



TECHNISCHE  
UNIVERSITÄT  
WIEN

DIPLOMARBEIT

# **Boltzmann-Sampling von kombinatorischen Strukturen**

ausgeführt am Institut für

Diskrete Mathematik und Geometrie  
der Technischen Universität Wien

unter der Anleitung von

Ao.Univ.Prof. Dipl.-Ing. Dr.techn. Bernhard Gittenberger

durch

Carina Ringelhahn, BSc

Untere Bergfeldgasse 1  
3385 Prinzersdorf

## Vorwort

Diese Arbeit behandelt das Problem der effizienten Erzeugung von Objekten kombinatorischer Strukturen. Diese Generatoren, sogenannte Boltzmann-Sampler, erzeugen Objekte einer Menge, die durch bestimmte Konstruktionen beschrieben werden. Die Basis-Konstruktionen, mithilfe deren man diese kombinatorischen Strukturen beschreiben kann und weitere Grundlagen, werden in Kapitel 2 behandelt.

Der Grundidee der Boltzmann-Sampler basiert auf einem Verfahren, das 1993 von Flajolet, Zimmermann und Van Cutsem in [22] eingeführt wurde. Diese rekursive Methode wird in Kapitel 3 beschrieben. Das Problem hierbei ist eine Vorberechnung der Anzahlen an kombinatorischen Objekten einer bestimmten Größen von  $\mathcal{O}(n^2)$ .

2003 wurde von Duchon, Flajolet, Louchard und Schaeffer in [16] die Methode namens Boltzmann-Sampling eingeführt. Dieser ist weit aus effizienter als die rekursive Methode, es werden dabei jedoch Einschränkungen bei der Genauigkeit der gewünschten Objektgröße gemacht. Je nach dem ob man Genauigkeitseinschränkungen toleriert oder nicht, nennt man die Boltzmann-Sampler entweder *Approximate-Size-* oder *Exact-Size-Sampler*.

Eine andere Frage tritt auf, bei der Wahl des Boltzmann-Parameter. Dies wird ebenfalls in Kapitel 4 behandelt. In Kapitel 5 werden zwei Anwendungen der Boltzmann-Sampler vorgestellt. Das erste behandelt die Zufallserzeugung von Wörtern linearer Sprachen mit linearem Mischen und letzteres die Erzeugung von zufälligen apollonischen Netzwerkstrukturen.

# Inhaltsverzeichnis

1	Einleitung	5
2	Grundlagen	6
2.1	Unmarkierte kombinatorische Strukturen	8
2.1.1	Disjunkte Vereinigung: $A + B$	8
2.1.2	Kartesisches Produkt: $A \times B$	8
2.1.3	Folge: $\text{SEQ}(A)$	9
2.1.4	Menge: $\text{SET}(A)$	9
2.1.5	Multimenge: $\text{MSET}(A)$	10
2.2	Markierte kombinatorische Strukturen	11
2.2.1	Disjunkte Vereinigung: $A + B$	11
2.2.2	Produkt: $A \star B$	11
2.2.3	Folge: $\text{SEQ}(A)$	12
2.2.4	Menge: $\text{SET}(A)$	12
2.2.5	Zyklus: $\text{CYC}(A)$	13
2.2.6	Auszeichnen: $\Theta A$	13
2.3	Standardbeschreibung	14
3	Rekursive Methode	19
3.1	Der Algorithmus	19
3.2	Der Syntaxbaum	22
3.3	Laufzeitanalyse	22
4	Boltzmann-Sampler	25
4.1	Wahl des Verteilungsparameter	26
4.2	Auswertung der erzeugenden Funktion	41
4.2.1	Iterative Auswertung	42
4.2.2	Newton-Methode	47
4.3	Gewöhnliche Boltzmann-Sampler	51
4.4	Exponentielle Boltzmann-Sampler	57

## *Inhaltsverzeichnis*

4.5	Toleranzbereich . . . . .	61
4.5.1	Approximate-Size-Sampler . . . . .	61
4.5.2	Exact-Size-Sampler . . . . .	63
4.5.3	Singuläre Sampler . . . . .	63
5	Anwendungen der Boltzmann-Sampler	65
5.1	Reguläre Sprachen mit linearer Mischung . . . . .	65
5.2	Apollonische Netzwerkstrukturen . . . . .	72
A	Anhang	76
A.1	Landau-Symbole . . . . .	76
A.2	Zufallsgeneratoren . . . . .	77
A.2.1	Inversionsmethode . . . . .	77
A.2.2	Bernoulli-Verteilung . . . . .	78
A.2.3	Geometrische Verteilung . . . . .	78
A.2.4	Poissonverteilung . . . . .	79
A.2.5	Logarithmische Verteilung . . . . .	80
A.2.6	Komplexitätsanalyse . . . . .	80
A.3	Laplace-Transformation und Faltung . . . . .	80
	Literaturverzeichnis	82

# 1 Einleitung

Viele kombinatorische Strukturen lassen sich durch wenige einfache Basisoperatoren beschreiben. Diese Strukturen nennen wir hier zerlegbar und die Basiskonstruktionen, mithilfe deren die Strukturen beschrieben werden, werden in Kapitel 2 vorgestellt. Aus dieser oft unendlichen Menge an Objekten von zerlegbaren Strukturen ist oft ein zufälliges Objekt einer bestimmten Größe gefragt. Die rekursive Methode, die 1993 eingeführt wurde und in Kapitel 3 erklärt wird, verwendet dabei erzeugende Funktionen um einen Algorithmus auf eine Standardbeschreibung der Struktur anzuwenden. Wie man Beschreibungen in eine Standardbeschreibung überführt wird am Ende von Kapitel 2 vorgestellt. In [22] wurden zwei Möglichkeiten der rekursiven Methode erläutert: der sequentielle und der boustrophedone Algorithmus. Der boustrophedone Algorithmus verwendet einen etwas effizienteren Suchalgorithmus. Beide Methoden benötigen jedoch die Anzahlen von Objekten einer bestimmten Größe, d.h. die Folge der Koeffizienten der erzeugenden Funktion bis hin zur gewünschten Größe.

Beim Boltzmann-Sampler wird diese teure Vorberechnung nicht benötigt. Hier hat man eine Verteilung auf der Menge aller Objekte, die mittels eines Parameters verändert werden kann. Dabei bleiben aber immer Objekte ein und derselben Größe immer gleichwahrscheinlich. Jedoch wird die gewünschte Größe oft nur annähernd getroffen. Wie man den Boltzmann-Parameter gut wählt und inwiefern dies von der Art der Struktur abhängt, wird in Kapitel 4 behandelt. Für diesen bestimmten Parameter muss die erzeugende Funktion ausgewertet werden. Dies kann iterativ oder mit der Newton-Methode realisiert werden.

Da es zwei Arten von Klassen von kombinatorischen Strukturen gibt, nämlich die markierten und unmarkierten, müssen für die Basiskonstruktionen dieser Strukturen die grundlegenden Algorithmen für beide Klassen von Strukturen vorgestellt werden. Dies passiert in Kapitel 4.

Am Ende von Kapitel 4 werden noch die Unterschiede und Laufzeiten bei der Wahl des Toleranzbereichs angeführt.

In Kapitel 5 werden zwei konkrete Probleme aus der Informatik vorgestellt und wie diese mithilfe eines Boltzmann-Samplers gelöst werden können.

## 2 Grundlagen

Spricht man von kombinatorischen Strukturen, benötigt man zuerst eine formale Definition davon. Meist werden diese durch gewisse Basiskonstruktionen aufgebaut, die in diesem Kapitel vorgestellt werden. Weiters wird auf eine Methode eingegangen um sogenannte Beschreibungen von zerlegbaren Strukturen zu vereinheitlichen.

**Definition 2.1.** Sei  $A$  eine höchstens abzählbare Menge und  $w : A \rightarrow \mathbb{N}$  eine Funktion, für die gilt, dass das Urbild jeder einelementigen Menge endlich ist. Dann nennt man das Paar  $(A, w)$  **kombinatorische Struktur**. Die Elemente von  $A$  bezeichnet man als **kombinatorische Objekte**.

Sei  $a \in A$  ein kombinatorisches Objekt. Dann wird  $w(a)$  als die **Größe** von  $a$  bezeichnet. Einfachheitshalber schreibt man oft auch  $|a|$ .

Ein kombinatorisches Objekt der Größe 1 nennt man **Atom**. Hier werden Atome mit  $\square$  bezeichnet.

Ein kombinatorisches Objekt der Größe 0 nennt man **neutrales Objekt** und bezeichnet man mit  $\epsilon$ .

**Definition 2.2.** Eine Reihe der Gestalt  $f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n$  heißt eine **Potenzreihe** mit Anschlußstelle  $z_0$ .

Es gilt: Es gibt eine Zahl  $R$ , der **Konvergenzradius**, mit  $0 \leq R \leq \infty$ , sodass für  $|z - z_0| < R$  die Reihe konvergiert und für  $|z - z_0| > R$  die Reihe divergiert.

Über das Verhalten der Reihe auf dem Konvergenzkreis kann jedoch keine allgemeine Aussage getroffen werden.

**Definition 2.3.** Unter der **erzeugenden Funktion** einer kombinatorischen Struktur  $(A, w)$  versteht man die formale Potenzreihe

$$A(x) = \sum_{a \in A} x^{w(a)} = \sum_{n=0}^{\infty} A_n x^n,$$

wobei  $A_n$  definiert ist als  $A_n := |w^{-1}(\{n\})|$ .

Unter der **exponentiellen erzeugenden Funktion** einer kombinatorischen Struktur  $(A, w)$  versteht man die formale Potenzreihe

$$A(x) = \sum_{a \in A} \frac{x^{w(a)}}{w(a)!} = \sum_{n=0}^{\infty} A_n \frac{x^n}{n!}.$$

## 2 Grundlagen

Ist man an mehreren Attributen der Objekte interessiert verwendet man sogenannt multivariate erzeugende Funktionen. Im Falle von zwei Attributen nennt man sie bivariate erzeugende Funktionen.

**Definition 2.4.** *Unter der **bivariaten erzeugenden Funktion** der doppelt indizierten Folge  $(A_{nk})_{n,k \geq 0}$  versteht man die formale Potenzreihe*

$$A(x, u) = \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} A_{nk} x^n u^k.$$

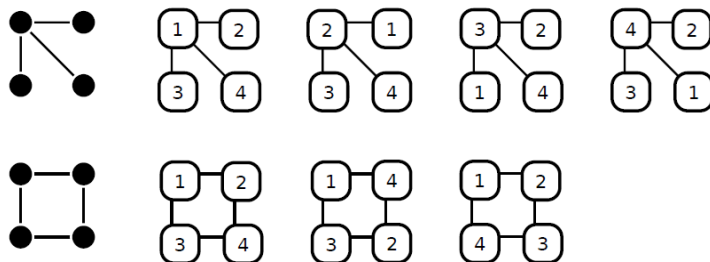
*Unter der **bivariaten exponentiellen erzeugenden Funktion** der doppelt indizierten Folge  $(A_{nk})_{n,k \geq 0}$  versteht man die formale Potenzreihe*

$$A(x, u) = \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} A_{nk} \frac{x^n}{n!} u^k.$$

Erzeugende Funktionen sind ein wichtiges Werkzeug um kombinatorische Probleme systematischer und eleganter zu lösen. Sie sind zudem in verschiedenen anderen Gebieten der Stochastik nützlich.

In weiterer Folge werden solche kombinatorische Strukturen betrachtet, die sich durch gewisse Konstruktionen aus Atomen beschreiben lassen.

