Gottfried Wilhelm
Leibniz Universität Hannover
Fakultät für Elektrotechnik und Informatik
Institut für Theoretische Informatik

Zufallsbegriffe in der Berechenbarkeitstheorie

Masterarbeit

im Studiengang M.Sc. Informatik

von

Sabrina Alexandra Gaube
Matrikelnr: 3069800

Prüfer: Prof. Dr. Vollmer
Zweitprüfer: PD Dr. Meier

Hannover, 24. August 2020

Erklärung der Selbstständigkeit

Hiermit versichere ich, dass ich die vorliegende Masterarbeit selbständig und ohne fremde Hilfe verfasst und keine anderen als die in der Arbeit angegebenen Quellen und Hilfsmittel verwendet habe. Die Arbeit hat in gleicher oder ähnlicher Form noch keinem anderen Prüfungsamt vorgelegen.

Hannover, den 24. August 2020

Sabrina Alexandra Gaube

Zusammenfassung

In dieser Arbeit wird die vorherige Bachelorarbeit der Autorin fortgeführt. Darin ging es um das Konzept der Kolmogorov-Komplexität im Zusammenhang mit Datenkompressionen. Im Zuge dessen wurde der Begriff von zufälligen Wörtern eingeführt, die mithilfe der Kolmogorov-Komplexität definiert wurden.
In der Literatur finden sich weitere Zufallsbegriffe in der Berechenbarkeitstheorie, die wir in dieser Arbeit betrachten wollen. Dazu werden zunächst in Kapitel 1 die wichtigsten Aussagen der Bachelorarbeit wiederholt und kurz aufgegriffen. Beweise sind an dieser Stelle weggelassen und es wird auf entsprechende Literatur bzw. die zurückliegende Bachelorarbeit verwiesen.
In Kapitel 2 geht es um einen weiteren Zufallsbegriff, nämlich den Zufallsbegriff nach Per Martin-Löf. Dieser beruht auf konstruktiven Nullmengen und bedarf einiger Grundlagen der Maßtheorie, sowie statistischer Tests. Der Zufallsbegriff wird zunächst für endliche Wörter definiert, die Äquivalenz zur Zufälligkeit nach Kolmogorov bewiesen und später auf unendliche Wörter erweitert.
In Kapitel 3 geht es um einen dritten Zufallsbegriff, der von Claus Schnorr eingeführt wurde und auf Martingalen beruht. Auch dieser wird zunächst für endliche Wörter, anschließend für unendliche Wörter definiert und die Äquivalenz zu den anderen beiden Zufallsbegriffen bewiesen. In dieser Arbeit wird nur die Äquivalenz zwischen dem Zufallsbegriff nach Per Martin-Löf und Claus Schnorr gezeigt, jedoch folgt aus Transitivitätsgründen direkt auch die Äquivalenz von Kolmogorovs und Schnorrs Begriff.
Abschließend wird in Kapitel 4 ein Zusammenhang zur Komplexitätstheorie, bzw. die Auswirkungen der „Resource-bounded Measuretheory“ auf die Komplexitätsklassen E und EXP, sowie ihrer schweren bzw. vollständigen Probleme gezeigt.

Inhaltsverzeichnis

1 Zufallsbegriff nach Kolmogorov

In diesem Kapitel wollen wir die wichtigsten Definitionen und Sätze der Bachelorarbeit der Autorin [Gau18] in Erinnerung rufen.

1.1 Motivation der Kolmogorov-Komplexität

Wir zitieren zunächst die ursprüngliche Motivation der Bachelorarbeit [Gau18], um im weiteren Verlauf eine bildlichere Beschreibung des Begriffs zu ermöglichen. Dort lag der Fokus auf Kolmogorov-Komplexität und Datenkompression und stellte in diesem Kontext einen Zusammenhang zu einem intuitiven Zufallsbegriff her, der später wohldefiniert, im Sinne von Äquivalenzklassen, bewiesen wurde. Dieser Gedankengang soll zur Illustration hier noch einmal aufgegriffen werden.

Beispiel 1.1. *Man kann das binäre Wort*

$$10000000000000000000000000000000000 = 10^{35}$$

kürzer darstellen durch „Schreibe 1 und dann 35 Nullen“. Diese Zeichenkette enthält 29 Zeichen.
Dadurch wird eine Kompression von 36 auf 29 Zeichen bewirkt.

Möchte man Daten komprimieren stellen sich zunächst folgende Fragen:

- Inwiefern ist es möglich Daten zu komprimieren?
- Kann man alle Wörter über dem binären Alphabet komprimieren?
- Wo liegen Grenzen der Datenkompression?

Als Erstes wollen wir uns daher systematisch anhand von Beispielen überlegen, welche Wörter man komprimieren kann, um damit eine Illustration des Zufallsbegriffs zu schaffen.

Beispiel 1.2. *(Siehe [Gau18] Beispiel 2 (gekürzt))*

1. *Betrachten wir das Wort*

$$00000000000000000000000000 = 0^{26}$$

so sehen wir, dass diese Zahl lediglich aus 26 Nullen besteht. Diese Folge ist komprimierbar durch einen zu Beispiel 1.1 analogem Algorithmus (z.B. „Print 26 times a 0“. Hierdurch erhalten wir eine Kompression auf 18 Zeichen).

2. *Das nächste zu betrachtende Wort ist das Wort*

$$01000110110000010100111001.$$

Das Wort sieht auf dem ersten Blick relativ willkürlich aus, jedoch kann man bei näherem Betrachten erkennen, dass es sich um eine Zeichenfolge mit folgendem Konstruktionsmuster handelt.
Wir haben 0,1,00,01,10,11,000,001,010,011,100,... *also Wörter über dem Alphabet* $\{0,1\}$ *lexikographisch aufgezählt und nach* 26 *Zeichen wurde die Zeichenkette abgebrochen. Das heißt, zuerst kommen die Wörter mit einem Zeichen, dem Wert aufsteigend sortiert, dann kommen die Wörter mit zwei Zeichen, dem Wert aufsteigend sortiert usw.*
Wir erkennen ein Muster und vermuten daher, dass es sich um ein komprimierbares Wort handelt.

Rein intuitiv kann man sagen, dass Wörter, die in gewisse Muster fallen, komprimierbar wirken. Die Wörter, die hingegen 'zufällig' wirken, scheinen nicht komprimierbar zu sein. Wir sehen also einen Zusammenhang zwischen Unkomprimierbarkeit und Zufälligkeit. Diese Anschauung soll uns im weiteren Verlauf der Arbeit immer wieder begegnen.
Diese Zufälligkeit soll nun durch den Begriff der Kolmogorov-Komplexität präziser formuliert werden.

1.2 Definition der Kolmogorov-Komplexität

Definition 1.3 (Kolmogorov-Komplexität)**.**
Die Kolmogorov-Komplexität $K_M(x)$ *eines Wortes* $x \in \{0,1\}^*$ *bezüglich einer Turingmaschine* M *ist definiert als die Länge der kürzesten Eingabe* $p \in \{0,1\}^*$ *mit* $M(p) = x$*.*

$$K_M(x) := \min\{\, |p| : p \in \{0,1\}^*, M(p) = x \,\}$$

Gibt es ein solches p für ein x nicht, so setzen wir

$$K_M(x) := \infty.$$

Definition 1.4. *Eine Turingmaschine M_2 heißt additiv optimal in Bezug zu einer weiteren Turingmaschine M_1, wenn es eine Konstante c so gibt, dass*

$$K_{M_1}(x) \leq K_{M_2}(x) + c,$$

für alle $x \in \{0,1\}^*$ gilt.

Satz 1.5. *Sei U eine universelle Turingmaschine und $x \in \{0,1\}^*$. Dann gilt*

$$K_U(x) \leq K_M(x) + c,$$

für jede Turingmaschine M. Hierbei hängt c nur von M und nicht von x ab.

Beweis. Siehe [Gau18]. □

Definition 1.6. *Es sei U eine fest gewählte universelle Turingmaschine. Dann definieren wir für $x \in \{0,1\}^+$ die Kolmogorov-Komplexität $K(x)$ durch*

$$K(x) := K_U(x).$$

Definition 1.7. *Die Kolmogorov-Komplexität $K(x)$ eines Wortes x ist die Länge der kürzesten Turingmaschinenbeschreibung, die x erzeugt.*

Definition 1.8.

- *Seien $x, y \in \{0,1\}^*$. Dann heißt*

  $$K(x \mid y)$$

  die bedingte Kolmogorov-Komplexität von x unter y und entspricht der Länge der kleinsten Turingmaschine, die unter Eingabe von y das Wort x erzeugt.

- *Die Kolmogorov-Komplexität einer natürlichen Zahl n ist definiert als Kolmogorov-Komplexität der Binärdarstellung dieser natürlichen Zahl.*

  $$K(n) = K(bin(n)).$$

Definition 1.9. *Die längenbedingte Kolmogorov-Komplexität eines Wortes x ist definiert als*

$$K(x \mid |x|).$$

Satz 1.10. *Es existiert eine Konstante $c \in \mathbb{N}$, sodass für alle x*

$$K(x) \leq |x| + c$$

gilt.

Satz 1.11. *Es existiert eine Konstante $c \in \mathbb{N}$, sodass für alle x und y*

$$K(x \mid y) \leq K(x) + c$$

gilt.

Diese beiden Sätze werden wir im nächsten Kapitel benötigen. Durch Einsetzen erhalten wir hierfür

$$K(x \mid y) \leq K(x) + c \leq |x| + c_1 + c =: |x| + c_2.$$

Theorem 1.12 (Invarianz-Theorem)**.**
Sei $x \in \{0,1\}^*$ und seien S, T universelle Turingmaschinen. Dann existiert eine Konstante c, so dass

$$|K_T(x) - K_S(x)| \leq c$$

gilt.

Mithilfe des Invarianz-Theorems lässt sich eine Äquivalenzrelation definieren, woraus sich Äquivalenzklassen ableiten lassen.

Definition 1.13. *Zwei Kolmogorov-Komplexitäten $K_A(x)$ und $K_B(x)$ heißen äquivalent, in Zeichen*

$$K_A(x) \equiv K_B(x),$$

wenn es eine Konstante $c \in \mathbb{N}$ gibt, so dass

$$|K_A(x) - K_B(x)| \leq c$$

gilt.

Dass diese Äquivalenzklassen aus Definition 2.37 auch wohldefiniert sind, in dem Sinne, dass $\equiv$ eine Äquivalenzrelation ist, wurde in [Gau18] bewiesen.

Wir haben die Kolmogorov-Komplexität ursprünglich eingeführt, um einen Zufallsbegriff zu definieren, nachdem ein Wort zufällig ist, wenn es unkomprimierbar ist. Daher definieren wir mithilfe der Kolmogorov-Komplexität die c-Komprimierbarkeit und die daraus resultierende Zufälligkeit.

Definition 1.14. *(nach [LV08])*
Sei $c \in \mathbb{N}$. Ein Wort $x \in \{0,1\}^+$ heißt c-komprimierbar, falls

$$K(x) < |x| - c$$

gilt.
Gilt $K(x) \geq |x| - c$, so heißt x c-unkomprimierbar oder zufällig.

Damit diese Definition des Begriffs zufällig überhaupt sinnvoll ist, sollten solche zufälligen Wörter auch existieren. Dies zeigt der folgende Satz:

Satz 1.15. *(nach [LV08])*
Es gibt mindestens

$$2^n - 2^{n-c+1} + 1$$

c-unkomprimierbare Wörter der Länge n.

Um diese Aussage noch einmal zu veranschaulichen, schauen wir uns nun dieses Beispiel an.

Beispiel 1.16. *(Aus [Gau18])*
Wir betrachten die Wörter, die 7-unkomprimierbar sind. Die 7-unkomprimierbaren Wörter sind nach Definition genau die Wörter, deren Kolmogorov-Komplexität geringer als $n - 7$ ist, also genau die Wörter, bei denen man durch eine andere Darstellung mindestens 7 Zeichen einspart.
Setzen wir diesen Wert für c in die Formel aus Satz 1.15 ein, erhalten wir

$$2^n - 2^{n-c+1} + 1 = 2^n - 2^{n-7+1} + 1 = 2^n - 2^{n-6} + 1$$

Wörter die 7-unkomprimierbar sind. Möchte man daraus nun den Anteil dieser Wörter unter den Wörtern der Länge n bestimmen, so berechnet man das grundsätzlich mit

$$\frac{2^n - 2^{n-6} + 1}{2^n} = 1 - \frac{1}{2^6} + \frac{1}{2^n} = \frac{63}{64} + \frac{1}{2^n}.$$

Da der Nenner mit wachsendem n immer größer wird, wird der hintere Bruch immer kleiner, also geht der Anteil der 7-unkomprimierbaren Wörter bei wachsendem n von oben gegen $\frac{63}{64} \approx 98,4375\%$. Für konkrete Werte für n erhalten wir

$n = 7$ *:* $\frac{63}{64} + \frac{1}{2^7} = \frac{127}{128} \approx 99,219\%$

$n = 10$*:* $\frac{63}{64} + \frac{1}{2^{10}} = \frac{1009}{1024} \approx 98,535\%$

$n = 15$: $\frac{63}{64} + \frac{1}{2^{15}} \approx 98,44\%$

$n = 16$: $\frac{63}{64} + \frac{1}{2^{16}} \approx 98,439\%$

Wir sehen also, dass bereits bei Wörtern der Länge 16 sich der Anteil der 7-unkomprimierbaren Wörtern vom Grenzwert erst in der dritten Nachkommastelle unterscheidet.

Außerdem haben wir weitere Eigenschaften der Kolmogorov-Komplexität bereits gesehen:

Satz 1.17. *Es gibt Wörter, die eine beliebig große Kolmogorov-Komplexität haben. Mit anderen Worten: Für alle $n \in \mathbb{N}$ existiert ein Wort $x \in \{0,1\}^*$ mit $K(x) \geq n$. Insbesondere ist $K(x)$ unbeschränkt.*

Beweis. Angenommen, es gibt ein $m \in \mathbb{N}$, so dass es kein Wort x gibt mit $K(x) \geq m$. Dann gibt es $\sum_{i=0}^{m-1} 2^i$ Wörter mit der Länge von maximal m.

$$\sum_{i=0}^{m-1} 2^i = 2^m - 1.$$

Allerdings gibt es genau 2^m Wörter der Länge m. Nach dem Schubfachprinzip wissen wir, dass es damit genau ein Wort gibt mit Kolmogorov-Komplexität größer m, da wir verlustfrei komprimieren wollen. Die Verlustfreiheit erzwingt, dass wir aus jeder komprimierten Darstellung eines Wortes nur genau ein ursprüngliches Wort erhalten können. Demnach haben wir einen Widerspruch, also ist unsere Annahme, dass es ein solches m gibt, falsch. □

Satz 1.18. *Es gibt maximal 2^n Wörter mit Kolmogorov-Komplexität n.*

Beweis. Es gibt genau 2^n verschiedene Wörter der Länge n über dem Alphabet $\{0,1\}$. Damit es mehr Worte x mit $K(x) = n$ gäbe, müsste mindestens eines dieser kürzeren Darstellungen mehr als ein Wort erzeugen, was der verlustfreien Kompression widerspricht, da diese eine eindeutige Zuordnung erzwingt. □

Damit haben wir gezeigt, dass es zufällige Wörter gibt und der Zufallsbegriff daher sinnvoll ist.

Theorem 1.19. *$K(x)$ ist nicht berechenbar.*

Beweis. Ein Beweis findet sich in [Gau18] oder [LV08]. □

Da wir später Äquivalenz zu anderen Zufallsbegriffen zeigen wollen, lässt diese Stelle bereits vermuten, dass es im Allgemeinen nicht entscheidbar ist, ob Wörter zufällig sind, da wir $K(x) < |x| - c$ nur entscheiden können, wenn wir $K(x)$ nicht berechnen können, bzw. keine obere Schranke dafür finden, welche kleiner $|x| - c$ ist. Wie man solche Oberschranken beispielsweise bestimmen kann, möchten wir uns an einem Beispiel abschließend ansehen. Weitere Beispiele hierzu finden sich in [Gau18].

Als Eingangsbeispiel (Beispiel 1.2) haben wir das Wort 0^{26} betrachtet und bereits erkannt, dass es komprimierbar ist. Dieses wollen wir hier nun wieder aufgreifen.

Beispiel 1.20. *Betrachten wir das Wort $x = 0^{26}$. Hierfür sei nun M eine Turingmaschine mit Eingabe 26 als Binärwort. Ersetze nun die Binärdarstellung von 26 in eine Unärdarstellung der Form 0^{26}.*
Wir benötigen also nur $\log 26$ des Platzes, den 0^{26} verbrauchen würde. Wenn c die Größe der Turingmaschine bezeichnet, dann gilt also

$$K(x) \leq \log 26 + c$$

als Oberschranke.
Vergleichen wir nun unsere anfangs bestimmte Komprimierung auf 18 Zeichen, so sehen wir, dass $\log 26 + c \approx 4,7 + c$ ist und damit unsere Komprimierung für kleine c relativ schlecht war.

Satz 1.21. *Sei x ein binäres Wort der Länge n. Hat x die Form 0^n oder 1^n, so gilt*

$$K(x) \leq \log n + c.$$

Alternativ könnte man den Satz auch für eine natürliche Zahl mit Binärdarstellung der Länge n, die nur aus Nullen oder nur aus Einsen besteht, formulieren, da $K(bin(x)) =: K(x)$. Insbesondere fallen für alle natürlichen Zahlen x alle Zahlen der Form $2^x - 1$ in diese Kategorie.

Beweis. Gehe hierbei wie im Beispiel vor und ersetze jede 26 durch ein n und ggf. jede 0 durch eine 1. □

2 Zufallsbegriff nach Per Martin-Löf

Dieses Kapitel richtet sich nach [ML66] und [LV08].
In diesem Kapitel geht es darum, den in Kapitel 1 eingeführten Begriff der Zufälligkeit nach Kolmogorov, siehe Definition 1.14, zu verfizieren. Das heißt konkret, dass wir die Korrespondenz zwischen Nichtzufälligkeit und dem Finden von regelmäßigen Strukturen in die Sprache der stochastischen Zufälligkeit übersetzen und auf unendlich lange Wörter erweitern wollen.

2.1 Maßtheoretische Grundlagen

Um keine stochastischen Grundlagen voraussetzen zu müssen, wollen wir an dieser Stelle zunächst die notwendigen stochastischen Begriffe definieren. Um dies formal korrekt tun zu können und auch um in Kapitel 4 zu der Ressource-bounded Measuretheory übergehen zu können, wollen wir an dieser Stelle zunächst maßtheoretische Grundlagen präsentieren. Dieses Unterkapitel richtet sich nach den Definitionen und Sätzen aus [AEE08].

Definition 2.1. *Sei $\Omega \neq \emptyset$ eine Menge. Eine Teilmenge $\mathcal{A}$ der Potenzmenge von Ω $\mathfrak{P}(\Omega)$ heißt σ-Algebra, wenn sie die folgenden Eigenschaften erfüllt:*

1. *$\Omega \in \mathcal{A}$*
2. *Gilt $A \in \mathcal{A}$, dann auch $\overline{A} \in \mathcal{A}$.*
3. *Gilt $(A_j) \in \mathcal{A}^{\mathbb{N}}$, so auch $\bigcup_{j\in\mathbb{N}} A_j \in \mathcal{A}$.*

Die oben definierten σ-Algebren sind als das zu deuten, was man aus der Wahrscheinlichkeitsrechnung aus der Schule als mögliche Ereignismengen kennt. Daher ist es insbesondere wichtig, dass die gesamte Ergebnismenge Ω in der Menge der möglichen Ereignisse liegt. Ebenso gilt dies für ein bestimmtes Ereignis. Ist dieses Ereignis A möglich, so soll auch das Gegenereignis $\overline{A}$ möglich sein. Der dritte

Unterpunkt bezieht sich nun darauf, dass wir die Ereignisse, welche möglich sind, beliebig vereinigen dürfen und sie dann immer noch möglich sein sollen.

