabealphabet $\{0, 1, \sqcup\}$ verwenden.

Für Strukturen über einer Signatur mit angeordnetem Universum und lediglich einer einstelligen Relation kann diese auch direkt als String verwendet werden. Dies könnte ohne weiteres als Spezialfall in obige Definition mit aufgenommen werden. Dieses Vorgehen wird als Struktur $c(I)$ für einen solchen String geschrieben. Aus technischen Gründen besteht diese Struktur sogar aus zwei Prädikaten, eines für die 0 und ein weiteres für die 1. Dies kann auf Strings mit einer Länge n^k ausgedehnt werden, indem k-stellige Prädikate und die lexikographische Anordnung der Tupel verwendet werden.

Definition 36
Sei C eine Komplexitätsklasse von Funktionen $A^* \to \mathbb{N}$ für endliche Alphabete A. Dies umfaßt die formalen Sprachen als charakteristische Funktionen. Dann ist die Klasse $\hat{C}$ von Funktionen, die Strukturen auf natürliche

Zahlen abbilden, folgendermaßen definiert: eine solche Funktion s ist in $\hat{C}$ genau dann, wenn es eine Funktion $f \in C$ gibt, die folgende Eigenschaften hat:

- *Eine Struktur hat denselben Wert wie ihre Standardkodierung:* $s = f \circ e$.
- *Für Strings, die keine Standardkodierung sind, gilt* $f \equiv 0$.
- *Für Standardkodierungen x und y isomorpher Strukturen gilt* $f(x) = f(y)$.

Im folgenden werden die Komplexitätsklassen C und $\hat{C}$ nur als C geschrieben, da aus dem Zusammenhang klar ist, welche gemeint ist, und hier auch die Maschinenklassen auf Standardkodierungen als Eingaben eingeschränkt werden.

Mit diesen Grundlagen ist es unter anderem möglich, die Klassen $\mathbf{P}$ und $\mathbf{NP}$ logisch zu charakterisieren. Die vorgestellten Beweise sind solchen aus [Pap94] nachempfunden. Dabei ist das Ergebnis für $\mathbf{P}$ nicht typisch für diese Arbeit, entsteht aber bei dem eingeschlagenen Beweisweg fast umsonst und wird daher auch exakt formuliert. Dieses Ergebnis geht eigentlich auf Grädel in [Grä92] zurück, dort wird es aber von einer ganz anderen Seite her bewiesen.

Folgendes Lemma stützt sich auf eine Idee aus [Pap94] und ist als technische Grundlage für andere Ergebnisse dieses Kapitels zu verstehen.

Lemma 37
Ein beliebiges als Formel der Prädikatenlogik erster Stufe darstellbares Entscheidungsproblem liegt in $\mathbf{P}$*:*

$$e(\mathrm{FO}[<]) \subset \mathbf{P}$$

Beweis: Da die Formel nicht zur Eingabe gehört, sind alle äquivalenten Formeln der Prädikatenlogik erster Stufe gleichwertig. O.E. sei also die Formel in Pränex-Normalform mit k quantifizierten Variablen gegeben. Dann gibt es n^k mögliche Belegungen für diese Variablen, wobei n die Größe des Universums angibt. Die Eingabe, in der die Prädikate dargestellt werden, ist mindestens n Zeichen lang.

Es wird also eine entsprechende Tabelle angelegt. Für jeden Eintrag muß im wesentlichen einmal die Eingabe gelesen, die aktuell betrachteten konstant vielen Wahrheitswerte gemerkt und gemäß der jetzt aussagenlogischen Formel ausgewertet werden. Dann muß diese Tabelle gemäß der

Quantorenstruktur zusammengefaßt werden. Dazu muß für jeden Quantor höchstens die ganze Tabelle einmal ausgelesen werden. Somit ist die Laufzeit dieser deterministischen Turingmaschine polynomiell beschränkt. ■

3.2 Turingmaschinen

Aus technischen Gründen wird hier ein etwas ungewöhnliches Maschinenmodell eingeführt, das aber zumindest für die vorkommenden Komplexitätsklassen vollkommen äquivalent zu den üblichen Definitionen ist. Auch die Einschränkung auf Standardkodierungen verschwindet durch eine berechnungsschwache Reduktion. Daher wird im folgenden davon ausgegangen, daß die Maschine auf solchen Standardkodierungen arbeitet.

Die hier verwendeten Turingmaschinen entsprechen den üblichen Definitionen, wie sie etwa in [Pap94] exakt angegeben sind, mit den folgenden Modifikationen:

Eine **Turingmaschine** ($\mathbf{TM}$) hat ein Arbeitsband und ein Auswahlband.

Das Arbeitsband ist einseitig beschränkt, hat ein beliebiges Alphabet, am linken Ende steht die Eingabe, eine Standardkodierung über dem Alphabet $\{\sqcup, 1\}$.

Zu Beginn der Berechnung steht der Arbeitskopf links von der Eingabe auf dem Symbol $\triangleright$, das dann dort unverändert stehen bleibt und nicht noch einmal besucht wird.

Auf dem Auswahlband steht ein Wort aus $\{0,1\}^*$. In jedem Schritt entscheidet das aktuelle Bit des Auswahlbandes, welcher Befehlssatz verwendet wird. Dies entspricht der Möglichkeit einer nichtdeterministischen Auswahl, bei der in jedem Schritt genau zwei Alternativen zur Verfügung stehen.

Wenn das Ende des Auswahlbandes erreicht ist, hält die Maschine. Der Inhalt des Arbeitsbandes wird als Ergebnis interpretiert. Es handelt sich dabei um einen Bitstring der Länge n^k, wobei ein $\sqcup$ als 0 und alle anderen Zeichen als 1 interpretiert werden.

In der Notation entspricht $B(x,y)$ dem Ausgabestring der Turingmaschine B bei Eingabe x und Auswahlbandinhalt y.

In diesem Rahmen kann man einige bekannte Komplexitätsklassen darstellen. Dabei ist zu beachten, daß dies nur für Eingaben gilt, die mindestens die Länge zwei haben. Da es sich bei den Eingaben um Standardkodierungen handelt, und diese für nichtleeres Universum mindestens Länge zwei haben, ergibt sich keine Einschränkung.

Dabei ist zahl die Abbildung der Strings in die natürlichen Zahlen, die den String mit niederwertigster Stelle links binär interpretiert.

NP Die Klasse der in polynomieller Zeit nichtdeterministisch berechenbaren Entscheidungsprobleme $A \subseteq \{0,1\}^*$.

$$A \in \mathbf{NP} \Leftrightarrow \text{es gibt } k, \mathbf{TM} \text{ B:} \quad x \in A \leftrightarrow \exists y \in \{0,1\}^{|x|^k}\, B(x,y) \in \{0\}^*$$

P Die Klasse der in polynomieller Zeit deterministisch berechenbaren Entscheidungsprobleme $A \subseteq \{0,1\}^*$.

$$A \in \mathbf{P} \;\Leftrightarrow\; \text{es gibt } k, \mathbf{TM} \text{ B:} \quad x \in A \;\leftrightarrow\; B(x, 0^{|x|^k}) \in \{0\}^*$$

#P Die Klasse der Funktionen $f\colon \{0,1\}^* \to \mathbb{N}$, die durch die Anzahl akzeptierender Pfade einer polynomialzeitbeschränkten nichtdeterministischen Berechnung darstellbar sind.

$$f \in \#\mathbf{P} \;\Leftrightarrow\; \text{es gibt } k, \mathbf{TM} \text{ B: } f(x) = \left|\left\{ y \in \{0,1\}^{|x|^k} \colon B(x,y) \in \{0\}^* \right\}\right|$$

optP Die Klasse der Funktionen $f\colon \{0,1\}^* \to \mathbb{N}$, die als Optimum der binär interpretierten Ausgaben aller akzeptierenden Pfade einer polynomialzeitbeschränkten nichtdeterministischen Berechnung darstellbar sind:

$$f \in \mathbf{optP} \;\Leftrightarrow\; \text{es gibt } k, \mathbf{TM} \text{ B: } f(x) = \operatorname*{opt}_{y \in \{0,1\}^{|x|^k}} \mathrm{zahl}(B(x,y))$$

mit $\mathrm{opt} \in \{\min, \max\}$.

optPB Die Klasse der Funktionen $f\colon \{0,1\}^* \to \mathbb{N}$, die durch das Optimum der unär interpretierten Ausgaben aller akzeptierenden Pfade einer polynomialzeitbeschränkten nichtdeterministischen Berechnung darstellbar sind:

$$f \in \mathbf{optPB} \;\Leftrightarrow\; \text{es gibt } k, \mathbf{TM} \text{ B: } f(x) = \operatorname*{opt}_{y \in \{0,1\}^{|x|^k}} |B(x,y)|_1$$

mit $\mathrm{opt} \in \{\min, \max\}$. Durch die unäre Ausgabe, der Anzahl von Einsen, die in der Ausgabe vorkommen, können genau die Funktionen aus **optP** dargestellt werden, deren Funktionswerte durch ein Polynom in der Eingabelänge beschränkt sind.

3.3 Berechnungstafel

Die Berechnung einer zeitbeschränkten Turingmaschine kann auch durch eine Berechnungstafel charakterisiert werden. Die hier verwendete Konstruktion ist [Pap94], Seite 166, nachempfunden.

Für ein festes $k \in \mathbb{N}$ bezeichnet man als polynomiell große Berechnungstafel ein Feld der Größe $n^k \times n^k$, das auf jeder Position durch ein Alphabetszeichen und möglicherweise einen Zustand belegt ist.

Interessant sind Berechnungstafeln, die gültige Berechnungen von Turingmaschinen darstellen.

Dazu muß der linke Rand mit $\triangleright$, der rechte Rand wegen der Zeitschranke für die Berechnung nur mit $\sqcup$ belegt sein.

Außerdem muß die erste Zeile die Eingabe wiedergeben. Es muß dort also die Standardkodierung der Eingabestruktur stehen. Diese beginnt auf dem zweiten Feld, die übrigen Felder bis zum rechten Rand sind mit $\sqcup$ belegt.

In dieser ersten Zeile hat lediglich das $\triangleright$ in der ersten Spalte auch den Startzustand, die übrigen Felder enthalten nur Alphabetszeichen.

Je zwei aufeinanderfolgende Zeilen müssen durch einen korrekten und dem Auswahlband entsprechenden Zustandsübergang der Turingmaschine auseinander hervorgehen. Da eine solche Maschine das Arbeitsband nur an der aktuellen Kopfposition verändern kann, läßt sich dies lokal überprüfen.

Der untere Rand enthält ab der zweiten Position den Ausgabestring.

Da die Berechnungstafel durch Eingabe und Auswahlband eindeutig festgelegt ist, überrascht es nicht, daß es möglich ist, die Korrektheit einer Berechnungstafel durch Hornklauseln auszudrücken.

Eine derartige Eindeutigkeit wird auch für andere Ergebnisse wichtig sein. Der Versuch, dies auszudrücken führt zu einer Schwierigkeit in der Notation. Zum einen ist es wünschenswert, das Prädikat, das – falls es existiert – eindeutig bestimmt ist, als solches durch den Quantor zu kennzeichnen. Auf der anderen Seite gibt es den sinnvoll und wohldefinierten Quantor der eindeutigen Existenz. Da die Eindeutigkeit des Prädikates durch die konstruierte Formel erzwungen wird, ist man in der Situation, daß

$$(\exists R\,\varphi) \equiv (\exists! R\,\varphi)$$

gilt. Dieser Sachverhalt soll im folgenden durch die abkürzende Schreibweise

$$(\exists{=}\exists!) R\,\varphi$$

dargestellt werden. Dieser Unterschied zu den anderen hier verwendeten ungewöhnlichen Operatoren und Modifikatoren sollte nicht übersehen werden und könnte zu Unklarheiten Anlaß geben. Da aber andererseits die redundante Wiederholung des gemeinsamen Formelteiles φ die Übersichtlichkeit der Darstellung stark einschränkt, erscheint diese abkürzende Darstellung gerechtfertigt.

3.4 Berechnungstafel in Hornklauseln

Im folgenden wird eine Menge von Hornklauseln vorgestellt, die genau von solchen Strukturen erfüllt wird, die eine gültige polynomielle Berechnungstafel darstellen. Dabei wird bereits durch die Kodierung des Auswahlbandes eine Rechenzeit von n^k festgelegt.

Um Schwierigkeiten mit der Definition von Hornformeln zu vermeiden, wird zum Kodieren eines Bitstrings für Auswahlband und Ausgabe in einer Relation A jeweils ein Paar von Relationen A_0 und A_1 verwendet. $A_0(\mathbf{x})$ gilt, falls an der durch $\mathbf{x}$ kodierten Stelle des Strings eine 0 steht, entsprechend $A_1(\mathbf{x})$, falls dort eine 1 steht. Durch den Begriff des Kodierens wird die Konsistenz der beiden Relationen zugesichert.

Lemma 38
Für jede Turingmaschine B gibt es eine Hornformel $\varphi \in \mathrm{FO}[\mathrm{succ}]$ über einer passenden Signatur mit Nachfolgerrelation, so daß für alle $\mathcal{I}$ und für alle y und z gilt:

$$z = B(e(\mathcal{I}), y) \qquad \Leftrightarrow \; (\mathcal{I}, c(y), c(z)) \models (\exists{=}\exists!)\mathbf{R}\,\forall\mathbf{u}\,\varphi$$

Beweis: Die im folgenden angegebenen Hornklauseln werden der Konvention entsprechend über alle vorkommenden Variablen der ersten Stufe all-quantifiziert. Dabei seien Auswahl und Ausgabe jeweils durch zwei k-stellige Relationen C_0 und C_1 sowie O_0 und O_1 dargestellt.

Zunächst muß man bis n^k zählen können. Dazu wird eine k-stellige Relation S_k verwendet, die die lexikographische Nachfolgerrelation auf k-Tupeln realisiert. Dies geschieht mit den in Satz 28 definierten Klauseln. Die entstehenden S_i für $i < k$ bleiben ungenutzt, werden aber zur induktiven Verifizierung bzw. Erzeugung von S_k benötigt.

Schließlich wird für jedes mögliche Symbol τ der Berechnungstafel eine $2k$-stellige Relation $T_\tau(\mathbf{x}, \mathbf{y})$ eingeführt, die genau dann wahr ist, wenn an der Stelle der Berechnungstafel, die durch die Tupel $\mathbf{x}$ und $\mathbf{y}$ festgelegt ist, das Symbol τ steht.

Der Prädikatenvektor $\mathbf{R}$ besteht aus den S_i, den T_τ und den später beschriebenen A_i.

Zunächst soll sichergestellt werden, daß die T_τ, falls sie eine Belegung der Berechnungstafel charakterisieren, diese eindeutig festlegen, also nicht an einer Stelle zwei verschiedene T_u gültig sind. Dazu werden für alle möglichen Kombinationen u, v von ungleichen Symbolen der Berechnungstafel die Klauseln

$$T_u(x_1, \ldots, x_k, y_1, \ldots, y_k) \wedge T_v(x_1, \ldots, x_k, y_1, \ldots, y_k) \quad \Rightarrow \quad \square$$

eingeführt. Damit müssen die übrigen Klauseln nur noch erzwingen, daß überall mindestens ein Symbol dargestellt wird, und daß diese Symbole eine für C_i, O_i und succ korrekte Berechnungstafel bilden.