Man unterscheidet bei kombinatorischen Strukturen zwischen **markierten** und **unmarkierten** Strukturen. Bei unmarkierten kombinatorischen Strukturen unterscheidet man zwischen den einzelnen darin vorkommenden Atomen nicht. Bei markierten Objekten der Größe  $n$  ist jedes Atom mit einem eindeutigen Wert aus  $\{1, \dots, n\}$  gekennzeichnet.



**Abbildung 2.1:** Unmarkierte Objekte mit allen entsprechenden markierten Objekten

Im folgenden Abschnitt werden einige der angesprochenen Konstruktionen von kombinatorischen Strukturen erläutert. Dabei muss zwischen den markierten und unmarkierten Strukturen unterschieden werden.

## 2.1 Unmarkierte kombinatorische Strukturen

Bei unmarkierten kombinatorischen Strukturen wird die gewöhnliche erzeugende Funktion verwendet. Angenommen, es seien  $(A, w_1)$  und  $(B, w_2)$  unmarkierte kombinatorische Strukturen. Dann können folgende kombinatorischen Strukturen daraus konstruiert werden:

### 2.1.1 Disjunkte Vereinigung: $A + B$

Seien  $A$  und  $B$  disjunkt. Dann ist die disjunkte Vereinigung  $A + B$  definiert als kombinatorische Struktur  $(C, w)$ , mit  $C := A \cup B$ ,  $w(c) = w_1(c)$ , wenn  $c \in A$ , und  $w(c) = w_2(c)$ , wenn  $c \in B$ . Da die Mengen disjunkt sein müssen, ist die Größenfunktion wohldefiniert. Für die erzeugende Funktion gilt nun:

**Satz 2.1.** *Seien  $(A, w_1)$  und  $(B, w_2)$  unmarkierte kombinatorische Strukturen, wobei  $A$  und  $B$  disjunkt sind. Die erzeugende Funktion von  $A \cup B$  ist die Summe der erzeugenden Funktionen von  $A$  und  $B$ .*

*Beweis.*

$$C(x) = \sum_{c \in A+B} x^{w(c)} = \sum_{a \in A} x^{w(a)} + \sum_{b \in B} x^{w(b)} = A(x) + B(x) \quad \square$$

### 2.1.2 Kartesisches Produkt: $A \times B$

Das Produkt zweier kombinatorischer Strukturen  $A \times B$  ist definiert als kartesisches Produkt von  $A$  und  $B$  gemeinsam mit der Funktion  $w((a, b)) = w(a) + w(b)$ . Die erzeugende Funktion ergibt sich nun aus:

**Satz 2.2.** *Seien  $(A, w_1)$  und  $(B, w_2)$  unmarkierte kombinatorische Strukturen. Die erzeugende Funktion von  $A \times B$  ist das Produkt der erzeugenden Funktionen von  $A$  und  $B$ .*

*Beweis.*

$$C(x) = \sum_{c \in A \times B} x^{w(c)} = \sum_{a \in A} \sum_{b \in B} x^{w(a)+w(b)} = \left( \sum_{a \in A} x^{w(a)} \right) \left( \sum_{b \in B} x^{w(b)} \right) = A(x)B(x) \quad \square$$



## 2 Grundlagen

### 2.1.3 Folge: SEQ(A)

Die Konstruktion der Folgen (Sequenzen) erzeugt aus einer kombinatorischen Struktur durch Anwendung von Vereinigung und Produkt eine Struktur der Form

$$\text{SEQ}(A) = \{\epsilon\} + A + A \times A + A \times A \times A + \dots$$

Um die disjunkten Summen bilden zu können, sei vorausgesetzt, dass die vorkommenden Summanden paarweise disjunkt sind und dass  $A$  kein neutrales Objekt enthält. Sei  $a = (a_1, a_2, \dots, a_n) \in \text{SEQ}(A)$ . Die Größe von  $a$  berechnet sich nun durch  $w(a) = \sum_{i=1}^n w_1(a_i)$ .

**Satz 2.3.** Sei  $(A, w)$  eine unmarkierte kombinatorische Struktur, wobei  $A, A \times A, \dots$  paarweise disjunkt sind und  $A$  kein neutrales Objekt enthält, und sei  $A(x)$  die erzeugende Funktion von  $A$  und  $C(x)$  die erzeugende Funktion von  $\text{SEQ}(A)$ . Dann gilt:  $C(x) = \frac{1}{1-A(x)}$

*Beweis.* Die Aussage folgt sofort aus dem Aufbau von  $\text{SEQ}(A)$  durch Summe und Produkt und mit den Sätzen 2.1 und 2.2:

$$C(x) = 1 + A(x) + A(x)^2 + A(x)^3 + \dots = \frac{1}{1-A(x)} \quad \square$$

### 2.1.4 Menge: SET(A)

Die Konstruktion der kombinatorischen Struktur der Menge, bezeichnet mit  $\text{SET}(A)$ , ist die Menge aller endlichen Teilmengen von  $A$ . Die Größe eines Objektes der Struktur  $\text{SET}(A)$  ist die Summe aller Größen (bzgl  $w_1$ ) der Elemente des Objekts. Die Struktur lässt sich durch Konstruktion mit Vereinigung und Produkt darstellen:

$$\text{SET}(A) = (\{\epsilon\} + \{a_1\}) \times (\{\epsilon\} + \{a_2\}) \times \dots,$$

wobei  $A = \{a_1, a_2, \dots\}$ .

Für die erzeugende Funktion gilt:

**Satz 2.4.** Sei  $(A, w)$  eine unmarkierte kombinatorische Struktur und sei  $A(x)$  die erzeugende Funktion von  $A$  und  $C(x)$  die erzeugende Funktion von  $\text{SET}(A)$ . Dann gilt:

$$C(x) = \exp \left( A(x) - \frac{A(x^2)}{2} + \frac{A(x^3)}{3} - \dots \right)$$

## 2 Grundlagen

*Beweis.* Die Aussage folgt aus dem Aufbau von  $\text{SET}(A)$  und den bisher bewiesenen Sätzen 2.1 und 2.2.:

$$\begin{aligned}
 C(x) &= \prod_{i=1}^{\infty} (1 + x^{w(a_i)}) = \prod_{N \geq 0} (1 + x^N)^{A_N} = \exp \left( \sum_{N \geq 0} A_N \log(1 + x^N) \right) = \\
 &= \exp \left( - \sum_{N \geq 0} A_N \sum_{k \geq 1} (-1)^k \frac{x^{Nk}}{k} \right) = \exp \left( - \sum_{k \geq 1} (-1)^k \frac{A(x^k)}{k} \right) = \\
 &= \exp \left( A(x) - \frac{A(x^2)}{2} + \frac{A(x^3)}{3} - \dots \right) \quad \square
 \end{aligned}$$

### 2.1.5 Multimenge: $\text{MSET}(A)$

Die Multimenge einer Struktur  $(A, w_1)$  ist die kombinatorische Struktur, die alle endlichen Teilmengen von  $A$ , mit erlaubter Wiederholung, enthält. Die Struktur lässt sich durch die bisher behandelten Konstruktionen darstellen:

$$\text{MSET}(A) = \text{SEQ}(a_1) \times \text{SEQ}(a_2) \times \dots,$$

wobei  $A = \{a_1, a_2, \dots\}$ .

Die Größe eines Objekts ergibt sich klarerweise wieder aus der Summe der Größen der einzelnen Elemente (entsprechend den Vielfachheiten gezählt), d.h. sei  $c = \{c_1^{j_1}, c_2^{j_2}, \dots\}$  ein Objekt der Struktur  $\text{MSET}(A)$ , dann gilt:  $w(c) = \sum_{i=1}^{\infty} j_i w(c_i)$ . Für die erzeugende Funktion gilt:

**Satz 2.5.** Sei  $(A, w)$  eine kombinatorische Struktur und sei  $A(x)$  die erzeugende Funktion von  $A$  und  $C(x)$  die erzeugende Funktion von  $\text{MSET}(A)$ . Dann gilt:

$$C(x) = \exp \left( A(x) + \frac{A(x^2)}{2} + \frac{A(x^3)}{3} + \dots \right)$$

*Beweis.* Durch den Aufbau von  $\text{MSET}(A)$  und den Sätzen 2.2 und 2.3 ergeben sich folgende Gleichungen:

$$\begin{aligned}
 C(x) &= \prod_{a \in A} \frac{1}{(1 - x^{w(a)})} = \prod_{N \geq 0} \frac{1}{(1 - x^N)^{A_N}} = \exp \left( \sum_{N \geq 0} A_N \log \frac{1}{1 - x^N} \right) = \\
 &= \exp \left( \sum_{N \geq 0} A_N \sum_{k \geq 1} \frac{x^{Nk}}{k} \right) = \exp \left( \sum_{k \geq 1} \frac{A(x^k)}{k} \right) = \exp \left( A(x) + \frac{A(x^2)}{2} + \frac{A(x^3)}{3} + \dots \right) \quad \square
 \end{aligned}$$

## 2.2 Markierte kombinatorische Strukturen

Bei markierten kombinatorischen Strukturen werden exponentielle erzeugende Funktionen verwendet. Es gibt wieder eine Reihe von Konstruktionen, die aus gegebenen markierten Strukturen andere markierte Strukturen erzeugen. Dabei ist zu beachten, dass die Markierungen entsprechend angepasst werden. Hier werden die für das Thema dieser Arbeit benötigten Konstruktionen behandelt, dazu wird angenommen, es seien  $(A, w_1)$  und  $(B, w_2)$  markierte kombinatorische Strukturen.

### 2.2.1 Disjunkte Vereinigung: $A + B$

Angenommen  $A$  und  $B$  sind disjunkt. Wie bei den unmarkierten kombinatorischen Strukturen ist die Summe von markierten kombinatorischen Strukturen, eine kombinatorische Struktur  $(C, w)$  mit  $C := A \cup B$  und die Größenfunktion bleibt unverändert. Für die exponentielle erzeugende Funktion gilt:

**Satz 2.6.** *Seien  $(A, w_1)$  und  $(B, w_2)$  markierte kombinatorische Strukturen, wobei  $A$  und  $B$  disjunkt sind. Die erzeugende Funktion von  $A + B$  ist die Summe der erzeugenden Funktionen von  $A$  und  $B$ .*

*Beweis.*

$$C(x) = \sum_{c \in A \cup B} \frac{x^{w(c)}}{w(c)!} = \sum_{a \in A} \frac{x^{w(a)}}{w(a)!} + \sum_{b \in B} \frac{x^{w(b)}}{w(b)!} = A(x) + B(x) \quad \square$$

### 2.2.2 Produkt: $A \star B$

Bei dieser Konstruktion wird aus zwei markierten kombinatorischen Strukturen eine kombinatorische Struktur erzeugt, die als Menge die geordneten Paare von  $A$  und  $B$  enthält. Die Markierungen der einzelnen Objekte werden nun folgendermaßen festgelegt:

Sei  $(a, b) \in A \times B$ . Dann werden die Marken  $\{1, 2, \dots, |a| + |b|\}$  auf dieses Objekt vergeben.

Man wählt zuerst  $|a|$  Elemente aus dieser Menge aus. Dies kann auf  $\binom{|a|+|b|}{|a|}$  verschiedene Arten geschehen.

Diese  $|a|$  Marken werden ordnungserhaltend auf  $a$  verteilt.

Nun bleiben  $|b|$  Marken übrig, die ebenfalls ordnungserhaltend auf  $b$  verteilt werden.

Auf diese Weise entstehen alle Objekte, die im Partitionsprodukt liegen.