Beispiel 2.2. *Für eine beliebige Menge Ω ist $\mathcal{A}_1 := \{\emptyset, \Omega\}$ die kleinstmögliche σ-Algebra und wird auch triviale σ-Algebra genannt. Die größtmögliche σ-Algebra mit Ω als Grundmenge ist $\mathcal{A}_2 := \mathfrak{P}(\Omega)$.*
Eine nichttriviale aber recht einfache σ-Algebra erhält man für eine beliebige Menge Ω und eine beliebige Teilmenge $A \subset \Omega$ mit $\{\emptyset, \overline{A}, A, \Omega\}$.

Definition 2.3. *Ein Tupel $(\Omega, \mathcal{A})$ heißt messbarer Raum, wenn Ω eine beliebige Grundmenge und $\mathcal{A}$ eine σ-Algebra ist.*

Bemerkung 2.4. *Einen messbaren Raum wollen wir im Folgenden dafür benutzen, um darauf ein Wahrscheinlichkeitsmaß definieren zu können, um dann auch mit Wahrscheinlichkeiten rechnen zu können.*

Definition 2.5. *Es sei $\mathcal{C}$ ein System von Teilmengen von Ω mit $\emptyset \in \mathcal{C}$ und J eine Indexmenge. Ferner sei ϕ eine Abbildung von $\mathcal{C}$ nach $[0, \infty]$ mit $\phi(\emptyset) = 0$. ϕ heißt σ-subadditiv, wenn für jede Folge $(A_j)_{j\in J}$ in $\mathcal{C}$ mit $\bigcup_{j\in J} A_j \in \mathcal{C}$ gilt:*

$$\phi\left(\bigcup_{j\in J} A_j\right) \leq \sum_{j\in J} \phi(A_j)$$

Die Abbildung ϕ mit $\phi(\emptyset) = 0$ heißt σ-additiv, wenn für jede disjunkte Folge (A_j) in $\mathcal{C}$ mit $\bigcup_j A_j \in \mathcal{C}$

$$\phi\left(\bigcup_{j} A_j\right) = \sum_{j} \phi(A_j)$$

gilt.

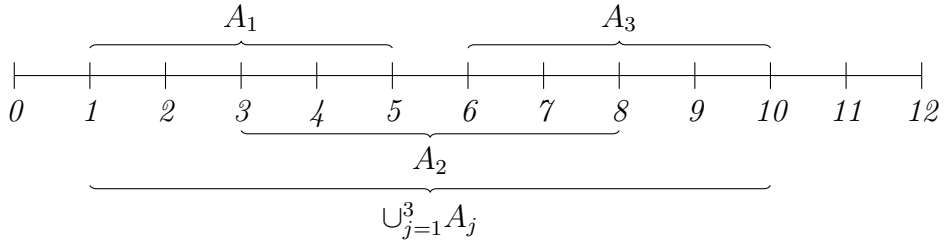
Diese Definition lässt sich durch das nachfolgende Beispiel veranschaulichen.

Beispiel 2.6. *Seien $\Omega = \mathbb{N}$, $\mathcal{A} = \mathcal{P}(\mathbb{N})$, $A_1 = \{1, \ldots, 5\}$, $A_2 = \{3, \ldots, 8\}$ und $A_3 = \{6, \ldots, 10\}$. Mit ϕ bezeichnen wir die Abbildung auf $\mathcal{C} = \{\emptyset, A_1, A_2, A_3, \Omega\}$, welche die Elemente jeder Menge zählt. Dann gilt*

$$\phi(A_1) = 5, \phi(A_2) = 6 \text{ und } \phi(A_3) = 5.$$

Es gilt $\phi(\emptyset) = 0$ und die Abbildung ϕ ist σ-subadditiv. Es gilt

$$\phi\left(\bigcup_{j=1}^{3} A_j\right) = 10 \leq \sum_{j=1}^{3} \phi(A_j) = 16.$$



Um die σ-Subadditivität zu zeigen, müssten wir dies noch für jede mögliche Vereinigung in $\mathcal{C}$ zeigen. Dieses möchten wir uns an dieser Stelle ersparen, da das Beispiel lediglich zur Anschauung dienen soll.

Die Abbildung ϕ ist ebenfalls σ-additiv. Es gilt

$$\phi(A_1 \cup A_3) = \phi(A_1) + \phi(A_3)$$

und $A_1 \cap A_3 = \emptyset$ und daher disjunkt.

Definition 2.7. *Es sei $\mathcal{A}$ eine σ-Algebra über Ω und $\mu \colon \mathcal{A} \to [0, \infty]$ sei σ-additiv. Dann heißt μ Maß auf Ω und $\mathcal{A}$ und $(\Omega, \mathcal{A}, \mu)$ heißt Maßraum. Gilt $\mu(\Omega) = 1$, so heißt μ auch Wahrscheinlichkeitsmaß und $(\Omega, \mathcal{A}, \mu)$ ist ein Wahrscheinlichkeitsraum.*

Als nächstes schauen wir uns einige Eigenschaften von Maßfunktionen (oder kurz: Maßen) an.

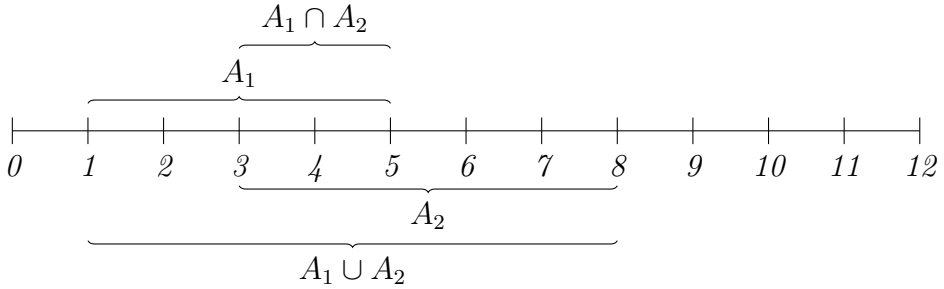
Bemerkung 2.8. *Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum. Für $A, B \in \mathcal{A}$ gilt:*

- $\mu(A \cap B) + \mu(A \cup B) = \mu(A) + \mu(B)$
- *Wenn $A \subset B$, dann $\mu(B \backslash A) = \mu(B) - \mu(A)$*
- *Wenn $A \subset B$, dann $\mu(B) \geq \mu(A)$*

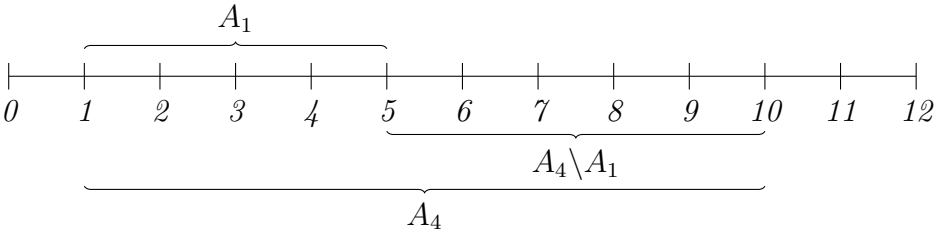
Diese Rechenregeln folgen direkt aus grundlegenden Regeln zur Mengenlehre.

Beispiel 2.9. *Wenn wir die Notationen und Mengen aus Beispiel 2.6 betrachten und nun eine Menge $A_4 = \{1, \ldots, 10\}$ hinzufügen, dann gelten*

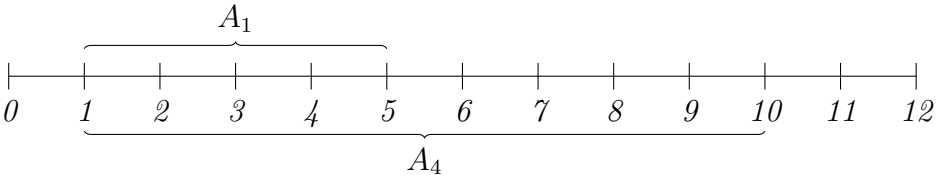
- $\phi(A_1 \cap A_2) + \phi(A_1 \cup A_2) = 3 + 8 = 11 = \phi(A_1) + \phi(A_2)$,



- $5 = \phi(A_4 \backslash A_1) = \phi(A_4) - \phi(A_1) = 10 - 5$ *und*



- $5 = \phi(A_1) \leq \phi(A_4) = 10$, *da* $A_1 \subset A_4$.



Definition 2.10. *Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum. Jedes $B \in \mathcal{A}$ mit $\mu(B) = 0$ heißt (μ-)Nullmenge.*

Bisher haben wir implizit nur diskrete Maße betrachtet. Eine Familie von Standardbeispielen für stetige Maße ist die Familie der Lebesgue-Maße. Um diese einführen zu können, benötigen wir zunächst einige Definitionen.

Definition 2.11. *Es sei $\Omega \neq \emptyset$ und $M \subset \mathfrak{P}(\Omega)$. Dann heißt*

$$\sigma(M) := \bigcap_{A \supset M, A \ \sigma\text{-Algebra}} A$$

die von M erzeugte σ-Algebra.

Die obige Definition liefert uns die kleinste σ-Algebra, die M enthält.

Definition 2.12. *Die Borel-σ-Algebra $\mathcal{B}(\mathbb{R}^n)$ auf $\mathbb{R}^n$ ist die von der Menge der sogenannten LORA-Intervallen (links offen und rechts abgeschlossen) in $\mathbb{R}^n$ erzeugte σ-Algebra.*

Definition 2.13. *Das Lebesgue-Maß λ auf der Borel-σ-Algebra $\mathcal{B}(\mathbb{R}^n)$ ist das eindeutige Maß mit der Eigenschaft, dass es n-dimensionalen Hyperrechtecken ihr n-dimensionales Volumen zuordnet*

$$\lambda([a_1, b_1] \times \cdots \times [a_n, b_n]) = (b_1 - a_1) \cdots (b_n - a_n).$$

Beispiel 2.14. *Das 3-dimensionale Lebesgue-Maß ist damit das, was wir als Volumen(-Maß) verstehen. Analog entspricht das 2-dimensionale Lebesgue-Maß dem Flächeninhalt eines ebenen Rechtecks.*

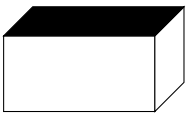


Abbildung 2.1: Beispielhafte Visualisierung einer Nullmenge

Es ist also wichtig, das zu betrachtende Maß genau zu spezifizieren, wenn man es berechnen möchte, da das schwarz markierte Rechteck ein positives 2-dimensionales Lebesgue-Maß besitzt, aber das 3-dimensionale Lebesgue-Maß dieser Fläche 0 *ist.*

Beispiel 2.15. *Ein typisches Beispiel für eine Nullmenge ist eine Fläche, wie das obige schwarz markierte Rechteck, wenn wir ein Volumenmaß betrachten. Flächen haben kein Volumen und sind daher im dreidimensionalen Lebesgue-Maß Nullmengen.*

Definition 2.16. *Wir sagen, eine Aussage gilt (μ-)fast überall, wenn sie für alle Mengen bis auf (μ-)Nullmengen gilt.*
Ebenso sagen wir, (μ-)fast alle Mengen haben eine Eigenschaft, wenn alle Mengen bis auf (μ-)Nullmengen diese Eigenschaft besitzen.

Beispiel 2.17. *Wenn wir Beispiel 2.15 erneut aufgreifen, so haben fast alle Mengen ein positives Volumen.*

Daher kommt auch die allgemeine Aussage, wenn wir „bis auf endlich viele “ durch „fast alle“ ersetzen, da man endlich viele Objekte als Nullmenge von einer Menge mit unendlich vielen Elementen betrachtet.

Satz 2.18. *Sei A eine μ-Nullmenge. Dann gibt es für jedes $\delta \geq 0$ eine offene Überdeckung $\mathcal{U}$ dieser Menge, so dass*

$$\mu(\mathcal{U}) \leq \delta$$

gilt.

Eine offene Überdeckung ist eine Überdeckung durch offene Mengen. Diese Mengen müssen nicht disjunkt sein und müssen in der Vereinigung mindestens A enthalten.

Beweis. Ein Beweis findet sich unter anderem in [AEE08]. □

2.2 Stochastische Grundlagen

In diesem Unterkapitel geht es um die notwendigen stochastischen Grundlagen, die wir im Verlauf des Kapitels zur Martin-Löf-Zufälligkeit benötigen werden. Die vorangegangene Maßtheorie wird zwar ebenfalls benötigt, um der Stochastik eine solide Grundlage zu geben, jedoch ist sie hauptsächlich dem Gebiet der Analysis anzuordnen. Dennoch haben wir in Kapitel 2.1 bereits gesehen, was ein Wahrscheinlichkeitsmaß und ein Wahrscheinlichkeitsraum sind.

Definition 2.19. *Sei P ein Wahrscheinlichkeitsmaß. Analog zu Definition 2.16 gilt eine Aussage (P-)fast sicher, wenn sie für alle Mengen bis auf Nullmengen gilt.*

Dieser Begriff hat sich in der Stochastik eingebürgert, wohingegen man in anderen Bereichen der Mathematik eher „fast alle“ oder „fast überall“ benutzt. Da es sich bei beiden Kontexten um Maßfunktionen handelt und nur das Wahrscheinlichkeitsmaß eine größere Einschränkung hat, ist „fast sicher “ ein Speziallfall von „fast alle“.

Satz 2.20 (Gesetz der Großen Zahlen)**.** *Seien $X_1, X_2, \ldots$ unabhängig gleichverteilte Zufallsvariablen mit Erwartungswert $E(X_1) = E(X_2) = \ldots = \mu$, dann gilt für $s_n = \sum_{i=1}^{n} X_i$*

$$\lim_{n\to\infty} \frac{s_n}{n} = \mu.$$

Satz 2.21 (Gesetz des iterativen Logarithmus)**.** *Sei $X_1, X_2, \ldots$ eine Folge unabhängiger, identisch verteilter Zufallsvariablen mit Erwartungswert 0 und Varianz 1, dann gelten*

$$\overline{\lim}_{N\to\infty} \frac{\sum_{k=1}^{N} X_k}{\sqrt{2N \log\log N}} = 1$$

fast sicher und

$$\underline{\lim}_{N\to\infty} \frac{\sum_{k=1}^{N} X_k}{\sqrt{2N \log\log N}} = -1$$

fast sicher.
Hierbei bezeichnet $\overline{\lim}$ den Limes superior und $\underline{\lim}$ den Limes inferior.

Außerdem wird, im Gegensatz zu vielen anderen Bereichen der theoretischen Informatik oder der Mathematik, in der Stochastik die charakteristische Funktion lieber Indikatorfunktion genannt.

Definition 2.22. *Sei $A \subset \Sigma^*$ eine Menge. Die Indikatorfunktion $\mathcal{I}_A$ zu A ist definiert als*

$$\mathcal{I}_A(x) = \begin{cases} 1 & , x \in A \\ 0 & , x \notin A \end{cases}.$$

Bemerkung 2.23. *Die Namensgebung ist historisch gewachsen, dadurch, dass der Name „charakteristische Funktion“ bereits belegt war. Die charakteristische Funktion einer Zufallsvariable X ist dabei definiert durch $\varphi_X(t) = E(\exp(itX))$ und wird unter anderem für die Momenterzeugung benutzt.*

2.3 Statistische Grundlagen

In der Statistik geht es darum, ausgehend von den Ergebnissen eines Zufallsexperiments, eine Aussage über die zugehörige Wahrscheinlichkeit, dass ein bestimmtes Ereignis eintrifft, zu treffen bzw. formaler gesagt, die Wahrscheinlichkeitsverteilung oder das Wahrscheinlichkeitsmaß einer (oder mehrerer) Zufallsvariablen zu bestimmen.
Ein intuitives Beispiel dazu ist das Folgende.

Beispiel 2.24. *Wir haben eine (nichtfaire) Münze* 30*-Mal geworfen. Nur* 7 *mal ist Kopf gefallen. Wir vermuten also eine Wahrscheinlichkeit von $\frac{7}{30}$, dass die Münze Kopf anzeigt und $\frac{23}{30}$, dass die Münze Zahl anzeigt.*

Definition 2.25. *Ein Stichprobenraum $(\Omega, \mathcal{A})$ ist ein messbarer Raum, der alle möglichen Datenwerte enthält.*

Auf einem Stichprobenraum $(\Omega, \mathcal{A})$ betrachtet man eine Familie $\mathcal{P}$ von Wahrscheinlichkeitsmaßen. Für die infrage kommenden Wahrscheinlichkeitsverteilungen der Daten wird dann das „beste “ Wahrscheinlichkeitsmaß gesucht.

Tests

Es sei $\mathcal{P}$ eine Familie von Wahrscheinlichkeitsmaßen auf einem Stichprobenraum $(\Omega, \mathcal{A})$. Oft soll anhand der Daten entschieden werden, ob die tatsächliche Verteilung P in einer vorgegebenen Teilfamilie P_0 von $\mathcal{P}$ liegt.
In der Fachsprache sagt man auch, man prüft die Hypothese $\mathrm{P} \in P_0$.

Beispiel 2.26. *Ein Beispiel für einen solchen Hypothesentest ist der χ^2-Test, welcher zu Hilfe genommen werden kann, wenn man beispielsweise eine Münze wirft und prüfen möchte, ob die Münze fair ist.*
Eine n-fache Wiederholung entspricht dann einer Stichprobe, mit der der χ^2-Test arbeiten soll.
Hält die Münze dem χ^2-Test stand, so wird die Hypothese nicht verworfen.

Definition 2.27. *Eine (messbare) Funktion $\phi : \Omega \to [0,1]$ heißt (randomisierte) Testfunktion zum Signifikanzniveau $\alpha \in [0,1]$, wenn gilt:*

$$E_P\phi(X) \leq \alpha,$$

für alle $P \in \mathcal{P}$. Die Abbildung $P \to E_P\phi(X)$ heißt Gütefunktion und beschreibt den Erwartungswert von $\phi(X)$ unter der Annahme, dass die zu grundeliegende Wahrscheinlichkeitsverteilung P ist.

Bei Vorliegen der Beobachtung x wird die Hypothese H mit Wahrscheinlichkeit $\phi(x)$ verworfen, also wird bei einem Test zum Signifikanzniveau α die Wahrscheinlichkeit für eine irrtümliche Ablehnung der Hypothese nicht größer als α. Bei Tests geht es also darum, eine vorgegebene Hypothese anhand der Stichprobendaten entweder zu verwerfen oder nicht zu verwerfen. Meist werden nichtrandomisierte Testfunktionen benutzt. Konkreter heißt dies, dass ϕ nur die Werte 0 und 1 annimmt.

Definition 2.28. *Die Menge $\{x \in \Omega \mid \phi(x) = 1\}$ heißt Ablehnungsbereich oder kritischer Bereich eines solchen Tests.*

2.4 Martin-Löf-Zufälligkeit für endliche Wörter

In der Praxis sind statistische Tests eine effektive Vorschrift, so dass wir für jedes gewünschte Signifikanzlevel berechnen können, für welche Wörter bzw. Ereignisse die Hypothese verworfen werden sollte.
Wir möchten, dass diese statistischen Tests nicht nur existieren, sondern auch effizient berechenbar sind.
Insgesamt wollen wir Tests konstruieren, die die Hypothese „Wort x hat folgende Form“ ablehnen und mitteilen, dass x zufällig ist und wir die Hypothese verwerfen, weil wir $x \in V_m$ dem kritischen Bereich vorfinden.