Der Ausdruck in φ muß folgendes beschreiben:

1. Die Ränder der Berechnungstafel entsprechen der Eingabe und der Ausgabe.

2. Die gesamte Berechnungstafel entspricht der Übergangsfunktion der Maschine und dem Auswahlband.

Forderung 1 ist für den rechten Rand sehr leicht zu überprüfen. Wegen der Zeitschranke kann die Maschine die letzten Felder des Turingbandes nicht benutzen. Diese sind also sicher mit dem Zeichen $\sqcup$ belegt. Dies führt zu der Klausel

$$\omega(x) \quad \Rightarrow \quad T_\sqcup(x, \ldots, x, \mathbf{y})$$

Ähnlich einfach ist der linke Rand festgelegt. Hier darf ab der zweiten Zeile nur ein $\triangleright$ stehen, also

$$\alpha(x) \wedge S_k(\mathbf{y}', \mathbf{y}) \quad \Rightarrow \quad T_\triangleright(x, \ldots, x, \mathbf{y})$$

In der linken oberen Ecke stehen ein $\triangleright$ und der Startzustand S, also das Symbol $(\triangleright, S)$, daher die Klausel

$$\alpha(x) \quad \Rightarrow \quad T_{(\triangleright, S)}(x, \ldots, x, x, \ldots, x)\,.$$

In der ersten Zeile wird die Standardkodierung der Eingabestruktur als Bitstring in der Darstellung der T_τ zugesichert. Hier ist die ansonsten etwas unhandliche Standardkodierung recht praktisch, da sie einer Kodierung des Bitstrings, der eine Eingabestruktur darstellt, genau entspricht.

Dazu muß noch innerhalb des Universums bis zur Anzahl der Eingabeprädikate m gezählt werden können. Dies wird mit den Prädikaten A_i aus Satz 34 durch die entsprechenden Klauseln

$$\begin{aligned} A_i(x) \wedge \mathrm{succ}(x, y) &\Rightarrow A_{i+1}(y) \\ x \neq y \wedge A_i(x) \wedge A_i(y) &\Rightarrow \square \end{aligned}$$

realisiert, wobei A_1 für das Grundprädikat α steht. Mit diesen Prädikaten können nach Satz 34 auch solche Eingabestrukturen richtig behandelt werden, deren Universum nicht mächtiger als m ist.

Mit Satz 34 ist auch klar, daß diese Anforderung an die Größe des Universums keine Einschränkung bedeutet.

Sei d die maximale Stelligkeit der Eingabeprädikate. Bei den folgenden Formeln wird jeweils das S_k verwendet, um die beschriebene Position um 1 zu verschieben, was wegen der $\triangleright$ in der ersten Spalte nötig ist.

Der erste Block der Standardkodierung, der die Größe des Universums unär angibt, kann als zweistelliges Prädikat dargestellt werden. Dies geschieht durch $I_0(x_1, x_2) := \alpha(x_2)$, da dies genau für die ersten n Elemente der n^2 Paare (x_1, x_2) in lexikographischer Ordnung gilt. Der Konvention der Standardkodierung entsprechend wird dieser String so oft wiederholt, bis er einen ganzen n^d-Block ausfüllt. Daher wird dieses I_0 im weiteren als abkürzende Schreibweise zu den Eingabeprädikaten hinzugenommen.

Dementsprechend werden für die Eingabeprädikate $I_0, \ldots, I_m$ die Klauseln

$$\alpha(y) \wedge S_k(x_1, \ldots, x_d, a_i, a_1, \ldots, a_1, \mathbf{z}) \wedge I_i(x_1, \ldots, x_{r(i)}) \quad \Rightarrow \quad T_1(\mathbf{z}, y, \ldots, y)$$

und

$$\alpha(y) \wedge S_k(x_1, \ldots, x_d, a_i, a_1, \ldots, a_1, \mathbf{z}) \wedge \neg I_i(x_1, \ldots, x_{r(i)}) \quad \Rightarrow \quad T_\sqcup(\mathbf{z}, y, \ldots, y)$$

hinzugenommen, die erzwingen, daß der a_i-te Block von n^d Bits das Prädikat I_i darstellt.

Schließlich müssen noch die übrigen Felder mit $\sqcup$ belegt werden. Die Klauseln

$$\begin{aligned} &\alpha(y) \wedge A_1(a_1) \wedge \cdots \wedge A_m(a_m) \wedge x \neq a_1 \wedge \cdots \wedge x \neq a_m \\ &\qquad \wedge S_k(u_1, \ldots, u_d, x, y, \ldots, y, \mathbf{z}) \quad \Rightarrow \quad T_\sqcup(\mathbf{z}, y, \ldots, y)\,, \end{aligned}$$

und für alle $p > d + 1$

$$\alpha(y) \wedge \neg\alpha(u_p) \wedge S_k(u_1, \ldots, u_k, \mathbf{z}) \quad \Rightarrow \quad T_\sqcup(\mathbf{z}, y, \ldots, y)\,,$$

erreichen genau alle $(u_1, \ldots, u_k)$, die „größer" als $m \cdot n^l$ sind, da an der entscheidenden Position $x = u_{d+1} > a_m$ gilt.

Am unteren Rand wird die Übereinstimmung des Arbeitsbandes mit dem Ausgabeprädikat überprüft. Dazu stehen für alle β, die einen beliebigen Zustand und das Bandsymbol $\sqcup$ darstellen, die Klauseln

$$\alpha(a) \wedge \omega(y) \wedge S_k(x, a, \ldots, a, \mathbf{x}) \wedge T_\beta(x, \ldots, x, y, \ldots, y) \quad \Rightarrow \quad O_0(x) \, ,$$

und für alle übrigen γ die Klauseln

$$\alpha(a) \wedge \omega(y) \wedge S_k(x, a, \ldots, a, \mathbf{x}) \wedge T_\gamma(x, \ldots, x, y, \ldots, y) \quad \Rightarrow \quad O_1(x) \, .$$

Mit Eigenschaft 2 geht das Programm der Turingmaschine ein. Es muß lediglich überprüft werden, ob jeweils drei in einer Zeile aufeinanderfolgende Symbole u, v, w zu dem in der Mitte darunterstehenden Symbol t paßt. Dadurch, daß die Ränder schon festgelegt sind, wird so die Korrektheit der gesamten Berechnungstafel zugesichert.

Dies kann für alle Kombinationen von u, v, w statisch ausgewertet werden. Kombinationen, die nicht auftreten dürfen, weil auf zwei Feldern auch ein Zustand dargestellt wird, können unberücksichtigt bleiben, da sie durch die Konstruktion ausgeschlossen sind.

Für alle diese u, v, w gibt es dann in Abhängigkeit von dem Auswahlband ein t, das in der nächsten Zeile unter dem v stehen muß. Dies führt zu den Klauseln

$$\begin{aligned} &S_k(\mathbf{x}', \mathbf{x}'') \wedge S_k(\mathbf{x}'', \mathbf{x}''') \wedge S_k(\mathbf{y}, \mathbf{y}') \wedge T_u(\mathbf{x}', \mathbf{y}) \\ &\qquad\qquad \wedge T_v(\mathbf{x}'', \mathbf{y}) \wedge T_w(\mathbf{x}''', \mathbf{y}) \wedge C_i(\mathbf{x}) \quad \Rightarrow \quad T_t(\mathbf{x}'', \mathbf{y}') \, , \end{aligned}$$

mit denen zugesichert ist, daß in der nächsten Zeile mindestens für das passende t die Relation T_t erfüllt ist, und so an dieser Stelle die Berechnungstafel richtig dargestellt ist. Zusammen mit der ersten Sorte Klauseln ist so, induktiv von der ersten bis zur letzten Zeile der Berechnungstafel, sichergestellt, daß nur die gültige Berechnungstafel dargestellt werden kann. ■

3.5 Der Satz von Fagin

Satz 39 (Fagin: NP $= \exists^2\mathrm{FO}[<]$)
Die Klasse der prädikatenlogischen Formeln mit vorangestellten existentiellen Quantoren zweiter Stufe entspricht genau **NP***.*

Beweis: Für eine beliebige solche Formel hat das existentiell quantifizierte Prädikat zweiter Stufe endlich viele Stellen. Eine nichtdeterministische Turingmaschine rät die passenden Prädikate, und prüft dann, ob die prädikatenlogische Formel erster Ordnung dadurch zu einer Tautologie geworden ist. Dies ist nach Lemma 37 in polynomieller Zeit auf der Eingabe $e(\mathcal{J})$ möglich.

Sei andererseits eine Menge M von Strukturen über einer festen Signatur gegeben, so daß es eine Sprache $A \in \mathbf{NP}$ gibt, für die gilt

$$e(\mathcal{I}) \in A \quad \Leftrightarrow \quad \mathcal{I} \in M .$$

Um die Behauptung des Satzes zu zeigen, muß eine Formel ϑ angegeben werden, für die gilt

$$\mathcal{I} \models \exists \mathbf{R}\, \vartheta .$$

Sei also B die Turingmaschine, die A entscheidet, und φ die durch sie festgelegte Formel aus Lemma 38.

Die Formel

$$\tau := \Big(C_0(\mathbf{x}) \vee C_1(\mathbf{x})\Big) \wedge \Big(\neg C_0(\mathbf{x}) \vee \neg C_1(\mathbf{x})\Big) \wedge O_0(\mathbf{x}) \wedge \neg O_1(\mathbf{x})$$

garantiert, daß die C_c eine korrekte Darstellung des Auswahlbandes sind, und daß die Ausgabe aus $\{\sqcup\}^*$ ist, also auf dem entsprechenden Berechnungspfad akzeptiert wird.

Weiterhin sei ρ die Formel aus Definition 24, die zusichert, daß succ die zu den geratenen Prädikaten α und ω passende Nachfolgerrelation ist.

Dann gilt mit $\mathbf{R} = (C_0, C_1, O_0, O_1, \sigma, \mathbf{R}')$, wobei $\mathbf{R}'$ der Prädikatenvektor aus Lemma 38 ist, und $\vartheta := \tau \wedge \rho \wedge \varphi$ die obige Darstellung. Die entstandene Formel liegt also sogar in der Klasse $\exists^2\Pi_2$. ■

Bemerkung: Es ist auch möglich, einen gleichwertigen Ausdruck mit Einschränkung für die Stelligkeit der Prädikate auf k anzugeben, wenn das eingebaute Prädikat $+$ zur Verfügung steht [Imm87].

Satz 40 ($\mathbf{P} = \exists^2\mathrm{H{\small ORN}}[=, \mathrm{succ}]$)
Die Klasse der prädikatenlogischen Hornformeln mit Nachfolgerrelation entspricht genau $\mathbf{P}$.

Beweis: „$\supseteq$“: Eine Menge Q sei gegeben durch eine quantorenfreie Formel φ in KNF aus Hornklauseln vermöge

$$\mathcal{I} \in Q \;\Leftrightarrow\; (\mathcal{I}, \sigma) \models \exists \mathbf{R}\, \forall \mathbf{x}\, \varphi .$$

Dann akzeptiert folgender $\mathbf{P}$-Algorithmus die Menge Q:

1. Alle möglichen Kombinationen für $\mathbf{x}$ werden in φ substituiert, d.h. die Konjugation
$$\bigwedge_{\mathbf{a}\in I^k} \varphi[\mathbf{x} \mapsto \mathbf{a}]$$
wird gebildet. So entsteht eine Klauselmenge der Größe $O(n^k)$ mit Konstanten aus $|\mathcal{I}|$.

2. Alle Literale, außer denen aus $\mathbf{R}$, werden ausgewertet; die Klauselmenge entsprechend modifiziert, das heißt, falls ein Literal gilt, wird die Klausel entfernt, falls nicht, das Literal aus der Klausel. So entsteht eine Hornklauselmenge der Aussagenlogik. Falls die leere Klausel entsteht, gilt $\mathcal{I} \notin Q$, der Algorithmus endet.

3. Beginnend mit den Mengen für alle Relationen aus $\mathbf{R}$, die sich aus den positiven Grundwahrheiten ergeben, werden den Implikationen von φ gemäß entsprechende Tupel aufgenommen. Dies wird solange wiederholt, bis sich die Relationen nicht mehr ändern.

4. Die rein negativen Klauseln werden überprüft. Genau dann, wenn alle erfüllt sind, akzeptiert der Algorithmus die Eingabe.

In 3 wird eine kleinste Menge von Tupeln in den Relationen bestimmt, die alle Implikationen berücksichtigt. Jede erfüllende Belegung muß mindestens diese umfassen. Wird in 4 abgelehnt, kann es daher keine erfüllende Belegung geben. Wird dort akzeptiert, dann ist die so gefundene Belegung eine erfüllende. Da in 3 bei jedem Durchlauf eines der n^k möglichen Tupel in eine der Relationen aufgenommen wird, wird auch hier polynomielle Laufzeit erreicht.

„$\subseteq$“: Sei andererseits eine Menge Q von Strukturen über einer Signatur gegeben, so daß es eine Sprache $A \in \mathbf{P}$ gibt, für die gilt
$$e(\mathcal{I}) \in A \;\Leftrightarrow\; \mathcal{I} \in Q\,.$$

Um nun die Behauptung des Satzes zu zeigen, muß eine Formel ϑ aus der Klasse $\exists^2$Horn$[=, \mathrm{succ}]$ angegeben werden mit
$$\mathcal{I} \in Q \quad \Leftrightarrow \quad (\mathcal{I}, \sigma) \models \vartheta.$$

Sei also B die Turingmaschine, die A entscheidet, und φ die durch sie festgelegte Formel aus Lemma 38.

Dann garantiert die Formel
$$\tau := \forall \mathbf{x}\, C_0(\mathbf{x}) \wedge \neg C_1(\mathbf{x}) \wedge O_0(\mathbf{x}) \wedge \neg O_1(\mathbf{x})\,,$$
daß die C_c ein Auswahlband mit Inhalt 0^{n^k} repräsentieren, und daß die Ausgabe aus $\{0\}^*$ ist.

Also ist auch $\varphi \wedge \tau$ in Horn-Form darstellbar, somit gilt die Behauptung. ■

4

Ergebnisse in der Quantorendarstellung

Nachdem im letzten Kapitel die Verbindung zu bekannten Komplexitätsklassen aufgezeigt wurde, wird hier die innere Struktur der Klassen bezüglich der Quantoren erster Stufe weiter untersucht. Dies ist vor allem als Illustration der vorgestellten Quantoren und der so entstehenden Hierarchien zu verstehen.

Dabei werden einige bekannte Ergebnisse, die auch schon von Thakur in seiner Dissertation zusammengestellt wurden in den neuen Rahmen gebracht und um Betrachtungen über den Einfluß der als Grundstruktur zur Verfügung gestellten Ordnung erweitert.

Außerdem werden in der sich durch die Darstellung ergebenden Systematik bisher nicht untersuchte Hierarchien angegeben. Diese helfen die zunächst unerwartete Struktur der bekannten Hierarchien zu verstehen.

Zunächst ein technisches Ergebnis, das bereits eine typische Verwendung der Quantorendarstellung zeigt.