Für die exponentielle erzeugende Funktion gilt nun:

**Satz 2.7.** *Seien  $(A, w_1)$  und  $(B, w_2)$  markierte kombinatorische Strukturen. Die exponentielle erzeugende Funktion von  $A \times B$  ist das Produkt der exponentiellen erzeugenden Funktionen von  $A$  und  $B$ .*

## 2 Grundlagen

*Beweis.*

$$\begin{aligned}
 C(x) &= \sum_{c \in A \times B} \frac{x^{w(c)}}{w(c)!} = \sum_{a \in A} \sum_{b \in B} \binom{w(a) + w(b)}{w(a)} \frac{x^{w(a) + w(b)}}{(w(a) + w(b))!} = \\
 &= \sum_{a \in A} \sum_{b \in B} \frac{(w(a) + w(b))!}{w(a)! \cdot w(b)!} \frac{x^{w(a)} \cdot x^{w(b)}}{(w(a) + w(b))!} = \left( \sum_{a \in A} x^{w(a)} \right) \left( \sum_{b \in B} x^{w(b)} \right) = A(x)B(x) \quad \square
 \end{aligned}$$

### 2.2.3 Folge: SEQ(A)

Die Folge von markierten kombinatorischen Strukturen unterscheidet sich von den Folgen der unmarkierten dadurch, dass statt dem kartesischen Produkt das Partitionsprodukt verwendet wird:

$$\text{SEQ}(A) = \{\epsilon\} + A + A \star A + A \star A \star A + \dots$$

Deshalb gilt für die exponentielle erzeugende Funktion:

**Satz 2.8.** Sei  $(A, w)$  eine markierte kombinatorische Struktur, wobei  $A, A \star A, \dots$  paarweise disjunkt sind und  $A$  kein neutrales Objekt enthält, und sei  $A(x)$  die erzeugende Funktion von  $A$  und  $C(x)$  die exponentielle erzeugende Funktion von  $\text{SEQ}(A)$ . Dann gilt:  $C(x) = \frac{1}{1 - A(x)}$

*Beweis.* Die Aussage folgt wieder aus dem Aufbau von  $\text{SEQ}(A)$  durch Summe und Partitionsprodukt und mit den Sätzen 2.6 und 2.7:

$$C(x) = 1 + A(x) + A(x)^2 + A(x)^3 + \dots = \frac{1}{1 - A(x)} \quad \square$$

### 2.2.4 Menge: SET(A)

Bei den Mengen von markierten Strukturen wird wieder vorausgesetzt, dass  $A, A \star A, \dots$  paarweise disjunkt sind. Dann hat die Menge der kombinatorische Struktur  $A$  folgende Form:

$$\text{SET}(A) = \{\epsilon\} + A + \frac{1}{2!}A \star A + \frac{1}{3!}A \star A \star A + \dots$$

Dies lässt sich dadurch erklären, dass bei der Menge, im Gegensatz zur Folge, Objekte, die sich nur durch die Reihenfolge der Elemente unterscheiden, als gleich gelten.

**Satz 2.9.** Sei  $(A, w)$  eine markierte kombinatorische Struktur und sei  $A(x)$  die exponentielle erzeugende Funktion von  $A$  und  $C(x)$  die exponentielle erzeugende Funktion von  $\text{SET}(A)$ . Dann gilt  $C(x) = \exp(A(x))$ .

## 2 Grundlagen

*Beweis.* Aus dem Aufbau von  $\text{SET}(A)$  und den bisher bewiesenen Sätzen 2.6 und 2.7. lässt sich die Aussage sofort beweisen:

$$C(x) = 1 + \sum_{k=1}^{\infty} \frac{A^k(x)}{k!} = \sum_{k=0}^{\infty} \frac{A^k(x)}{k!} = \exp(A(x)) \quad \square$$

### 2.2.5 Zyklus: $\text{CYC}(A)$

Bei den Zyklen werden Folgen von Objekten aus  $A$  nicht unterschieden, wenn diese durch zyklische Verschiebung auseinander hervorgehen. Deshalb müssen wieder  $A, A \star A, \dots$  paarweise disjunkt sein. Dann haben die Zyklen folgende Form:

$$\text{CYC}(A) = A + \frac{1}{2}A \star A + \frac{1}{3}A \star A \star A + \dots$$

**Satz 2.10.** Sei  $(A, w)$  eine markierte kombinatorische Struktur und sei  $A(x)$  die exponentielle erzeugende Funktion von  $A$  und  $C(x)$  die exponentielle erzeugende Funktion von  $\text{CYC}(A)$ . Dann gilt  $C(x) = \log\left(\frac{1}{1-A(x)}\right)$ .

*Beweis.* Durch den Aufbau und der Potenzreihenentwicklung des natürlichen Logarithmus ergeben sich folgende Gleichungen:

$$C(x) = \sum_{k=1}^{\infty} \frac{A^k(x)}{k} = -\log(1 - A(x)) = \log\left(\frac{1}{1 - A(x)}\right) \quad \square$$

### 2.2.6 Auszeichnen: $\Theta A$

Die Konstruktion des Auszeichnen spielt eine wichtige Rolle bei der Erzeugung von zufälligen kombinatorischen Strukturen und ist definiert als die Struktur der Form

$$\Theta A = \sum_{n \geq 1} (\mathcal{A}_n \times \{1, 2, \dots, n\}),$$

wobei  $\mathcal{A}_n$  die Menge der Objekte  $a$  mit  $w(a) = n$  bezeichnet. Dass das Auszeichnen bei der Zufallserzeugung so hilfreich ist, liegt daran, dass es das kombinatorische Analogon zur Differentiation ist. Dies sieht man wenn man die exponentielle erzeugende Funktion betrachtet:

**Satz 2.11.** Sei  $(A, w)$  eine markierte kombinatorische Struktur und sei  $A(x)$  die exponentielle erzeugende Funktion von  $A$  und  $C(x)$  die exponentielle erzeugende Funktion von  $\text{CYC}(A)$ . Dann gilt  $C(x) = x \cdot A'(x)$ .

*Beweis.*

$$C(x) = \sum_{n=1}^{\infty} \frac{A_n}{n!} \cdot x^n \cdot n = x \cdot \sum_{n=1}^{\infty} \frac{A_n}{n!} \cdot x^{n-1} \cdot n = x \cdot A'(x) \quad \square$$

## 2 Grundlagen

Bei den Konstruktionen der Folge, der Menge, der Multimenge und des Zyklus lässt sich eine weitere Einschränkung bezüglich der Kardinalität bestimmen.  $\text{SET}_{<k}(\mathcal{A})$  bedeutet beispielsweise, dass nur Mengen mit einer Kardinalität kleiner als  $k$  beinhaltet sind.  $\text{SEQ}_k(\mathcal{A})$  wären Folgen mit genau  $k$  Elementen.

### 2.3 Standardbeschreibung

Sogenannte zerlegbare kombinatorische Strukturen lassen sich nun mithilfe der definierten Konstruktionen erzeugen. Eine Beschreibung einer Struktur ist der Aufbau unter Zuhilfenahme von erlaubten Konstruktionen. Oder formaler:

**Definition 2.5.** Eine **Beschreibung** für ein  $n$ -Tupel  $\mathcal{A} = (\mathcal{A}^{(1)}, \dots, \mathcal{A}^{(n)})$  von kombinatorischen Strukturen ist eine Menge von  $n$  Gleichungen, sodass die  $i$ -te Gleichung von der Form

$$\mathcal{A}^{(i)} = \Psi_i(\square, \mathcal{A}^{(1)}, \dots, \mathcal{A}^{(n)}),$$

ist, wobei  $\Psi_i$  ein Term ist, der aus Atomen  $\square$ , neutralen Objekten, den Strukturen  $\mathcal{A}^{(j)}$  und den vorgestellten Konstruktionen aufgebaut ist.

Eine Struktur nennt man **zerlegbar**, wenn eine Beschreibung dafür existiert.

#### Beispiel: Ebener Wurzelbaum

Eine wichtige Klasse von Bäumen stellen die ebenen Wurzelbäume dar, da diese für Suchalgorithmen und die Untersuchung der Komplexität von Algorithmen eine bedeutende Rolle spielen. Der ebene Wurzelbaum lässt sich rekursiv definieren: Ein ebener Wurzelbaum kann ein Atom sein. Weiters kann ein ebener Wurzelbaum ein Wurzelknoten mit einer beliebigen Anzahl von Nachfolgern sein, die wiederum ebene Wurzelbäume sind und deren Reihenfolge relevant ist. Dadurch ergibt sich folgende formale Beschreibung:

$$\{\mathcal{A} = \{\square\} \star \text{SEQ}(\mathcal{A})\}.$$

Es ist nun möglich eine Beschreibung zu standardisieren, indem man alle Konstruktionen in eine Form überführt, die mit Atomen, dem neutralen Objekt und den Basiskonstruktionen (Summe, Produkt, Auszeichnung) auskommt.

**Definition 2.6.** Eine **Standardbeschreibung von unmarkierten kombinatorischen Strukturen** ist eine Beschreibung in Form von Gleichungen, die alle eine der folgenden Formen besitzen müssen:

## 2 Grundlagen

- $\mathcal{A}^{(i)} = \{\epsilon\}$ ;
- $\mathcal{A}^{(i)} = \{\square\}$ ;
- $\mathcal{A}^{(i)} = \mathcal{B}^{(j)} + \mathcal{B}^{(k)}$ ;
- $\mathcal{A}^{(i)} = \mathcal{B}^{(j)} \times \mathcal{B}^{(k)}$ ,

wobei  $\mathcal{B}^{(j)}, \mathcal{B}^{(k)} \in \{\{\epsilon\}, \{\square\}, \mathcal{A}^{(1)}, \dots, \mathcal{A}^{(n)}\}$ .

Ganz ähnlich lässt sich eine Standardbeschreibung für markierte Strukturen definieren. Dabei ist aber zusätzlich die Konstruktion des Auszeichnens zu beachten.

**Definition 2.7.** *Eine Standardbeschreibung von markierten kombinatorischen Strukturen ist eine Beschreibung in Form von Gleichungen, die alle eine der folgenden Formen besitzen müssen:*

- $\mathcal{A}^{(i)} = \{\epsilon\}$ ;
- $\mathcal{A}^{(i)} = \{\square\}$ ;
- $\mathcal{A}^{(i)} = \mathcal{B}^{(j)} + \mathcal{B}^{(k)}$ ;
- $\mathcal{A}^{(i)} = \mathcal{B}^{(j)} \star \mathcal{B}^{(k)}$ ;
- $\Theta \mathcal{A}^{(i)} = \mathcal{B}^{(j)} \star \mathcal{B}^{(k)}$ ,

wobei  $\mathcal{B}^{(j)}, \mathcal{B}^{(k)} \in \{\{\epsilon\}, \{\square\}, \mathcal{A}^{(1)}, \dots, \mathcal{A}^{(n)}, \Theta \mathcal{A}^{(1)}, \dots, \Theta \mathcal{A}^{(n)}\}$ .

Da alle Konstruktionen von kombinatorischen Strukturen sich aus Atomen, dem neutralen Objekt und den Basiskonstruktionen der disjunkten Vereinigung, des (kartesischen) Produkts und des Auszeichnens aufbauen, ist es naheliegend, dass sich alle Beschreibungen in eine äquivalente Standardbeschreibung überführen lassen. Dazu wird bei markierten Strukturen wie folgt vorgegangen:

### Standardisierung einer Beschreibung

1. Benenne Folgen, Mengen und Zyklen als eigene Strukturvariablen.
2. Gleichungen, die mehr als einen Summen- oder Produktoperator enthalten (Polynome) werden auf mehrere Gleichungen aufgeteilt. Dabei werden neue Variablen eingeführt.

Damit haben alle Gleichungen der Beschreibung eine der Formen in der Definition 2.4 oder  $\mathcal{A} = \text{SEQ}(\mathcal{B})$  oder  $\mathcal{A} = \text{SET}(\mathcal{B})$  oder  $\mathcal{A} = \text{CYC}(\mathcal{B})$  oder einer dieser mit zusätzlicher Kardinalitätseinschränkung.