Wie im einführenden Beispiel in [Gau18] und dieser Arbeit können wir auch hier endliche 01-Folgen der Länge n zunächst mit einem n-fachen Münzwurf erzeugen.
Wenn wir nun mit dem Erwartungswert einer Binomialverteilung, also die Anzahl der Versuche (hier: n) multipliziert mit der Wahrscheinlichkeit, dass eine Eins geworfen wird (hier: $0,5$), argumentieren, so würden wir bei n Würfen in etwa $n \cdot \frac{1}{2}$ Fällen Einsen erwarten, in etwa $n \cdot \frac{1}{4}$ Fällen eine Teilfolge bestehend aus 00, in etwa $n \cdot \frac{1}{8}$ Fällen eine Teilfolge bestehend aus 111 usw. erwarten.
Wir wollen nun also Tests betrachten, die dabei die Hypothese (nicht zufällig)

verwerfen, wenn das Verhältnis von Nullen uns Einsen zu sehr von $\frac{1}{2}$ abweicht. Anders ausgedrückt, wenn

$$\left| n - 2\sum_{i=1}^{n} x_i \right| > f(m,n)$$

gilt. Die Variable m bezeichnet dabei das Signifikanzniveau und das n die Länge des Wortes $x = x_1 \cdots x_n$. Der Test f soll hierbei nur dadurch bestimmt sein, dass bei Wörtern der Länge n weniger gleich 2^{n-m} diese Ungleichung gilt. Dieses liegt in den nichtkomprimierbaren Wörtern nach Kolmogorov begründet. Im weiteren Verlauf werden wir sehen, dass die genaue Definition dieses Tests nicht weiter entscheidend ist.

Wir können also beliebige Theoreme der Wahrscheinlichkeitstheorie über zufällige, endliche Folgen direkt für unkomprimierbare Wörter über einem Alphabet gleicher Mächtigkeit übernehmen. Daher können wir uns an der Statistik bedienen und eine Teststatistik zugrundelegen, welche verwirft, wenn die relative Häufigkeit der Einsen zu sehr von $\frac{1}{2}$ abweicht. Diese Abweichung, oder auch dieses Signifikanzniveau, bezeichnen wir im Folgenden mit α. Welche Teststatistik wir genau zugrunde legen ist aufgrund der Äquivalenz, wie wir im Verlauf dieser Arbeit sehen werden, nicht weiter wichtig.
Wählen wir nun $\alpha = 2^{-m}$ für $m = 1, 2, ...$, so können wir unseren Ablehnungsbereich $U \subseteq \mathbb{N} \times \Omega$ beschreiben. Hierbei ist Ω die Menge der binären Zeichenketten.
Wir definieren außerdem $U_m = \{x \mid (m,x) \in U\}$. Dann gilt $U_{m+1} \subseteq U_m$, für $m = 1, 2, ...$
Damit U_m nun ein Ablehnungsbereich zum Signifikanzniveau $\alpha = 2^{-m}$ sein kann, muss die Anzahl der Wörter in U_m der Länge n kleiner gleich 2^{n-m} sein.
Durch die These von Church wissen wir, dass die Menge U insbesondere rekursiv aufzählbar ist. Daher wollen wir im Folgenden beschreiben, was eine Teststatistik ist, die auf eine rekursiv aufzählbare Menge angewendet wird.

Lemma 2.29. *Für jede Teststatistik V gibt es eine Teststatistik U, so dass*

$$V_{m+c} \subseteq U_m,$$

$m = 1, 2, ...$, wobei c eine Konstante ist, welche von U und V abhängt.

Bemerkung 2.30. *Die Teststatistiken V und U werden in der Literatur auch Tests genannt.*

Bemerkung 2.31. *Das Komplement eines kritischen Bereichs zum Signifikanzlevel*

α ist ggf. aus der Schule oder einer Stochastikvorlesung besser bekannt als $(1-\alpha)$-Konfidenzintervall.

Bemerkung 2.32. *In statistischen Tests kann die Mitgliedschaft von (m, x) in V gewöhnlicherweise in Polynomialzeit (in $\mathcal{O}(|x| + |m|)$) getestet werden.*

Lemma 2.33. *Es gibt eine rekursiv aufzählbare Menge $T \subseteq \mathbb{N} \times \mathbb{N} \times \Omega$, so dass U eine Teststatistik ist, genau dann wenn*

$$U = \{(m, x) \mid (i, m, x) \in T\},$$

für gewisse $i = 1, 2, ...$

Für den Beweis zeigen wir zunächst, dass die Menge der Teststatistiken effektiv aufzählbar ist und beweisen wie folgt beide Lemmata zusammen. Mit effektiv aufzählbar meinen wir, dass wir einen Algorithmus angeben können, der die Menge aufzählt und uns nicht auf die reine Existenz eines solchen Algorithmus verlassen wollen.

Beweis. Es ist bekannt, dass die Menge aller rekursiv aufzählbaren Teilmengen von $\mathbb{N} \times \Omega$ effektiv aufzählbar ist. Wir nutzen diesen Fakt aus, indem wir eine partielle rekursive Funktion f der Form $f\colon \mathbb{N} \times \mathbb{N} \to \mathbb{N} \times \Omega$, mit der Eigenschaft, dass wenn sie für (i, j) definiert ist auch $(i, 1), (i, 2), (i, \dots), (i, j-1)$ zum Definitionsbereich zählen. Außerdem ist eine Teilmenge von $\mathbb{N} \times \Omega$ rekursiv aufzählbar, genau dann, wenn sie die Form $\{f(i, j) \mid j = 1, 2, ...\}$ hat, für ein $i = 1, 2, ...$ Diese Mengen passen wir nun an, so dass sie den Bedingungen von statistischen Tests entsprechen. Wir erinnern uns dabei daran, dass eine rekursiv aufzählbare Menge $U \subseteq \mathbb{N} \times \Omega$, die zu einem Test gehört, zuallererst die Bedingung $U_m \supseteq U_{m+1}$, für $m = 1, 2, ...$ und als zweites die Bedingung erfüllt, dass die Anzahl der Elemente der Länge n in U_m weniger gleich 2^{n-m} sind.
Wir wählen nun ein festes aber beliebiges $i = 1, 2, ...$ Ist $f(i, j)$ für alle j undefiniert, so ist die korrespondierende rekursive aufzählbare Menge die leere Menge. Wenn dies nicht der Fall ist, so sei $f(i, 1) = (m_1, x_1)$. Falls die Menge aller (m, x_1), mit $m \leq m_1$, den Bedingungen eines Tests genügen $(m_1 \leq |x_1|)$, fügen wir (i, m, x_1) zur Menge T hinzu, für alle $m \leq m_1$. Falls nicht, so bleibt die Stelle i in T leer und die Modifikation ist für dieses i zuende. Im ersten Fall berechnen wir nun $f(i, 2) = (m_2, x_2)$, falls es definiert ist und analog zu $j = 1$ fügen wir (i, m, x_2) zu T hinzu, für alle $m \leq m_2$, falls die gewünschten Testeigenschaften erfüllt sind. Ansonsten ist die Modifikation für das i fertig. Diese Konstruktion wird so weitergeführt. Der Bereich von T mit

dem Eintrag i entspricht dann der Menge $\{f(i,j) \mid j = 1,2,...\}$ und durch die Konstruktion erfüllt er auch die Testeigenschaften.

Den universellen Test U erhalten wir als Bild von T unter der Abbildung

$$(i, m+i, x) \mapsto (m, x).$$

Nehmen wir nun einen beliebigen Test V, so gilt für ein i

$$V = \{x \mid (i, m+i, x) \in T\}.$$

Damit folgt

$$V_{m+i} = \{x \mid (i, m+i, x) \in T\} \subseteq \{x \mid (m, x) \in U\} = U_m,$$

für alle $m \in \mathbb{N}$.

□

Wir verwenden nun statt des Signifikanzniveaus das kritische Level bezüglich eines Tests U.

Definition 2.34. *Das kritische Level bezüglich eines Tests U ist definiert als*

$$m_U(x) = \max_{x \in U_m} m.$$

Da $m_U(x)$ für jedes x definiert ist, definieren wir zusätzlich mit U_0 die Menge aller Binärwörter, die $0 \leq m_U(x) \leq |x|$ genügen.

Dieses $m_U(x)$ spielt beim Begriff der Martin-Löf-Zufälligkeit in etwa die Rolle, welche die Kolmogorov-Komplexität im vorherigen Kapitel spielt und besitzt daher (bis auf Vorzeichen) ähnliche Eigenschaften, wie wir im Laufe dieses Unterkapitels noch sehen werden. Der wesentliche Unterschied hierbei liegt darin begründet, dass sich der Zufallsbegriff nach Kolmogorov nur sehr schwierig auf unendlich lange Wörter übernehmen lässt, da Rechnen im Unendlichen immer eigene Schwierigkeiten mit sich bringt. Wir hatten an der Stelle von Zufälligkeit via c-Komprimierbarkeit gesprochen, wobei $c \in \mathbb{N}$ eine Konstante ist. Jedoch gilt

$$\infty - c = \infty,$$

weshalb dieser Begriff sich nur schwer übertragen lässt.
Bei dem in diesem Kapitel verwendeten Begriff jedoch handelt es sich in der

Grundlage um Teststatistiken, für welche es eine ausgiebige Theorie auch für unendliche Tests gibt, sodass wir uns dabei problemlos der Resultate daraus bedienen können. Daher ist es an dieser Stelle sinnvoll, eine zweite Art der Betrachtung der Kolmogorov-Komplexität zu untersuchen.

Bemerkung 2.35. *Es gibt für jeden universellen Test U einen Test V und eine Konstante $c \in \mathbb{N}$, sodass folgende Ungleichung für alle Wörter x erfüllt ist:*

$$m_V(x) \leq m_U(x) + c$$

Eine Ungleichung der Form hatten wir bereits in der Definition der Kolmogorov-Komplexität für universelle Turingmaschinen vorgestellt, vergleiche 1.4, welche aus dem Invarianz-Theorem resultierte, siehe 1.12. Dabei war allerdings die Ungleichung genau andersrum definiert, so dass das universelle Maß kleiner gleich dem beliebigen Maß plus einer Konstante ist. Dies haben wir vorher mit „(bis auf Vorzeichen) ähnliche Eigenschaften“ gemeint. Daher wollen wir nun das kritische Level bezüglich eines universellen Tests möglichst analog zur Kolmogorov-Komplexität im voherigen Kapitel definieren.

Definition 2.36. *Das kritische Level bezüglich eines universellen Tests, der aus dem Kontext eindeutig ist, bezeichnen wir mit*

$$m(x) := m_U(x).$$

Eine solche Konstruktion haben wir bereits im vorherigen Kapitel gesehen, in Definition 1.6, als wir bei der Kolmogorov-Komplexität die entsprechende universelle Turingmaschine in der Notation $K(x)$ wegließen. Wir wollen ähnlich zu [Gau18] nun die Wohldefiniertheit zeigen.

Definition 2.37. *Zwei kritische Level universelle Tests $m_A(x)$ und $m_B(x)$ heißen äquivalent, in Zeichen*

$$m_A(x) \equiv m_B(x),$$

wenn es eine Konstante $c \in \mathbb{N}$ gibt, so dass

$$|m_A(x) - m_B(x)| \leq c$$

gilt.

Bemerkung 2.38. *Die Relation $\equiv$ aus Definition 2.37 ist eine Äquivalenzrelation.*

Beweis. Wir müssen zeigen, dass $\equiv$ reflexiv, symmetrisch und transitiv ist. Seien A, B und C universelle Tests und seien c, c_1 und $c_2 \in \mathbb{N}$ Konstanten.

- symmetrisch: Es gilt $|m_A(x) - m_A(x)| = 0 \leq c$, für alle natürlichen Zahlen c.

- reflexiv: Angenommen, es gilt $|m_A(x) - m_B(x)| \leq c$. Dann gilt

$$|m_A(x) - m_B(x)| = |m_B(x) - m_A(x)|,$$

  da die Differenz von zwei natürlichen Zahlen innerhalb eines Betrags kommutativ ist. Demnach ist auch

$$|m_B(x) - m_A(x)| \leq c.$$

- transitiv: Angenommen, es gelten sowohl $|m_A(x) - m_B(x)| \leq c_1$ als auch $|m_B(x) - m_C(x)| \leq c_2$. Dann gilt

$$\begin{aligned} |m_A(x) - m_C(x)| &= |m_A(x) - m_B(x) + m_B(x) - m_C(x)| \\ &= |m_A(x) - m_B(x) - (m_C(x) - m_B(x))| \\ &\leq |m_A(x) - m_B(x)| + |m_C(x) - m_B(x)|, \end{aligned}$$

  wobei die letzte Zeile mittels Dreiecksungleichung folgt.
  Wir hatten bereits bei der Reflexivität gesehen, dass wir die Differenz im Betrag ganz rechts umdrehen dürfen. Daher folgt

$$\begin{aligned} &|m_A(x) - m_B(x)| + |m_C(x) - m_B(x)| \\ &= |m_A(x) - m_B(x)| + |m_B(x) - m_C(x)| \\ &\leq c_1 + c_2. \end{aligned}$$

  Wenn wir unsere neue Konstante auf

$$c := c_1 + c_2$$

  setzen, dann erhalten wir das gewünschte:

$$|m_A(x) - m_C(x)| \leq c$$

□

Definition 2.39.
- *Durch die Relation $\equiv$ aus Definition 2.37 werden die Äquivalenzklassen*
$$[m_A] := \{m_B \mid m_B \equiv m_A\}$$
induziert.

- *Die Ordnung $\leq$ auf den Äquivalenzklassen definieren wir wie folgt.*
Für Äquivalenzklassen $[m_A]$ und $[m_B]$ gilt $[m_A] \leq [m_B]$, wenn für den kanonischen Vertreter
$$m_A(x) \leq m_B(x) + c$$
gilt.

Bemerkung 2.40. *Die Ordnung aus Definition 2.39 ist eine partielle Ordnung auf den Äquivalenzklassen. Insbesondere gibt es ein minimales Element m_{A_0}, so dass für alle m_B die Ungleichung*
$$[m_{A_0}] \leq [m_B]$$
gilt.

Beweis. Eine Ordnung heißt partiell, wenn sie reflexiv, antisymmetrisch und transitiv ist. Reflexivität und Transitivität lassen sich ähnlich zu Bemerkung 2.38 zeigen.

- reflexiv: $m_A(x) \leq m_A(x) + c$, gilt für alle Konstanten natürlichen Zahlen c.

- antisymmetrisch: Angenommen, es gelten sowohl $m_A(x) \leq m_B(x) + c_1$ als auch $m_B(x) \leq m_A(x) + c_2$. Subtrahiert man in der ersten Ungleichung beide Seiten mit $m_B(x)$ und in der zweiten Ungleichung beide Seiten mit $m_A(x)$, so erhält man
$$m_A(x) - m_B(x) \leq c_1 \text{ und } m_B(x) - m_A(x) \leq c_2.$$
Insgesamt erhält man damit, da sowohl $m_B(x)$ als auch $m_A(x)$ positiv sind
$$|m_A(x) - m_B(x)| \leq \max\{c_1, c_2\} =: c.$$
Damit liegen $m_A(x)$ und $m_B(x)$ in der selben Äquivalenzklasse und sind dementsprechend äquivalent.

- transitiv: Angenommen, es gilt $m_A(x) \leq m_B(x) + c_1$ und außerdem noch $m_B(x) \leq m_C(x) + c_2$. Dann lässt sich $m_B(x)$ in der ersten Ungleichung durch die rechte Seite der zweiten Ungleichung ersetzen, da $\geq$ dann immer noch gilt und man erhält
$$m_A(x) \leq (m_C(x) + c_2) + c_1.$$

Setzt man die neue Konstante auf $c := c_1 + c_2$, so erhält man das gewünschte

$$m_A(x) \leq m_C(x) + c.$$

Wir haben also eine partielle Ordnung. Da eine partielle Ordnung auf einer Menge, die nach unten beschränkt ist, ein minimales Element besitzt, können wir dieses m_{A_0} nennen und haben unsere Aussage gezeigt. Die Beschränkung nach unten kommt durch die natürliche Grenze der 0 zustande. □

Im folgenden Satz wollen wir nun einen Zusammenhang zwischen Kolmogorov-Komplexität und dem kritischen Level eines Tests herstellen.

Satz 2.41. *Es gibt eine Konstante $c \in \mathbb{N}$, sodass folgende Ungleichung für alle Binärwörter x gilt*

$$||x| - K(x \mid |x|) - m(x)| \leq c.$$

Beweis. Wir definieren zunächst

$$\begin{aligned} V &= \{(m, x) \mid K(x \mid |x|) < |x| - m\} \\ &= \{(m, x) \mid (\exists p)(|p| < |x| - m \text{ und } M(p, |x|) = x)\} \subseteq \mathbb{N} \times \Omega, \end{aligned}$$

wobei M eine universelle Turingmaschine ist. V ist ein Test und

$$m_V(x) = |x| - K(x \mid |x|) - 1,$$

also

$$|x| - K(x \mid |x|) - 1 \leq m_V(x) + c.$$

Um die Ungleichung in der anderen Richtung zu zeigen, sei U ein universeller Test, der das kritische Level definiert und wir wählen eine rekursiv aufzählbare Funktion f der Form $f\colon \mathbb{N} \to \mathbb{N} \times \Omega$, die U ohne Wiederholungen aufzählt.
Mittels f konstruieren wir nun folgenden Algorithmus M von $\Omega \times \mathbb{N}$ nach Ω:

- Falls $f(1) = (m_1, x_1)$, sei $M(\underbrace{00\ldots00}_{|x_1|-m_1}, |x_1|) = x_1$.
- Falls $f(2) = (m_2, x_2)$ und $(m_1, |x_1|) = (m_2, |x_2|)$, dann sei

$$M(\underbrace{00\ldots0}_{|x_2|-m_2-1}\ 1, |x_2|) = x_2.$$

- Falls $f(2) = (m_2, x_2)$ und $(m_1, |x_1|) \neq (m_2, |x_2|)$, dann sei

$$M(\underbrace{00\ldots00}_{|x_2|-m_2}, |x_2|) = x_2.$$

Da U ein Test ist und wir die Äquivalenzklassen oben als wohldefiniert bewiesen haben, können wir obige Konstruktionen ohne Mehrdeutigkeiten anwenden. Damit gilt nun

$$K_M(x \mid |x|) = |x| - m(x),$$

also insgesamt

$$K(x \mid |x|) \leq |x| - m(x) + c.$$

□

Wir haben also insgesamt eine zu Satz 1.11 ähnliche Aussage gesehen.

Definition 2.42. *Ein Binärwort $w_1w_2\cdots w_n$ heißt zufällig im Martin-Löf-Sinn, wenn*

$$m(w_1w_2\cdots w_n) < c,$$

für eine Konstante $c \in \mathbb{N}$ gilt.

Damit können wir nun die Äquivalenz von Kolmogorov-Zufälligkeit und Martin-Löf-Zufälligkeit im endlichen Fall zeigen.