Beobachtung 41
Die Quantoren zweiter Stufe $\max^2$ *und* $\#$ *sind invariant unter eindeutiger Erweiterung.*

Sei φ *eine Formel mit Bereich 1 und* $(\exists{=}\exists!)R\,\varphi(R)$, *dann gilt*

$$\exists! R\,\varphi(R) \quad\equiv\quad \exists R\,\varphi(R) \quad\equiv\quad \# R\,\varphi(R) \quad\equiv\quad \max_R{}^2\ \varphi(R)\,.$$

Beweis: Klar nach Definition der Quantoren. ■

4.1 Fagin-Hierarchie

Die hier vorgestellte Hierarchie hat keine bekannten besonderen Eigenschaften, wegen der sie erforscht wäre. Vielmehr ergibt sie sich beim konsequenten Ausleuchten des hier eingegrenzten Gebietes als einfachste Situation, in der sich aber bereits die typischen Vorgehensweisen umsetzen lassen. Auf dieser Grundlage können dann die weiteren Ergebnisse leichter dargestellt werden.

Aus dem Beweis des Satzes von Fagin (Satz 39) geht durch die angegebene Formel hervor, daß in jedem Fall in der ersten Stufe eine Π_2-Formel genügt. Dieses Resultat läßt sich auch viel direkter durch eine Skolemisierung gewinnen. Da sehr ähnliche Überlegungen hinter allen hier vorgestellten Resultaten über kollabiernde Hierarchien stehen, soll dieses Argument im hier verwendeten definitorischen Rahmen exakt formuliert werden.

Dabei wird auch die Anwesenheit einer totalen Ordnung immer wieder verwendet werden. Es zeigt sich, daß diese nur dann sinnvoll verwendet werden kann, wenn tatsächlich gezählt wird.

Beobachtung 42
Es gilt

$$\exists^2 < (\mathrm{Ordnung}\ <)\, \Sigma_1 \subseteq \Sigma_1 \,.$$

Beweis: Sei $\exists^2 < (\mathrm{Ordnung}\ <)\, \exists \mathbf{x}\, \varphi$ eine entsprechende Formel. Dann kann φ in disjunktive Normalform gebracht werden, also $\varphi = \vee_{i=1}^{N} \varphi_i$, wobei jedes φ_i die Form $\varphi_i = A_{i,1} \wedge \cdots \wedge A_{i,N_i}$ hat. (Falls in einem φ_i ein Zyklus mit $<$ besteht, kann dieses φ_i entfallen.)

Seien in einem φ_i die $x_1, \ldots, x_n$ die Variablen, die in $<$ eingesetzt werden. Dann gibt es nur endlich viele Ordnungstypen, also Möglichkeiten, diesen Variablen Werte zuzuweisen, die bezüglich $<$ unterschieden werden können. Für jede solche generiere man eine Formel ψ_j, die erzwingt, daß genau die dem Ordnungstyp entsprechenden Gleichheiten gelten.

Behauptung: Die Formel $\vee_j \psi_j$ ist äquivalent zu φ. Falls φ erfüllbar war, dann gibt es auch ein erfülltes φ_i. Die dann existenten x_j waren angeordnet, also gibt es ein erfülltes ψ_j.

Sei ein so entstandenes ψ_j erfüllbar. Dann gibt es eine Ordnung, die die x_i der Entstehung des ψ_j entsprechend anordnet. Mit dieser Ordnung wird ein φ_i erfüllbar, also auch φ. ∎

Die hier vorgestellte Skolemisierung besteht darin, daß existentielle Quantoren der ersten Stufe mit Hilfe von in der zweiten Stufe existentiell-quantifizierten Prädikaten mit Allquantoren der ersten Stufe vertauscht werden können.

Es kann sogar erreicht werden, daß diese neuen Prädikate eindeutig bestimmt sind, falls es sie gibt. Diese können dann auch durch Zählquantoren ersetzt werden, ohne die Semantik zu verändern, so daß sich dieses Resultat auch in den später vorgestellten Zählklassen verwenden läßt.

Satz 43
Für jedes ganzzahlige $k \geq 0$ gelten

$$\Pi_{k+3}[<] \subseteq \exists^2 \Pi_{k+2}[<]$$

und

$$\Pi_{k+3} \subseteq \exists^2(\text{Funktion})\Pi_{k+1} ,$$

sowie

$$\Pi_{k+3}[<] \subseteq (\exists{=}\exists!)^2 \Pi_{k+2}[<] .$$

Beweis: Sei $\psi = \exists \mathbf{R}\, \forall \mathbf{x}\, \exists \mathbf{y}\, \forall \mathbf{z}\, \varphi(\mathbf{R}, \mathbf{x}, \mathbf{y}, \mathbf{z})$ eine Formel mit $\varphi \in \Sigma_k$, wobei $\mathbf{x}$, $\mathbf{y}$ und $\mathbf{z}$ als Variablenvektoren zu verstehen sind bzw. jeweils für mehrere durch gleichartige Quantoren gebundene Variablen stehen.

Dann ist bekanntermaßen die Reihenfolge der Quantoren und der durch sie gebundenen Variablen für die Semantik der Formel ausschlaggebend. So wird im allgemeinen die Formel mit dem Quantorenpräfix $\forall \mathbf{x}\, \forall \mathbf{z}\, \exists \mathbf{y}$ zwar von allen Strukturen erfüllt, die die Originalformel erfüllen, aber zusätzlich noch von anderen. Dies liegt daran, daß in der Originalformel zunächst ein beliebiges $\mathbf{x}$ festgelegt wird, und dann ein $\mathbf{y}$ existieren muß, das für alle $\mathbf{z}$ die Formel φ erfüllt. Bei der anderen Quantorenreihenfolge genügt es, wenn es für jedes $\mathbf{z}$ jeweils ein $\mathbf{y}$ gibt, das die Formel erfüllt. Diese zusätzliche Auswahlfreiheit kann durch ein Prädikat der zweiten Stufe ausgeglichen werden, indem für jedes $\mathbf{x}$ jeweils nur ein $\mathbf{y}$ das Prädikat erfüllt.

Daher sind

$$\exists P\, \forall \mathbf{x} \forall \mathbf{z} \forall \mathbf{y}' \forall \mathbf{y}''\, \exists \mathbf{y}\, P(\mathbf{x}, \mathbf{y}) \wedge \varphi(\mathbf{x}, \mathbf{y}, \mathbf{z}) \wedge \Big((P(\mathbf{x}, \mathbf{y}') \wedge P(\mathbf{x}, \mathbf{y}'')) \rightarrow \mathbf{y}' = \mathbf{y}'' \Big)$$

und

$$\exists P\, (\text{Funktion } P)\, \forall \mathbf{x} \forall \mathbf{y} \forall \mathbf{z}\, P(\mathbf{x}, \mathbf{y}) \rightarrow \varphi(\mathbf{x}, \mathbf{y}, \mathbf{z})$$

zu ψ äquivalente Formeln. So folgen die ersten beiden Teilbehauptungen.

So ist P aber noch nicht eindeutig festgelegt. Dies wird dadurch erreicht, daß für P zusätzlich gefordert wird, daß es ausschließlich für das kleinstmögliche solche $\mathbf{y}$ gilt. Dementsprechend wird die obige Formel ergänzt um

$$\forall \mathbf{x} \forall \mathbf{y}' \forall \mathbf{y}''\;\; P(\mathbf{x}, \mathbf{y}') \wedge (\mathbf{y}'' < \mathbf{y}') \rightarrow \neg(\forall \mathbf{z}' \varphi(\mathbf{x}, \mathbf{y}'', \mathbf{z}')) .$$

Somit ergibt sich durch Herausziehen der Quantoren und wegen der Eindeutigkeit eines existierenden P die Darstellung

$$\begin{array}{lll} (\exists{=}\exists!) P\, \forall \mathbf{x} \forall \mathbf{y}' \forall \mathbf{y}'' \forall \mathbf{z}\, \exists \mathbf{y} \exists \mathbf{z}' & & P(\mathbf{x}, \mathbf{y}) \wedge \varphi(\mathbf{x}, \mathbf{y}, \mathbf{z}) \\ & \wedge & \Big(P(\mathbf{x}, \mathbf{y}') \wedge P(\mathbf{x}, \mathbf{y}'') \rightarrow \mathbf{y}' = \mathbf{y}'' \Big) \\ & \wedge & \Big(P(\mathbf{x}, \mathbf{y}') \wedge (\mathbf{y}'' < \mathbf{y}') \rightarrow \neg\varphi(\mathbf{x}, \mathbf{y}'', \mathbf{z}') \Big) \end{array}$$

Dabei ist $\varphi \in \Sigma_k$ und $\neg\varphi \in \Pi_k$. Im ganzen ergeben sich also $k+2$ Quantorenwechsel in der ersten Stufe, wobei der erste Quantor ein Allquantor ist. Damit ist die Formel in der Klasse $(\exists^2 = \exists^2!)\Pi_{k+2}[<]$.

Da das $\mathbf{y}$ im allgemeinen für mehr als eine Variable erster Stufe steht, muß $\mathbf{y}'' < \mathbf{y}'$ durch eine Formel ersetzt werden, die die lexikographische Ordnung überprüft. Dies ist quantorenfrei möglich durch

$$x_1 < y_1 \vee \cdots \vee (x_1 = y_1 \wedge \cdots \wedge x_{n-1} = y_{n-1} \wedge x_n < y_n)\,. \qquad \blacksquare$$

Daher kollabieren die Quantorenhierarchien der ersten Stufe auf die entsprechenden Klassen mit $k = 0$. Der Satz liefert durch endliche Wiederholung dieser Umformung sogar konstruktiv eine Formel in der entsprechenden Klasse.

Falls die Ordnung des Universums durch eine Nachfolgerrelation gegeben ist, fällt dieser Kollaps noch deutlicher aus:

Satz 44
In Anwesenheit eines existentiellen Quantors zweiter Stufe und einer Logik mit Nachfolgerrelation kollabiert die Quantorenhierarchie erster Stufe auf die Klasse Π_1*.*

$$\mathrm{FO}[=,\mathrm{succ}] \subseteq (\exists{=}\exists!)^2\Pi_1[=,\mathrm{succ}]$$

Beweis: Falls nur die Nachfolgerrelation zur Verfügung steht, kann nach Definition 22 und Lemma 26 durch Π_1-Formeln eindeutig die passende Ordnungsrelation beschrieben werden.

Mit Satz 43 genügt es zu zeigen, daß eine beliebige Formel der Gestalt $\psi = \forall\mathbf{x}\,\exists\mathbf{y}\,\varphi(\mathbf{x},\mathbf{y})$ mit quantorenfreiem φ auch in $\Pi_1[\mathrm{succ}]$ dargestellt werden kann.

Der Existenzquantor $\exists\mathbf{y}$ wird durch ein Schwellenprädikat P ersetzt. Dieses $P(\mathbf{x},\mathbf{y})$ ist für alle kleinen $\mathbf{y}$, für die φ nicht gilt, nicht erfüllt, ab dem ersten φ erfüllenden $\mathbf{y}$ (einschließlich) gilt $P(\mathbf{x},\mathbf{y})$ für alle größeren $\mathbf{y}$. Aus dieser Idee ergibt sich folgende Formel

$$\begin{array}{lll}
(\exists{=}\exists!)P\,\forall\mathbf{x}\forall\mathbf{y}\forall\mathbf{y}'\forall a\forall o & & \alpha(a)\wedge\omega(o) \;\rightarrow\; P(\mathbf{x},\mathbf{o})\wedge(\neg P(\mathbf{x},\mathbf{a})\vee\varphi(\mathbf{x},\mathbf{a}))\\
& \wedge & \mathrm{succ}(\mathbf{y},\mathbf{y}')\wedge P(\mathbf{x},\mathbf{y}) \;\rightarrow\; P(\mathbf{x},\mathbf{y}')\\
& \wedge & \mathrm{succ}(\mathbf{y},\mathbf{y}')\wedge\neg P(\mathbf{x},\mathbf{y})\wedge P(\mathbf{x},\mathbf{y}') \;\rightarrow\; \varphi(\mathbf{x},\mathbf{y}')\\
& \wedge & \neg P(\mathbf{x},\mathbf{y}) \;\rightarrow\; \neg\varphi(\mathbf{x},\mathbf{y})\,.
\end{array}$$

Dabei steht succ für das entsprechende S_k aus Satz 28, die Formel muß also noch um entsprechende Prädikate und die dort angegebene Hornformel erweitert werden. Diese zu ψ äquivalente Formel ist in der gewünschten Klasse. $\blacksquare$

Satz 45
In Anwesenheit eines existentiellen Quantors zweiter Stufe und einer Logik mit Funktionen kollabiert die Hierarchie auf die Stufe Π_1.

$$\mathrm{FO}[<] \subseteq (\exists{=}\exists!)^2(\mathrm{Funktion})\Pi_1[<]$$

Beweis: Diese Verstärkung von Satz 43 kann hier recht einfach gezeigt werden, indem ausgenutzt wird, daß die Nachfolgerrelation eindeutig bestimmt ist.

Hier im Beweis werden der Übersichtlichkeit wegen Abkürzungen wie $\leq$ oder $x = y = z$ verwendet, die offenbar quantorenfrei aufgelöst werden können. Für funktionale Prädikate φ wird statt $\varphi(x, y)$ auch $y = \varphi(x)$ geschrieben. Nach Lemma 20 könnte sogar der gesamte Funktorenkalkül verwendet werden, hier soll der Klarheit halber jedoch nur diese intuitive Abkürzung verwendet werden.

Zunächst wird eine nullstellige Funktion bestimmt, die das kleinste Element des Universums repräsentiert. Dies kann durch

$$(\exists{=}\exists!)a()\ (\mathrm{Funktion}\ a)\ \forall x \forall y\ \Big(y = a() \rightarrow y \leq x\Big)$$

erreicht werden. Eine analoge Formel erzeugt das größte Element $o()$ des Universums.

Die Nachfolgerrelation wird durch eine Funktion ersetzt, die Nachfolgerfunktion ist und lediglich das letzte Element auf sich selber abbildet. Dementsprechend übernimmt $y = \sigma(x) \wedge x \neq y$ die Rolle von $\mathrm{succ}(x, y)$. Dies wird durch die Formel

$$\begin{aligned}
&(\exists{=}\exists!)\sigma\ (\mathrm{Funktion}\ \sigma)\ \forall x \forall x' \forall y \\
&\qquad\qquad\qquad y = \sigma(x) \ \rightarrow\ y > x \vee x = \sigma(x) = o() \\
&\qquad\quad \wedge\quad y = \sigma(x) = \sigma(x') \ \rightarrow\ x = x' \vee x = o() \vee x' = o() \\
&\qquad\quad \wedge\quad x < y \ \rightarrow\ \sigma(x) < \sigma(y) \vee y = o()
\end{aligned}$$

erreicht. Diese wird von der oben beschriebenen totalen Nachfolgerfunktion erfüllt. Außerdem wird σ eindeutig bestimmt. Würde σ ein Element x auf ein $y = \sigma(x)$, das größer als der eigentliche Nachfolger ist, abbilden, so würde dies wegen der letzten Zeile auch für alle größeren Elemente gelten. Dann würde aber auch mindestens das drittgrößte Element auf das größte abgebildet werden, was wegen der letzten Zeile der Formel (mit diesem und dem vorletzten als x und y) unmöglich ist.