## 2 Grundlagen

3. Um zu einer Standardbeschreibung zu kommen müssen diese Typen von Gleichungen in eine konforme Form gebracht werden. Dazu ersetzt man die Gleichung durch eine oder mehrere Gleichungen, die dazu äquivalent sind.

- $\mathcal{A} = \text{SEQ}(\mathcal{B})$

Diese Konstruktion kann auch rekursiv ausgedrückt werden. Entweder es besteht aus einem Atom oder aus einer Folge mit einem Element aus  $\mathcal{B}$

$$\Rightarrow \mathcal{A} = \{\square\} + (\mathcal{A} \star \mathcal{B})$$

Dies kann erweitert werden auf zwei konforme Gleichungen, indem man eine neue Variable einführt.

$$\Rightarrow \mathcal{A} = \{\square\} + \mathcal{C}, \mathcal{C} = \mathcal{A} \star \mathcal{B}$$

- $\mathcal{A} = \text{SEQ}_k(\mathcal{B})$

Für  $k = 0$  gilt  $\mathcal{A} = \emptyset$ . Für  $k = 1$  enthält  $\mathcal{A}$  alle Folgen der Länge 1, d.h.  $\mathcal{A} = \mathcal{B}$ . Hier kann in den Gleichungen einfach  $\mathcal{B}$  durch  $\mathcal{A}$  ersetzt werden.

Da für  $k \geq 2$  endliche Folgen betrachtet werden ist diese Gleichung äquivalent zu einem Polynom k-ten Grades

$$\Rightarrow \mathcal{A} = \underbrace{\mathcal{B} \star \dots \star \mathcal{B}}_{k \text{ mal}}$$

Dieses Polynom kann nun auf  $k - 1$  Gleichungen aufgeteilt werden.

$$\Rightarrow \mathcal{A}_1 = \mathcal{B} \star \mathcal{B}, \mathcal{A}_2 = \mathcal{A}_1 \star \mathcal{B}, \mathcal{A}_3 = \mathcal{A}_2 \star \mathcal{B}, \dots, \mathcal{A} = \mathcal{A}_{k-2} \star \mathcal{B}$$

- $\mathcal{A} = \text{SEQ}_{\leq k}(\mathcal{B})$

Für  $k = 0$  und  $k = 1$  gilt selbiges wie im vorherigen Punkt.

Für  $k \geq 2$  werden wieder endliche Folgen betrachtet, die sich auch durch Polynome k-ten Grades ausdrücken lassen.

$$\Rightarrow \mathcal{A} = \mathcal{B} + \text{SEQ}_2(\mathcal{B}) + \dots + \text{SEQ}_{k-1}(\mathcal{B}) + \text{SEQ}_k(\mathcal{B}),$$

Diese Gleichung lässt sich dann entsprechend umwandeln in konforme Gleichungen. Dabei wird der letzte Punkt angewendet.

- $\mathcal{A} = \text{SEQ}_{\geq k}(\mathcal{B})$

Für  $k = 0$  und  $k = 1$  gilt  $\mathcal{A} = \text{SEQ}(\mathcal{B})$ . Für  $k \geq 2$  werden unendliche Folgen betrachtet, die mindestens Länge  $k$  haben.

$$\Rightarrow \mathcal{A} = \text{SEQ}_k(\mathcal{B}) \star \text{SEQ}(\mathcal{B}),$$

Mit den letzten Punkten lässt sich dies in eine Standardform bringen.

- $\mathcal{A} = \text{SET}(\mathcal{B})$

Bei der Menge kann man die Konstruktion des Auszeichnens verwenden.

$$\Rightarrow \Theta \mathcal{A} = \mathcal{A} \star \Theta \mathcal{B}$$



## 2 Grundlagen

- $\mathcal{A} = \text{SET}_k(\mathcal{B})$

Mengen, bei denen eine Einschränkung bezüglich der Kardinalität besteht, lassen sich rekursiv mit dem Auszeichnungsoperator beschreiben. Eine markierte Menge ist dabei eine Menge gemeinsam mit einem markierten Objekt von  $\mathcal{B}$ .

$$\Rightarrow \mathcal{A}_1 = \mathcal{B}, \Theta\mathcal{A}_2 = \mathcal{A}_1 \star \Theta\mathcal{B}, \Theta\mathcal{A}_3 = \mathcal{A}_2 \star \Theta\mathcal{B}, \dots, \Theta\mathcal{A} = \mathcal{A}_{k-1} \star \Theta\mathcal{B}$$

- $\mathcal{A} = \text{SET}_{\leq k}(\mathcal{B})$

Analog zu  $\text{SEQ}_{\leq k}(\mathcal{B})$  lässt sich  $\mathcal{A}$  schreiben als Summe

$$\Rightarrow \mathcal{A} = \text{SET}_1(\mathcal{B}) + \text{SET}_2(\mathcal{B}) + \dots + \text{SET}_{k-1}(\mathcal{B}) + \text{SET}_k(\mathcal{B})$$

Mit dem letzten Punkt lässt sich diese Gleichung dann standardisieren.

- $\mathcal{A} = \text{SET}_{\geq k}(\mathcal{B})$

$$\Rightarrow \Theta\mathcal{A}_0 = \mathcal{A}_0 \star \Theta\mathcal{B}, \Theta\mathcal{A}_1 = \mathcal{A}_0 \star \Theta\mathcal{B}, \Theta\mathcal{A}_2 = \mathcal{A}_1 \star \Theta\mathcal{B}, \dots, \\ \Theta\mathcal{A} = \mathcal{A}_{k-1} \star \Theta\mathcal{B}$$

- $\mathcal{A} = \text{CYC}(\mathcal{B})$

Zyklen können auf Folgen zurückgeführt werden. Dabei ist ein markierter Zyklus eine Folge gemeinsam mit einer markierten Komponente, an dem der Zyklus 'geöffnet' werden kann.

$$\Rightarrow \Theta\mathcal{A} = \mathcal{C} \star \Theta\mathcal{B}, \mathcal{C} = \text{SEQ}(\mathcal{B})$$

Die Zyklen mit Einschränkung der Kardinalität lassen sich nun sehr einfach daraus ableiten.

- $\mathcal{A} = \text{CYC}_k(\mathcal{B})$

$$\Rightarrow \Theta\mathcal{A} = \text{SEQ}_{k-1}(\mathcal{B}) \star \Theta\mathcal{B}$$

- $\mathcal{A} = \text{CYC}_{\leq k}(\mathcal{B})$

$$\Rightarrow \Theta\mathcal{A} = \text{SEQ}_{\leq k-1}(\mathcal{B}) \star \Theta\mathcal{B}$$

- $\mathcal{A} = \text{CYC}_{\geq k}(\mathcal{B})$

$$\Rightarrow \Theta\mathcal{A} = \text{SEQ}_{\geq k-1}(\mathcal{B}) \star \Theta\mathcal{B}$$

### Beispiele

**Nicht ebene Wurzelbäume** sind Bäume, bei denen die Reihenfolge der Unterbäume irrelevant ist.

Das bedeutet ein nicht ebener Wurzelbaum kann ein Atom (Knoten) sein. Weiters kann ein nicht ebener Wurzelbaum ein Knoten zusammen mit einer Menge an Unterbäumen sein, die alle nicht ebene Wurzelbäume sind. Auf die Beschreibung wendet man nun den Algorithmus zur Standardisierung an.

$$\{\mathcal{A} = \{\square\} \star \text{SET}(\mathcal{A})\} \\ \Rightarrow \{\mathcal{A} = \{\square\} \star \mathcal{C}, \mathcal{C} = \text{SET}(\mathcal{A})\} \\ \Rightarrow \{\mathcal{A} = \{\square\} \star \mathcal{C}, \Theta\mathcal{C} = \mathcal{C} \star \Theta\mathcal{A}\}$$

## 2 Grundlagen

Die Beschreibung der **ebenen Wurzelbäume** wurde schon angeführt. Wendet man darauf den Standardisierungsalgorithmus erhält man:

$$\begin{aligned} & \{\mathcal{A} = \{\square\} \star \text{SEQ}(\mathcal{A})\} \\ \Rightarrow & \{\mathcal{A} = \{\square\} \star \mathcal{C}, \mathcal{C} = \text{SEQ}(\mathcal{A})\} \\ \Rightarrow & \{\mathcal{A} = \{\square\} \star \mathcal{C}, \mathcal{C} = \{\square\} + \mathcal{D}, \mathcal{D} = \mathcal{C} \star \mathcal{A}\} \end{aligned}$$

Als Untergruppe der ebenen Wurzelbäume wird noch der **ebene Binärbaum** erwähnt. Der ebene Binärbaum ist ein ebener Wurzelbaum, wobei jeder Knoten zwei Nachfolger hat.

$$\begin{aligned} & \{\mathcal{A} = \{\square\} + \{\square\} \star \mathcal{A} \star \mathcal{A}\} \\ \Rightarrow & \{\mathcal{A} = \{\square\} + \mathcal{B}, \mathcal{B} = \{\square\} \star \mathcal{A} \star \mathcal{A}\} \\ \Rightarrow & \{\mathcal{A} = \{\square\} + \mathcal{B}, \mathcal{B} = \{\square\} \star \mathcal{C}, \mathcal{C} = \mathcal{A} \star \mathcal{A}\} \end{aligned}$$

Die **Hierarchien** sind ein wichtiges Beispiel der kombinatorischen Strukturen, wenn man von Zufallsgenerierung spricht. Hierarchien haben folgende Beschreibung:

$$\begin{aligned} & \{\mathcal{A} = \{\square\} + (\text{SET}_{\geq 2}(\mathcal{A}))\} \\ \Rightarrow & \{\mathcal{A} = \{\square\} + \mathcal{B}, \Theta \mathcal{B} = \mathcal{C} \star \Theta(\mathcal{A}), \Theta \mathcal{C} = \mathcal{D} \star \Theta(\mathcal{A}), \Theta \mathcal{D} = \mathcal{D} \star \Theta(\mathcal{A})\} \end{aligned}$$

Die **Mengenpartition** der Größe  $n$  ist ein kombinatorisches Objekt, das das Integer-Intervall  $[1, n]$  in nicht-leere Integer-Intervalle, auch Blöcke genannt, aufteilt. Die markierte Struktur der Mengenpartitionen ist gegeben durch

$$\mathcal{P} = \text{SET}(\text{SET}_{\geq 1}(\{\square\}))$$

Standardisiert man diese Beschreibung erhält man:

$$\begin{aligned} & \{\mathcal{P} = \text{SET}(\text{SET}_{\geq 1}(\{\square\}))\} \\ \Rightarrow & \{\mathcal{A} = \{\square\}, \Theta \mathcal{B} = \mathcal{B} \star \Theta \mathcal{A}, \Theta \mathcal{C} = \mathcal{B} \star \Theta \mathcal{A}, \Theta \mathcal{D} = \mathcal{C} \star \Theta \mathcal{A}, \Theta \mathcal{P} = \mathcal{P} \star \Theta \mathcal{D}\} \end{aligned}$$

## 3 Rekursive Methode

Eine Methode, die für eine Zahl  $n \in \mathbb{N}$  ein zufälliges markiertes kombinatorisches Objekt der Größe  $n$  erzeugt, ist die rekursive Methode, die 1993 von Flajolet, Zimmermann und Cutsem in [22] eingeführt wurde. Sie lässt sich auf alle zerlegbaren, markierten Strukturen anwenden.