Satz 2.43. *Kolmogorov-Zufälligkeit und Martin-Löf-Zufälligkeit sind für endliche Wörter äquivalent.*

Beweis. Wir wissen aus [Gau18], dass

$$K(x \mid |x|) \leq K(x) \leq K(x \mid |x|) + 2K(|x| - K(x \mid |x|))$$

gilt. Nehmen wir an, x sei unkomprimierbar, dann sind somit $K(x) = K(x \mid |x|)$. Wir erhalten mithilfe von Satz 2.41 nun

$$||x| - K(x) - m(x)| \leq c.$$

Da wir x unkomprimier annehmen, gilt $|x| - K(x) = -c_1$, für eine Konstante $c_1 \in \mathbb{N}$. Insgesamt liefert uns dies

$$|-c_1 - m(x)| \leq c.$$

Sowohl c_1 als auch $m(x)$ sind nicht negative Zahlen, also ergibt sich

$$\begin{aligned} |-(c_1 + m(x))| &\leq c \\ c_1 + m(x) &\leq c \\ m(x) &\leq c - c_1 < c \end{aligned}$$

und damit haben wir genau Martin-Löf-Zufälligkeit, wenn Kolmogorov-Zufälligkeit vorliegt.

Andersrum, wenn $m(x) < c_1$ gilt, liefert uns Satz 2.41

$$\begin{aligned} ||x| - c_1 - K(x \mid |x|)| &\leq c \\ ||x| - (c_1 + K(x \mid |x|)| &\leq c \end{aligned}$$

Auf der linken Seite können wir die Dreiecksungleichung anwenden und erhalten

$$||x| - (c_1 + K(x \mid |x|)| \leq ||x|| + |c_1 + K(x \mid |x|)|.$$

Auf der rechten Seite lässt sich $K(x \mid |x|)$ von oben abschätzen durch $K(x)$ und damit erhalten wir

$$K(x) \leq |x| + c_1,$$

also c_1-Unkomprimierbarkeit und damit Zufälligkeit. □

2.5 Martin-Löf-Zufälligkeit für unendliche Wörter

Im Fall endlicher Binärwörter haben wir gesehen, dass die Einführung eines universellen Tests nichts weiter als eine Umformulierung des Komplexitätsmaßes der Kolmogorov-Komplexität ist. Wir hatten angesprochen, dass diese Umformulierung sinnvoll ist, da sie sich sehr viel einfacher auch auf Wörter unendlicher Länge erweitern lässt. Wir werden nun im folgenden Unterkapitel auf analogem Weg einen universellen sequentiellen Test für unendlich lange Binärwörter definieren, um dieses Komplexitätsmaß auf unendliche Wörter zu verallgemeinern. Mit unendlichen Wörtern meinen wir hier unendlich lange Folgen über einem endlichen Alphabet.

Wir stellen uns daher nun vor, wir haben ein Zufallsexperiment, wie zum Beispiel einen (unendlich oft wiederholten) fairen Münzwurf, der uns möglicherweise unendlich lange Binärwörter $w_1 w_2 \cdots w_n \cdots$ erzeugt. Dabei soll das Gesetz der

Großen Zahlen (siehe 2.2), also

$$\lim_{n\to\infty} \frac{s_n}{n} = \frac{1}{2}$$

oder sogar das Gesetz des iterierten Logarithmus (siehe 2.21)

$$\overline{\lim}_{n\to\infty} \frac{2s_n - n}{\sqrt{2n \log\log n}} = \pm 1$$

gelten.
Wir wollen die maßtheoretische Wahrscheinlichkeitstheorie benutzen. Dies ist darin motiviert, zu zeigen, dass alle Mengen, für die dieses Gesetz nicht gilt, Maß null haben, also Nullmengen sind. Nach Satz 2.18 bedeutet dies, dass es für jedes $\delta \geq 0$ eine offene Überdeckung $\mathcal{U}$ dieser Menge gibt, so dass

$$\mu(\mathcal{U}) \leq \delta$$

gilt. Hierbei ist μ das Wahrscheinlichkeitsmaß auf $\{0, 1\}$ mit $\mu(1) = \frac{1}{2} = \mu(0)$, bei dem alle Ereignisse (0 und 1) unabhängig und gleichverteilt mit Wahrscheinlichkeit $\frac{1}{2}$ sind.

Definition 2.44. *$J(w_1 w_2 \cdots w_n)$ bezeichnet die Menge aller unendlichen Binärwörter, die mit dem Teilwort $w_1 w_2 \cdots w_n$ beginnen.*

Wir wollen im Folgenden wieder einen kritischen Bereich, nämlich in diesem Fall eine Nullmenge konstruieren, so dass alle Wörter, die dem Test nicht standhalten Maß 0 haben. Diese Menge heißt im Folgenden $\mathcal{U}$.

Wir betrachten nun die Menge

$$U = \{x \mid J(x) \subseteq \mathcal{U}\},$$

also die Menge aller Wortanfänge aus $\mathcal{U}$.
Da diese Wortanfänge stets endlich sind, können wir hierauf unsere Erkenntnisse aus Kapitel 2.4 anwenden.

Bemerkung 2.45. *Es gilt $\mathcal{U} = \bigcup_{x\in U} J(x)$ genau dann, wenn $\mathcal{U}$ offen ist.*

Die Menge U hat die Eigenschaft, dass sie alle möglichen Erweiterungen von ihren Elementen beinhaltet.
Mit anderen Worten: U kann als kritischer Bereich eines sequentiellen Tests zum Signifikanzlevel $\alpha \geq 0$ betrachtet werden.
Nun können wir den Argumentationen des vorherigen Kapitels (Kapitel 2.4) folgen.

Wir nehmen nun wieder an, die Familie der kritischen Bereiche, bzw. hier der offenen Überdeckungen $U \subseteq \mathbb{N} \times \Omega$, sei rekursiv aufzählbar. U muss folgende Einschränkungen erfüllen:

- Sei $(m, x) \in U$, so auch $(n, y) \in U$, für $n \leq m$ und $y \geq x$.
- Die Anzahl der Wörter der Länge n, welche in $U_m := \{x \mid (m, x) \in U\}$ enthalten sind, ist kleiner gleich 2^{n-m}, für alle n und m.

Damit können wir die Hauptaussage dieses Kapitels beweisen, nämlich, dass die Menge aller sequentiellen Tests bzw. dazu korrespondierenden offenen Überdeckungen effektiv aufzählbar ist:

Satz 2.46. *Es gibt eine rekursiv aufzählbare Menge $T \subseteq \mathbb{N} \times \mathbb{N} \times \Omega$, so dass U ein sequentieller universeller Test ist, genau dann wenn*

$$U = \{(m, x) \mid (i, m, x) \in T\},$$

für bestimmte $i \in \mathbb{N}$

Beweis. Der Beweis unterscheidet sich nur minimal von dem Analogon im vorherigen Kapitel (siehe Lemma 2.33).
Wir wählen eine partiell rekursive Funktion f wie oben und fixieren ein beliebiges $i = 1, 2....$ Anschließend berechnen wir $f(i, 1) = (m_1, x_1)$, falls es definiert ist. Dann fügen wir, sofern mit dem Signifikanzlevel zulässig (also hierbei $m_1 \leq |x_1|$), (i, m, x) zur Menge T hinzu, für alle $m \leq m_1$ und $x \geq x_1$. Andernfalls bleibt T an i leer. Wenn wir zu diesem Zeitpunkt noch nicht fertig sind, berechnen wir $f(i, 2) = (m_2, x_2)$, falls definiert und fügen (i, m, x) zur Menge T hinzu, für alle $m \leq m_2$ und $x \geq x_2$. Die im obigen Beweis erwähnten Eigenschaften gelten dann auch für die Menge T hier. □

Satz 2.47. *Es existiert ein universeller sequentieller Test U, so dass für jeden sequentiellen Test V gilt:*

$$V_{m+c} \subseteq U_m, \text{ für } m= 1,2,...,$$

wobei c eine Konstante ist, die von U und V abhängig ist.

Auch diese Aussage haben wir bereits im endlichen Fall gesehen, vergleiche hierzu Lemma 2.29.
Da wir gleiche Voraussetzungen für U (für ein $n \in \mathbb{N}$) in diesem Abschnitt vorfinden, wie im vorherigen, übernehmen wir Definitionen wie $m_U(x)$ analog.

Beobachtung 2.48. *Auch hier kann, wie im endlichen Fall, U als Bild von T unter der Abbildung*

$$(i, m+i, x) \mapsto (m, x)$$

verstanden werden.

Das kritische Level $m_U(x) = \max_{x \in U_m} m$ bezüglich eines sequentiellen Tests U erfüllt

$$0 \leq m_U(x) \leq |x|$$

und

$$m_U(x) \leq m_U(y),$$

für alle $x \leq y$.

Daher können wir nun ein kritisches Level eines unendlichen Binärwortes definieren.

Definition 2.49. *Das kritsche Level eines unendlichen Binärwortes ist definiert als*

$$m_U(w_1 w_2 \cdots w_n \cdots) = \lim_{n \to \infty} m_U(w_1 w_2 \cdots w_n).$$

Fixieren wir nun einen universellen Test U, lassen wir wie zuvor zur Vereinfachung der Schreibweise das U im Index weg.

Beobachtung 2.50. *Es gilt*

$$0 \leq m_U(w_1 w_2 \cdots w_n \cdots) \leq \infty.$$

Definition 2.51. *Ein unendliches Binärwort $w_1 w_2 \cdots w_n \cdots$ heißt zufällig, wenn*

$$m(w_1 w_2 \cdots w_n \cdots) < \infty$$

gilt.

Diese Definition ist analog zum endlichen Fall mit dem Unterschied, dass die obere Schranke angepasst werden muss, da die Länge des Wortes hier nun unendlich lang werden kann.

Bemerkung 2.52. *Definition 2.51 ist unabhängig von der Wahl des universellen Tests bezüglich welcher das kritische Level definiert ist.*

Konstruktive Nullmengen

In der Literatur fällt häufig der Begriff der konstruktiven Nullmengen, wenn es um Martin-Löf-Zufälligkeit geht. Diesem Begriff wollen wir uns in diesem Unterkapitel zuwenden.
Hierzu bedarf es zunächst einiger Definitionen.

Definition 2.53.
- *Eine offene Menge unendlicher Binärwörter $\mathcal{U}$ heißt konstruktiv offen, falls $\{x \mid J(x) \subseteq \mathcal{U}\}$ rekursiv aufzählbar ist.*
- *$\mathcal{U}_1, \mathcal{U}_2, \ldots$ heißt konstruktive Folge konstruktiv offener Mengen, wenn $\{(m,x) \mid J(x) \subseteq \mathcal{U}_m\}$ rekursiv aufzählbar ist.*
- *$\mathcal{A}$ heißt konstruktive Nullmenge, falls $\mathcal{A} \subseteq \mathcal{U}_m$, für $m = 1, 2, ...,$ wobei $\mathcal{U}_1, \mathcal{U}_2, \ldots$ eine konstruktive Folge konstruktiv offener Mengen ist mit $\mu(\mathcal{U}_m) \to 0$, konstruktiv schnell wie $m \to \infty$. Damit ist gemeint, dass $\mu(\mathcal{U}_m) \leq 2^{-k}$, für alle $m \geq h(k)$, für eine rekursive Funktion h ist.*

Bemerkung 2.54. *Die Menge der nichtzufälligen Folgen erzeugen eine inklusionsmaximale konstruktive Nullmenge.*

Sei $\mathcal{B}$ eine beliebige konstruktive Nullmenge und $\mathcal{V}_1, \mathcal{V}_2, \ldots$ die assoziierten Überdeckungen. Ohne Beschränkung der Allgemeinheit können wir annehmen, dass $\mathcal{V}_1 \supseteq \mathcal{V}_2 \supseteq \ldots$ und $\mu(\mathcal{V}_m) \leq 2^{-m}$, so dass

$$V = \{(m,x) \mid J(x) \subseteq \mathcal{V}_m\}$$

ein sequentieller Test ist. Auch hier gilt für einen universellen sequentiellen Test U

$$V_{m+c} \subseteq U_m, \text{ für m= 1,2,...,}$$

wobei c eine Konstante abhängig von U und V ist. Dementsprechend gilt

$$\mathcal{B} \subseteq \bigcap_{m=1}^{\infty} \mathcal{V}_m = \bigcap_{m=1}^{\infty} \mathcal{V}_{m+1} \subseteq \bigcap_{m=1}^{\infty} \mathcal{U}_m = \mathcal{A},$$

für $\mathcal{U}_m = \bigcup_{x \in U_m} J(x)$, wie zuvor. Wie auch in [Gau18] fragen wir uns nun, ob nach diesem Begriff zufällige unendliche Binärwörter existieren.

Satz 2.55. *Fast alle unendlichen Binärwörter sind zufällig.*

Beweis. Sei $\mathcal{U}_m = \bigcup_{x \in U_m} J(x)$, für $m = 1, 2, ...$ Da U ein sequentieller Test ist, gelten $\mathcal{U}_1 \supseteq \mathcal{U}_2 \supseteq ...$ und $\mu(\mathcal{U}_m) \leq 2^{-m}$, für $m = 1, 2, ...$ Die Menge der nichtzufälligen

Wörter ist dann exakt die Nullmenge $\bigcap_{m=1}^{\infty} \mathcal{U}_m$, da U universell gewählt wurde und der Schnitt von Nullmengen wieder eine Nullmenge ergibt und nach Definition einer σ-Algebra auch wieder in der σ-Algebra enthalten ist. □

Wir haben in diesem Kapitel also gesehen, dass sich die Kolmogorov-Komplexität mithilfe einer anderen Betrachtungssichtweise, wie Martin-Löf sie präsentierte, auf unendlich lange Wörter ausweiten lässt.
Im Wesentlichen entsprachen die universellen sequentiellen Tests hierbei den universellen Turingmaschinen und die Nullmengen entsprechen Mengen nichtzufälliger, also komprimierbaren, Wörter im Kolmogorov-Sinn.

3 Zufallsbegriff nach Claus Schnorr

Wir wollen zunächst Martingale für unseren Kontext definieren. Wie auch im Kapitel 2 müssen wir hier wieder unseren Kontext in eine neue Sprache übersetzen. Wir zeigen also, wie sich Martin-Löf-Zufälligkeit und damit Kolmogorov-Zufälligkeit in die Sprache der Martingale übersetzen lässt, da wir für den endlichen Fall gesehen haben, dass die beiden Zufälligkeiten dort äquivalent sind. Darauf aufbauend definieren wir weitere Zufallsbegriffe und zeigen anschließend Aussagen über die verschiedenen Zufallsbegriffe mithilfe der arithmetischen Hierarchie. Dieses Kapitel richtet sich nach [Sch71].

3.1 Martingale

Unser bisheriger Stand ist der Folgende: Die intuitive Idee besteht darin, dass ein unendliches Binärwort x zufällig ist, wenn es allen konstruktiven stochastischen Tests standhält. Ein solcher stochastischer Test kann durch eine Funktion $f\colon \Sigma^* \to \mathbb{R}_{\geq 0}$ ausgedrückt werden. Dabei zeigt $f(x)$ an, wie anfällig x bezüglich des stochastischen Tests ist. Es scheint natürlich zu sein, dass $f(x)$ einen großen Wert annimmt, wenn x bezüglich eines Tests anfällig ist und niedrig, wenn nicht.

Um diese Beispiele oben zu verallgemeinern, nennen wir eine Funktion $f\colon \Sigma^* \to \mathbb{R}_{\geq 0}$ einen konstruktiven Test, wenn f folgende informelle Eigenschaften besitzt:

1. f kann konstruiert werden, d.h. f lässt sich durch einen Algorithmus berechnen.

2. Es gibt eine Regel, welche Nullmengen bezüglich f bestimmt, so dass diese Nullmenge die Teilmenge der unendlich langen Wörter ist, die diesem Test nicht standhalten.

Für den zweiten Unterpunkt gibt es zwei Möglichkeiten, die sich betrachten lassen um Nullmengen zu erzeugen, nämlich:

1. $\mathcal{N}_f := \{z \in \Sigma^\infty \mid \underline{\lim}_{n\to\infty} f(z(n)) = \infty\}$ und
2. $\mathcal{N}_f := \{z \in \Sigma^\infty \mid \overline{\lim}_{n\to\infty} f(z(n)) = \infty\}$.

Wir schauen uns nun zunächst anhand eines Beispiels an, was ein Martingal ist. Das typische Einstiegsbeispiel ist hierfür ein Münzwurf oder beim Roulette das Setzen auf schwarz oder rot, wobei man gerne die grüne Null vernachlässigt, so dass man Wahrscheinlichkeiten von je $\frac{1}{2}$ für beide Ereignisse hat.

Beispiel 3.1. *Nehmen wir an $f\colon \Sigma^* \to \mathbb{R}_{\geq 0}$ sei eine Funktion, die das Kapital eines Spielers angibt, welcher auf Binärwörtern spielt. Dabei steht $f(x)$ für das Kapital nach dem $|x|$-ten Versuch, wenn die Spielfolge das Startwort x hat.*
Bei einem fairen Spielsystem sollte der Zugewinn des Spielers folgende Gleichung erfüllen:

$$f(x) = \frac{1}{2}\left(f(x1) + f(x0)\right) \tag{3.1}$$

Die Schreibweise $x1$ steht dafür, dass an ein endliches Wort x das Zeichen 1 angehängt wird. Es ist zu erwarten, dass dabei $\overline{\lim}_{n\to\infty} f(z(n)) < \infty$ gilt, sofern z zufällig ist. Das heißt also auch, dass ein hoher Wert von $f(z(n))$ bedeutet, dass das Wort $z(n)$ anfällig für den Test f ist.
Man kann außerdem zeigen, dass für jede Funktion f, die die Gleichung 3.1 erfüllt, die Menge

$$\{z \in \Sigma^\infty \mid \overline{\lim}_{n\to\infty} f(z(n)) = \infty\}$$

eine Nullmenge ist.

Bemerkung 3.2. *Im Wesentlichen kann man sich Martingale in der allgemeinen Literatur als faire Glücksspiele vorstellen. Diese heißen in der Stochastik fair, weil man nicht mithilfe einer bestimmten Spielstrategie das Spiel so beeinflussen kann, dass man selbst gewinnt (Submartingal) oder die Bank beeinflussen könnte, dass ein Spieler verliert (Supermartingal).*

Dennoch sollte im Rahmen einer wissenschaftlichen Arbeit eine formale Definition aus der mathematischen Stochastik genannt werden, welche im nächsten Kapitel auf den aktuellen Kontext, mithilfe des obigen Beispiels, bzw. deren Gleichung, angepasst wird, so dass die ursprüngliche Bedeutung aber noch (zumindest ansatzweise) beibehalten wird.

Daher geben wir nun zur Vervollständigung die mathematisch-stochastische Definition nach [Grü14] hier ohne weitere Erwähnung kurz an:

Definition 3.3. *Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum $T \subset \mathbb{R}$ eine Menge von Zeitpunkten, $(\mathcal{F}_t)_{t\in T}$ eine Filtration, also eine aufsteigende (bzgl.*

Teilmengenbeziehung) Folge von Unter-σ-Algebren von $\mathcal{A}$ und eine hierzu adaptierte Familie von Zufallsvariablen $(X_t)_{t\in T}$ mit $E(|X_t|) < \infty$, für alle Zeitpunkte $t \in T$. Dann heißt $(X_t, \mathcal{F}_t)_{t\in T}$ Martingal, wenn gilt

$$X_s = E([X_t \mid \mathcal{F}_t]),$$

für alle $s \leq t$ und $s, t \in T$. Also wenn sich der Erwartungswert (bspw. des Gesamtgewinns) nach einem oder mehreren Spieldurchläufen nicht verändert, dadurch dass wir nach einigen Durchläufen mehr Informationen vorfinden.