Dann ergibt sich durch Satz 44 die gewünschte Aussage für beliebige Formeln der ersten Stufe. Die dort beschriebene Formel wird lediglich um das oben angegebene Stück ergänzt und verwendet $y = \sigma(x) \wedge x \neq y$ statt $\mathrm{succ}(x, y)$. ∎

Satz 46
Für die existentielle Prädikatenlogik zweiter Stufe gilt somit

$$\exists^2\mathrm{FO}[<] = \exists^2\Pi_2[<] = \exists^2\Pi_1[\mathrm{succ}] = \exists^2(\mathrm{Funktion})\Pi_1[<]$$

und

$$\exists^2\mathrm{FO} = \exists^2\Pi_2 = \exists^2(\mathrm{Funktion})\Pi_1 \, .$$

Diesen Sätzen entsprechend ist im hier betrachteten Zusammenhang eine Logik, die lediglich die Ordnungsrelation zur Verfügung hat, interessanter, da die betrachtete Hierarchie der Quantoren erster Stufe nicht unbedingt vollständig kollabiert.

Zusammenfassend entsteht somit folgende echte Hierarchie:

$$\begin{tikzcd}
\exists^2\Pi_1 \arrow[dr, no head] & & \\
& \exists^2\Sigma_2 \arrow[r, no head] & \exists^2\Pi_2 \\
\exists^2\Sigma_1 \arrow[ur, no head] & &
\end{tikzcd}$$

Die Klassen mit eingebauter Ordnung müssen nach Beobachtung 42 nicht gesondert behandelt werden. Beachtenswert ist, daß keines der hier vorgestellten Gegenbeispiele zur Darstellung in der Klasse den Quantor zweiter Stufe tatsächlich benötigt. Dadurch können Zählvarianten dieser Gegenbeispiele zum Trennen der Stufen einer Zählhierarchie benutzt werden.

Eine Struktur erfüllt FORALL, wenn das einzige einstellige Prädikat für alle Elemente des Universums gilt.

Lemma 47
FORALL *liegt in* $\exists^2\Pi_1$*, aber nicht in* $\exists^2\Sigma_1$.

Beweis: Dies entspricht der Formel $\forall x P(x)$, diese ist in $\Pi_1 \subset \exists^2\Pi_1$.

Angenommen, FORALL liegt in $\exists^2\Sigma_1$. Dann gibt es eine quantorenfreie Formel ψ mit:

$$\mathrm{FORALL}(\mathcal{A}) \;\Leftrightarrow\; \mathcal{A} \models \exists^2\mathbf{T}\,\exists\mathbf{x}\,\psi(\mathbf{T},\mathbf{x}) \, .$$

Für die Struktur $\mathcal{J} = (\{A\}, P = \{A\})$ gilt offenbar FORALL$(\mathcal{J})$. Sei also $\mathbf{T}$ der Prädikatenvektor, für den $\mathcal{J} \models \exists\mathbf{x}\,\psi(\mathbf{T},\mathbf{x})$ gilt.

Für die Struktur $\mathcal{K} = (\{A, B\}, P = \{A\})$ gilt offenbar FORALL nicht. Da aber $\langle\mathcal{K},\mathbf{T}\rangle$ eine Erweiterung von $\langle\mathcal{J},\mathbf{T}\rangle$ ist, ergibt sich mit Lemma 17 der Widerspruch, daß $\mathcal{K} \models \exists\mathbf{x}\,\psi(\mathbf{T},\mathbf{x})$ gilt. ∎

Daß die Klassen tatsächlich unvergleichbar sind, folgt analog, kann aber wegen des $\exists^2$-Quantors nicht trivial gefolgert werden. Man betrachtet das Problem EXISTS, das genau dann für eine Struktur gilt, wenn das einzige einstellige Prädikat der Struktur nicht die leere Menge ist.

Lemma 48
EXISTS *liegt in* $\exists^2\Sigma_1$*, aber nicht in* $\exists^2\Pi_1$.

Beweis: Dies entspricht der Formel $\exists x P(x)$, diese ist in $\Sigma_1 \subset \exists^2\Sigma_1$.

Angenommen, EXISTS liegt in $\exists^2\Pi_1$. Dann gibt es eine quantorenfreie Formel ψ mit:

$$\text{EXISTS}(\mathcal{A}) \;\Leftrightarrow\; \mathcal{A} \models \exists^2\mathbf{T}\,\forall\mathbf{x}\,\psi(\mathbf{T},\mathbf{x})$$

Für die Struktur $\mathcal{K} = (\{A,B\}, P = \{B\})$ gilt EXISTS. Sei also $\mathbf{T}$ der Prädikatenvektor, für den $\mathcal{K} \models \forall\mathbf{x}\psi(\mathbf{T},\mathbf{x})$ gilt.

Für die Struktur $\mathcal{J} = (\{A\}, P = \emptyset)$ gilt EXISTS offenbar nicht.

Da aber $\langle\mathcal{J},\mathbf{T}\rangle$ eine Substruktur von $\langle\mathcal{K},\mathbf{T}\rangle$ ist, ergibt sich mit Lemma 16 $\mathcal{J} \models \forall\mathbf{x}\psi(\mathbf{T},\mathbf{x})$, ein Widerspruch. ∎

Damit folgt auch $\exists^2\Pi_1 \subsetneq \exists^2\Sigma_2$ und $\exists^2\Sigma_1 \subsetneq \exists^2\Sigma_2$.

Für einen Graphen, in dem von jeder Ecke eine Kante ausgeht, also in jedem Knoten einen Außengrad von mindestens 1 hat, gilt OUTEDGE.

Lemma 49
OUTEDGE *liegt in* $\exists^2\Pi_2$*, aber nicht in* $\exists^2\Sigma_2[<]$.

Beweis: OUTEDGE wird durch die Formel $\forall x \exists y\; e(x,y)$ beschrieben, die zu $\Pi_2 \subset \exists^2\Pi_2$ gehört.

Im Widerspruch zur Behauptung des Satzes wird angenommen, OUTEDGE sei in der Klasse $\exists^2\Sigma_2[<]$. Dann existiert eine quantorenfreie Formel ψ mit

$$\text{OUTEDGE}(\mathcal{G}) \;\Leftrightarrow\; \mathcal{G} \models \exists^2\mathbf{T}\,\exists\mathbf{x}\,\forall\mathbf{y}\,\psi(\mathbf{x},\mathbf{y},\mathbf{T})\,.$$

Sei t die Stelligkeit von $\mathbf{x}$. Wir betrachten den Graphen G, der ein Kreis aus $n = t+1$ Ecken ist. Für diesen gilt OUTEDGE, also gibt es $\mathbf{T}'$ und $\mathbf{x}'$ mit $(G,\mathbf{T}') \models \forall\mathbf{y}\,\psi(\mathbf{x}',\mathbf{y},\mathbf{T}')$. Dann gibt es eine Ecke a, die keine Komponente von $\mathbf{x}'$ ist. Sei G_1 der Subgraph von G, der durch Löschen der Ecke a entsteht. Dieser hat eine Ecke, von der keine Kante ausgeht und ist Substruktur von G. Da auch $\mathbf{T}'_1 \subset \mathbf{T}'$ für das von der Eckenmenge von G_1 induzierte T'_1 gilt, folgt mit Lemma 16 $(G_1,\mathbf{T}'_1,\mathbf{x}') \models \forall\mathbf{y}\,\psi(\mathbf{x}',\mathbf{y},\mathbf{T}'_1)$, und so die Formel, also ein Widerspruch. Somit liegt OUTEDGE nicht in $\exists^2\Sigma_2[<]$. ∎

Auch wenn die hier vorgestellten trennenden Probleme sehr einfach und uninteressant waren, gilt dies nicht unbedingt für die betrachteten Klassen. So ist etwa die Erfüllbarkeitsfrage von 3CNF echt in $\exists^2\Pi_1$, sowie 3DNF echt in $\exists^2\Sigma_1$ und HAMILTONIAN echt in $\exists^2\Sigma_2$ ist. Allerdings zeigen auch die angegebenen Gegenbeispiele die Qualität der auf diese Art gebildeten Klassen.

4.2 Charakterisierung von $\#\mathrm{P}$

Dieses Kapitel gibt von Saluja, Subrahmanyam und Thakur in [SST92] vorgestellte Resultate wieder, die eine wesentliche Motivation für die vorgestellten Definitionen waren. Von dort stammen auch die Probleme, die die verschiedenen Klassen trennen. Wo dies möglich war, wurden diese vereinfacht, da hier die Struktur der Klassen und nicht die Komplexität bestimmter Probleme im Vordergrund stehen soll. Die Kollapsresultate sind in dem neuen definitorischen Rahmen etwas allgemeiner formulierbar, die Beweise verwenden aber keine neuen Ideen.

Zunächst zeigt sich, daß die in [SST92] verwendete Notation über Mächtigkeiten der Mengen erfüllender Belegungen in der Quantorendarstellung aufgeht.

Beobachtung 50
Für eine Formel φ über einer Signatur mit $<$, dem Vektor von freien Prädikaten $\mathbf{R} = R_1, \ldots, R_k$ und freien Variablen der ersten Stufe $\mathbf{x} = x_1, \ldots, x_j$ gilt für passende Strukturen $\mathcal{J}$

$$\mathrm{val}(\,\mathcal{J} \models \#R_1 \cdots \#R_k \#x_1 \cdots \#x_j \varphi) = \Big|\{\langle \mathbf{R}, \mathbf{x}\rangle \colon (\mathcal{I}, \mathbf{R}) \models \varphi(\mathbf{x})\}\Big|\,.$$

Beweis: Für jedes Element der Menge gibt es genau einen Summanden 1 in der Summe. ∎

Um sinnvoll zählen zu können, scheinen angeordnete Universen unentbehrlich. So kann etwa mit Hilfe eines Zählquantors keine Ordnung erzeugt werden, da alle $n!$ möglichen Anordnungen zunächst gleichberechtigt sind. Die Struktur dieser Problematik wird an den Ergebnissen über die Berechnungskraft der Ordnung aus Abschnitt 4.4 deutlich werden.

Satz 51 ($\#\mathbf{P} = e(\#\mathrm{FO}[<])$)
Die Klasse der Funktionen, die durch den Zähloperator, angewendet auf eine Formel der Prädikatenlogik erster Stufe mit geordnetem Universum, dargestellt werden kann, entspricht genau $\#\mathbf{P}$.

Beweis: „$\supseteq$“ Die Funktion f sei gegeben durch

$$f(\mathcal{I}) = \mathrm{val}(\,\mathcal{I} \models \#\mathbf{R}\#\mathbf{x}\,\varphi).$$

Es muß eine Turingmaschine B angegeben werden mit

$$f(\mathcal{I}) = \Big|\{y \in \{0,1\}^{|e(\mathcal{I})|^k} \colon B(e(\mathcal{I}), y) \in \{0\}^*\}\Big|$$

für ein festes k.

Die Maschine rät nichtdeterministisch alle möglichen Belegungen für die Prädikate und Belegungen der in der ersten Stufe durch einen Zählquantor gebundenen Variablen und überprüft dann, ob so die Formel φ erfüllt wird. Genau in diesem Fall akzeptiert die Maschine, indem sie das Arbeitsband löscht. Dies ist nach Lemma 37 innerhalb der erlaubten Berechnungszeit möglich.

Da alle geratenen Prädikatenvektoren durch die eingebaute Ordnungsrelation verschieden (nicht isomorph) sind, und alle möglichen geraten werden, ist die Anzahl der akzeptierenden Pfade genau gleich der Anzahl der passenden Prädikatenvektoren.

„$\subseteq$“ Sei f gegeben durch

$$f(\mathcal{I}) = \left|\{y \in \{0,1\}^{|e(\mathcal{I})|^k} : B(e(\mathcal{I}), y) \in \{0\}^*\}\right|.$$

Es wird ein ϑ mit

$$f(\mathcal{I}) = \mathrm{val}(\mathcal{I} \models \#\mathbf{R}\,\vartheta)$$

angegeben. Dabei wird kein Zählquantor der ersten Stufe verwendet.

Sei also B die entsprechende Turingmaschine und φ die dadurch festgelegte Formel aus Lemma 38.

Die Formel

$$\tau := (C_0(\mathbf{x}) \vee C_1(\mathbf{x})) \wedge (\neg C_0(\mathbf{x}) \vee \neg C_1(\mathbf{x})) \wedge O_0(\mathbf{x}) \wedge \neg O_1(\mathbf{x})$$

garantiert, daß die C_c für eine korrekte Darstellung des Auswahlbandes stehen, und daß die Ausgabe aus $\{0\}^*$ ist, die Maschine also akzeptiert.

Weiterhin sei ρ die Formel aus Lemma 26, die zusichert, daß succ die zu der totalen Ordnung und den geratenen Prädikaten α und ω passende Nachfolgerrelation ist.

Dann gilt mit $\mathbf{R} := (C_0, C_1, O_0, O_1, \sigma, \mathbf{R}')$, wobei $\mathbf{R}'$ der Prädikatenvektor aus Lemma 38 ist, und $\vartheta := \tau \wedge \rho \wedge \varphi$ die obige Darstellung. Da so alle Prädikate außer den C_c durch die Eingabe eindeutig festgelegt sind, ist die Anzahl der erfüllenden Prädikatenvektoren gleich der Anzahl der akzeptierenden Pfade der Turingmaschine. ∎

Satz 52
Die Hierarchie der logisch definierten Zählklassen kollabiert je nach Signatur:

$$\#\mathrm{FO}[<] = \#\Pi_2[<] = \#\Pi_1[=, \mathrm{succ}] = \#(\mathrm{Funktion})\Pi_1[<]\,.$$

Beweis: Da die im letzten Satz angegebenen Formeln in der Klasse Π_2 sind, kann jede Formel auf diese Form gebracht werden.

Auch Satz 43 liefert dieses Ergebnis, da aus $(\exists{=}\exists!)P\varphi$ auch $\exists P\varphi \equiv \#P\varphi$ folgt. Dies liefert mit Satz 44 und Satz 45 den Rest der Behauptung. ∎

Man könnte aus den bisherigen Resultaten erwarten, daß die gleiche Hierarchie wie bei $\exists^2$ entsteht. Dies erweist sich durch die Definition, die auch Zählquantoren der ersten Stufe erlaubt, als falsch.

Lemma 53
Es gilt

$$\exists^1\mathrm{QF}[<] \subset \#^1\forall^1\mathrm{QF}[<]\,.$$

Beweis: Der $\exists$-Quantor kann durch einen $\#$-Quantor ersetzt werden, wenn existierende Objekte eindeutig bestimmt sind. Dies kann durch die Ordnung erzwungen werden, indem jeweils nur das kleinste solche Objekt als gültig betrachtet wird.