Für eine gegebene Größe  $n \geq 0$  liefert der Algorithmus ein zufälliges Objekt unter der Gleichverteilung der Größe  $n$ . Die Wahrscheinlichkeit für ein bestimmtes Objekt, ausgewählt zu werden, muss somit  $\frac{1}{A_n}$  betragen.

### 3.1 Der Algorithmus

Der Algorithmus nimmt Gebrauch von der Tatsache, dass sich jede zerlegbare kombinatorische Struktur mit einer Standardbeschreibung definieren lässt. Dies bedeutet, dass als erster Schritt die Beschreibung der gewünschten Struktur mithilfe des in Kapitel 2.3 vorgestellten Algorithmus in eine Standardbeschreibung übergeführt werden muss.

Für diese Beschreibung muss nun  $\frac{A_0}{0!}, \frac{A_1}{1!}, \dots, \frac{A_n}{n!}$  berechnet werden. Dies geschieht vor der eigentlichen Generierung einmal.

Nun kann davon ausgegangen werden, dass sich die Beschreibung in Standardform befindet und die ersten  $n$  Koeffizienten der exponentiellen erzeugenden Funktion bekannt sind. Nach Erzeugung des vorerst unmarkierten Objekts können die Markierungen verteilt werden, indem eine Zufallspermutation von  $\{1, \dots, n\}$  den einzelnen Atomen zugeordnet wird. Dieser Markierungsvorgang benötigt eine lineare Laufzeit.

Eine rekursiver Algorithmus liefert dann das Zufallsobjekt:

---

```
1: procedure GENERATOR( $\mathcal{A}, n$ )
2:   if  $\mathcal{A} = \{\epsilon\}$  then
3:     if  $n = 0$  then
4:       return  $\epsilon$ 
5:     else
6:       return NULL
```

### 3 Rekursive Methode

```
7:     end if
8: end if
9:
10: if  $\mathcal{A} = \{\square\}$  then
11:     if  $n = 1$  then
12:         return  $\square$ 
13:     else
14:         return NULL
15:     end if
16: end if
17:
18: if  $\mathcal{A} = \mathcal{B} + \mathcal{C}$  then
19:      $z :=$  gleichverteilte Zufallszahl  $\in [0, 1]$ 
20:     if  $z < \frac{B_n}{A_n}$  then
21:         return GENERATOR( $\mathcal{B}, n$ )
22:     else
23:         return GENERATOR( $\mathcal{C}, n$ )
24:     end if
25: end if
26:
27: if  $\mathcal{A} = \mathcal{B} \star \mathcal{C}$  then
28:      $z :=$  gleichverteilte Zufallszahl  $\in [0, 1]$ 
29:      $i := 0$ 
30:      $w := \frac{B_0 \cdot C_n}{A_n}$ 
31:     while  $z > w$  do
32:          $i := i + 1$ 
33:          $w := w + \binom{n}{i} \cdot \frac{B_i \cdot C_{n-i}}{A_n}$ 
34:     end while
35:     return [GENERATOR( $\mathcal{B}, i$ ), GENERATOR( $\mathcal{C}, n - i$ )]
36: end if
37:
38: if  $\mathcal{A} = \ominus \mathcal{B}$  then
39:      $z :=$  gleichverteilte Zufallszahl  $\in [0, 1]$ 
40:      $\mathcal{C} :=$  GENERATOR( $\mathcal{B}, n$ )
41:      $\mathcal{D} := \mathcal{C}$ , wobei das  $\lfloor n \cdot z \rfloor$ -te Atom markiert ist
42:     return  $\mathcal{D}$ 
43: end if
44:
```

### 3 Rekursive Methode

```

45:   if  $\Theta A = \mathcal{B}$  then
46:       if  $n=0$  and  $A_0 \neq 0$  then
47:           return  $\epsilon$ 
48:       end if
49:       if  $n \geq 1$  then
50:            $\mathcal{C} := \text{GENERATOR}(\mathcal{B}, n)$ 
51:            $\mathcal{D} := \mathcal{C}$ , wobei die Markierung entfernt wurde
52:           return  $\mathcal{D}$ 
53:       end if
54:   end if
55: end procedure

```

Der Algorithmus unterscheidet zwischen den einzelnen Formen, die die Gleichung der Standardbeschreibung der Struktur annehmen kann.

- **Zeile 2:** Wenn die Struktur nur das neutrale Objekt  $\epsilon$  enthält, ist klar, dass hier nur für  $n = 0$  ein Objekt zurückgegeben werden kann. Und hier existiert nur eines, nämlich  $\epsilon$ . Für  $n \geq 1$  existieren keine Objekte der Struktur.
- **Zeile 10:** Wenn die Struktur nur Atome  $\square$  enthält, wird für  $n = 1$  das Atom zurückgegeben. Für  $n = 0$  oder  $n \geq 2$  existieren keine Objekte in der Struktur.
- **Zeile 18:** Die Zeilen 19 - 24 werden ausgeführt, wenn die Struktur die Form  $\mathcal{A} = \mathcal{B} + \mathcal{C}$  hat. Die Anzahl der Objekte von  $\mathcal{B}$  mit Gewicht  $n$  ist  $B_n$ , die Anzahl der Objekte von  $\mathcal{C}$  mit Gewicht  $n$  ist  $C_n$ . Dies bedeutet, dass die Anzahl der Objekte von  $\mathcal{A}$  mit Gewicht  $n$  gleich  $B_n + C_n$  ist. Die Wahrscheinlichkeit, dass ein Objekt von  $\mathcal{B}$  stammt, ist somit  $\frac{B_n}{B_n + C_n} = \frac{B_n}{A_n}$ .
- **Zeile 27:** Dieser Fall tritt ein, wenn die Struktur durch das Produkt zweier Strukturen definiert ist. Sei  $B_n$  bzw.  $C_n$  die Anzahl der Objekte von  $\mathcal{B}$  bzw.  $\mathcal{C}$  der Größe  $n$ . Dann ist die Anzahl der Objekte der Größe  $n$  in  $\mathcal{C}$  gegeben durch

$$A_n = \sum_{k=0}^n \binom{n}{k} \cdot B_k \cdot C_{n-k}.$$

Die Wahrscheinlichkeit, dass ein Objekt der Größe  $n$  von  $\mathcal{A}$  ein Produkt eines Objekts der Größe  $k$  von  $\mathcal{B}$  und eines Objekts der Größe  $n - k$  von  $\mathcal{C}$  ist, ist

$$\binom{n}{k} \cdot \frac{B_k \cdot C_{n-k}}{A_n}.$$

Der Algorithmus findet das kleinste  $i$ , sodass die Wahrscheinlichkeit, dass eine Komponente von  $\mathcal{B}$  der Größe  $i$  oder kleiner generiert wird, kleiner ist als eine zuvor erzeugte Zufallszahl zwischen 0 und 1.

- **Zeile 38:** Bei der Zufallserzeugung eines Objektes von  $\Theta\mathcal{B}$  wird ein zufälliges Objekt von  $\mathcal{B}$  erzeugt mit einer zufälligen Markierung der  $n$  Atome.
- **Zeile 45:** Kommt in der Beschreibung die Gleichung  $\Theta\mathcal{A} = \mathcal{B}$  vor, dann werden dafür die Zeilen 46 - 53 ausgeführt. Für  $n = 0$  kann nur das leere Objekt zurückgegeben werden, wenn dieses in der Struktur vorhanden ist. Wenn  $n \geq 1$  wird ein Zufallsobjekt von  $\mathcal{B}$  erzeugt, bei dem die Markierung gelöscht wird.

## 3.2 Der Syntaxbaum

Der Syntaxbaum ist ein Wurzelbaum, der zur Visualisierung von Wörtern einer Grammatik dient. Auch hier lässt sich ein Baum anwenden, um zu Veranschaulichen auf welche Weise ein kombinatorisches Objekt durch den vorgestellten Algorithmus entstanden ist.

**Beispiel:** Der Syntaxbaum wird anhand der Erzeugung eines Binärbaums erläutert. Die Gleichungen der Standardbeschreibung eines Binärbaums  $\mathcal{A}$  lauten:

$$\begin{aligned}\mathcal{A} &= \mathcal{B} + \mathcal{C}, \\ \mathcal{B} &= \{\square\}, \\ \mathcal{C} &= \mathcal{B} \star \mathcal{D}, \\ \mathcal{D} &= \mathcal{A} \star \mathcal{A}.\end{aligned}$$

Mithilfe der rekursiven Methode möchte man einen Binärbaum mit 5 Knoten erzeugen. Beim Durcharbeiten des Algorithmus könnte der Syntaxbaum aus Abbildung 3.1 entstehen. Dabei bedeuten die tiefgestellten Zahlen, die Einschränkungen bezüglich der Kardinalität. Zu Beginn wird also ein Objekt der Struktur  $\mathcal{A}$  mit der Kardinalität 5 gesucht. Der Algorithmus entscheidet weiters, ein Objekt von  $\mathcal{C}$  der Kardinalität 5 zu suchen. Weiters wird durch Zufallsberechnung ein Objekt von  $\mathcal{B}$  der Größe 1 und ein Objekt von  $\mathcal{D}$  der Größe 4 gesucht. Dies geschieht nun so lange bis nur mehr Terminalsymbole bzw. Knoten zurückgegeben werden. Der Syntaxbaum aus Abbildung 3.1 würde den Binärbaum aus Abbildung 3.2 erzeugen.

## 3.3 Laufzeitanalyse

Nachteil der rekursiven Methode ist mitunter der hohe Aufwand der Vorberechnung der  $n$  Koeffizienten der erzeugenden Funktionen. Diese werden rekursiv berechnet und dadurch kann der nächste Satz, der auch in [21] zu finden ist, durch Induktion bewiesen werden.

### 3 Rekursive Methode

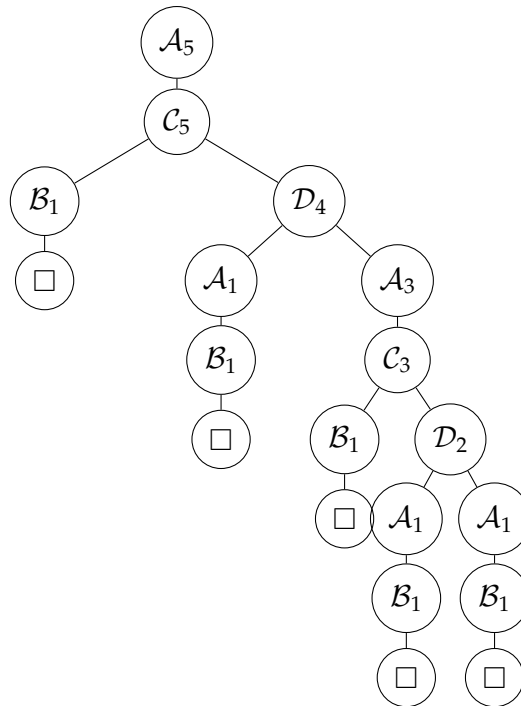


Abbildung 3.1: Beispiel eines Syntaxbaums

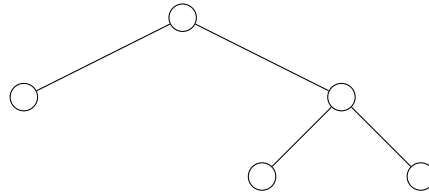


Abbildung 3.2: Binärbaum zum Syntaxbaum aus Abbildung 3.1

**Satz 3.1.** Für eine gegebene Beschreibung mit zugehöriger erzeugender Funktion  $\sum_{n=0}^{\infty} A_n x^n$  bzw.  $\sum_{n=0}^{\infty} A_n \frac{x^n}{n!}$  lassen sich die ersten  $n$  Folgenglieder der Folge  $(A_i)_{i=0}^{\infty}$  mit  $\mathcal{O}(n^2)^1$  arithmetischen Operationen berechnen.