Beispiel 3.4. *Wenn wir bei mehrfachem fairen Würfelwurf beobachten, dass lange keine 6 gefallen ist, so bringt uns das keine neuen Erkenntnisse, wenn wir vorhersagen wollen, welche Zahl als nächstes fällt.*

Im folgenden Kapitel wollen wir zunächst eine Spielstrategie als Martingal bezeichnen, die Bit für Bit des Wortes lesen soll und auf das Zeichen wetten/ das Zeichen richtig erraten soll. Die Martingalcharakterisierung dieser Spielstrategie zeigt dann, dass wir bei zufällig erzeugten Wörtern, keinen Gewinn durch immer mehr Informationen erlangen können und damit letztendlich Martingale im ursprünglichen Sinn vorfinden.

3.2 (1)-Zufälligkeit und Zusammenhang zur Martin-Löf-Zufälligkeit

Anknüpfend an Kapitel 2 wollen wir nun den dort definierten Zufallsbegriff durch die Sprache der Martingale ausdrücken.

Bemerkung 3.5. *Analog zur Menge Σ^*, die alle (endlichen) Wörter über dem Alphabet Σ enthält, wollen wir in diesem Kapitel mit Σ^∞ die Menge aller (unendlich langen) Zeichenketten/Wörter über dem Alphabet Σ bezeichnen.*

Definition 3.6. *Wir betrachten im Folgenden die Abbildung $\varphi : \Sigma^* \to \Sigma^\infty$, $\varphi(A) = A\Sigma^\infty$. Das heißt, φ bildet ein endliches Wort A auf die Menge der unendlichen Wörter mit endlichem Wortanfang A ab.*

Definition 3.7. *Das Maß μ, welches wir in diesem Kapitel betrachten wollen ist ein Maß auf $\{0,1\}^\infty$, so dass an jeder Stelle des (unendlichen) Wortes 0 und 1 gleichwahrscheinlich sind.*

Definition 3.8. *Eine totale Funktion $f\colon \Sigma^* \to \mathbb{R}_{\geq 0}$ heißt schwach berechenbar, wenn es eine rekursive Funktion $g : \mathbb{N} \times \Sigma^* \to \mathbb{Q}$ gibt, sodass*

$$\begin{aligned} g(i,x) &\leq g(i+1,x) \text{ für } i \in \mathbb{N} \text{ und } x \in \Sigma^* \\ \lim_{i\to\infty} g(i,x) &= f(x) \text{ für } x \in \Sigma^*. \end{aligned}$$

Eine totale Funktion $f\colon \Sigma^* \to \mathbb{R}_{\geq 0}$ heißt berechenbar, falls f und $-f$ schwach berechenbar sind.
Eine totale Funktion $f\colon \Sigma^* \to \mathbb{R}_{\geq 0}$ hat die Martingaleigenschaft bezüglich der Wahrscheinlichkeit $\frac{1}{2}$ für 0 und 1, falls sie folgende Gleichung erfüllt:

$$f(x) = \frac{1}{2}f(x0) + \frac{1}{2}f(x1), \text{ für } x \in \Sigma^*. \tag{3.2}$$

Funktionen mit dieser Eigenschaft wollen wir nun im Rahmen dieser Arbeit als Martingale bezeichnen.

Lemma 3.9. *Falls $f : \Sigma^* \to \mathbb{R}_{\geq 0}$ die Martingaleigenschaft (Gleichung 3.2) erfüllt, dann ist die Menge*

$$\mathcal{N} = \{z \in \Sigma^\infty \mid \overline{\lim}_{n\to\infty} f(z(n)) = \infty\}$$

eine Nullmenge.

Beweis. Wir definieren zunächst

$$\begin{aligned} F_k &= \{x \in \Sigma^* \mid f(x) > k\} \\ \overline{F}_k &= \{x \in F_k \mid x \notin F_k\Sigma\Sigma^*\}. \end{aligned}$$

Mit der Schreibweise $F_k\Sigma\Sigma^*$ soll die Konkatenation der Mengen F_k, Σ und Σ^* gemeint sein. Daraus folgt, dass der Schnitt $\overline{F}_k \cap \overline{F}_k\Sigma\Sigma^*$ leer ist. $\overline{F}_k$ besteht aus allen solchen Folgen/Wörtern aus F_k, die kein Präfix (also Anfangswort) in F_k haben.
Es gilt $\mu\varphi(F_k) = \mu\varphi(\overline{F}_k) = \sum_{x\in F_k} 2^{-|x|}$. Mit der Martingaleigenschaft 3.2 folgt

$$f(\varepsilon) \geq \sum_{x\in\overline{F}_k} f(x)2^{-|x|} \geq k \sum_{s\in\overline{F}_k} 2^{-|x|} \geq k\mu\varphi(F_k),$$

wobei ε das leere Wort bezeichnet. Damit gilt nun auch

$$f(\varepsilon)k^{-1} \geq \mu\varphi(F_k).$$

Da $\mathcal{N} \subset \bigcap_{k\in N} \varphi(F_k)$ gilt, folgt also auch, dass $\mathcal{N}$ eine Nullmenge ist. □

Mit diesem Vorwissen können wir nun das erste Konzept einer Testfunktion definieren.

Definition 3.10. *Eine totale Funktion $f : \Sigma^* \to \mathbb{R}_{\geq 0}$ heißt (1)-Test, falls f schwach berechenbar ist und die Martingaleigenschaft (3.2) besitzt.*

Definition 3.11. *Die Menge der unendlichen Binärwörter, die dem (1)-Test f nicht standhalten, definieren wir als*

$$\mathcal{N}_f = \{z \in \Sigma^\infty \mid \overline{\lim}_{n\to\infty} f(z(n)) = \infty\}.$$

Um die Hauptaussage dieses Abschnittes (Satz 3.13) zu beweisen, benötigen wir zunächst folgendes Lemma.

Lemma 3.12. *Für jede rekursiv aufzählbare Menge $A \subset \Sigma^*$ gibt es eine rekursive Menge $B \subset \Sigma^*$, so dass $A\Sigma^* = B\Sigma^*$ und $B \cap B\Sigma\Sigma^* = \emptyset$, also B präfixfrei ist.*

Beweis. Sei A definiert durch eine rekursive Funktion $h\colon N \to \Sigma^* \cup \emptyset$, so dass $A = h(N) \cap \Sigma^*$. Wir benutzen hierfür die Schreibweise $A_n = \bigcup_{i\leq n} h(i) \cap \Sigma^*$. Dabei bezeichnen wir mit $r(n)$ die maximale Wortlänge in A_n und sagen $r(n) = 0$, falls $A_n = \emptyset$.
Wir definieren die Menge B nun anhand ihrer Indikatorfunktion, siehe Definition 2.22. Sei nun $\mathcal{I}_B(x) = 0$, für alle $x \in \Sigma^*$ mit $|x| \neq r(n) + n$, für alle $n \in \mathbb{N}$. Für die restlichen $x \in \Sigma^*$ mit $|x| = r(n) + n$ definieren wir die Indikatorfunktion induktiv wie folgt. Für $|x| = r(0)$ sei

$$\mathcal{I}_B(x) = \begin{cases} 1, & x = h(0) \\ 0, & \text{sonst} \end{cases}$$

und für $|x| = r(n) + n$, und $r(n) \geq 1$ sei

$$\mathcal{I}_B(x) = \begin{cases} 1, & x \in h(n)\Sigma^* - A_{n-1}\Sigma^* \\ 0, & \text{sonst} \end{cases}.$$

Nach Konstruktion ist B dann präfixfrei. Außerdem gilt

$$A_n\Sigma^* = (B \cap \{x \mid |x| \leq r(n) + n\})\Sigma^*$$

und damit, wie gewünscht, auch $A\Sigma^* = B\Sigma^*$. □

Satz 3.13. *Ein unendliches Binärwort hält allen (1)-Tests stand genau dann, wenn es zufällig im Martin-Löf-Sinn ist.*

Beweis. $\Longrightarrow$: Sei zunächst U ein rekursiver sequentieller Test, welcher durch eine rekursiv aufzählbare Menge $V \subset \mathbb{N} \times \Sigma^*$ gegeben ist. Wir definieren den zugehörigen (1)-Test $f : \Sigma^* \to \mathbb{R}_{\geq 0}$ wie folgt:

$$f(x) = \sum_{i \in \mathbb{N}} i \left(\sum_{(x,y) \in V_i} 2^{-|y|} + \sum_{x(n) \in V_i, n < |x|} 1 \right).$$

f erfüllt dabei die Martingaleigenschaft (Gleichung 3.2).
Wir müssen also lediglich $F(x)$, $F(x0)$ und $F(x1)$ betrachten, welche aus $y \in V_i$ resultieren. Da V_i präfixfrei ist, gilt $f(\varepsilon) = \sum_{i \in \mathbb{N}} \mu\varphi(V_i)$. Das heißt insbesondere, dass $f(\varepsilon)$ beschränkt ist. Daher ist $f : \Sigma^* \to \mathbb{R}_{\geq 0}$ eine Funktion. Aus der Definition der Funktion f folgt direkt, dass f schwach berechenbar ist.
Wir nehmen nun an, z sei in $\varphi(V_i)$. Dann gibt es ein n, so dass $z(n) \in V_i$. Daher folgt $f(z(n)) \geq i$ und damit auch $\bigcap_{i \in \mathbb{N}} \varphi(V_i) \subset \mathcal{N}_f$.

$\Longleftarrow$: Sei z nun zufällig im Martin-Löf-Sinn und sei $f : \Sigma^* \to \mathbb{R}_{\geq 0}$ ein (1)-Test. Wir wählen nun ein $k > F(\varepsilon)$ und definieren die Menge $V \subset \mathbb{N} \times \Sigma^*$ via

$$V_i = \{x \in \Sigma^* \mid f(x) > 2^i k\}.$$

Da f schwach berechenbar ist, ist V rekursiv aufzählbar und daher gilt auch $\mu\varphi(V_i) \leq 2^{-i}$. Also ist U ein rekursiver universeller sequentieller Test und kann definiert werden mithilfe von $U_i = V_i\Sigma^*$.
Mit $z \notin \bigcap_{i \in \mathbb{N}} \varphi(U_i)$ folgt nun, dass z dem Test f standhält und damit (1)-zufällig ist.

□

Die Existenz eines universellen (1)-Test folgt bereits aus der Existenz von Martin-Löf-zufälligen Wörtern, aber lässt sich auch unabhängig davon mittels eines einfachen Arguments zeigen.

Bemerkung 3.14. *Sei $(f_i \mid i \in \mathbb{N})$ eine rekursive Aufzählung aller (1)-Tests mit $f_i(\varepsilon) \leq 1$. Also ist $f = \sum_{i \in \mathbb{N}} 2^{-i} f_i$ ein universeller (1)-Test.*

Außerdem könnten wir „$\underline{\lim}$“ anstelle von „$\overline{\lim}$“ in der Definition von $\mathcal{N}_f$ für (1)-Tests benutzen, wie folgendes Lemma zeigt.

Lemma 3.15. *Sei f ein universeller (1)-Test. Dann sind folgende Aussagen äquivalent:*

1. $\underline{\lim}_{n\to\infty} f(z(n)) = \infty$

2. $\overline{\lim}_{n\to\infty} f(z(n)) = \infty$

Beweis. Sei $U \subset \mathbb{N} \times \Sigma^*$ ein universeller rekursiver Test, welcher durch eine rekursiv aufzählbare Menge $V \subset \mathbb{N} \times \Sigma^*$, wie in der Hinrichtung vom Beweis von Satz 3.13. Ebenso betrachten wir den (1)-Test, wie in eben diesem Beweis. Dann reicht es zu zeigen, dass $z \in \bigcap_{i\in\mathbb{N}} \varphi(V_i)$ durch $\lim_{n\to\infty} f(z(n)) = \infty$ impliziert wird.
Angenommen $z \in \bigcap_{i\in\mathbb{N}} \varphi(V_i)$. Dann gibt es für ein beliebiges $i \in \mathbb{N}$ ein $n \in \mathbb{N}$, so dass $z(n) \in V_i$. Dies impliziert $f(z(m)) \geq i$, für alle $m \geq n$. □

3.3 (2)-Zufälligkeit

Die algorithmische Struktur eines (1)-Tests f ist nicht symmetrisch. Außerdem gibt es keinen Grund, weshalb wir für ein Martingal f schwache Berechenbarkeit fordern und nicht aber $-f$ schwach berechenbar fordern. Daher definieren wir nun ein zweites Testkonzept.

Definition 3.16. *Eine totale Funktion $f : \Sigma^* \to \mathbb{R}_{\geq 0}$ heißt (2)-Test, falls sie die Martingaleigenschaft (3.2) hat und falls $-f$ schwach berechenbar ist. Die Menge der unendlichen Binärwörter, die dem (2)-Test f nicht standhalten, definieren wir analog als*

$$\mathcal{N}_f = \{z \in \Sigma^\infty \mid \overline{\lim}_{n\to\infty} f(z(n)) = \infty\}.$$

Nun möchten wir die beiden Testkonzepte bezüglich ihrer Äquivalenz in Beziehung setzen und betrachten daher zunächst ein Hilfslemma, bevor wir den entscheidenden Satz, der einen Unterschied zwischen (1)-Tests und (2)-Tests aufzeigt, diesbezüglich beweisen können:

Lemma 3.17. *Sei f ein (2)-Test und $a \in \mathbb{Q}_{>0}$. Dann gibt es ein rekursives $z \in \Sigma^\infty$, sodass*

$$f(z(n)) < f(\varepsilon) + a, \text{ für } n \in \mathbb{N}.$$

Das Lemma sagt insbesondere aus, dass das entsprechende z sich nicht in der Nullmenge $\mathcal{N}_f$ befindet.

Beweis. Sei f nun der (2)-Test, welcher durch ein rekursives $g : \mathbb{N} \times \Sigma^* \to \mathbb{Q}$ mit $g(i, x) \geq g(i + 1, x)$ und $\overline{\lim}_{n\to\infty} g(i, x) = f(x)$ gegeben ist. Sei $b \in \mathbb{Q}$, sodass

$$f(\varepsilon) - \frac{a}{2} < b < F(\varepsilon).$$

Wir konstruieren nun das Wort z induktiv.
Wir nehmen an, $z(n) = z_1 \dots z_n$ (die ersten n Buchstaben des Wortes z) wurde so konstruiert, dass

$$f(z(n)) \leq b + a \sum_{j=0}^{n} 2^{-j-1}, \text{ für } i \leq n.$$

Dieses ist für $n = 0$ offensichtlich, welches den Induktionsanfang darstellt. Dann existiert ein $x \in \Sigma$, sodass

$$f(z(n)x) \leq b + a \sum_{j=0}^{n} 2^{-j-1}, \text{ für } i \leq n.$$

Hierbei soll $z(n)x$ wieder für die Konkatenation von $z(n)$ und x stehen. Wir können also i und x so bestimmen, dass

$$g(i, z(n)x) \leq b + a \sum_{j=0}^{n+1} 2^{-j-1}, \text{ für } i \leq n$$

gilt. Dann definieren wir $z(n+1) = z(n)x$. Nach Konstruktion folgt dann

$$f(z(n)) \leq b + a \leq f(\varepsilon) + a, \text{ für } n \in \mathbb{N}.$$

□

Nun werden wir sehen, dass es einen wesentlichen Unterschied zwischen (1)-Tests und (2)-Tests gibt.

Satz 3.18. *Es gibt Binärwörter, die* (2)*-zufällig sind, aber nicht* (1)*-zufällig.*

Beweis. Sei $(f_i \mid i \in \mathbb{N})$ eine Aufzählung aller (2)-Tests mit $f_i(\varepsilon) \leq 1$. Es reicht nun ein Wort z zu konstruieren, welches nicht (1)-zufällig ist, aber $\overline{\lim}_{n\to\infty} f_i(z(n)) < \infty$, für alle $i \in \mathbb{N}$ erfüllt.
Sei f ein universeller (1)-Test, also $\Sigma^* \backslash \mathcal{N}_f$ besteht exakt aus allen (1)-zufälligen Wörtern. Sei nun $z \in \Sigma^*$, welches wir wie folgt induktiv defineren. Dazu nehmen wir wie im vorherigen Beweis an, dass $z(n_k)$ bereits definiert ist und $f(z(n_k)) \geq k$ und dass

$$\sum_{i=0}^{k} 2^{-n_i - i} f_i(z(j)) \leq \sum_{i=0}^{k} 2^{-i+1}, \text{ für } j \leq n_k$$

gilt, für $n_k \in \mathbb{N}$. Auch hier ist diese Behauptung für $k = 0 = n_0$ und $f(\varepsilon) \leq 1$ offensichtlich und liefert uns hierbei den Induktionsanfang.
Für den Induktionsschritt betrachten wir nun f_{k+1}. Aus vorheriger Gleichung, $f(\varepsilon) \leq$

1 und der Martingaleigenschaft wissen wir, dass $2^{-n_k-k} f_{k+1}(z(n_k)) \leq 2^{-k}$ gilt. Das heißt, es gibt ein rekursives $y \in \Sigma^\infty$, das $y(n_k) = z(n_k)$ die Gleichung

$$\sum_{i=0}^{k+1} 2^{-n_i-i} f_i(y(j)) \leq \sum_{i=0}^{k+1} 2^{-i+1}, \text{ für } j \in \mathbb{N} \text{ und } n_{k+1} \geq n_k$$

erfüllt. Dies folgt direkt aus der Konstruktion im Beweis des vorherigen Hilfslemmas. Da y rekursiv ist, existiert ein $n_{k+1} > n_k$, so dass $f(y(n_{k+1})) > k+1$. Und damit definieren wir $z(n_{k+1}) = y(n_{k+1})$.
Nach Konstruktion gilt $z \in \mathcal{N}_f$. Andererseits gilt

$$\overline{\lim}_{j\to\infty} \sum_{i=0}^{\infty} 2^{-n_i-i} f_i(z(j)) \leq 4,$$

für $i \in \mathbb{N}$. Das heißt insbesondere, dass $z \notin \mathcal{N}_{f_i}$. □

Bemerkung 3.19. *Jedes* (1)*-zufällige Binärwort ist auch* (2)*-zufällig. Es ist also entscheidend für ein Martingal f, ob f oder $-f$ schwach berechenbar ist.*

3.4 (3)-Zufälligkeit und die arithmetische Hierarchie

Wir sehen insbesondere, dass beide Testkonzepte noch nicht symmetrisch sind und wollen daher ein Testkonzept finden, welches auf Martingalen basiert und symmetrische algorithmische Struktur aufweist.

Definition 3.20. *Ein Martingal $f\colon \Sigma^* \to \mathbb{R}_{\geq 0}$ ist ein* (3)*-Test, falls es eine rekursive Funktion $g\colon \mathbb{N} \times \Sigma^* \to \mathbb{Q}$ gibt, sodass*

$$\lim_{i\to\infty} g(i,x) = f(x),$$

für $x \in \Sigma^*$. Die Menge der unendlichen Binärwörter, die dem (3)*-Test f nicht standhalten, definieren wir als*

$$\mathcal{N}_f = \{z \in \Sigma^\infty \mid \overline{\lim}_{n\to\infty} f(z(n)) = \infty\}.$$

Jeder (3)-Test ist also auch ein (1)-Test, da (3)-Tests mehr Anforderungen besitzen als (1)-Tests. Daher sind alle (3)-zufälligen Wörter auch (1)-zufällig.
Wir wollen nun mithilfe der arithmetischen Hierarchie zeigen, dass die Umkehrung nicht gilt.