Daher sei $\vartheta(x, y)$ eine quantorenfreie Formel, die genau dann gilt, wenn x lexikographisch kleiner oder gleich y ist. Im folgenden steht $x \leq y$ für $(x < y \vee x = y)$. Also

$$\begin{aligned}\vartheta(x_n, \ldots, x_1, y_n, \ldots, y_1) \quad &\triangleq \quad (x_n \leq y_n) \vee (x_n = y_n \wedge x_{n-1} \leq y_{n-1}) \vee \cdots \\ &\qquad \cdots \vee (x_n = y_n \wedge x_{n-1} = y_{n-1} \wedge \cdots \wedge x_1 \leq y_1)\end{aligned}$$

Dann gilt für quantorenfreies $\psi(\mathbf{x})$

$$\exists\mathbf{x}\,\psi(\mathbf{x}) \equiv \#\hat{\mathbf{x}}\,\forall\mathbf{x}\,\psi(\hat{\mathbf{x}}) \wedge (\psi(\mathbf{x}) \to \vartheta(\hat{\mathbf{x}}, \mathbf{x}))\,.$$

Falls die linke Seite zu 0 auswertet, so gilt dies auch für die rechte. Falls sie zu 1 auswertet, gilt die Formel mit freiem $\hat{\mathbf{x}}$ genau für das lexikographisch kleinste solche. ∎

Satz 54
Es gilt

$$\#\Sigma_1[<] \subset \#\Pi_1[<]\,.$$

Beweis: Dies ergibt sich direkt aus dem letzten Lemma. ∎

Die anderen sich sofort ergebenden Inklusionen der Zählklassen sind echt. Dies zeigen die folgenden Sätze, die typisch für derartige Argumentationen sind.

Satz 55
Es gilt

$$\#\Pi_0[<] \subsetneq \#\Sigma_1[<]\,.$$

Beweis: Nach [Tha92]. Man betrachtet die Zählfunktion „Anzahl der Knotenpaare in einem gerichteten Graphen, die durch einen Weg der Länge genau zwei verbunden sind.“ (DIST2GRAPH). Diese ist gegeben durch

$$f(G) = \mathrm{val}\,(G \models \#x\#y\,\exists z\, x < y \wedge z \neq x \wedge z \neq y \wedge \neg e(x,y) \wedge e(x,z) \wedge e(z,y)\,)$$

für einen gerichteten Graphen G, der durch seine Pfeilmenge als zweistellige Relation e im Universum der Knoten dargestellt ist. Also ist $f \in \#\Sigma_1[<]$.

Im Widerspruch zur Behauptung sei ψ eine quantorenfreie Formel mit

$$f(G) = \mathrm{val}(G \models \#\mathbf{T}\#\mathbf{z}\,\psi)\,.$$

Es wird der Graph $G = (V, E)$ mit $V = \{1,2,3\}$ und $E = \{(1,2),(2,3)\}$ betrachtet. Es gilt $f(G) = 1$.

Mit der Charakterisierung durch ψ gibt es genau ein Tupel $\langle \mathbf{T}', \mathbf{z}' \rangle$, so daß $(G, \mathbf{T}') \models \psi(\mathbf{T}', \mathbf{z}')$ gilt.

Dann bezeichne $Z = \{z'_1, \ldots, z'_m\}$ mit $\mathbf{z}' = (z'_1, \ldots, z'_m)$ die Menge der so „geratenen“ Ecken, G_Z den durch Z induzierten Subgraphen und $\mathbf{T}'_Z$ die auf Z eingeschränkten Prädikate von $\mathbf{T}'$.

<u>Fall 1:</u> $|Z| \leq 2$
Dann gilt $G_Z \models \psi(\mathbf{T}'_Z, \mathbf{z}')$, da nach Lemma 16 eine solche Formel φ auch von Substrukturen erfüllt wird. Also gilt $\psi(\mathbf{T}', \mathbf{z}')$ bei Eingabe G_Z, also $f(G_Z) \geq 1$, was bei einem Graphen, der nur aus zwei Ecken besteht, nicht der Fall sein kann.

<u>Fall 2:</u> $|Z| > 2$
Dann ist $Z = V = \{1,2,3\}$. Man betrachtet den Graphen $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ mit $\mathcal{V} = \{1,2,3,4\}$ und den Kanten $\{(1,2),(2,3),(3,4),(1,3),(2,4)\}$, dargestellt in Abbildung 4.1.

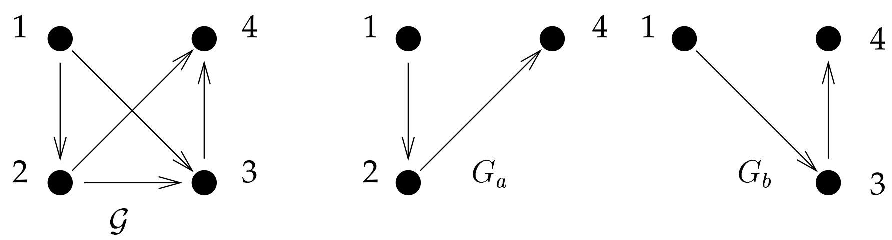

Abbildung 4.1: Graph $\mathcal{G}$ aus Satz 55, sowie die im Beweis auftretenden Subgraphen

Da bis auf 1 und 4 alle Knoten direkt verbunden sind, gilt $f(\mathcal{G}) = 1$, da nicht die verschiedenen Wege, sondern die verschiedenen Endpunkte von

Wegen gezählt werden. Es werden die von $V_a = \{1,2,4\}$ und $V_b = \{1,3,4\}$ induzierten Subgraphen $G_a = \mathcal{G}_{V_a}$ und $G_b = \mathcal{G}_{V_b}$ betrachtet. Für $i \in \{a,b\}$ ist G_i isomorph zu G vermöge der Abbildung $\varphi_i\colon V \to V_i$. Dann entstehen mit $T'_i = \varphi_i(T')$ und $\mathbf{z}'_i = \varphi_i(\mathbf{z}')$ (komponentenweise) zwei verschiedene Tupel $\langle T'_i, \mathbf{z}'_i\rangle$, da in z'_a wegen $Z = V$ eine 2 vorkommen muß, was in z'_b unmöglich ist. Die „Zeugen“ von G werden also nach $\mathcal{G}$ auf verschiedene Strukturen übertragen. Aus Satz 19 ergibt sich $G_i \models \psi(T'_i, z'_i)$. Mit Lemma 17 folgt $\mathcal{G} \models \psi(T'_i, z'_i)$. Also ist $\mathrm{val}(\mathcal{G} \models \#\mathbf{T}\#\mathbf{z}\psi) \geq 2$, im Widerspruch zur Annahme, daß ψ im angenommenen Sinn f charakterisiert. f ist also nicht in der Klasse $\#\mathrm{FO}[<]$ darstellbar. ■

Satz 56
Es gilt

$$\#\Sigma_1[<] \subsetneq \#\Pi_1[<]\,.$$

Beweis: Man betrachtet die Funktion

$$\textsc{ForAll}^{\#}(\mathcal{J}) := \begin{cases} 2^{||\mathcal{J}||} & \text{falls } \textsc{ForAll}(\mathcal{J}) \\ 0 & \text{sonst} \end{cases}\,.$$

So gilt für einstelliges R:

$$\textsc{ForAll}^{\#} \equiv \#R\,\forall x\, P(x),$$

die Funktion ist also in $\#\Pi_1[<]$. Angenommen, $\textsc{ForAll}^{\#}$ sei in $\#\Sigma_1[<]$. Dann gibt es eine quantorenfreie Formel φ mit

$$\textsc{ForAll}^{\#} \equiv \#\mathbf{R}\,\#\mathbf{z}\,\exists x\,\varphi.$$

Dann gilt aber auch

$$\textsc{ForAll} \equiv (\geq 1)\#\mathbf{R}\,\#\mathbf{z}\exists x\,\varphi,$$

und so

$$\textsc{ForAll} \equiv \exists\mathbf{R}\,\exists\mathbf{z}\,\exists x\,\varphi.$$

Somit ist $\textsc{ForAll} \in \exists^2\Sigma_1[<] \subseteq \exists^2\Sigma_1$, ein Widerspruch zu Lemma 47. ■

Bemerkung: Dieser Satz wird auch in [SST92] und [Tha92] formuliert, dort wird $\#3\mathrm{CNF} \in \#\Pi_1[<] \setminus \#\Sigma_1[<]$ gezeigt.

Satz 57
Es gilt

$$\#\Pi_1[<] \subsetneq \#\Sigma_2[<]\,.$$

Beweis: Nach [Tha92]. Man betrachtet das Problem DEG1NGB: In einem ungerichteten Graphen wird die Anzahl der Ecken bestimmt, die einen Nachbarn vom Grad 1 haben, also einen, der nur über diese Ecke mit dem Graphen verbunden ist. Für einen als Kantenrelation gegeben Graphen wird also f definiert durch

$$f(\mathcal{G}) = \mathrm{val}(\mathcal{G} \models \#z\, \exists y\, \forall x\, ((x = z) \vee \neg e(x,y)) \wedge e(z,y)\,)\,.$$

Im Widerspruch zur Behauptung des Satzes wird angenommen, dieses f sei in der Klasse $\#\Pi_1[<]$. Es existiert also eine quantorenfreie Formel ψ mit

$$f(\mathcal{G}) = \mathrm{val}(\mathcal{G} \models \#T\, \#z\, \forall \mathbf{x}\, \psi)\,.$$

Dann ist die Anzahl der Komponenten von **z** festgelegt. Diese wird im weiteren als m bezeichnet.

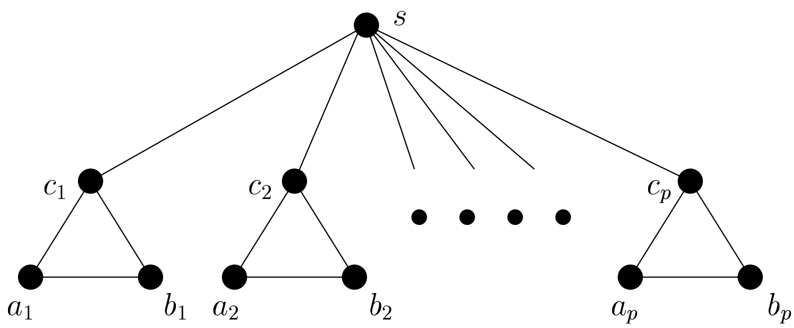

Abbildung 4.2: Der Graph G aus dem Beweis von Satz 57

Für ein festes $p > 2m + 1$ betrachtet man den folgenden Graphen, dargestellt in Abbildung 4.2.

$$G = (V, E), \quad V = \{a_1, \ldots, a_p, b_1, \ldots, b_p, c_1, \ldots, c_p, s\}\,,$$

angeordnet als

$$a_1 < b_1 < c_1 < a_2 < b_2 < c_2 < \cdots < a_p < b_p < c_p < s$$

und mit folgenden Kanten

$$E = \{\,(s, c_i) \mid i \in I\,\} \cup \{\,(a_i, b_i) \mid i \in I\,\} \cup \{\,(b_i, c_i) \mid i \in I\,\} \cup \{\,(c_i, a_i) \mid i \in I\,\}\,,$$

wobei $I = \{1, \ldots, p\}$ gilt.

Für $i \in I$ definiert man $G_i = G_{V \setminus \{a_i, b_i\}}$ als den geordneten, induzierten Subgraphen. In diesem Graphen ist c_i die einzige Ecke mit Grad 1, also ist

$f(G_i) = 1$. Seien also $\langle T_i', \mathbf{z}_i' \rangle$ die Tupel, für die $(G_i, T_i') \models \forall \mathbf{x}\, \psi(\mathbf{x}, \mathbf{T}_i', \mathbf{z}_i')$ gilt.

In jedem $\mathbf{z}_i'$ müssen sowohl s als auch c_i enthalten sein. Wäre dem nicht so, dann betrachtet man den Subgraphen H_i von G_i, der durch Entfernen dieser nicht vorkommenden Ecken entsteht. Dieser hat in beiden Fällen keine Ecke vom Grad 1. Wenn man die diesem Subgraphen entsprechende Substruktur $T_i'' \subset T_i'$ betrachtet, gilt mit Lemma 16 $(H_i, T_i'') \models \forall \mathbf{x} \psi(\mathbf{x}, \mathbf{T}_i'', \mathbf{z}_i')$, und mit $f(H_i) \geq 1$ ein Widerspruch.

Es gibt $u, v \in I$, $u \neq v$, so daß sowohl $\mathbf{z}_u'$ zu den Ecken $\{a_v, b_v, c_v\}$ als auch $\mathbf{z}_v'$ zu $\{a_u, b_u, c_u\}$ disjunkt ist. Dies beweist man wie folgt durch Widerspruch.

Man betrachtet für $i \in I$ die Mengen $K_i \subset I$, die durch $k \in K_i :\Leftrightarrow$ (Vektor $\mathbf{z}_i'$ enthält eine der Ecken $\{a_k, b_k, c_k\}$) definiert sind. In K_i sind also alle von dem Vektor $\mathbf{z}_i'$ „besuchten Dreiecke" dargestellt.

Die Widerspruchsannahme lautet dann, daß für beliebiges $u \neq v$ zumindest $u \in K_v$ oder $v \in K_u$ gilt.

Es gilt für alle $i \in I$: $|K_i| \leq m$, da der Vektor $\mathbf{z}_i'$ nur soviele Stellen hat.

Für die punktierte Vereinigung $K := \{\, (i, a) \mid i \in I, a \in K_i \,\}$ der K_i gilt $mp \geq |K|$. Faktorisiert man K über die Äquivalenzrelation $(a, b) \sim (b, a)$ und $(a, b) \sim (a, b)$, so entsteht die Menge $\hat{K} := {}^{K}\!/_{\sim}$, mit $|K| \geq |\hat{K}|$.

Sei nun $\{u, v\}$ eine beliebige zweielementige Teilmenge von I, also $u \neq v$. Dann gilt nach Widerspruchsannahme $u \in T_v$ oder $v \in T_u$. Also ist mindestens eines von (u, v) und (v, u) ein Element von K. Da beide im Sinne der Faktorisierung äquivalent sind, ist $[(v, u)] \equiv \{u, v\}$ ein Element von $\hat{K}$. Damit ist $\hat{K}$ die Menge aller zweielementigen Teilmengen von I, es gilt

$$|\hat{K}| = \binom{p}{2} = \frac{p\,(p-1)}{2} > \frac{p\,(2m+1-1)}{2} = mp$$

und somit

$$mp \geq |K| \geq |\hat{K}| > mp,$$

also ein Widerspruch.

Seien also $u, v \in I$, $u \neq v$, mit $\mathbf{z}_u'$ disjunkt zu $\{a_v, b_v, c_v\}$ und $\mathbf{z}_v'$ disjunkt zu $\{a_u, b_u, c_u\}$, im folgenden fest. Man definiert $G_{u,v}$ als den Subgraphen von G, der durch Löschen der Ecken a_u, b_u, a_v und b_v entsteht. Offenbar gilt $G_{u,v} = G_{v,u}$.

Seien $\mathcal{R} := \langle T_{u,v}', \mathbf{z}_{u,v}', E_{u,v} \rangle$ die von $G_{u,v}$ in $\langle T_u', \mathbf{z}_u' \rangle$ induzierte Substruktur und $\mathcal{S} := \langle T_{v,u}', \mathbf{z}_{v,u}', E_{u,v} \rangle$ die von $\langle T_v', \mathbf{z}_v' \rangle$. Dies ist durch die Wahl von u und v auch für die Konstantenvektoren möglich. Außerdem sind $\mathcal{S}$ und $\mathcal{R}$ verschieden, da einerseits $\mathbf{z}_u'$ zwar c_u aber nicht c_v, und andererseits $\mathbf{z}_v'$ zwar c_v aber nicht c_u enthält.