*Beweis.* Jede Beschreibung lässt sich laut Kapitel 2.3 in eine Standardbeschreibung überführen. Da die erzeugenden Funktionen von äquivalenten Strukturen gleich sind, kann angenommen werden, dass es sich um eine Standardbeschreibung handelt.

Eine gesonderte Betrachtung der Formen, die eine Beschreibungsgleichung der Standardbeschreibung annehmen kann, führt uns zum Beweis. Dabei wird angenommen, dass bei der Berechnung des Folgenglieds  $A_n$  alle vorhergehenden Folgenglieder  $A_i$ ,  $i < n$  und alle Folgenglieder  $B_1, \dots, B_n$  bzw.  $C_1, \dots, C_n$  der erzeugenden Funktionen der Unterstrukturen bereits bekannt sind.

<sup>1</sup> Landau-Symbole werden verwendet, um das asymptotische Verhalten von Funktionen zu beschreiben. Zur Erläuterung der einzelnen Symbole siehe Anhang A.1.

### 3 Rekursive Methode

- **neutrales Objekt:** Wenn  $\mathcal{A} = \{\epsilon\}$ , dann gilt  $A_n = 1$  für  $n = 0$ , sonst  $A_n = 0$ . Dies bedeutet einen Rechenaufwand von  $\mathcal{O}(1)$ .
- **Atom:** Wenn  $\mathcal{A} = \{\square\}$ , dann gilt  $A_n = 1$  für  $n = 1$ , sonst  $A_n = 0$ . Ein zusätzlicher Rechenaufwand von  $\mathcal{O}(1)$  besteht.
- **Vereinigung:** Wenn  $\mathcal{A} = \mathcal{B} + \mathcal{C}$ , berechnet sich der Koeffizient  $A_n$  durch  $B_n + C_n$ . Deshalb ist der zusätzliche arithmetische Aufwand  $\mathcal{O}(1)$ .
- **(Kartesisches) Produkt:** Wenn  $\mathcal{A} = \mathcal{B} \times \mathcal{C}$ , errechnet sich  $A_n$  durch die Summe  $\sum_{k=0}^n B_k C_{n-k}$ . Dadurch ergibt sich ein zusätzlicher Rechenaufwand von  $\mathcal{O}(n)$ .
- **Auszeichnen:** Wenn  $\mathcal{A} = \Theta\mathcal{B}$ , dann gilt  $A_n = n \cdot B_n$ . Deshalb ergibt sich ein zusätzlicher Aufwand von  $\mathcal{O}(1)$ .

Die Berechnung eines Folgenglieds benötigt somit  $\mathcal{O}(n)$  arithmetische Operationen. Bei der inkrementellen Berechnung der ersten  $n$  Folgenglieder von  $(A_i)_{i=0}^{\infty}$  ergibt sich somit ein Gesamtaufwand von  $\mathcal{O}(n^2)$  arithmetischen Operationen.  $\square$

Zu Beginn der Erzeugung eines kombinatorischen Objektes der Größe  $n$  müssen also diese ersten  $n$  Koeffizienten der erzeugenden Funktion berechnet werden. Dies bedeutet einen einmaligen Aufwand von  $\mathcal{O}(n^2)$  arithmetischen Operationen. Der Speicheraufwand dafür beträgt  $\mathcal{O}(n)$ .

Die eigentliche Erzeugung eines Objektes benötigt  $\mathcal{O}(n^2)$  Operationen. Dies lässt sich zeigen, indem man den zugehörigen Syntaxbaum betrachtet. Die Pfadlänge des Syntaxbaums ergibt die Komplexität der Erzeugung, welche im schlechtesten Fall  $\mathcal{O}(n^2)$  ergibt. Einen detaillierten Beweis findet man in [22].

Mit weiteren Überlegungen lässt sich dieser Algorithmus optimieren, indem man die Berechnungen so anordnet, dass nur  $\mathcal{O}(n \cdot \log n)$  arithmetische Operationen benötigt werden. Genauer dazu in [22],[11].



## 4 Boltzmann-Sampler

Boltzmann-Sampler sind eine Familie von Algorithmen zur Generierung von zufälligen kombinatorischen Objekten. Im Gegensatz zur rekursiven Methode konzentrieren sie sich nicht auf Objekte fixer Größe. Es wird ein Maß auf allen Objekten der Struktur definiert und bei Opferung der genauen Größe des Objekts lässt sich somit ein schnellerer und leichter zu implementierender Algorithmus angeben. Zugrunde liegt eine Wahrscheinlichkeitsverteilung, die abhängig von einem Parameter  $x$  ist und bei der Objekte gleicher Größe gleiche Wahrscheinlichkeiten besitzen. Erstmals wurde diese Methode der Zufallserzeugung von Duchon, Flajolet, Louchard und Schaeffer in [16] vorgestellt.

**Definition 4.1.** Sei  $\mathcal{A}$  eine kombinatorische Struktur mit ihrer erzeugenden Funktion  $A(x) = \sum_{n=0}^{\infty} A_n \frac{x^n}{\omega_n}$ , wobei bei unmarkierten Strukturen  $\omega_n = 1$  und bei markierten Strukturen  $\omega_n = n!$ . Sei  $x$  eine positive reelle Zahl, sodass  $A(x)$  konvergiert. Dann ordnet die **Boltzmann-Verteilung auf  $\mathcal{A}$**  jedem Objekt  $\alpha \in \mathcal{A}$  die Wahrscheinlichkeit

$$\mathbb{P}_x(\alpha) = \frac{x^{|\alpha|}}{\omega_{|\alpha|}} \frac{1}{A(x)}$$

zu. Ein **Boltzmann-Sampler**  $\Gamma_{\mathcal{A}}(x)$  für eine Struktur  $\mathcal{A}$  ist eine Prozedur, welche Objekte von  $\mathcal{A}$  entsprechend der Boltzmann-Verteilung generiert.

Objekte einer bestimmten Größe sind bei der Boltzmann-Verteilung gleich wahrscheinlich. D.h. beinhaltet die Struktur nur Objekte einer bestimmten Größe, ist die Boltzmann-Verteilung äquivalent zur Gleichverteilung. Die Boltzmann-Verteilung ist abhängig von einem Parameter  $x$ . Damit die Wahrscheinlichkeit  $\mathbb{P}_x(\alpha)$  auch berechenbar ist, muss dieser Parameter beschränkt werden. Wenn  $R_{\mathcal{A}}$  den Konvergenzradius von der erzeugenden Funktion bezeichnet, muss der Parameter  $x$  aus  $(0, R_{\mathcal{A}})$  gewählt werden, damit die Boltzmann-Verteilung sinnvoll definiert werden kann.

Das allgemeine Ziel ist es nun, effiziente Boltzmann-Sampler für möglichst viele kombinatorische Strukturen zu erstellen.

Bei gegebener Beschreibung einer kombinatorischen Struktur  $\mathcal{A}$ , von welcher man sehr große Zufallsobjekte generieren möchte, kann die Methode des Boltzmann-Sampling folgendermaßen zusammengefasst werden:

## 4 Boltzmann-Sampler

1. Grundvoraussetzung um die Boltzmann-Methode anzuwenden ist es, dass die gegebene Struktur  $\mathcal{A}$  zerlegbar ist. Ist dies der Fall, kann eine Beschreibung dafür angegeben werden. Die wichtigsten Konstruktionen und wie man daraus eine Beschreibung entwickelt, wurden in Kapitel 2 behandelt.
2. Mithilfe der Beschreibung lässt sich eine Gleichung für die (exponentielle) erzeugende Funktion angeben. Dies geschieht indem man die Konstruktionen der Beschreibung mithilfe der Sätze 2.1. - 2.11. in die erzeugenden Funktionen übersetzt.
3. Lässt sich ein expliziter Ausdruck für die (exponentielle) erzeugende Funktion aus dieser Gleichung ableiten, dann kann der Parameter so gewählt werden, dass der Erwartungswert der Größe des generierten Objekts  $n$  beträgt. Dies wird in Kapitel 4.1 beschrieben.
4. Alle erzeugenden Funktionen der Strukturen, die in der Beschreibung vorkommen müssen an der Stelle  $x$  ausgewertet werden. Ist dies nicht exakt möglich, werden Approximationen berechnet. Die Vorgehensweise, um zu solchen Näherungen zu kommen wird im Kapitel 4.2 erläutert.
5. Einen Boltzmann-Sampler erhält man nun dadurch, dass man die in Kapitel 4.3 und 4.4 vorgegebenen Sampler der einfachen Konstruktionen verwendet, um daraus einen Sampler für  $\mathcal{A}$  zu entwickeln.
6. Hat man eine bestimmte Toleranz  $\epsilon$  bezüglich der Größe  $n$  des Objekts  $\alpha$ , d.h.  $|\alpha| \in [n - \epsilon, n + \epsilon]$ , dann wird Schritt 5 so lange ausgeführt, bis die Größe des generierten Objekts in diesen Toleranzbereich fällt. Dieser Schritt wird in Kapitel 4.5 behandelt.

### 4.1 Wahl des Verteilungsparameter

Die Verteilung der Objekte der kombinatorischen Struktur hängen von einem Parameter  $x$  ab. Dieser Parameter wird zwischen 0 und dem Konvergenzradius der Potenzreihe  $\sum_{n \geq 0} a_n z^n$  gewählt. Der Konvergenzradius hängt hier nun aber von der kombinatorischen Struktur ab. Man könnte nun einen zufälligen Parameter wählen. Wenn man nun aber ein kombinatorisches Objekt erhalten möchte, das einer Größe möglichst nahe kommt, kann man diesen Parameter optimieren, indem man den Erwartungswert betrachtet.

Die folgenden Sätze belegen Aussagen über den Erwartungswert und die Varianz der Größe des kombinatorischen Objekts unter der Boltzmann-Verteilung.

**Satz 4.1.** *Sei  $\mathcal{A}$  eine kombinatorische Struktur,  $A(z)$  die erzeugende Funktion von  $\mathcal{A}$ ,  $R_A$  der Konvergenzradius von  $A(z)$  und  $x \in (0, R_A)$ . Sei  $\mathcal{A}$  boltzmann-verteilt. Dann ist die zu erwartende Größe des kombinatorischen Objekts*

$$\mathbb{E}_x(|\alpha|) := \frac{x A'(x)}{A(x)}.$$

#### 4 Boltzmann-Sampler

*Beweis.* Die Wahrscheinlichkeit, dass ein Objekt  $\alpha$  der Größe  $n$  aus einer kombinatorischen Struktur  $\mathcal{A}$  gezogen wird beträgt

$$\mathbb{P}_x(|\alpha| = n) = \sum_{\beta \in \mathcal{A}, |\beta|=n} \frac{x^{|\beta|}}{\omega_{|\beta|}} \frac{1}{A(x)} = \sum_{\beta \in \mathcal{A}, |\beta|=n} \frac{x^n}{\omega_n} \frac{1}{A(x)} = \frac{A_n x^n}{\omega_n A(x)}.$$

Für die erste Ableitung der erzeugenden Funktion gilt

$$A'(x) = \sum_{n \geq 1} \frac{n A_n x^{n-1}}{\omega_n}.$$

Somit errechnet sich der Erwartungswert aus

$$\mathbb{E}_x(|\alpha|) = \sum_{n \geq 0} n \cdot \mathbb{P}_x(|\alpha| = n) = \sum_{n \geq 0} \frac{n A_n x^n}{\omega_n A(x)} = \frac{x A'(x)}{A(x)}. \quad \square$$