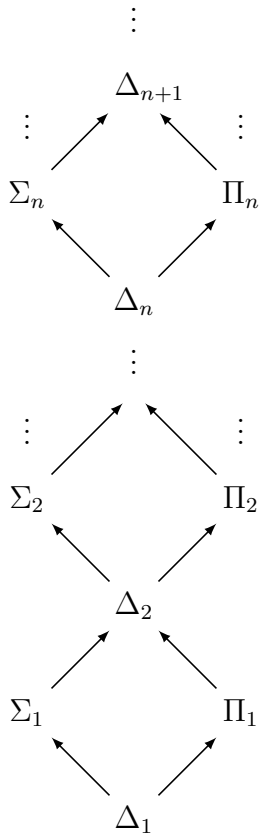


Abbildung 3.1: Die Arithmetische Hierarchie

Definition 3.21. *Bezeichne nun z_n die n-te Stelle des Wortes z. Ein Wort $z \in \Sigma^\infty$ ist in Σ_n, falls $\{n \mid z_n = 1\} \in \Sigma_n$.*
Ein Wort $z \in \Sigma^\infty$ ist in Π_n, falls $\{n \mid z_n = 1\} \in \Pi_n$.

Zunächst wiederholen wir dazu ein paar Eigenschaften der arithmetischen Hierarchie aus der Vorlesung Berechenbarkeit und Logik [Vol19].

Bemerkung 3.22. *1.* $\Sigma_0 = \Pi_0 = \Sigma_1 \cap \Pi_1 = \Delta_1$

2. $A \in \Sigma_n \iff \overline{A} \in \Pi_n$

3. $\Sigma_n \cup \Pi_n \subset \Sigma_{n+1} \cap \Pi_{n+1} =: \Delta_{n+1}$

4. $A \in \Sigma_{n+1} \iff$ *A ist rekursiv aufzählbar in einer Menge* $B \in \Pi_n$

5. $A \in \Sigma_{n+1} \cap \Pi_{n+1} = \Delta_{n+1} \iff$ *A ist rekursiv relativ zu einer Menge* $B \in \Pi_n$.

Der einzige Unterschied zur Definition aus [Vol19] ist hier, dass wir $\Delta_n := \Sigma_n \cap \Pi_n$ zusätzlich definieren.

Lemma 3.23. *Es gibt ein Wort in $\Delta_2 = \Sigma_2 \cap \Pi_2$, das (1)-zufällig ist.*

Beweis. Sei $f\colon \Sigma^* \to \mathbb{R}_{\geq 0}$ ein universeller (1)-Test, welcher durch eine rekursive Funktion $g\colon \mathbb{N} \times \Sigma^* \to \mathbb{Q}$ gegeben ist. Wir nehmen an, dass $f(\varepsilon) < 1$ gilt. Dann existiert ein $P \in \Pi_1$ mit $x \in P \iff \forall i \in \mathbb{N} g(i, x) < 1$. Mit Eigenschaft 5, aus Bemerkung 3.22 folgt nun, dass $z \in \Delta_2$ ist. Nach Konstruktion ist z allerdings nicht in $\mathcal{N}_f$. □

Satz 3.24. *Es gibt kein Wort in $\Sigma_2 \cup \Pi_2$, das (3)-zufällig ist.*

Beweis. Sei zunächst z ein Wort Σ_2 gemäß Definition 3.21. Dann existiert ein rekursives P, sodass für alle $n \in \mathbb{N}$ gilt

$$z_n = 1 \iff \exists j \in \mathbb{N} \forall i \in \mathbb{N} (j, i, n) \in P.$$

Wir definieren nun einen (3)-Test f, mit $z \in \mathcal{N}_f$ durch eine rekursive Funktion $g\colon \mathbb{N} \times \Sigma^* \to \mathbb{Q}$. Wir schreiben dabei

$$f(i, n) = \{j \mid \forall r \leq i (j, r, n) \in P, j \leq i\}.$$

Die endliche Menge $f(i, n)$ kann dabei bestimmt werden. Danach berechnen wir $g(i, x)$ wie folgt:

$$g(i, \varepsilon) = 1, \text{ für } i \in \mathbb{N}.$$

Falls $(f(i, n) \neq \emptyset \wedge y_n = 1) \vee (f(i, n) = \emptyset \wedge y_n = 0)$, definieren wir

$$g(i, y(n)) = 2g(i, y(n-1)), \text{ für } y \in \Sigma^\infty.$$

Falls $(f(i, n) \neq \emptyset \wedge y_n = 0) \vee (f(i, n) = \emptyset \wedge y_n = 1)$, definieren wir

$$g(i, y(n)) = 0, \text{ für } y \in \Sigma^\infty.$$

Dies liefert uns

$$\overline{\lim}_{i \to \infty} g(i, y(n)) = \begin{cases} 2 \lim_{i \to \infty} g(i, y(n-1)), & z_n = y_n \\ 0, & \text{sonst.} \end{cases}$$

Also ist f ein (3)-Test und es folgt

$$f(y(n)) = \begin{cases} 2^n, & y(n) = z(n) \\ 0, & \text{sonst.} \end{cases}$$

Dies zeigt die Aussage für $z \in \Sigma_2$. Für $z \in \Pi_2$ heißt dies, dass wir $\{n \mid z_n = 0\} \in \Sigma_2$

vorfinden und daher gilt selbige Argumentation, um für $z \in \Pi_2$ zu zeigen, dass es nicht (3)-zufällig ist.

□

Insgesamt haben wir also gesehen

1. Es gibt (1)-zufällige Wörter in $\Delta_2 = \Sigma_2 \cap \Pi_2$.
2. Es gibt keine (3)-zufälligen Wörter in $\Sigma_2 \cup \Pi_2$.
3. Nach Bemerkung 3.22 und Abbildung 3.1 gilt $\Delta_2 = \Sigma_2 \cap \Pi_2 \subseteq \Sigma_2 \cup \Pi_2$.
4. Nicht jedes (1)-zufällige Wort ist auch (3)-zufällig.

Wir haben in diesem Kapitel also gesehen, dass wir die Martin-Löf-Zufälligkeit auch in die Sprache der Martingale übersetzen können und mithilfe dessen dem Konstrukt eine symmetrische algorithmische Struktur verleihen zu können, wenn wir weitere zusätzliche Annahmen erzwingen. Zudem lassen sich mithilfe verschiedener Zufälligkeitsbegriffe nach Schnorr verschieden Aussagen über die arithmetische Hierarchie treffen, was dem Themengebiet eine Anwendung in einem anderen Kontext anbietet.

4 Resource-bounded Measuretheory

Dieses Kapitel richtet sich nach [Lut96]. In diesem Kapitel geht es darum, mithilfe der in den vorherigen Kapiteln beschriebenen (Wahrscheinlichkeits-)Maßtheorie Komplexitätsklasseninklusionen genauer zu betrachten und Aussagen über schwere und vollständige Probleme zu formulieren. Wir haben gesehen, dass vergleichbar kleine oder vernachlässigbare Mengen Nullmengen sind. Ebenso haben deren Komplemente, also große Mengen, Maß 1.

Dieses Kapitel soll in erster Linie einen kurzen Einblick in die Theorie geben und verweist daher bei einigen Beweisen nur auf die entsprechende Literatur, um den Rahmen nicht zu sprengen und dennoch wesentliche Aussagen präsentieren zu können. Da wir hier außerdem den Martingalbegriff wieder verwenden werden, bietet sich dieser Ausblick thematisch an.

4.1 Notationen und Grundlagen

In diesem Abschnitt wollen wir bisher nicht definierte Begriffe definieren und einige Grundlagen schaffen, welche wir im nachfolgenden Abschnitt benötigen werden.

Definition 4.1. *Eine Sprache A heißt sparse, wenn es ein Polynom $q(n)$ gibt, so dass $|A_{\leq n}| \leq q(n)$, wobei $|A_{\leq n}|$ die Anzahl der Wörter aus der Sprache A sind, die weniger gleich n Zeichen besitzen.*
Eine Sprache A heißt dicht, falls es eine reelle Zahl $\delta > 0$ gibt, so dass $|A_{\leq n}| > 2^{n^\delta}$.

Definition 4.2. *Für eine Funktion $f\colon \{0,1\}^* \to \{0,1\}^*$ und eine Zahl $i \in \mathbb{N}$ definieren wir die Funktion $f_i\colon \{0,1\}^* \to \{0,1\}^*$ mit $f_i(x) = f(\langle 0^i, x\rangle)$. Damit können wir f als uniforme Aufzählung der Funktionen $f_0, f_1, \ldots$ betrachten.*

Die Komplexitätsklassen, die wir aus der Vorlesung „Komplexität von Algorithmen“ (siehe [MSV20]) kennen, sind lediglich für Sprachen definiert. Um aus ihnen Funktionsklassen zu machen, hängen wir als Präfix an die entsprechende Klasse ein F, um zu zeigen, dass es die zugehörige Funktionsklasse ist.

Beispiel 4.3. *Sei $t : \mathbb{N} \to \mathbb{N}$. FTIME(t) ist dann die Klasse aller Funktionen $f : \{0,1\}^* \to \{0,1\}^*$, so dass $f(x)$ berechenbar in Zeit $\mathcal{O}(t(|x|))$ ist.*

Wir wollen im folgenden Kapitel die Komplexitätsklassen E und EXP untersuchen, die wir hier einführen wollen.

Definition 4.4.

- $E := TIME(2^{\mathcal{O}(n)})$
- $EXP := \bigcup_{k\in\mathbb{N}} TIME(2^{n^k})$

4.2 Ressource-bounded Measuretheory für E und EXP

Wir sagen, fast alle Sprachen einer Komplexitätsklasse $\mathcal{C}$ haben eine bestimmte Eigenschaft, wenn die Menge der Sprachen, die in dieser Klasse $\mathcal{C}$ diese Eigenschaft haben, Maß 1 über $\mathcal{C}$ haben (siehe Definition 2.16).

Bemerkung 4.5. *Anders ausgedrückt: Für Mengen X von Sprachen sind die folgenden Aussagen äquivalent.*

- $P(X) = 0$
- *Es gibt ein Martingal d, welches auf allen Elementen von X erfolgreich ist/ standhält.*

Vom Standpunkt der klassischen Berechenbarkeitstheorie sind Klassen, wie $\mathrm{P}, \mathrm{NP}, \mathrm{PH}, \mathrm{PSPACE}$ vernachlässigbar klein und sind daher schwierig zu unterscheiden. Daher wollen wir uns im Folgenden die Klassen E und EXP in dieser Theorie ansehen.

Bemerkung 4.6. *P hat Maß 0 in EXP.*

Im folgenden Abschnitt wollen wir nun die interne Maßstruktur von $E_1 = \mathrm{E}$ und $E_2 = \mathrm{EXP}$ untersuchen.

Definition 4.7.

$$\begin{aligned} P = p_1 &= \{f \mid f \text{ ist polynomialzeitberechenbar}\} \\ p_2 &= \{f \mid f \text{ ist in Zeit } n^{(\log n)^{\mathcal{O}(1)}} \text{ berechenbar}\} \end{aligned}$$

Definition 4.8. *Ein Martingal d ist p_i-berechenbar, falls eine Funktion $\hat{d} : \mathbb{N} \times \{0,1\}^* \to \mathbb{Q}$ existiert, sodass $\widehat{d} \in p_i$ und für alle $r \in \mathbb{N}$ und $w \in \{0,1\}^*$ gilt*

$$\left|\widehat{d}_r(w) - d(w)\right| \leq 2^{-r}.$$

Ein p_i-Martingal ist ein Martingal, welches p_i-berechenbar ist.

Die Schlüsselidee ist nun die Folgende:

Definition 4.9. *Eine Menge von Sprachen X hat p_i-Maß 0 ($\mu_{p_i}(X) = 0$), wenn es ein p_i-Martingal d gibt, dass auf allen Sprachen aus X erfolgreich ist bzw. stand hält.*
Eine Menge von Sprachen X hat p_i-Maß 1 ($\mu_{p_i}(X) = 1$), wenn $\mu(\overline{X}) = 0$.

Wenn wir also zeigen wollen, dass eine Eigenschaft gilt, so bietet sich die „Probabilistische Methode“ aus [AZ14] an. Diese besagt, dass wenn auf einer Menge von Objekten die Wahrscheinlichkeit, dass ein Objekt bestimmte Eigenschaften hat echt kleiner 1 und echt größer 0 hat, dann muss es ein solches Objekt mit diesen Eigenschaften geben. Auf unserem Kontext angepasst, bedeutet dies, wenn die Wahrscheinlichkeit, dass ein Objekt dieser Menge eine Eigenschaft hat 1 ist, dann haben fast alle Objekte (bei einer endlichen Grundmenge alle Objekte) diese Eigenschaft.

Bemerkung 4.10. *Um die p_i-Konstruktion nun auf unsere Klassen E und EXP anwenden zu können, betrachten wir $E = E_1$ und $EXP = E_2$, also die ersten zwei Klassen der natürlichen Exponentialhierarchie. Die Idee lässt sich prinzipiell auf weitere Stufen dieser Hierarchie ausweiten, würde im Rahmen dieser Arbeit aber zu wenig weiterer Veranschaulichung der Arbeitsweise in diesem Themengebiet führen, sowie den Rahmen dieser Arbeit sprengen.*

Definition 4.11. *Eine Menge X hat Maß 0 in E_i und wir schreiben dazu*

$$\mu(X \mid E_i) = 0, \text{ falls } \mu_{p_i}(E_i \cap X) = 0.$$

Eine Menge X hat Maß 1 in E_i und wir schreiben dazu

$$\mu(X \mid E_i) = 1, \text{ falls } \mu(\overline{X} \mid E_i) = 0.$$

In diesem Fall sagen wir auch, fast jede Sprache aus E_i gehört auch zu X.

Bemerkung 4.12. *Für alle $X \subseteq \{0,1\}^\infty$ und $A \subseteq \{0,1\}^*$ zufällig gewählt, gilt*

- $\mu_p(X) = 0 \Rightarrow \mu_{p_2}(X) = 0 \Rightarrow Pr[A \in X] = 0$
- $\mu_p(X) = 0 \Rightarrow \mu(X \mid E) = 0$
- $\mu_{p_2}(X) = 0 \Rightarrow \mu(X \mid EXP) = 0$

Die Notation $Pr[A \in X] = 0$ bedeutet, dass die Wahrscheinlichkeit, dass $A \in X$ gilt, 0 ist.

Beweis.
- $\mu_p(X) = 0$ bedeutet, es gibt ein p_1-Martingal, das auf allen Sprachen von X stand hält. Ein p_1-Martingal ist ein Martigal, welches p_1-berechenbar, also polynomialzeitberechenbar ist. Wenn d dieses p_1-Martingal ist, so ist d auch in Zeit $n^{(\log n)^{\mathcal{O}(1)}}$ berechenbar und damit auch ein p_2-Martingal für X. Also gilt auch $\mu_{p_2}(X) = 0$. Wenn X eine Nullmenge ist, so ist auch die Wahrscheinlichkeit, dass eine Sprache in ebendieser Nullmenge liegt, gleich 0.
- $\mu(X \mid \mathrm{E}) = \mu_p(\mathrm{E} \cap X)$. Da $\mathrm{E} \cap X \subset X$, gilt nach den Rechenregeln für Maßfunktionen, siehe Bemerkung 2.8, dass $\mu_p(E \cap X) \leq \mu_p(X) = 0$. Da Maße aber minimal den Wert 0 annehmen, folgt aus $\mu_p(X) = 0$ bereits $\mu(X \mid \mathrm{E}) = 0$.
- Nach obiger Begründung mit p_2 anstelle von p und EXP anstelle von E folgt die Aussage.

□

Satz 4.13. *Es gelten $\mu(E \mid E) \neq 0$ und $\mu(EXP \mid EXP) \neq 0$.*

Beweis. Ein Beweis findet sich in [Lut89]. □

Satz 4.14. *Sei $\mathcal{C}$ eine Menge von Sprachen, die entweder unter symmetrischer Differenz oder unter (endlicher) Vereinigung und Schnitt abgeschlossen ist, so gilt*

1. $\mu(\mathcal{C} \mid E) = 1 \Rightarrow E \subseteq \mathcal{C}$
2. $\mu(\mathcal{C} \mid EXP) = 1 \Rightarrow EXP \subseteq \mathcal{C}$

Beweis. Ein Beweis findet sich hierzu in [SCR02]. □

4.3 Inkompressibilität

Viele Resultate über die Struktur von E und EXP unter $\leq_m^{\mathrm{P}}$-Reduktion betrachten Sprachen, welche unkomprimierbar unter $\leq_m^{\mathrm{P}}$-Reduktionen sind. Dies führt uns zu folgenden Definitionen.

Definition 4.15. *Die Kollisionsmenge einer Funktion* $f\colon \{0,1\}^* \to \{0,1\}^*$ *ist*

$$C_f = \{x \in \{0,1\}^* \mid \exists y < x : f(y) = f(x)\}.$$

Hierbei nutzen wir die gewöhnliche lexikographische Ordnung auf $\{0,1\}^*$.

Bemerkung 4.16. *f ist injektiv genau dann, wenn ihre Kollisionsmenge $C_f = \emptyset$ ist.*

Beweis. Eine Funktion $f\colon \{0,1\}^* \to \{0,1\}^*$ ist injektiv genau dann, wenn für alle x und alle y mit $x \neq y$ gilt, dass $f(x) \neq f(y)$. Dieses ist lediglich eine andere Formulierung für eine leere Kolisionsmenge C_f. □

Dies wollen wir nun in die Sprache der Maßtheorie übersetzen bzw. verallgemeinern.

Definition 4.17. *Eine Funktion $f\colon \{0,1\}^* \to \{0,1\}^*$ ist fast überall injektiv, wenn ihre Kollisionsmenge C_f endlich ist.*

Diese Definition lässt sich damit begründen, dass endliche Mengen Maß null haben und wir in Definition 2.16 den Begriff „fast überall“ mit „bis auf Nullmengen“ gleichgesetzt haben.

Wir wollen nun einen neuen Reduktionsbegriff einführen.

Definition 4.18. *Seien A, B Sprachen über dem Alphabet $\{0,1\}$ und sei $t\colon \mathbb{N} \to \mathbb{N}$. Eine $\leq_m^{TIME(t)}$-Reduktion von A auf B ist gegeben durch eine Funktion $f \in FTIME(t)$, so dass $A = f^{-1}(B)$.*
Mit anderen Worten:

$$A \leq_m^{TIME(t)} B :\Leftrightarrow (\forall x \in \{0,1\}^* : x \in A \Leftrightarrow f(x) \in B).$$

Eine $\leq_m^{TIME(t)}$-Reduktion von A auf eine Funktion f ist eine $\leq_m^{TIME(t)}$-Reduktion von A auf $f(A)$.

In Kapitel 1 hatten wir bereits einen Begriff von Unkomprimierbarkeit gesehen. Jetzt wollen wir einen Begriff bzgl. unserer neuen Reduktionen definieren.