Nach Lemma 16 gilt $\mathcal{R} \models \forall\mathbf{x}\,\psi$ und $\mathcal{S} \models \forall\mathbf{x}\,\psi$, somit $f(G_{u,v}) \geq 2$. Da in $G_{u,v}$ die Ecke s die einzige mit zwei Nachbarn vom Grad 1 ist, gilt $f(G_{u,v}) = 1$. Dieser Widerspruch zeigt, das f nicht in der Klasse $\#\Pi_1[<]$ liegt. ∎

Satz 58
Es gilt

$$\#\Sigma_2[<] \subsetneq \#\Pi_2[<]\,.$$

Beweis: Nach [Tha92]. Man betrachtet die Zählfunktion f „Anzahl der Hamiltonschen Kreise in einem Graphen", diese ist in $\#\mathbf{P}$. Dazu rät sich eine geeignete Turingmaschine jeweils eine Kantenauswahl und überprüft dann, ob es sich tatsächlich um einen Kreis handelt, der alle Knoten genau einmal berührt. Nach Satz 51 und Satz 52 gilt $f \in \#\Pi_2[<]$.

Im Widerspruch zur Behauptung des Satzes wird angenommen, dieses f sei in der Klasse $\#\Sigma_2$. Dann existiert eine quantorenfreie Formel ψ mit

$$f(G) = \mathrm{val}(G \models \#T\#z\,\exists\mathbf{x}\,\forall\mathbf{y}\,\psi)\,.$$

Sei m die Stelligkeit von $\mathbf{z}$, t die von $\mathbf{x}$. Wir betrachten den Graphen G, der ein Kreis aus $n = m + t + 1$ Ecken ist. Da $f(G) = 1$ ist, gibt es $\langle\mathbf{T}', \mathbf{z}'\rangle$ und $\mathbf{x}'$ mit $(G, \mathbf{T}') \models \forall\mathbf{y}\psi(\mathbf{x}', \mathbf{y}, \mathbf{T}', \mathbf{z}')$. Dann gibt es ein Ecke a, die weder in $\mathbf{z}'$ noch in $\mathbf{x}'$ Komponente ist. Sei G_1 der Subgraph von G, der durch Löschen der Ecke a entsteht. Dieser hat keinen Hamiltonschen Kreis, aber mit der passenden Substruktur $\mathbf{T}'_1 \subset \mathbf{T}'$ gilt $(G_1, \mathbf{T}'_1) \models \forall\mathbf{y}\,\psi(\mathbf{x}', \mathbf{y}, \mathbf{T}'_1, \mathbf{z}')$, also

$$f(G_1) = \mathrm{val}(G_1 \models \#T\,\#z\,\exists\mathbf{x}\,\forall\mathbf{y}\,\psi) \geq 1\,.$$

Dieser Widerspruch zeigt, daß f nicht in $\#\Sigma_2[<]$ liegt. ∎

Somit ergibt sich

Satz 59
Die syntaktisch definierten Zählklassen ergeben folgende echte lineare Hierarchie:

$$\#\Sigma_0[<] = \#\Pi_0[<] \subsetneq \#\Sigma_1[<] \subsetneq \#\Pi_1[<] \subsetneq \#\Sigma_2[<] \subsetneq \#\Pi_2[<] = \#\mathrm{FO}[<]\,.$$

4.3 Die $\#^2\mathrm{FO}[<]$-Hierarchie

Auch diese Hierarchie wurde wohl in der Literatur noch nicht betrachtet. Sie stellt die Verbindung zwischen den letzten beiden Kapiteln her und zeigt, daß die anscheinend vor allem handliche und nicht weiter überdachte Vermischung von erster und zweiter Stufe der Prädikatenlogik in der Definition von $\#\mathcal{FO}$ in [SST92] durchaus wesentlichen Einfluß auf die Struktur der Hierarchie hat.

Die Linearisierung der gerade betrachteten Hierarchie hängt offenbar an der Definition der Klassen, die es erlaubt, Zählquantoren der ersten und zweiten Stufe zu verwenden. Daher bietet es sich an anders definierte Zählklassen zu betrachten, in denen lediglich ein Zählquantor der zweiten Stufe erlaubt ist.

Dementsprechend betrachtet man die Klasse $\#^2\mathrm{FO}[<]$, in der nur Zählquantoren der zweiten Stufe erlaubt sind. Die Beweise von Satz 43, Satz 44 und Satz 45 liefern

$$\#^2\mathrm{FO}[<] = \#^2\Pi_2[<] = \#^2\Pi_1[=,\mathrm{succ}] = \#^2(\mathrm{Funktion})\Pi_1[<]\,.$$

In der Klasse $\#^2\mathrm{QF}[<]$ können keine sinnvollen Formeln gebildet werden. Für die übrigen Klassen ergibt sich die Inklusionsstruktur der $\exists^2$-Hierarchie. Dies ist kein Zufall, da sich sogar die trennenden Probleme von dort auf die Situation des Zählens übertragen lassen.

Lemma 60
*Die Probleme aus Lemma 47 (*FORALL*), Lemma 48 (*EXISTS*) und Lemma 49 (*OUTEDGE*) für $\exists^2$ trennen auch die Klassen der $\#^2$-Hierarchie.*

Beweis: Analog zu Satz 56. Aus den trennenden Entscheidungsproblemen läßt sich durch einen nicht weiter benutzten Zählquantor der zweiten Stufe über einem einstelligen Prädikat ein Zählproblem erzeugen, das, falls das Entscheidungsproblem gültig war, den Wert $2^{||\mathcal{A}||}$ hat, ansonsten den Wert 0. Diese Probleme liegen dann in den entsprechenden Klassen. Dazu wird der Formel für das Entscheidungsproblem ein Zählquantor, der sich auf ein neues einstelliges Prädikat bezieht, vorangestellt. Da die zitierten Probleme den Existenzquantor der zweiten Stufe nicht benötigen, entsteht eine Formel in der entsprechenden Zählklasse, die genau das hier beschriebene Zählproblem darstellt.

Auch die Trennungseigenschaft der Probleme in den $\exists^2$-Klasse überträgt sich auf die Zählklassen. Dazu nimmt man an, das Problem läge doch in der Zählklasse. Dann ist das Entscheidungsproblem insbesondere durch $(\geq 1)\#^2\varphi$ darstellbar. Dies entspricht genau dem $\exists^2$, also ist das Entscheidungsproblem im Widerspruch zu den zitierten Sätzen in der entsprechenden $\exists^2$-Klasse. ■

4.4 Maximierungsklassen

Die hier vorgestellte Hierarchie ist wohl der Ausgangspunkt aller derartigen Betrachtungen, wie er von Papadimitriou und Yannakakis durch die Klassen MAXNP und MAXSNP in [PY91] gelegt wurde. Diese Klassen lassen sich auch sehr natürlich in der Quantorendarstellung ausdrücken. Es werden auch hier nur bekannte Ergebnisse wiedergegeben, zum Teil sind Sätze und Beweise den Notationen deutlich angepaßt, und sind so vor allem als Illustration der Notation zu verstehen.

Dabei entsteht zunächst die Frage, ob man hier ohne weiteres nur Strukturen mit geordnetem Universum zulassen soll, oder ob dies hier die Ergebnisse beeinflußt. In den Klassen, in denen überprüft werden kann, ob eine Relation eine Ordnung darstellt, ist dies nicht der Fall. Es zeigt sich, daß es in den anderen Klassen Probleme gibt, die ohne Ordnung nicht entschieden werden können.

Dies führt wieder zu in der Literatur wohl noch nicht untersuchten Klassen, die sich von den bekannten unterscheiden.

Zunächst soll aber die bekannte und in [Tha92] verwendete Darstellung über Mächtigkeiten von Mengen in den hier vorgestellten definitorischen Rahmen gebracht werden.

Beobachtung 61
Für eine Formel φ mit freien Prädikaten $\mathbf{R} = R_1, \ldots, R_k$ und freien Variablen der ersten Stufe $\mathbf{x} = x_1, \ldots, x_j$ gilt für passende Strukturen $\mathcal{J}$

$$\mathrm{val}(\,\mathcal{J} \models \max_{R_1}{}^2 \cdots \max_{R_k}{}^2 \;\#x_1 \cdots \#x_j\, \varphi) = \max_{\mathbf{R}} \left| \{\mathbf{x}\colon (\mathcal{I}, \mathbf{R}) \models \varphi\} \right| .$$

Beweis: Genau für jedes Element der Menge gibt es jeweils einen Summanden 1 in der Summe. Somit ist der Zahlenwert in Abhängigkeit von $\mathbf{R}$ gleich, da über diese maximiert wird, ergibt sich derselbe Wert. ■

Es ergibt sich auch hier die Frage, welchen Einfluß eine Ordnung, die wegen der Vergleichbarkeit der Eingabe vorhanden ist, auf die Ausdrucksstärke jenseits der Eingabe, sozusagen beim Berechnen, hat.

Lemma 62
Es gilt

$$\max{}^2\,(\mathrm{Ordnung}\,)\#^1\Sigma_0 \not\subset \max{}^2\,\#^1\Sigma_1\,.$$

Beweis: Man betrachtet das Problem zu einem gerichteten Graphen die größtmögliche Anzahl an topologisch sortierbaren Pfeilen anzugeben, also die Funktion die durch

$$f(G) = \mathrm{val}(G \models \max_{<}{}^2\,(\mathrm{Ordnung}\;\;<)\#x\#y\, x < y \wedge e(x,y))$$

gegeben ist. Diese Funktion ist also in $\max^2$ (Ordnung)$\#^1\Sigma_0$.

Im Widerspruch zur Behauptung wird angenommen, die Funktion sei in $\max^2 \#^1\Sigma_1$ darstellbar. Sei also φ mit

$$f(G) = \mathrm{val}(G \models \max_{\mathbf{R}}^2 \#\mathbf{x}\,\exists\mathbf{y}\,\varphi)$$

die entsprechende Formel. Im folgenden bezeichne j die Stelligkeit von $\mathbf{x}$, k die von $\mathbf{y}$ und h die Anzahl der Prädikate in $\mathbf{R}$, und als Abkürzung $d = k+j$ und $D = 2d + 3$. Abbildungen werden auf Vektoren ausgedehnt, indem sie Komponentenweise agewendet werden.

Da G keine Schlingen hat, gibt es höchstens 2^h verschiedene einelementig Substrukturen von $(G, \mathbf{R})$. In einer beliebigen, aber ausreichend langen Folge solcher Substrukturen wird es also zwei gleiche geben, die mindestens $2D$ Positionen voneinander entfernt sind. Sei dementsprechend ein solches ungerades $n > 2D2^h + 1$ im folgenden fest. Betrachte dort nur jeden $2D$-ten Knoten, das sind $2^h + 1$ Stück.

Definiere den Graphen $G = (V, E)$ mit der Knotenmenge $V := \{1, \ldots, n\}$ und $E := \{\,(i, i+1) \mid 1 \leq i < n\,\} \cup \{\,(i, i+2) \mid i$ ungerade und $1 \leq i < n\,\}$ als Pfeilmenge. Offenbar gilt $f(G) = n - 1 + \frac{n-1}{2}$. Gemäß der angenommen Darstellung gibt es also Prädikate $\mathbf{R}$, so daß die Menge der Zeugen

$$K := \{\mathbf{c}_1, \ldots, \mathbf{c}_{n-1}\} = \{\,(\mathbf{x}, \mathbf{y}) \mid (G, \mathbf{R}, \mathbf{x}, \mathbf{y}) \models \varphi\,\}$$

diese Mächtigkeit hat.

Angenommen es gäbe ein $\mathbf{c} = (\mathbf{a}, \mathbf{b})$ in K, bei dem zwei Einträge $a_\mu < a_\nu$ in $\mathbf{a}$ eine Differenz $a_\nu - a_\mu \geq 2d + 3 = D$ haben. Da es in einem Intervall der Länge $2d + 3$ mindestens $d + 1$ ungerade Zahlen gibt, besteht der von $\mathbf{c}$ (als Menge) induzierten Subgraphen aus mindestens zwei Zusammenhangskomponenten, und es gibt ein ungerades g, daß diese trennt, das also zwischen den beiden Komponenten liegt und nicht in $\mathbf{c}$ enthalten ist.

Betrachte den Graphen G', der aus zwei Kopien von G besteht mit Knotenmenge $\{1, 2\} \times \{1, \ldots, n\}$, $f(G') = 2f(G)$ und Pfeilen $(\{1\} \times G) \cup (\{2\} \times G)$. Mit

$$\begin{aligned} R'_\lambda := \{\,(\,(t_1, x_1), \ldots, (t_l, x_l)\,) \mid \quad & (t_1 = \cdots = t_l = 1 \text{ und } \mathbf{x} \in R_\lambda) \\ \text{oder } \quad & (t_1 = \cdots = t_l = 2 \text{ und } \mathbf{x} \in R_\lambda) \\ \text{oder } \quad & (\mathbf{x} \in R_\lambda \text{ und } t_i \hat{=} (x_i < g)\,)\,\} \end{aligned}$$

sind Prädikate festgelegt, mit denen K' aus 2 Bildern von K besteht und zusätzlich aus dem neu zu zählenden $\mathbf{c}'$ bei dem alle Einträge kleiner als g mit 1 und die übrigen mit 2 markiert sind. Da g nicht in $\mathbf{c}'$ vorkommt, hat es keinen Einfluß auf die Auswertung von φ auf den c_i', daß es die mit g inzidenten Pfeile in G' nicht gibt, so daß nach Konstruktion von $\mathbf{R}'$ auch $\mathbf{c}'$ die Formel φ erfült. Wegen $a_1 < g < a_2$ ist $\mathbf{c}'$ auch von den anderen Elementen von K' im Sinne der Zählung in der Formel unterscheidbar. Somit ergibt sich mit $\mathrm{val}((\mathbf{R}', G) \models \#\mathbf{x}\,\exists\mathbf{y}\,\varphi) \geq 2f(G) + 1$ ein Widerspruch.

Seien die C_i die kleinsten Intervalle, die die Einträge der $\mathbf{x}_i$ mit $\mathbf{c}_i = (\mathbf{x}_i, \mathbf{y}_i)$ und die Zusammenhangskomponente im von c_i induzierten Subgraphen umfassen. Diese enthalten dann weniger als D Elemente.