**Satz 4.2.** Sei  $\mathcal{A}$  eine kombinatorische Struktur,  $A(z)$  die erzeugende Funktion von  $\mathcal{A}$ ,  $R_A$  der Konvergenzradius von  $A(z)$  und  $x \in (0, R_A)$ . Sei  $\mathcal{A}$  boltzmann-verteilt. Dann ist die Varianz der Größe des kombinatorischen Objekts

$$\mathbb{V}_x(|\alpha|) = x^2 \cdot \frac{A''(x)}{A(x)} + x \cdot \frac{A'(x)}{A(x)} - \left( x \cdot \frac{A'(x)}{A(x)} \right)^2.$$

*Beweis.* Für die zweite Ableitung der erzeugenden Funktion gilt

$$A''(x) = \sum_{n \geq 2} \frac{n(n-1) A_n x^{n-2}}{\omega_n}.$$

Für die Varianz gilt somit aufgrund des Verschiebungssatz und Satz 4.1

$$\begin{aligned} \mathbb{V}_x(|\alpha|) &= \mathbb{E}_x(|\alpha|^2) - (\mathbb{E}_x(|\alpha|))^2 = \\ &= \sum_{n \geq 0} n^2 \cdot \mathbb{P}_x(|\alpha| = n) - \left( \frac{x A'(x)}{A(x)} \right)^2 = \\ &= \sum_{n \geq 0} (n^2 - n) \cdot \mathbb{P}_x(|\alpha| = n) + \sum_{n \geq 0} n \cdot \mathbb{P}_x(|\alpha| = n) - \left( \frac{x A'(x)}{A(x)} \right)^2 = \\ &= \sum_{n \geq 2} \frac{n(n-1) A_n x^n}{\omega_n A(x)} + \frac{x A'(x)}{A(x)} - \left( \frac{x A'(x)}{A(x)} \right)^2 = \\ &= \frac{x^2 A''(x)}{A(x)} + \frac{x A'(x)}{A(x)} - \left( \frac{x A'(x)}{A(x)} \right)^2. \quad \square \end{aligned}$$

#### 4 Boltzmann-Sampler

Wenn nun ein Objekt der Größe  $n$  gesucht wird, wird der Parameter  $x$  so gewählt, damit der Erwartungswert  $n$  möglichst nahe kommt. Je kleiner die Varianz, um so wahrscheinlicher ist es dann auch, dass das zufällige Objekt eine Größe besitzt, die zumindest in der Nähe von  $n$  ist. Ziel ist es daher die Gleichung  $\frac{x A'(x)}{A(x)} = n$  nach  $x$  aufzulösen.

Im Normalfall ist die Funktion  $f(x) := \mathbb{E}_x(|\alpha|)$  eine streng monoton steigende und konvexe Funktion auf dem Intervall  $[0, R_A)$  und die Gleichung  $\frac{x A'(x)}{A(x)} = n$  hat maximal eine Lösung. Wenn  $A'(x)$  gegen unendlich strebt, wenn  $x$  sich  $R_A$  nähert, dann existiert für jedes  $n$  eine eindeutige Lösung  $x_n$ . Setzt man den Parameter  $x$  der Boltzmann-Verteilung auf diesen Wert  $x_n$ , dann resultiert daraus ein Boltzmann-Sampler, der Objekte mit erwarteter Größe  $n$  zurückliefert. Der selbe Wert maximiert darüber hinaus auch die Wahrscheinlichkeit, dass das generierte Objekt von der Größe  $n$  ist.

#### Beispiel: Kombinationen ohne Wiederholung

Die Struktur der Kombinationen ohne Wiederholung von  $m$  Elementen  $e_1, e_2, \dots, e_m$  wird beschrieben durch

$$\mathcal{A} = (\{\epsilon\} + \{e_1\}) \times (\{\epsilon\} + \{e_2\}) \times \dots \times (\{\epsilon\} + \{e_m\}).$$

Daraus ergibt sich die erzeugende Funktion  $A(x) = (1 + x)^m$ .

Der Erwartungswert der Größe der generierten Kombination  $\alpha$  ist

$$\mathbb{E}_x(\alpha) = \frac{x A'(x)}{A(x)} = \frac{xm}{1+x}$$

Diese Gleichung lässt sich einfach nach  $x$  auflösen und der Parameter wird optimal gewählt als

$$x_n := \frac{n}{m-n}.$$

Die Varianz beträgt

$$\mathbb{V}_x(|\alpha|) = x^2 \cdot \frac{A''(x)}{A(x)} + x \cdot \frac{A'(x)}{A(x)} - \left( x \cdot \frac{A'(x)}{A(x)} \right)^2 = \frac{xm}{(1+x)^2}$$

Mit dem optimalen Parameter  $x_n$  beträgt die Varianz

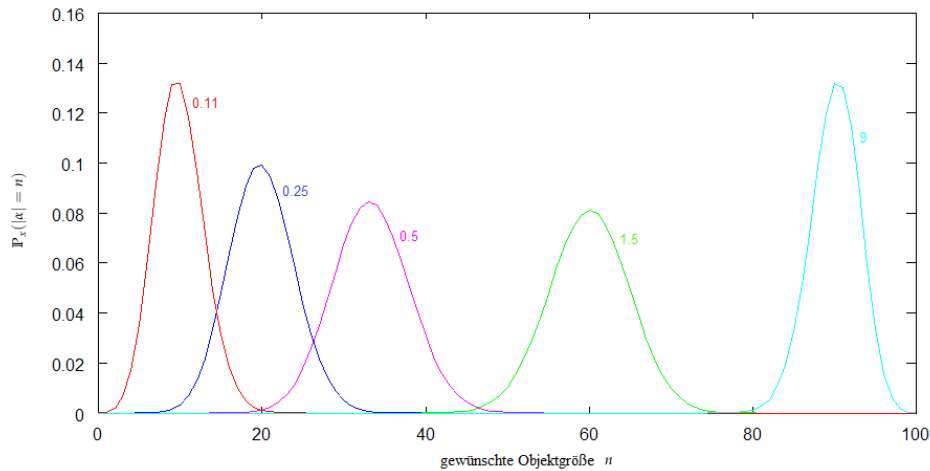
$$\mathbb{V}_{x_n}(|\alpha|) = n - \frac{n^2}{m}$$

Als konkretes einfaches Beispiel möchte man eine Kombination ohne Wiederholung von  $m = 100$  Elementen erzeugen. Die Größe des Objekts soll  $n = 10$  betragen. Mithilfe obiger Berechnungen wäre hier der optimale Parameter  $x_n \approx 0,11$  und die Varianz  $\mathbb{V}_{x_n}(|\alpha|) = 9$ . Dies bedeutet eine Standardabweichung von

## 4 Boltzmann-Sampler

$$\sigma = \sqrt{\mathbb{V}_{x_n}(|\alpha|)} = 3.$$

Möchte man ein Objekt der Größe  $n = 60$  erzeugen, dann wird als optimaler Parameter  $x_n = 1,5$  gewählt. Die Varianz beträgt  $\mathbb{V}_{x_n}(|\alpha|) = 24$  und die Standardabweichung  $\sigma = \sqrt{\mathbb{V}_{x_n}(|\alpha|)} \approx 4,9$ .



**Abbildung 4.1:** Größenverteilung der Kombinationen ohne Wiederholung mit entsprechendem Parameter  $x$

In Abbildung 4.1 ist die Größenverteilung der Kombinationen ohne Wiederholung unter dem Boltzmannmodell veranschaulicht. Je nach gewünschter Größe wird der Parameter  $x$  so gewählt, dass dort die Wahrscheinlichkeit am höchsten ist. Möchte man eine Kombination der Größe 20, dann wäre der optimale Parameter 0,25. Dies ist auch durch die Formel für den Erwartungswert zu berechnen.

### Beispiel: Permutationen

Die Permutationen der Zahlen  $1, \dots, n$  sind eine markierte kombinatorische Struktur mit exponentieller erzeugender Funktion

$$P(x) = \sum_{n \geq 0} n! \frac{x^n}{n!} = \frac{1}{1-x}$$

Der Konvergenzradius dieser Reihe beträgt 1. Daher muss für den Parameter ein Wert kleiner als 1 gewählt werden. Möchte man eine Permutation der Zahlen  $1, \dots, n$ , kann man sich den optimalen Wert wieder durch die Formel für den Erwartungswert errechnen:

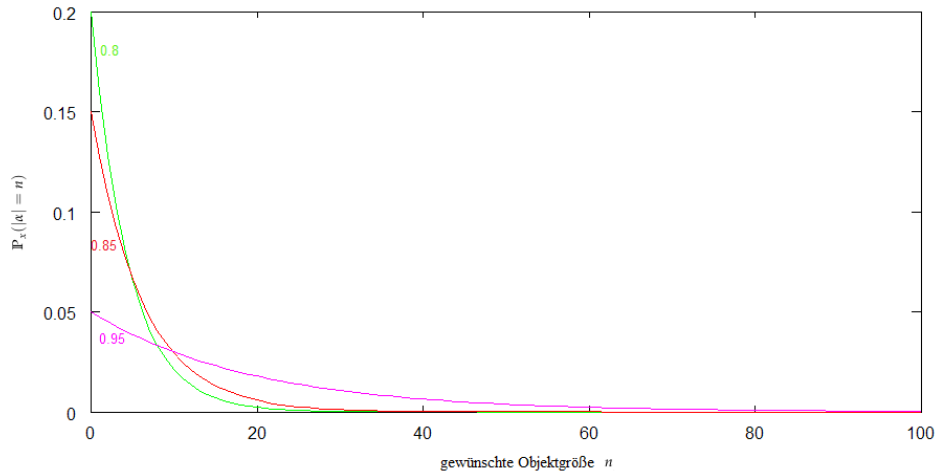
$$\mathbb{E}_x(\alpha) = \frac{xP'(x)}{P(x)} = \frac{x}{1-x}$$

$$\implies x_n := \frac{n}{1+n}.$$

In Abbildung 4.2 sieht man die Größenverteilung mit 3 verschiedenen Parametern. Je näher der Parameter an 1 herankommt, desto flacher wird die Kurve der Größenverteilung. Dies bedeutet, bei einer

## 4 Boltzmann-Sampler

Permutation der Zahlen  $1, \dots, n$ , wobei  $n$  sehr groß sein soll, wäre ein Parameter sehr nahe bei 1 idealer.



**Abbildung 4.2:** Größenverteilung der Permutationen mit entsprechendem Parameter  $x$

### Beispiel: Ebene Wurzelbäume

Als letztes Beispiel wird noch die Parameterwahl bei ebenen Wurzelbäumen diskutiert. Wie schon erwähnt haben die ebenen Wurzelbäume folgende Beschreibung:

$$\{\mathcal{A} = \{\square\} \star \text{SEQ}(\mathcal{A})\}.$$

Die erzeugende Funktion lautet

$$A(x) = \frac{1 - \sqrt{1 - 4x}}{2}.$$

Die Anzahl der Wurzelbäume mit  $n$  Knoten beträgt

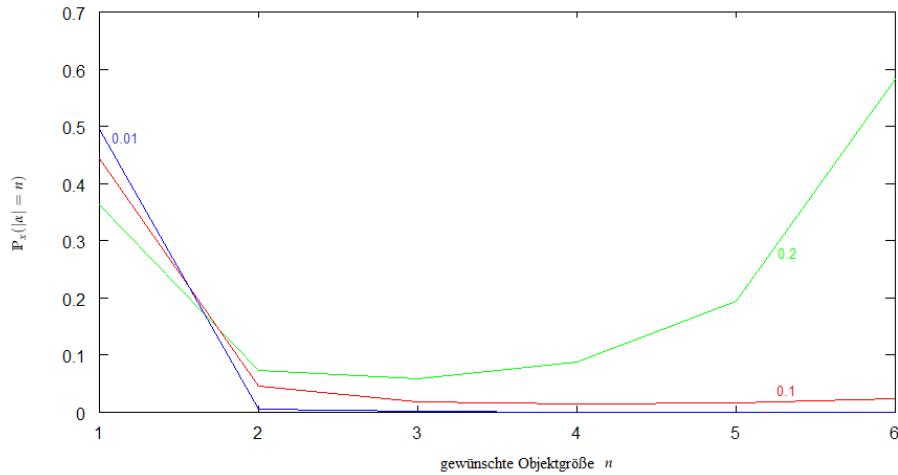
$$a_n = \frac{1}{n} \binom{2n-2}{n-1}.$$

Man kann sich den optimalen Wert wieder durch die Formel für den Erwartungswert errechnen:

$$\begin{aligned} \mathbb{E}_x(\alpha) &= \frac{x A'(x)}{A(x)} = \frac{-4x}{\sqrt{1-4x} - 1 + 4x} \\ \implies x_n &:= \frac{n}{4(n-1)^2} \end{aligned}$$

In Abbildung 4.3 ist zu erkennen, dass es hohe Spitzen bei Objekten geringer Größe gibt. Dadurch entstehen auch große Standardabweichungen. Wie man Größenverteilungen dieser Art umgehen kann, wird im Folgenden behandelt.