Definition 4.19. *Sei $t\colon \mathbb{N} \to \mathbb{N}$. Eine Sprache $A \subset \{0,1\}^*$ heißt unkomprimierbar durch $\leq_m^{TIME(t)}$-Reduktionen, falls jede $\leq_m^{TIME(t)}$-Reduktion von A fast überall injektiv ist.*

Bemerkung 4.20. *Für die bekannten $\leq_m^P$-Reduktionen bedeutet dies, dass eine Sprache A über dem Alphabet $\{0,1\}$ unkomprimierbar durch $\leq_m^P$-Reduktionen ist, wenn es unkomprimierbar durch $\leq_m^{TIME(q)}$-Reduktionen ist, für alle Polynome q.*

Intuitiv bedeutet dieser Komprimierbarkeitsbegriff, dass wenn f eine $\leq_m^{\mathrm{TIME}(t)}$-Reduktion von A auf B definiert und die entsprechende Kollisionsmenge C_f groß genug ist, dann komprimiert f viele Fragen der Form „$x \in A$?“ zu weniger Fragen der Form „$f(x) \in B$?“.
Ist A jedoch unkomprimierbar durch $\leq_m^{\mathrm{P}}$-Reduktionen, so können nur kleine solcher Kompressionen auftauchen.

Satz 4.21. *E enthält alle Sprachen, die unkomprimierbar durch $\leq_m^P$-Reduktionen sind.*

Beweis. Ein Beweis findet sich in [Mey76]. □

Das folgende Theorem soll nun zeigen, dass die Gegenrichtung für fast alle Sprachen aus E ebenfalls gilt.

Theorem 4.22. *Sei $c \in \mathbb{Z}_{\geq 0}$. Wir definieren die Mengen*

$$X := \{A \subseteq \{0,1\}^* \mid A \text{ ist unkomprimierbar durch } \leq_m^{TIME(2^{cn})} \text{-Reduktionen}\}$$
$$Y := \{A \subseteq \{0,1\}^* \mid A \text{ ist unkomprimierbar durch } \leq_m^{TIME(2^{c^n})} \text{-Reduktionen}\}.$$

Dann gilt für die entsprechenden Maße $\mu_{p_1}(X) = \mu_{p_2}(Y) = 1$.

Dies bedeutet also, dass fast alle Sprachen in E unkomprimierbar für $\leq_m^{\mathrm{TIME}(2^{cn})}$-Reduktionen und fast alle Sprachen in EXP unkomprimierbar für $\leq_m^{\mathrm{TIME}(2^{c^n})}$-Reduktionen sind. Hierfür wollen wir nun eine Beweisskizze angeben.

Beweis. Beweisskizze für $\mu_{p_1}(X) = 1$:
Es genügt für die Aussage ein p_1-Martingal $d\colon \{0,1\}^* \to [0,\infty)$ anzugeben, das allen Elementen von $\overline{X}$ standhält.
Sei nun $f \in \mathrm{TIME}(2^{(c+1)n})$, so dass $\mathrm{TIME}(2^{cn}) = \{f_i \mid i \in \mathbb{N}\}$. Für jedes $i \in \mathbb{N}$ definieren wir nun eine Menge Z_i wie folgt:
Wenn die Kollisionsmenge C_{f_i} endlich ist, so ist $Z_i = \emptyset$. Andernfalls, also wenn die Kollisionsmenge C_{f_i} nicht endlich ist, dann ist Z_i die Menge aller Sprachen A, so dass f_i eine $\leq_m^{\mathrm{TIME}(2^{cn})}$-Reduktion von A beschreibt. Dann gilt $X = \bigcup_i Z_i$ und wir definieren das Martingal d durch $d(w) = \sum_{i=0}^{\infty} 2^{-i} d_i(w)$. Hierbei benutzen wir $d_i\colon \{0,1\}^* \to [0,\infty)$, welche wie folgt induktiv definiert werden. Seien $i \in \mathbb{N}, w \in$

$\{0,1\}^*$ und $b \in \{0,1\}$. Sei nun die Folge $s_0, s_1, \ldots$ die lexikographische Aufzählung von $\{0,1\}^*$. Dann ist

1. $d_i(\varepsilon) = 1$.
2. Wenn $s_{|w|} \notin C_{f_i}$, dann $d_i(wb) = d_i(w)$.
3. Wenn $s_{|w|} \in C_{f_i}$, dann fixieren wir das letzte $j \in \mathbb{N}$, so dass $f_i(s_j) = f_i(s|w|)$ und definieren $d_i(wb) = 2 \cdot d_i(w) \cdot [\![b = w(j)]\!]$.

Hierbei meint die Notation $[\![x]\!]$ den Wahrheitswert, der von x angenommen wird.
Jedes d_i ist ein Martingal, wenn d ein Martingal ist.
Wenn wir diese d_i nach obiger Definition intuitiv beschreiben wollen, ergibt sich hierfür folgende intuitive Interpretation: Wir betrachten nun die Mitgliedschaft eines Wortes in einer Sprache als Zufallsexperiment. Dann wettet jedes d_i darauf, dass ein Wort in der Sprache liegt. $d_i(\varepsilon) = 1$ bedeutet, dass d_i zunächst mit einem Euro beginnt. Wenn $s_{|w|} \notin C_{f_i}$, dann heißt das, dass d_i bei solchen Wörtern, für die $x \notin C_f$ liegt, nicht wettet. Und schließlich, falls $s_{|w|} \in C_{f_i}$ gilt, so setzt d_i alles Geld darein, dass $x \in A \Leftrightarrow y \in A$, wobei y das erste Wort ist, welches $f_i(x) = f_i(y)$ erfüllt.
Wenn $A \in Z_i$, dann ist die Wette korrekt und das Geld von d_i wird unendlich oft verdoppelt. Daher hält d_i auf jedem Element von Z_i stand und damit auch d auf jedem Element von $\overline{X}$.
Jetzt fehlt noch, dass wir zeigen, dass d p_i-berechenbar ist. Dafür definieren wir

$$\widehat{d}\colon \mathbb{N} \times \{0,1\}^* \to \mathbb{Q} \text{ mit } \widehat{d}(r,w) = \sum_{i=0}^{r+|w|} 2^{-i} d_i(w).$$

Da $f \in \mathrm{TIME}(2^{(c+1)n})$ und die Berechnung der $d_i(w)$ lediglich Werte von $f_i(u)$ für $|u| \in \mathcal{O}(\log|w|)$ nimmt, folgt daraus, dass $d \in \mathrm{P}$. Also

$$\left|\widehat{d}(r,w) - d(w)\right| = \sum_{i=r+|w|+1}^{\infty} 2^{-i} d_i(w) \leq \sum_{i=r+|w|+1}^{\infty} 2^{|w|-i} = 2^{-r},$$

für alle $r \in \mathbb{N}$ und $w \in \{0,1\}^*$. Daher ist d p_1-berechenbar. □

Korollar 4.23. *Fast jede Sprache in E und fast jede Sprache in EXP sind unkomprimierbar durch $\leq_m^P$-Reduktionen.*

Dieses Korollar ist eine Umformulierung der Aussage von Theorem 4.22, da „fast jede“ (alle bis auf Nullmengen/ alle bis auf endlich viele) das gleiche besagt wie,

dass die Menge der unkomprimierbaren Sprachen Maß 1 hat.

Insgesamt ergibt sich hierdurch ein Rückgriff zum eigentlichen Thema, nämlich der Zufallsbegriffe in der Berechenbarkeitstheorie, die wir auch durch einen Unkomprimiertheitsbegriff dargestellt haben. Ebenso haben wir gesehen, dass fast alle Wörter dabei zufällig, also unkomprimierbar waren. Hier unterscheidet sich der Komprimierbarkeitsbegriff und die zugehörige Menge auf denen er definiert ist. Das Resultat bleibt jedoch das gleiche, so dass man in die Richtung weiter überlegen kann, ob man die Zufallsbegriffe in dem Kontext auf Sprachen erweitern möchte.

4.4 Bi-Immunität

In diesem Abschnitt soll es um P-Bi-Immunität gehen, bevor wir im nächsten Abschnitt zu den Komplexitätskernen kommen.
Dieser Abschnitt folgt dem Vorgehen von [Lut96].

Definition 4.24. *Eine Sprache $A \subseteq \{0,1\}^*$ heißt P-Immun, wenn für alle Sprachen $B \subseteq A$ und $B \in P$ folgt, dass B endlich ist und damit Maß 0 besitzt.*
Eine Sprache $A \subseteq \{0,1\}^*$ heißt P-Bi-Immun, falls A und $\overline{A}$ jeweils P-Immun sind.

Bemerkung 4.25. *Intuitiv bedeutet dies, dass eine P-Bi-Immune Sprache weder von Sprachen aus P von innen noch von außen trivial approximiert werden kann.*

Satz 4.26. *Jede Sprache, die unkomprimierbar durch $\leq_m^P$-Reduktionen ist, ist P-Bi-Immun.*

Dies lässt uns das Theorem 4.22 nocheinmal anders deuten, nämlich dass fast alle Sprachen aus E und fast alle Sprachen aus EXP P-Bi-Immun sind.

4.5 Komplexitätskerne

In diesem Kapitel geben wir eine Definition von Komplexitätskernen an und erwähnen, dass fast jede Sprache in E und fast jede Sprache in EXP sehr große Komplexitätskerne hat. Dieser Abschnitt richtet sich nach [Lut96].
Intuitiv geht es bei dem Komplexitätskern einer Sprache A um eine feste Eingabemenge K, so dass jede Maschine, die konsistent (siehe Definition 4.28) für A ist, bei allen (aber endlich vielen) Elementen von K scheitert, sie effizient zu entscheiden. Effizient ist hierbei ein kontextbezogener Parameter.

Definition 4.27. *Wenn eine Turingmaschine M bei Eingabe x hält, bezeichnen wir die Anzahl der Schritte, die $M(x)$ in der Berechnung benötigt, mit $time_M(x)$.*
Wenn $M(x) = \uparrow$ gilt, also M bei Eingabe x in eine Endlosschleife gerät, so ist $time_M(x) = \infty$.

Definition 4.28. *Eine Turingmaschine M heißt konsistent zu einer Sprache $A \subseteq \{0,1\}^*$, wenn $M(x) \leq [\![x \in A]\!]$, für alle $x \in \{0,1\}^*$.*
Hierfür wollen wir eine partielle Ordnung auf der Menge $\{0, 1, \uparrow\}$ mit $\uparrow < 0$ und $\uparrow < 1$, aber 0 und 1 unvergleichbar annehmen.

Der Begriff konsistent erweitert den Begriff, dass eine Turingmaschine eine Sprache entscheidet, in dem Sinne, dass auch das Nichthalten, bzw. das Gelangen in eine Endschlossschleife dazu zählt. Eine konsistente Turingmaschine für eine Sprache kann also für ein Wort der Sprache halten und akzeptieren oder nicht halten.

Definition 4.29. *Sei $t\colon \mathbb{N} \to \mathbb{N}$ und seien $A, K \subseteq \{0,1\}^*$. Dann ist K ein TIME($t(n)$)-Komplexitätskern von A, wenn für jedes $c \in \mathbb{N}$ und jeder Turingmaschine M, welche konsistent zu A ist, die schnellen Mengen*

$$F := \{x \mid time_M(x) \leq c \cdot t(|x|) + c\}$$

der Ungleichung $|F \cap K| < \infty$ genügen. F (für engl. fast) ist dann die Menge aller Wörter, die M effizient entscheidet.

Bemerkung 4.30. *Sei $A \subseteq \{0,1\}^*$ eine Sprache. Jede Teilmenge eines TIME($t(n)$)-Komplexitätskerns von A ist wieder ein TIME($t(n)$)-Komplexitätskern von A.*
Wenn $s(n) \in \mathcal{O}(t(n))$, dann ist jeder TIME($s(n)$)-Komplexitätskern von A auch ein TIME($t(n)$)-Komplexitätskern von A. Dies liegt in der $\mathcal{O}$-Notation aus [MSV20] begründet.

Wir wollen diese Begriffe nun genauer für die Komplexitätsklassen P, E und EXP ansehen.

Definition 4.31. *Seien $A, K \subseteq \{0,1\}^*$.*
K ist ein polynomieller Komplexitätskern (oder auch P-Komplexitätskern) von A, wenn K ein TIME(n^k)-Komplexitätskern von A für alle $k \in \mathbb{N}$ ist.
K ist ein exponentieller Komplexitätskern von A, wenn es eine reelle Zahl $\delta > 0$ gibt, so dass K ein TIME(2^{n^δ})-Komplexitätskern von A ist.

Lemma 4.32. *Sei $t\colon \mathbb{N} \to \mathbb{N}$ zeitkonstruierbar. Dann hat jede Sprache, die unkomprimierbar durch $\leq_m^{TIME(t)}$-Reduktionen ist, $K = \{0,1\}^*$ als TIME($t(n)$)-Komplexitätskern.*

Korollar 4.33. *1. Jede Sprache, die unkomprimierbar durch $\leq_m^P$-Reduktionen ist, hat $K = \{0,1\}^*$ als P-Komplexitätskern.*

2. Jede Sprache, die unkomprimierbar durch $\leq_m^{TIME(2^{cn})}$-Reduktionen ist, hat $K = \{0,1\}^*$ als TIME(2^{cn})-Komplexitätskern.

3. Jede Sprache, die unkomprimierbar durch $\leq_m^{TIME(2^{c^n})}$-Reduktionen ist, hat $K = \{0,1\}^*$ als TIME(2^{c^n})-Komplexitätskern.

Zusammen mit Korollar 4.33 liefert Theorem 4.22, dass fast jede Sprache, die in Exponentialzeit entscheidbar ist, den größtmöglichsten Komplexitätskern hat.

Korollar 4.34. *Sei $c \in \mathbb{Z}_{\geq 0}$.*

1. Fast jede Sprache in E hat $K = \{0,1\}^*$ als TIME(2^{cn})-Komplexitätskern.

2. Fast jede Sprache in EXP hat $K = \{0,1\}^*$ als TIME(2^{c^n})-Komplexitätskern.

Nun wollen wir Schlüsselaspekte der Strukturen von E und EXP unter Polynomialzeit many-one Reduktion untersuchen. Dazu definieren wir zunächst den $\leq_m^{\mathrm{P}}$-Span einer Sprache, um uns darauf aufbauend verschiedene Span-Theoreme anzusehen.

Definition 4.35. *Der untere $\leq_m^P$-Span einer Sprache $A \subseteq \{0,1\}^*$ ist definiert als*

$$P_m(A) = \{B \subseteq \{0,1\}^* \mid B \leq_m^P A\}.$$

Der obere $\leq_m^P$-Span einer Sprache $A \subseteq \{0,1\}^*$ ist definiert als

$$P_m(A)^{-1} = \{B \subseteq \{0,1\}^* \mid A \leq_m^P B\}.$$

Damit können wir mithilfe der $\leq_m^{\mathrm{P}}$-Reduktionen eine Struktur auf den Sprachen aufbauen, so dass $P_m(A)$ alle Sprachen sind, die „unter“ A liegen und $P_m(A)^{-1}$ alle Sprachen sind, die „über“ A liegen.
Abbildung 4.1 soll dies veranschaulichen.

Wir wollen uns nun im Folgenden um die Größe, also die resource-bounded Maße eines unteren Spans und eines oberen Spans verschiedener Sprachen kümmern.

Lemma 4.36. *$P_m^{-1}(A)$ hat Maß 0, für $A \notin P$.*

Lemma 4.37. *Sei A eine Sprache, die unkomprimierbar durch $\leq_m^P$-Reduktionen ist.*

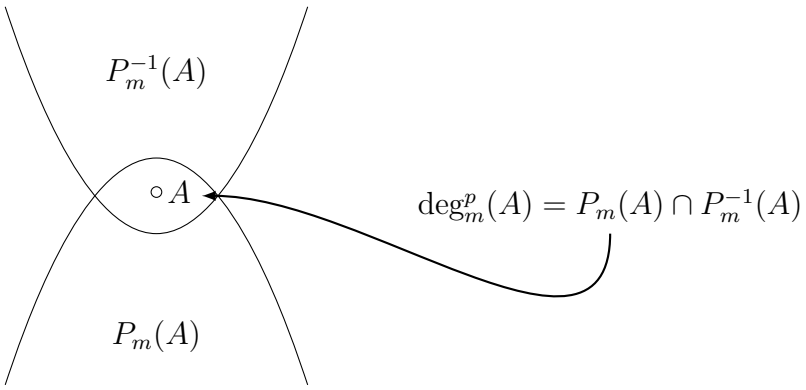


Abbildung 4.1: oberer Span, unterer Span und Grad von A

1. *Wenn $A \in E$, dann ist $\mu_{p_1}(P_m^{-1}(A)) = 0$ und damit auch $\mu(P_m^{-1}(A) \mid E) = 0$.*
2. *Wenn $A \in EXP$, dann ist $\mu_{p_2}(P_m^{-1}(A)) = 0$ und damit auch $\mu(P_m^{-1}(A) \mid EXP) = 0$.*

Dieses Lemma benötigen wir nun um das folgende Theorem zu beweisen.

Theorem 4.38 (Theorem vom kleinen $\leq_m^{\mathrm{P}}$-Span)**.** *1. Für jedes $A \in E$ gilt $\mu(P_m(A) \mid E) = 0$ oder $\mu_{p_1}(P_m^{-1}(A)) = \mu_{p_1}(P_m^{-1}(A) \mid E) = 0$.*

2. Für jedes $A \in EXP$ gilt $\mu(P_m(A) \mid EXP) = 0$ oder $\mu_{p_1}(P_m^{-1}(A)) = \mu_{p_2}(P_m^{-1}(A) \mid EXP) = 0$.

Das Theorem vom kleinen $\leq_m^{\mathrm{P}}$-Span sagt aus, dass wenn A in E oder EXP liegt, mindestens eine der Mengen $P_m(A)$ und $P_m^{-1}(A)$ klein ist, also nur kleine Spanmengen in E oder EXP auftauchen können. Abbildung 4.2 soll mögliche Konfigurationen mit kleinen Spannmengen veranschaulichen.

Beweis. Auch hier ist der Beweis für E und EXP jeweils analog, weshalb wir an dieser Stelle nur den Beweis für E angeben.
Sei nun also $A \in \mathrm{E}$ und sei weiterhin X die Menge aller Sprachen, die unkomprimierbar durch $\leq_m^{\mathrm{P}}$-Reduktionen sind. Wir betrachten nun zwei Fälle.

Fall 1: Wenn $P_m(A) \cap \mathrm{E} \cap X = \emptyset$ folgt mit Korollar 4.23 $\mu(P_m(A) \mid \mathrm{E}) = 0$.

Fall 2: Wenn $P_m(A) \cap \mathrm{E} \cap X \neq \emptyset$, dann halten wir eine Menge $B \in P_m(A) \cap \mathrm{E} \cap X$ fest. Da $B \in \mathrm{E} \cap X$, gilt $\mu_{p_1}(P_m^{-1}(B)) = \mu(P_m^{-1}(B) \mid \mathrm{E}) = 0$ nach Lemma 4.37. Es gilt $P_m^{-1}(A) \subseteq P_m^{-1}(B)$ und damit folgt auch $\mu_{p_1}(P_m^{-1}(A)) = \mu(P_m^{-1}(A) \mid \mathrm{E}) = 0$.