Seien die nach Konstruktion existenten, ununterscheidbaren einelementigen Substrukturen von G mit Abstand größer als $2D$ die Knoten $p < q$, beide ungerade. Dann kann es gibt es nicht C_i und C_j mit $C_i \cap C_j \neq \emptyset$, die p und q umfassen, $\{p, q\} \subset C_i \cup C_j$. Konstruiere nun einen Graphen $G'' = (V'', E'')$ der aus $f(G)$ Kopien von G und einer Schleife aus G besteht, in der p mit q identifiziert ist. Dazu wird die Abbildung π eingeführt, die alle Zahlen auf sich und lediglich q auf p abbildet. Knotenmenge ist also

$$V'' = \{\, (t, x) \mid t \in \{0, \ldots, f(G)\}, x \in \{1, \ldots, n\}, (t, x) \neq (0, q) \,\}$$

und Pfeilmenge

$$\begin{aligned} E'' &= \{\, ((t, x), (t, y)) \mid t > 0, (x, y) \in G \,\} \\ &\cup \{((0, x), (0, \pi(y))) \mid (x, y) \in G \,\}. \end{aligned}$$

Für diesen Graphen gilt $f(G'') = f(G)^2 + f(G) - 2$, da sich in den Kopien nichts geändert hat, und in der Schleife an einer Stelle zwei Pfeile entfernt werden müssen, um etwas Kreisfreies zu erhalten.

Definiere eine Abbildung

$$\tau\colon \{1, \ldots, f(G)\} \times V \to V''$$

wie folgt:

$$\tau(i, x) = \begin{cases} (i, x) & \text{falls } x \notin C_i \\ (0, \pi(x)) & \text{falls } x \in C_i \end{cases}$$

Für $\mathbf{x} \neq \mathbf{y} \in V^r$ und $i, j \in \{1, \ldots, f(G)\}$ beliebig mit $\tau(i, \mathbf{x}) = \tau(j, \mathbf{y})$ gilt $i \neq j$ und $\{\mathbf{x}, \mathbf{y}\} = \{(p, \ldots, p), (q, \ldots, q)\}$.

Wäre $i = j$, so muß $\pi(\mathbf{x}) = \pi(\mathbf{y})$ gelten. Dazu müssen p und q in den Vektoren vorkommen und in $C_i = C_j$ liegen, was nicht möglich ist.

Also gilt $i \neq j$. Dann sind beide Vektoren ganz in C_i bzw. C_j, da sonst schon in der ersten Komponente keine Gleichheit entstehen kann. Angenommen, es gibt eine Komponente $x_r \notin \{p, q\}$, so würde diese unverändert bleiben, also $y_r = x_r$, läge also zugleich in C_i und C_j was nicht möglich ist, da $\{p,q\} \subset C_i \cup C_j$ gilt. Also ist $i \neq j$ und $\{\mathbf{x},\mathbf{y}\} = \{(p,\ldots,p),(q,\ldots,q)\}$.

Die Tupel

$$c_i^{(k)} := \{k\} \times c_i$$

sind für $1 \leq i,k \leq f(G)$ genau $f(G)^2$ verschiedene Zeugen, die

$$c_i^{(0)} := \tau(i,c_i)$$

sind $f(G)$ Zeugen. Diese sind von den ersten verschieden, da alle gezählten x_i mit Marke 0 versehen sind.

Es könnten jetzt noch ursprünglich verschiedene gleich werden. Dies passiert höchstens für Zeugen mit $\mathbf{x}_i = (p,\ldots,p)$ und $\mathbf{x}_j = (q,\ldots,q)$. Es können also so höchstens zwei verschiedene Zeugen zu einem werden, es entstehen auf diese Art mindestens $f(G) - 1$ Zeugen.

Es muß noch gezeigt werden, daß alle diese Zeugen φ erfüllen. Dazu wird gezeigt, daß die atomaren Formeln, also die Prädikate auf den Elementen der Zeugen, genauso belegt sind bzw. belegt werden können wie in G.

Da die $c_i^{(0)}$ durch eine Abbildung entstehen, werden gleiche Elemente auf gleiche Bilder abgebildet. Zwei Elemente x und y in $c_i^{(0)}$, die gleich würden, also $\tau(i,x) = \tau(i,y)$ wären beide in C_i, das dann p und q umfassen müßte.

Für zwei Elemente x und y in c_i die mit einem Pfeil verbunden sind, ist es wegen der Definition der C_i nicht möglich, daß $x \in C_i$ und $y \notin C_i$ gilt. Falls beide nicht in C_i sind, gibt es den Pfeil zwischen (i,x) und (i,y), ansonsten zwischen $(0,\pi(x))$ und $(0,\pi(y))$.

Angenommen, es gäbe x und y in c_i, die nicht durch einen Pfeil verbunden sind, aber es gibt einen Pfeil zwischen $\tau(i,x)$ und $\tau(i,y)$. Dann müssen x und y entweder beide in C_i liegen, oder keines. Wenn sie nicht in C_i liegen, so gäbe es einen Pfeil zwischen (i,x) und (i,y), was nicht der Fall ist. Also sind beide in C_i.

Also sind $(0,\pi(x))$ und $(0,\pi(y))$ durch einen Pfeil verbunden. Also gilt nicht $x = \pi(x)$ und $y = \pi(y)$. Genau eines von beiden wird also von q nach p verändert, das andere auf sich selbst abgebildet. Also sind x und y mindestens $q - p - 2$ Positionen voneinander entfernt, ein Widerspruch zu $\{x,y\} \subseteq C_i$.

Außerdem wählt man $\mathbf{R}''$ so, daß die Zeugen auch im Sinne der $\mathbf{R}$ passen:

Setze $\tilde{R}_\lambda = \{\, ((i,z_1),\ldots,(i,z_r)) \mid 1 \leq i \leq f(G), (z_1,\ldots,z_r) \in R_\lambda \,\}$. und

$$\tau(\mathbf{z}) \in \hat{R}_\lambda \;\Leftrightarrow\; \mathbf{z} \in R_\lambda$$

und $R''_\lambda = \hat{R}_\lambda \cup \tilde{R}_\lambda$.

Dies könnte für gleiche Bilder verschiedener Vektoren zu widersprüchlichen Definitionen führen. Da derartige Urbilder nur die Vektoren $(p, \dots, p)$ und $(q, \dots, q)$ sind, und auf diesen alle R_λ gleichartig belegt sind, ist diese Definition statthaft. Da $\tilde{R}_\lambda \cap (\{1, \dots, f(G)\} \times V) \subset \hat{R}_\lambda$ ist, sind diese Mengen auf den normalen Kopien des Graphen konsistent.

Somit erfüllen alle diese Zeugen die Formel φ, es entsteht der Widerspruch $f(G)^2 + f(G) - 2 = f(G') \geq f(G)^2 + f(G) - 1$. ∎

Die Klassen $\max^2 \#^1\Sigma_1$ und $\max^2 \#^1\Sigma_0[<]$ sind unvergleichbar, und dadurch echte Teilmengen von $\max^2 \#^1(\text{Ordnung})\Sigma_1$. Dies ergibt sich aus

Lemma 63

$$\#^1\Sigma_1 \not\subset \max^2 \#^1(\text{Ordnung})\Sigma_0$$

Beweis: Das Problem DIST2GRAPH aus Satz 55 trennt auch diese Klassen. Dazu muß in dem dort gegeben Beweis lediglich die Anzahl der Prädikate unberücksichtigt bleiben, dies beeinträchtigt die Argumentation aber nicht. ∎

Lemma 64
Es gilt

$$\max^2 \#^1(\text{Ordnung})\Pi_1 \subseteq \max^2 \#^1\Pi_1 \, .$$

Beweis: Dies gilt mit der Formel aus Definition 22. ∎

So ergibt sich folgende echte Hierarchie der Maximierungsklassen mit und ohne Ordnung, die die später vorgestellte bekannte Hierarchie ohne Ordnung um die Klasse $\max^2 \#^1\Sigma_0[<]$ erweitert.

$$\begin{array}{ccccc} & \diagup & \max^2 \#^1\Sigma_1 & \diagdown & \\ \max^2 \#^1\Sigma_0 & & & & \max^2 \#^1\Pi_1 = \max^2 \#^1\Pi_1[<] \\ & \diagdown & \max^2 \#^1\Sigma_0[<] & \diagup & \end{array}$$

Auch hier ergibt sich ein Zusammenhang zum Maschinenmodell und mit der Faginschen Formel bzw. der Skolemisierung ein Kollaps der Hierarchie.

Satz 65

$$\mathbf{maxPB} = e(\max^2 \#^1\text{FO})$$

Beweis: Eine entsprechende Maschine erzeugt nichtdeterministisch alle in Frage stehenden Prädikate, wertet in polynomieller Zeit (Lemma 37) die gesamte Tabelle aller Belegungen der ersten Ordnung aus und kann so auch einen Zählquantor auswerten. Den so ermittelten Wert gibt sie unär aus. Durch den Auswertungsmodus solcher Maschinen ergibt sich die Behauptung.

Die für die andere Richtung benötigte Formel besteht zunächst aus der Beschreibung φ der gültigen Berechnungstafeln aus Lemma 38.

$$\max_{\mathbf{R}}{}^{2}\ \#w\ \varphi \wedge (C_c \text{ ist Auswahl}) \wedge (O \text{ ist String}) \wedge O_1(w)$$

Falls in $\mathbf{R}$ eine korrekte Berechnungstafel kodiert wird, wertet das $\#w$ zur von der Maschine ausgegebenen Zahl aus. Falls eine fehlerhafte Berechnungstafel kodiert wird, ergibt sich $\#w$ zu 0, der Wert wird ignoriert. So wird das Maximum über alle Ausgaben der Maschine gebildet. ■

Auch hier kollabieren die Hierarchien unterschiedlich, je nachdem welche syntaktischen Möglichkeiten erlaubt sind. Dabei wird ausgenutzt, daß ein Maximierungsquantor gleichzeitig ein Existenzquantor ist.
Dementsprechend gilt nach Satz 43:

$$\begin{aligned}\max{}^2\ \#^1\mathrm{FO}[<] &= \max{}^2\ \#^1\Pi_2[<] = \max{}^2\ \#^1\Pi_1[=,\mathrm{succ}] \\ &= \max{}^2\ \#^1(\mathrm{Funktion})\Pi_1[=]\,.\end{aligned}$$

Außerdem stellt sich heraus, daß der Maximierungsteil der zweiten Stufe den der ersten aufnehmen kann. Auch hier handelt es sich um eine Skolemisierung. Der weitergehende Kollaps entsteht dadurch, daß das neue Prädikat lediglich eine partielle Funktion sein dürfte. Dadurch, daß nach einem Maximum gesucht wird, ist diese „automatisch“ an den entscheidenden Stellen definiert.

Lemma 66
Es gilt

$$\#^1\exists^1 \subseteq \max{}^2\ (\mathrm{injektiv})\#^1\,.$$

Beweis: Es gilt

$$\mathrm{val}\Big(\#\mathbf{w}\ \exists\mathbf{x}\ \varphi(\mathbf{w},\mathbf{x})\Big) \equiv \mathrm{val}\Big(\max_{R}{}^2\ (\mathrm{injektiv}\ R)\#\mathbf{w}\#\mathbf{x}'\ \varphi(\mathbf{w},\mathbf{x}')\wedge R(\mathbf{w},\mathbf{x}')\Big)\,.$$

Wegen des Maximierungsquantors genügt es, solche R zu betrachten, die tatsächlich injektiv sind. Für festes $\mathbf{w}$ wertet jeweils der Rest der Formel gleich aus. Wenn $\exists\mathbf{x}\varphi(\mathbf{w},\mathbf{x})$ nicht gilt, so ist auch $\#\mathbf{x}'\ldots$ gleich 0. Gilt

$\exists \mathbf{x}\varphi(\mathbf{w}, \mathbf{x})$, dann ist genau ein solches $(\mathbf{w}, \mathbf{x}) \in R$, die Formel wertet zu 1 aus. ∎

In einer Logik mit $<$ liefert auch Lemma 53 ein ähnliches Ergebnis, dort wird aber anstelle des Quantors zweiter Stufe einen Allquantor der ersten Stufe benötigt.

Lemma 67
Es gilt für $k \geq 1$

$$\#^1\Sigma_{k+1} \subseteq \max{}^2\#^1\Pi_k \,.$$

Beweis: Es genügt zu zeigen, daß ein Block von Existenzquantoren durch einen von Allquantoren ersetzt werden kann.

Da im letzten Satz die Eigenschaft von R, injektiv zu sein, unabhängig von w und x' ist, können die entsprechenden Quantoren vertauscht werden. Dabei kann die Formel mit $\wedge$ um das Stück

$$\forall w' \forall x_1 \forall x_2((R(w', x_1) \wedge R(w', x_2)) \to x_1 = x_2)$$

erweitert, und so die Injektivität erzwungen werden. ∎

Damit ergibt sich für die Klassen, die $\max^2 \#^1\Pi_1$ umfassen, in denen die Ordnung keine Rolle mehr spielt:

Satz 68
Es gilt

$$\begin{aligned}\max^2 \#^1\mathrm{QF} \subsetneq \max^2 \#^1\Sigma_1 \subsetneq \max^2 \#^1\Sigma_2 = \max^2 \#^1\Pi_1 & \\ \subsetneq \max^2 \#^1\Pi_2 = \max^2 \#^1\mathrm{FO} = \mathbf{maxPB}.\end{aligned}$$

Die Echtheit der letzten beiden Inklusion gilt nach Thakur.

Die von Thakur untersuchte Klasse sogenannter „feasible“ (machbarer) Logik, in der die Mächtigkeiten von Relationen maximiert werden, ergibt sich in der Quantorendarstellung zu

Beobachtung 69

$$\max \mathrm{F}\,\mathrm{FO} = \max_{\mathbf{R}}{}^2 \ \max_{\mathbf{S}}{}^2 \ |S|_0\,\mathrm{FO}.$$

4.5 Minimierungsklassen

Zunächst erscheinen Maximierung und Minimierung als derselbe Vorgang, lediglich mit unterschiedlichen Vorzeichen. Es zeigt sich, daß dies insbesondere für die Struktur der Hierarchie nicht gilt. In der Situation hier erweist sich der Minimierungsquantor als stärkeres Strukturmittel, er kann mehr Funktionalität von den Quantoren erster Stufe aufnehmen. Zu dieser Asymmetrie kommt es wohl durch die Art und Weise, wie der Zählquantor Existenzquantoren ersetzen kann.

Zunächst stellt man fest, daß auch hier die von Thakur in [Tha92] verwendete Definition über Mächtigkeiten von Mengen in den Quantoren darstellbar ist.

Beobachtung 70
Für eine Formel φ mit freien Prädikaten $\mathbf{R} = R_1, \ldots, R_k$ und freien Variablen der ersten Stufe $\mathbf{x} = x_1, \ldots, x_j$ gilt für passende Strukturen $\mathcal{J}$

$$\mathrm{val}(\mathcal{J} \models \min_{R_1}^2 \cdots \min_{R_k}^2 \#x_1 \cdots \#x_j \varphi) = \min_{\mathbf{R}} \left| \{\mathbf{x} \colon (\mathcal{I}, \mathbf{R}) \models \varphi\} \right| .$$

Satz 71
Es gilt für $k \geq 0$

$$\mathbf{minPB} = e(\min^2 \#^1 \mathrm{FO}) = e(\min^2 \#^1 \Sigma_2) .$$

Beweis: Die Auswertung einer solchen Formel funktioniert mit demselben Programm wie beim Maximieren, wobei ablehnende Pfade einen sehr großen Wert ausgeben, so daß der veränderte Auswertungsmechanismus für das korrekte Ergebnis sorgt.