□

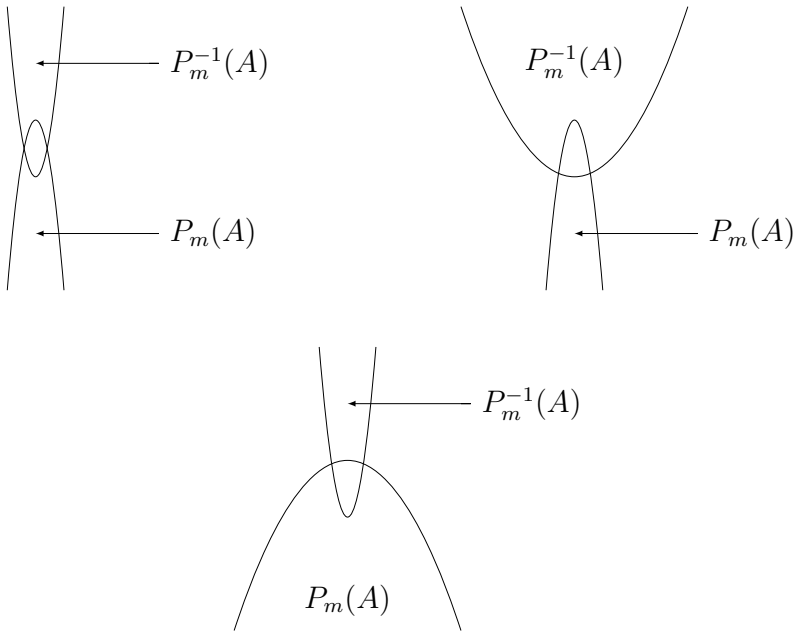


Abbildung 4.2: Mögliche Konfigurationen, bei denen kleine Spannmengen auftauchen

Mithilfe des Theorems vom kleinen $\leq_m^{\mathrm{P}}$-Span lässt sich also ebenfalls bemerken, dass es nur sehr wenige $\leq_m^{\mathrm{P}}$-schwere Probleme für E bzw. EXP geben kann.

Theorem 4.39. *Sei $\mathcal{H}_m(E)$ die Menge aller Sprachen, die $\leq_m^P$-schwer für E sind. Dann gilt $\mu_{p_1}(\mathcal{H}_m(E)) = 0$.*

Beweis. Sei A eine Sprache in E, die unkomprimierbar durch $\leq_m^{\mathrm{P}}$-Reduktionen ist. Dann sagt Lemma 4.37, dass $\mu_{p_1}(\mathcal{H}_m(\mathrm{E})) = \mu_{p_1}(P_m^{-1}(A)) = 0$. □

Um weitere Struktureigenschaften für E und EXP aus dem Theorem vom kleinen $\leq_m^{\mathrm{P}}$-Span zu folgern, benötigen wir im Folgenden die Definition des $\leq_m^{\mathrm{P}}$-Grads.

Definition 4.40. *Sei $A \subseteq \{0,1\}^*$ eine Sprache. Der $\leq_m^P$-Grad von A ist definiert als*

$$\deg_m^p(A) = P_m(A) \cap P_m^{-1}(A).$$

Da wir mit $P_m(A)$ die Menge der Sprachen identifizieren, die sich via $\leq_m^{\mathrm{P}}$ „unter“ A befinden und mit $P_m^{-1}(A)$, die Menge der Sprachen, die sich „über“ A befinden, entspricht $\deg_m^p(A)$ also der Menge der Sprachen, die sich auf „gleicher Ebene“ wie A via $\leq_m^{\mathrm{P}}$ befinden. Etwas formaler ausgedrückt, sind in $P_m(A)$ die Sprachen B mit $B \leq_m^{\mathrm{P}} A$ und in $P_m^{-1}(A)$ die Sprachen B mit $A \leq_m^{\mathrm{P}} B$ und damit also in $\deg_m^p$ die Sprachen mit $B \leq_m^{\mathrm{P}} A$ und $A \leq_m^{\mathrm{P}} B$, also $A \equiv_m^p B$. Abbildung 4.1 soll dies veranschaulichen.

Dieser Begriff ist daher also das Äquivalent zum Turinggrad auf $\leq_m^{\mathrm{P}}$-Reduktionen angepasst, welchen wir aus der Vorlesung Berechenbarkeit und Logik (siehe [Vol19]) für $\leq_T$ kennen.

Theorem 4.41. *Für alle Sprachen $A \subseteq \{0,1\}^*$ gilt*

$$\mu(\deg_m^p(A) \mid E) = \mu(\deg_m^p(A) \mid EXP) = 0.$$

Beweis. Dieses Resultat folgt direkt aus Theorem 4.38. □

Definition 4.42. *Wir bezeichnen mit $\mathcal{C}_m(C)$ die Menge der $\leq_m^P$-vollständigen Sprachen für eine Komplexitätsklasse C.*

Theorem 4.43 (Mayordomo)**.** *Es gilt $\mu(\mathcal{C}_m(E) \mid E) = \mu(\mathcal{C}_m(EXP) \mid EXP) = 0$.*

Beweis. Dieses Resultat folgt durch Anwendung von Theorem 4.41 und Theorem 4.39. □

Im Folgenden wollen wir uns nun schwache schwere Probleme von E und EXP ansehen. Ein Problem, das $\leq_m^{\mathrm{P}}$-schwer für E ist, ist provably intractable durch den Zeithierarchiesatz von Hartmanis und Stearns (siehe [MSV20]). Daher ist bekannt, dass solche Probleme sehr starke intractability-Eigenschaften haben.
Wir wollen nun die Klasse der provably intractable Probleme erweitern. Hierfür benötigen wir eine maßtheoretische Verallgemeinerung von $\leq_m^{\mathrm{P}}$-Schwere.

Definition 4.44. *Eine Sprache $A \subseteq \{0,1\}^*$ ist schwach $\leq_m^P$-schwer für E bzw. EXP, wenn $\mu(P_m(A) \mid E) \neq 0$ bzw. $\mu(P_m(A) \mid EXP) \neq 0$.*

Das heißt, eine Sprache A ist schwach $\leq_m^{\mathrm{P}}$-schwer für E, wenn es eine nicht vernachlässigbare Teilmenge von Sprachen in E gibt, die sich auf A $\leq_m^{\mathrm{P}}$-reduzieren lassen.

Bemerkung 4.45. *Jede Sprache, die $\leq_m^P$-schwer für E bzw. EXP ist, ist auch schwach $\leq_m^P$-schwer für E bzw. EXP.*

Für andere Reduktionsbegriffe, wie bspw. Turingreduktion ($\leq_T^P$) definieren wir schwache $\leq_T^P$-Schwere analog, in dem wir zunächst die Mengen $P_T(A)$ und $P_T^{-1}(A)$ analog zu Definition 4.35 definieren und anschließend den Begriff der schwachen $\leq_T^P$-Schwere analog zu Definition 4.44 definieren.

Bemerkung 4.46. *Wenn A schwach $\leq_T^P$-Schwer für E ist, gilt $A \notin P$.*

Dies lässt sich damit begründen, dass $\mu(\mathrm{P} \mid \mathrm{EXP}) = \mu(\mathrm{P} \mid \mathrm{E}) = 0$ ist. Gleichzeitig bedeutet dies auch, dass Sprachen, die schwach $\leq_m^{\mathrm{P}}$-schwer für E sind, intractable in einem stärkeren Sinn sind.

Theorem 4.47. *1. Jede Sprache, die $\leq_m^P$-schwer für E ist, hat einen dichten P-Komplexitätskern.*

2. Jede Sprache, die schwach $\leq_m^P$-schwer für E oder EXP ist, hat einen dichten exponentiellen Komplexitätskern.

Um die Begriffe schwach $\leq_m^{\mathrm{P}}$-schwer und $\leq_m^{\mathrm{P}}$-schwer zu unterscheiden, bietet sich der allgemeine Weg an, Sprachen zu suchen, die schwach $\leq_m^{\mathrm{P}}$-schwer aber nicht $\leq_m^{\mathrm{P}}$ -schwer sind. Dafür führen wir, wie üblich, zunächst einen Vollständigkeitsbegriff, bzw. in diesem Fall einen schwachen Vollständigkeitsbegriff, ein.

Definition 4.48. *Eine Sprache $A \subseteq \{0,1\}^*$ heißt schwach $\leq_m^P$-vollständig für E bzw. EXP, wenn A schwach $\leq_m^P$-schwer für E bzw. EXP und $A \in E$ bzw. $A \in EXP$ ist.*

Wir wollen nun die Notationen $\mathcal{C}_m(C)$ und $\mathcal{H}_m(C)$ auf schwache $\leq_m^{\mathrm{P}}$-Schwere erweitern.

Definition 4.49. *Sei C eine Komplexitätsklasse. Mit $\mathcal{WH}_m(C)$ bezeichnen wir die Menge der schwach $\leq_m^P$-schweren Probleme für C und mit $\mathcal{WC}_m(C)$ bezeichnen wir die Menge der Sprachen, die schwach $\leq_m^P$-vollständig für C sind.*

Wir wollen uns zunächst die Inklusionen dieser Mengen ansehen, bevor wir mit den Nichtinklusionen fortfahren, um darauf hin auf echte Teilmengenbeziehungen zu untersuchen.
Es ist bekannt, dass die $\leq_m^{\mathrm{P}}$-schweren Probleme für E dieselben sind wie für EXP. Damit gilt zunächst

$$\mathcal{H}_m(\mathrm{E}) = \mathcal{H}_m(\mathrm{EXP}).$$

Außerdem ist klar, da $\mathrm{EXP} = P_m(\mathrm{E})$, dass die $\leq_m^{\mathrm{P}}$-vollständigen Probleme für E die $\leq_m^{\mathrm{P}}$-vollständigen Probleme von EXP sind, welche in E liegen, also

$$\mathcal{C}_m(\mathrm{E}) = \mathrm{E} \cap \mathcal{C}_m(\mathrm{EXP}).$$

Bemerkung 4.45 liefert uns

$$\mathcal{H}_m(\mathrm{E}) \subseteq \mathcal{WH}_m(\mathrm{E}) \text{ und } \mathcal{H}_m(\mathrm{EXP}) \subseteq \mathcal{WH}_m(\mathrm{EXP}).$$

Lemma 4.50. *Sei X eine Menge von Sprachen.*

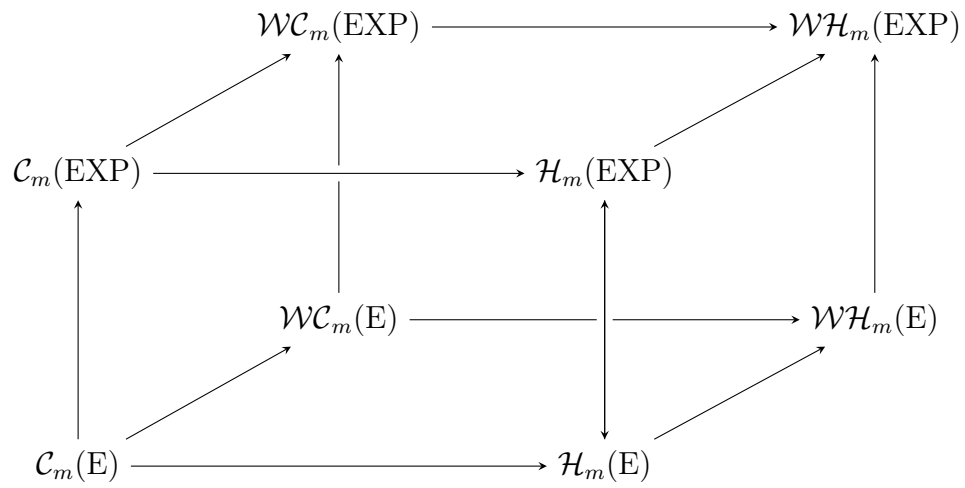


Abbildung 4.3: Inklusionsstruktur der betrachteten (schwachen) Schwere-Klassen

1. *Wenn* $\mu_{p_2}(P_m(X)) = 0$, *dann ist* $\mu_{p_1}(X) = 0$.
2. *Wenn* $\mu(P_m(X) \mid EXP) = 0$, *dann ist* $\mu(X \mid E) = 0$.

Dieses führt uns zu folgendem Theorem:

Theorem 4.51. $\mathcal{WH}_m(E) \subseteq \mathcal{WH}_m(EXP)$

Beweis. Sei $H \in \mathcal{WH}_m(\mathrm{E})$. Dann gilt $\mu(P_m(H) \mid \mathrm{E}) \neq 0$, nach Lemma 4.50, wenn wir für X $P_m(H)$ wählen. Denn $\mu(P_m(H) \mid \mathrm{EXP}) = \mu(P_m(P_m(H)) \mid \mathrm{EXP}) \neq 0$ und damit gilt $H \in \mathcal{WH}_m(\mathrm{EXP})$. □

Abbildung 4.3 fasst die Inklusionsstruktur zusammen.

Im Folgenden wollen wir uns die Nichtinklusionen genauer anschauen. Dazu entwickelte Lutz in [Lut95] die sogenannte Martingaldiagonalisierungstechnik und benutzte sie um folgendes Theorem zu beweisen.

Theorem 4.52. $\mathcal{C}_m(E) \subsetneq \mathcal{WC}_m(E)$

Korollar 4.53. $\mathcal{C}_m(EXP) \subsetneq \mathcal{WC}_m(EXP)$

Beweis. $\mathrm{E} \cap \mathcal{C}_m(\mathrm{EXP}) = \mathcal{C}_m(\mathrm{E}) \subsetneq \mathcal{WC}_m(\mathrm{E}) \subseteq \mathcal{WC}_m(\mathrm{EXP})$ □

Theorem 4.52 ist insofern interessant, als dass es besagt, dass die Klasse der schwachen $\leq_m^{\mathrm{P}}$-schweren Probleme für E eine strikt größere Klasse Probleme, welche beweisbar strongly intractable sind, ist, als die Klasse der $\leq_m^{\mathrm{P}}$-schweren Probleme für E.

Theorem 4.54. $\mu_{p_1}(\mathcal{WH}_m(E)) = 1$

Korollar 4.55. $\mu_{p_2}(\mathcal{WH}_m(EXP)) = 1$

Insgesamt bedeutet das also, dass fast jede Sprache in E schwach $\leq_m^{\mathrm{P}}$-vollständig ist, aber nicht $\leq_m^{\mathrm{P}}$-vollständig für E.

Theorem 4.56. $E \cap \mathcal{WC}_m(EXP) \not\subseteq \mathcal{WC}_m(E)$

Abschließend wollen wir uns nun obere Schranken für schwere Probleme anschauen.
Wir haben in Theorem 4.39 gesehen, dass $\leq_m^{\mathrm{P}}$-schwere Probleme für E ziemlich selten sind. In diesem Teil der Arbeit werden wir nun sehen, dass dies an einer nichttrivialen oberen Schranke bezogen auf die Größe der Komplexitätskerne solcher Sprachen liegt.

Theorem 4.57. *Für jede $\leq_m^P$-schwere Sprache H für E existieren $B, D \in TIME(2^{4n})$, so dass D dicht ist und $B = H \cap D$.*

Dies bedeutet, dass jede $\leq_m^{\mathrm{P}}$-schwere Sprache für E in Zeit $\mathcal{O}(2^{4n})$ auf einer dichten Menge von Instanzen entschieden werden kann, welche wiederum selbst in Zeit $\mathcal{O}(2^{4n})$ entschieden werden kann.

Theorem 4.58. *Jeder $TIME(2^{4n})$-Komplexitätskern jeder $\leq_m^P$-schweren Sprache für E hat ein dichtes Komplement.*

Nach Korollar 4.34 hat jede Sprache in E $\{0,1\}^*$ als $\mathrm{TIME}(2^{4n})$-Komplexitätskern. Theorem 4.58 sagt nun aus, dass $\leq_m^{\mathrm{P}}$-schwere Sprachen für E unnormal einfach sind in dem Sinne, dass sie unnormal kleine Komplexitätskerne für alle Sprachen in E enthalten. Dies erklärt zudem intuitiv auch Theorem 4.39 und Theorem 4.43.

Mithilfe der Zufallsbegriffe der Berechenbarkeitstheorie, bzw. mithilfe der Konstruktionen aus denen sie resultieren, lassen sich auch Aussagen auf die Komplexitätstheorie verallgemeinern, indem man sich die Maßtheorie und Martingale zunutze macht.

5 Ausblick

Wir haben in dieser Arbeit gesehen, dass wir drei äquivalente Definitionen desselben Zufälligkeitsbegriff vorliegen haben. Nun stellt sich die Frage, was man damit nun noch untersuchen kann.

Ein mögliches Gebiet, welches ähnliche Konzepte verwendet, haben wir bereits in Kapitel 4 gesehen. Dort haben wir einen kleinen Einblick gegeben, wie man Komplexitätstheorie mit Maßtheorie oder Martingalen betreiben kann. Wir haben uns nur Aussagen über E und EXP angesehen, allerdings kann man das auch auf andere Klassen der Hierarchie oder andere Klassen, wie NP oder P/POLY anwenden.

Ein weiteres angrenzendes Gebiet ist das Gebiet der algorithmischen Informationstheorie. Dabei verbindet man die vorgestellte Theorie mit der Informationstheorie, welche man von Shannon kennt und kommt dabei unter anderem mit der Chaitin Konstante in Berührung, von der man dann zeigen kann, dass sie zufällig ist.

Abbildungsverzeichnis

Literaturverzeichnis

[AEE08] Amann, Herbert ; Escher, Joachim ; Escher, Joachim: *Analysis III* -. 2. Aufl. Berlin Heidelberg New York : Springer-Verlag, 2008. – ISBN 978–3–764–38884–3

[AZ14] Aigner, Martin ; Ziegler, Günter M.: *Das BUCH der Beweise*. Springer Spektrum, 2014. – ISBN 3662444569

[Gau18] Gaube, S.: *Kolmogorov Komplexität und Datenkompression*, Leibniz Universität Hannover, Bachelorarbeit, 2018

[Grü14] Grübel, Rudolf: *Stochastik*. Vorlesungsskript, 2014

[HB76] Hartmanis, J. ; Berman, L.: On Isomorphisms and Density of NP and Other Complete Sets. In: *SIAM J. Comput.* 6 (1976), 01, S. 30–40. http://dx.doi.org/10.1145/800113.803628. – DOI 10.1145/800113.803628

[Lut89] Lutz, J.H.: Almost everywhere high nonuniform complexity. (1989), 07, S. 37 – 53. http://dx.doi.org/10.1109/SCT.1989.41813. – DOI 10.1109/SCT.1989.41813. ISBN 0–8186–1958–9

[Lut95] Lutz, Jack H.: Weakly hard problems. In: *SIAM Journal on Computing , 24* (1995)

[Lut96] Lutz, Jack H.: The Quantitative Structure of Exponential Time. TR96-08 (1996). https://lib.dr.iastate.edu/cs_techreports/115/

[LV08] Li, Ming ; Vitányi, Paul M.: *An Introduction to Kolmogorov Complexity and Its Applications (Texts in Computer Science)*. Springer, 2008. – ISBN 0387339981

[Mey76] Meyer, A. R.: (1976). – Publiziert in [HB76]

[ML66] Martin-Löf, Per: The definition of random sequences. (1966). https://doi.org/10.1016/S0019-9958(66)80018-9

[MSV20] Meier, Arne ; Schöning, Uwe ; Vollmer, Heribert: *Komplexität von Algorithmen - Mathematik für Anwendungen Band 4*. 2. Auflage. Berlin : Lehmanns Media, 2020. – ISBN 978–3–96543–137–9

[Sch71] Schnorr, Claus P.: A unified approach to the definition of random sequences. (1971). https://doi.org/10.1007/BF01694181

[SCR02] Sivakumar, D. ; Cai, J. ; Regan, K. W.: Pseudorandom Generators, Measure Theory, and Natural Proofs. In: *Annual Symposium on Foundations of Computer Science - Proceedings* (2002), 03

[Vol19] Vollmer, Heribert: *Berechenbarkeit und Logik*. Vorlesungsskript, 2019