Die für die andere Richtung benötigte Formel besteht zunächst aus der Beschreibung φ der gültigen Berechnungstafeln aus Lemma 38.

$$\min_{\mathbf{R}}^2 \#w(\varphi \wedge (C_c \text{ ist Auswahl}) \wedge (O \text{ ist String})) \to O_1(w)$$

Falls in $\mathbf{R}$ eine korrekte Berechnungstafel kodiert wird, wertet das $\#w$ zur von der Maschine ausgegebenen Zahl aus. Falls eine fehlerhafte Berechnungstafel kodiert wird, gilt die Implikation für alle w, der Zählquantor wertet zu n^k aus. Also wird tatsächlich das Minimum über alle Ausgaben der Maschine gebildet.

Da die bereits vorgestellten Formelteile im negierten Teil der Implikation stehen, handelt es sich um eine Σ_2-Formel. ∎

Analog folgt der Kollaps mit succ und $(\mathrm{Funktion})$ auf $\min^2 \Sigma_1$, mit dem folgenden sogar auf $\min^2 \Sigma_0$.

Lemma 72
Es gilt für $k \geq 0$

$$\#^1\Sigma_{k+1} \subseteq \min^2 \#^1\Pi_k .$$

Beweis: Nach [Tha92]. Ohne Einschränkung sei in der folgenden Darstellung die Stelligkeit von $\mathbf{w}$ und $\mathbf{x}$ wegen $\mathrm{val}(\#x\,\varphi) \equiv \mathrm{val}(\#x\#y\,\varphi \wedge x = y)$ gleich. Dann gilt mit $\varphi' := \Big(\mathbf{w} = \mathbf{x} \wedge R(\mathbf{w})\Big) \vee \Big(\varphi \wedge \neg R(\mathbf{w})\Big)$

$$\mathrm{val}\Big(\#\mathbf{w}\,\exists\mathbf{x}\,\varphi\Big) \equiv \mathrm{val}\Big(\min_R{}^2\,\#\mathbf{w}\#\mathbf{x}\,\varphi'\Big) .$$

Sei im folgenden $\hat{R} := \{\, \mathbf{w} \mid \exists\mathbf{x}\varphi \,\}$.

„$\geq$“ gilt mit $R = \hat{R}$. Hier gilt für festes w: $\mathrm{val}(\exists\mathbf{x}\varphi) \equiv \mathrm{val}(\#\mathbf{x}\varphi')$. Wenn $\exists\mathbf{x}\varphi$ nicht gilt, ist dies richtig (der erste Teil ist wegen R sicher falsch, der zweite wegen φ). Wenn $\exists\mathbf{x}\varphi$ gilt, gilt φ' genau für $x = w$, denn $\neg R(\mathbf{w})$ ist sicher falsch.

„$\leq$“ Sei R' eine Relation, für die das Minimum angenommen wird. Dann gilt $f(R') = f(R' \cup \hat{R})$. Dabei ist klar, daß der Wert für die Vereinigung mindestens so groß ist wie der der Relation R', da dieses das Minimum darstellt. Angenommen, dies sei echt kleiner. Dann gibt es beim sukzessiven Einfügen der Elemente von R zu R' ein erstes w, bei dem Ungleichheit auftritt. Durch Konstruktion kann sich lediglich $\#\mathbf{x}\varphi'$ für dieses w ändern, und zwar von 0 auf 1. Es gibt also links einen neuen Beitrag, rechts gab es keinen. Dann gilt aber $\exists x\varphi$ nicht, $w \notin R$.

Es gilt $f(R) \leq f(R \cup R')$, und somit die Behauptung. ∎

Satz 73
Es gilt

$$\min^2 \#^1\mathrm{QF} = \min^2 \#^1\Sigma_1 \subsetneq \min^2 \#^1\Pi_1 = \min^2 \#^1\mathrm{FO} = \mathbf{minPB} .$$

Beweis: [Tha92], Seite 45. ∎

Es zeigt sich, daß die vorgestellten Quantoren flexibel genug sind, auch die anderen von Thakur untersuchten Klassen zu beschreiben, etwa:

Beobachtung 74
Es gilt

$$\mathrm{MIN\,F\,FO} = \min_{\mathbf{R}}{}^2\ \min_{\mathbf{S}}{}^2\ |\mathbf{S}|_\infty\,\mathrm{FO}\ .$$

4.6 Optimieren modulo n

In den letzten beiden Abschnitten ergaben sich in scheinbar symmetrischen Situationen unterschiedliche Ergebnisse. Diese Symmetrie zeigt sich in der Quantorendarstellung. Durch den geeignet definierten Negationsmodifikator können tatsächlich Minimum- und Maximumbildung aufeinander zurückgeführt werden.

Zunächst ergibt sich durch $N - x$ mit dem Bereich N der Formel der bereits bekannte Modifikator, mit dem in Analogie zu Existenz- und Allquantor, die bekanntlich auch als Maximum und Minimum darstellbar sind, im folgenden Sinn:

$$\neg \max \varphi = \min \neg\varphi$$

Interessant ist das Verhalten des $\#$-Quantors. Für diesen gilt:

$$\neg\#\varphi = \#\neg\varphi$$

Dabei ist zu beachten, daß genau in dieser Situation der Bereich des Modifikators verändert wird.

Dabei stellt sich heraus, daß dieses Rechnen mit jeweils angepaßtem Bereich innerhalb der Formel korrekt bleibt.

Allerdings lassen sich die Resultate, die das eigentliche Ziel dieser Dualität sind, über Klassen von Formeln und den Kollaps von Hierarchien nicht so einfach übertragen. Dies liegt daran, daß durch entsprechende Veränderungen im allgemeinen der Bereich der Formel verändert wird. So wird z.B. aus dem Maximierungsproblem MAXCLIQUE mit Bereich n durch einfaches Anwenden des Negationsmodifikators das Minimierungsproblem MINVERTEXCOVER. Dieses kann in der Klasse $\min \Pi_0$ dargestellt werden. Dabei verändert sich allerdings der Bereich auf n^2, so daß durch nochmaliges Anwenden des Negationsmodifikators nicht mehr MAXCLIQUE sondern nur MAXCLIQUE $+ n^2 - n$ dargestellt werden kann.

Dieses Verhalten kann man vermeiden, wenn man den Wert der Formeln modulo n betrachtet. Dann sind zwar nur noch Probleme erlaubt, die im wesentlichen Teilmengen des Universums optimieren, dies trifft aber auf alle hier vorgestellten zu. Dies kann in der verwendeten Darstellung wieder leicht durch einen Modifikator geschehen.

Definition 75
Der mod*-Modifikator hat die Semantik*

$$\mathrm{val}(\mathcal{J} \models \mathrm{mod}\,\varphi) := \mathrm{val}(\mathcal{J} \models \varphi) \bmod |\mathcal{J}|\,.$$

Lemma 76
Für eine Formel φ mit Bereich n^k für festes $k \geq 1$ und n als Größe des Universums gilt

$$f = \mathrm{mod}\varphi \quad \Leftrightarrow \quad \neg f = \mathrm{mod}\neg\varphi,$$

wobei $\neg f$ für $n - f$ steht.

Beweis: Da $n^k \bmod n = 0$ gilt, folgt mit der Definition des Negationsmodifikators die Behauptung. ∎

Dabei ist zu beachten, daß dieser Modifikator im allgemeinen nicht vertauscht, es also wichtig ist, daß zuerst optimiert und dann modulo gerechnet wird.

Lemma 77
Es gilt für $k \geq 1$

$$\mathrm{mod}\,\mathrm{min}^2\,\#\Pi_{k+1} \subseteq \mathrm{mod}\,\mathrm{min}^2\,\#\Sigma_k\,.$$

Beweis: Sei $f(\mathcal{J}) := \mathrm{val}(\mathcal{J} \models \mathrm{mod}\,\mathrm{min}^2_R \#\mathbf{x}\varphi)$ mit $\varphi \in \Pi_{k+1}$. Dann ist

$$\neg f = \neg\mathrm{mod}\,\min_R{}^2\,\#\mathbf{x}\varphi = \mathrm{mod}\,\max_R{}^2\,\#\mathbf{x}\neg\varphi.$$

Da $\neg\varphi \in \Sigma_{k+1}$ ist, gibt es nach Lemma 67 ein $\psi \in \Pi_k$ mit

$$\neg f = \mathrm{mod}\,\max_{R'}{}^2\,\#\mathbf{x}'\psi.$$

Setze $\varphi' := \neg\psi$. Dann gilt

$$f = \mathrm{mod}\,\min_{R'}{}^2\,\#\mathbf{x}'\varphi'\,,$$

wobei $\varphi' \in \Sigma_k$ ist. ∎

Satz 78
Die Quantorenhierarchie der ersten Stufe in den Klassen $\mathrm{mod}\,\mathrm{min}\,\mathrm{FO}$ *und* $\mathrm{mod}\,\mathrm{max}\,\mathrm{FO}$ *kollabieren auf die Stufe* $\Pi_0 = \Sigma_0 = QF$.

Beweis: Iteriertes Anwenden von Lemma 72 und Lemma 77 liefert das Ergebnis für min^2 . Die Aussage für max^2 entsteht durch Anwenden des Negationsmodifikators. ∎

Dieses Ergebnis paßt zur ursprünglichen Verwendung der vorgestellten Klassen, der Untersuchung von Approximationseigenschaften **NP**-harter Optimierungsprobleme. Auch dort ist eine Veränderung des Bereiches, in dem sich die Lösung befinden darf, unter Umständen gefährlich, da die Approximationsgüte im allgemeinen als Verhältnis gemessen wird.

Auch dort müssen Reduktionen nicht nur sehr leicht berechenbar sein, sondern auch bestimmte Verhältnisse erhalten. Dies wird von einer modulo Rechnung geradezu typisch verletzt.

4.7 Charakterisierung von optP

Im Gegensatz zu den gerade vorgestellten Klassen, die nur polynomielle Optimalwerte bestimmen können, gibt es auch die bereits definierte Klasse $\mathbf{optP}$, in der die verschiedenen Ausgaben von nichtdeterministischen Turingmaschinen ausgewertet werden, indem das Optimum bestimmt wird. Dabei wird die Ausgabe binär oder dyadisch interpretiert und diese Werte verglichen. Hier sind also auch in der Eingabelänge exponentiell große Ergebnisse möglich.

Dieses Vorgehen läßt sich wieder mit einem Modifikator darstellen. Da dieser eine Relation, die die Darstellung einer Zahl als String codiert, auswerten soll, benötigt dieser im Gegensatz zum $|\cdot|$-Modifikator ein geordnetes Universum.

Definition 79
Die Formel $\mathrm{Bin}(S)\varphi$ *über einer Signatur mit Ordnung* $<$ *hat den Wert*

$$\mathrm{Bin}(S)\varphi := \begin{cases} M & \textit{falls } \varphi \neq 0 \\ 0 & \textit{falls } \varphi = 0 \end{cases}$$

wobei M *für die Zahl steht, die durch* S *repräsentiert wird. Dazu wird aus der* k*-stelligen Relation* S *ein binärer String der Länge* n^k *erzeugt, der gerade die charakteristische Funktion der Relation über die durch* $<$ *lexikographisch geordneten* k*-Tupel angibt, indem also an einer Stelle im String genau dann eine 1 steht, wenn das der Position entsprechende Tupel in der Relation ist. Dieser String wird dann als Binärzahl interpretiert.*

Satz 80

$$\begin{aligned} \mathbf{optP} &= e(\mathrm{opt}^{\,2}\,\mathrm{Bin}()\mathrm{FO}[<]) = e(\mathrm{opt}^{\,2}\,\mathrm{Bin}()\Pi_2[<]) \\ &= e(\mathrm{opt}^{\,2}\,\mathrm{Bin}()\Pi_2[\mathrm{succ}]) = e(\mathrm{opt}^{\,2}\,\mathrm{Bin}()(\mathrm{Funktion})\Pi_2) \end{aligned}$$

Beweis: Dies ergibt sich wie Satz 51 und durch die passende Definition der Turingmaschinen. ∎

Da in den entsprechenden Teilklassen nicht gezählt werden kann, und daher auch die Mächtigkeit einer Relation nicht in Dualdarstellung umgewandelt werden kann, dürfte es dort kaum natürliche Probleme geben. Nichtsdestoweniger sind Trennungsresultate zu erwarten, die den bereits vorgestellten sehr ähnlich sind.

5

Zusammenfassende Betrachtung

Die vorgestellten Quantoren geben eine Systematik an, in der noch einige weitere Klassen gebildet und untersucht werden können. Man könnte die hier vorgestellten Quantoren der zweiten Stufe noch auf verschiedene Art verknüpfen. Dabei gewinnt man zum einen Berechnungskraft und erhält immer flachere Hierarchien, da diese Quantorenkombinationen eher noch mehr Funktionalität der ersten Stufe übernehmen könnten.

Beliebige All- und Existenzquantoren in der zweiten Stufe führen zur Polynomialzeithierarchie. N. Immerman stellt in [Imm87] auch einige Resultate in dieser Richtung vor. Dort wird auch darauf eingegangen, daß die Stelligkeit der Prädikate mit der Laufzeit der Maschinen in Zusammenhang steht.

Diese Erweiterungen verändern aber vermutlich den Charakter der Ergebnisse nicht. Man wird weiterhin sehr feinfühlige Reaktionen der Ergebnisse auf minimale Veränderungen der Definitionen, insbesondere der Eingabekonventionen, erwarten müssen. Dies zwingt zu einer sehr ausführlichen Darstellung, in der es bekanntlich schwierig ist, grundsätzliche Zusammenhänge auf den Punkt zu bringen. Dies verwundert nicht, da man gerade die elementaren Möglichkeiten Formeln zu bilden einschränkt.

Index

Literaturverzeichnis

[EFT92] H.-D. Ebbinghaus, J. Flum, W. Thomas. *Einführung in die mathematische Logik*. BI-Wiss.-Verl., 3. Auflage, 1992.

[Fag74] Ronald Fagin. Generalized first-order spectra and polynomial-time recognizable sets. *SIAM—AMS Proceedings*, 7:43–73, 1974.

[Grä92] Erich Grädel. Capturing complexity classes by fragments of second-order logic. *Theoretical Computer Science*, (101):35–57, 1992.

[Imm87] Neil Immerman. Languages that capture complexity classes. *SIAM J. Comput.*, 16(4):760–778, August 1987.

[Pap94] Christos H. Papadimitriou. *Computational Complexity*. Addison-Wesley Publishing Company, Inc., 1994.

[PY91] Christos H. Papadimitriou Mihalis Yannakakis. Optimization, approximation, and complexity classes. *Journal of computer and system sciences*, 43:425–440, 1991.

[SST92] Sanjeev Saluja, K.V. Subrahmanyam, Madhukar N. Thakur. Descriptive complexity of #P functions. In *Proceedings of the 7th Annual Conference on Structure in Complexity Theory (Boston University, Boston, Massachusetts, June 22–25, 1992)*, pages 169–184, Los Alamitos-Washington-Brussels-Tokyo, 1992. IEEE Computer Society Press.

[Tha92] Madhukar Narayan Thakur. *Descriptive Complexity of Optimization and Counting Problems*. Dissertation, University of California, Santa Cruz, 1992.