## 4 Boltzmann-Sampler



**Abbildung 4.3:** Größenverteilung der ebenen Wurzelbäume mit entsprechendem Parameter  $x$

Wie man anhand der letzten drei Beispiele sieht, gibt es sehr unterschiedliche Formen von Größenverteilungen. Bei den Kombinationen ohne Wiederholung ist die Kurve sehr uneben. Sie weist eine Stelle auf, an der die Wahrscheinlichkeit, ein Objekt dieser Größe zu erhalten, sehr hoch ist. Je nach Wahl des Parameters liegt dieses Maximum an einer anderen Stelle (siehe Abbildung 4.1). Bei den Permutationen wird die Kurve der Größenverteilung immer flacher, je näher sich der Parameter  $x$  dem Konvergenzradius nähert (siehe Abbildung 4.2). Und bei ebenen Wurzelbäumen ist die Wahrscheinlichkeit, ein Objekt von sehr kleiner Größe zu erhalten, sehr hoch (siehe Abbildung 4.3). Man kann die Wahrscheinlichkeitsverteilungen der Größe des zufälligen Objekts nun in drei Typen einteilen. Wie man gesehen hat, ist dies von der kombinatorischen Struktur abhängig. Je nach Typ der Größenverteilung besitzt die Erzeugung solcher kombinatorischen Objekte eine unterschiedliche Laufzeit. Dazu mehr in Kapitel 4.6.

Die Einteilung der Boltzmann-Verteilungen erfolgt über Bedingungen, die die Singularitäten der erzeugenden Funktion verwenden. Deshalb werden hier einige Begriffe aus der komplexen Analysis vorgestellt.

**Definition 4.2.** Ein **Gebiet** ist eine zusammenhängende offene Menge der komplexen Ebene.

**Definition 4.3.** Sei  $D \subseteq \mathbb{C}$  offen,  $z_0 \in D$ ,  $f : D \rightarrow \mathbb{C}$ . Die Funktion  $f$  heißt in  $z_0$  **komplex differenzierbar**, wenn der Grenzwert

$$\lim_{h \rightarrow 0} \frac{f(z_0 + h) - f(z_0)}{h}$$

in  $\mathbb{C}$  existiert. Falls dieser Grenzwert existiert, nennt man ihn **Ableitung** von  $f$  in  $z_0$  und bezeichnet ihn mit  $f'(z_0)$ .

**Definition 4.4.** Eine komplexe Funktion mit Definitionsbereich  $f(z), z \in D(f)$ , wobei  $D(f)$  den Definitionsbereich der Funktion bezeichnet, heißt **analytisch** in  $D(f)$ , falls die beiden folgenden Bedingungen erfüllt sind.

#### 4 Boltzmann-Sampler

- $D(f)$  ist ein Gebiet;
- $f$  ist in jedem Punkt  $z \in D(f)$  komplex differenzierbar.

**Satz 4.3.** Sei  $G \subseteq \mathbb{C}$  offen, und  $f : G \rightarrow \mathbb{C}$ . Dann sind die folgenden Aussagen äquivalent:

- Die Funktion  $f$  ist analytisch in  $G$ .
- Für jeden Punkt  $w \in G$  gibt es eine Potenzreihe  $\sum_{n=0}^{\infty} a_n(z-w)^n$  mit Konvergenzradius  $R > 0$ , und einen Radius  $r \in (0, R]$ , sodass die Kreisscheibe  $U_r(w) := \{z \in \mathbb{C} : |z-w| < r\}$  ganz in  $G$  liegt und  $f$  sich auf dieser Kreisscheibe darstellen lässt als

$$f(z) = \sum_{n=0}^{\infty} a_n(z-w)^n, \quad z \in U_r(w).$$

Die Analytizität einer Funktion ist mit vielen weiteren Aussagen äquivalent. Diese und einen Beweis des letzten Satzes findet man unter anderem in [8].

**Definition 4.5.** Eine stetige Abbildung  $\phi : [a, b] \rightarrow \mathbb{R}^n$  heißt ein **Weg** im  $\mathbb{R}^n$ .

Das Bild eines Weges nennt man **Kurve** im  $\mathbb{R}^n$ .

Die Punkte  $\phi(a)$  bzw.  $\phi(b)$  heißen **Anfangs-** bzw. **Endpunkt** der Kurve.

$\phi$  heißt **geschlossener Weg**, wenn  $\phi(a) = \phi(b)$ . In diesem Fall bezeichnet man das Bild als **geschlossene Kurve**.

**Definition 4.6.** Sei  $G$  ein Gebiet und sei  $\phi : [a, b] \rightarrow G$  eine stetige, differenzierbare Funktion. Weiters sei  $f : G \rightarrow \mathbb{C}$  eine Funktion. Dann nennt man das Integral

$$\int_{\Gamma} f(z) d(z) := \int_a^b f(\phi(t)) \phi'(t) dt$$

das **Kurvenintegral**  $f$  entlang der Kurve  $\Gamma = \phi([a, b])$ .

**Satz 4.4. (Cauchysche Integralformel)** Sei  $f(z) = \sum_{n \geq 0} a_n z^n$  eine Potenzreihe, die analytisch ist in einem Kreis  $U_r(0) := \{z \in \mathbb{C} : |z| < r\}$ , und sei  $\phi : [a, b] \rightarrow G$  ein Weg, sodass  $\Gamma := \phi([a, b]) \subseteq U_r(0)$  und die Kurve den Punkt 0 einmal in positiver Orientierung umrundet, dann kann der Koeffizient  $a_n$  der Potenzreihe, mithilfe des Kurvenintegrals berechnet werden durch

$$[z^n]f(z) = a_n = \frac{1}{2\pi i} \int_{\Gamma} \frac{f(z)}{z^{n+1}} dz.$$



#### 4 Boltzmann-Sampler

Was es heißt, dass eine Kurve einen Punkt umrundet gibt die Umlaufzahl an. In dieser Arbeit werden nur Kurven und Punkte mit Umlaufzahl 1 betrachtet, d.h. die einmalige Umrundung in positiver Orientierung. Deshalb wird hier darauf nicht näher eingegangen. Näheres zur Umlaufzahl und wie diese berechnet wird in [5]. Auch ein Beweis der Cauchyschen Integralformel ist darin zu finden.

**Definition 4.7.** Sei  $f(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n$  eine analytische Funktion, sei  $R$  der Konvergenzradius der Potenzreihe und sei  $w$  ein Punkt am Rand des Konvergenzkreises, d.h.  $|w - z_0| = R$ .

Wenn es eine Umgebung  $U$  um  $w$  gibt und eine in  $U$  analytische Funktion  $\tilde{f}$ , sodass diese auf dem Konvergenzkreis übereinstimmen, d.h.

$$\tilde{f}|_{U \cap U_R(z_0)} = f|_{U \cap U_R(z_0)},$$

dann nennt man  $\tilde{f}$  eine **analytische Fortsetzung** von  $f$  in  $w$ .

Wenn es keine analytische Fortsetzung von  $f$  in  $w$  gibt, dann nennt man  $w$  **Singularität** von  $f$ .

Die betragsmäßig kleinsten Singularitäten einer Funktion nennt man **dominante Singularitäten**.

**Satz 4.5.** Sei  $f(z)$  eine bei 0 analytische Funktion. Sei  $f(z) = \sum_{n \geq 0} a_n z^n$  die Potenzreihe der Funktion mit endlichem Konvergenzradius  $R$ . Dann existiert eine Singularität  $w$  mit  $|w| = R$ .

*Beweis.* Angenommen es existiert keine Singularität  $w$  mit  $|w| = R$ . Dann kann  $f$  auf jeden Punkt des Randes des Konvergenzkreises analytisch fortgesetzt werden. Das bedeutet, man findet um jeden Punkt  $x$  mit  $|x| = R$  eine Umgebung  $U_x$  und eine auf  $U_x$  analytische Funktion  $f_x$ , die  $f$  fortsetzt.

Die Umgebungen  $U_x$  überdecken nun den Rand des Konvergenzkreises. Es überdecken nun sogar endlich viele dieser Umgebungen den Rand des Konvergenzkreises, da dieser kompakt ist.

Nun ist die Funktion  $f$  auf einem größeren Kreis mit Konvergenzradius  $\tilde{R}$  mit  $R < \tilde{R}$  analytisch und deshalb gilt laut der Cauchyschen Integralformel

$$a_n = \frac{1}{2\pi i} \int_{\Gamma} \frac{f(z)}{z^{n+1}} dz.$$

$\Gamma$  stellt hier die Kurve des Kreises mit Radius  $r$  dar, wobei  $R < r < \tilde{R}$ . Damit lässt sich der Koeffizient  $a_n$  abschätzen durch

$$|a_n| = \frac{1}{|2\pi i|} \left| \int_{\Gamma} \frac{f(z)}{z^{n+1}} dz \right| \leq \frac{2\pi r}{2\pi r^{n+1}} \cdot \max_{|z|=r} |f(z)| = \frac{1}{r^n} \cdot \max_{|z|=r} |f(z)|.$$

Da der Rand des Kreises mit Radius  $r$  kompakt ist besitzt die Menge  $\{|f(z)| : |z| = r\}$  ein Maximum. Deshalb gilt für  $|z| < \tilde{r} < r$

$$|a_n| \leq \frac{C}{r^n} \implies \sqrt[n]{|a_n z^n|} \leq \frac{\tilde{r}}{r} \sqrt[n]{C} \implies \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n z^n|} < 1$$

#### 4 Boltzmann-Sampler

Aufgrund des Wurzelkriteriums konvergiert  $f(z)$  für alle  $\tilde{r}$  mit  $R < \tilde{r} < r$ . Da aber eine Reihe strikt außerhalb ihres Konvergenzkreises divergiert und  $R < \tilde{r}$ , führt dies zu einem Widerspruch.  $\square$

Dieser Satz kann verschärft werden, wenn bekannt ist, dass die Koeffizienten  $a_n$  alle nichtnegativ sind. Da hier nur erzeugende Funktionen von kombinatorischen Strukturen betrachtet werden, müssen die Koeffizienten größer oder gleich 0 sein.

**Satz 4.6.** *Sei  $f(z)$  eine bei 0 analytische Funktion. Sei  $f(z) = \sum_{n \geq 0} a_n z^n$  die Potenzreihe der Funktion mit Konvergenzradius  $R$ . Wenn  $a_n \geq 0$  für alle  $n \geq 0$ , dann ist  $R$  eine Singularität.*

*Beweis.* Angenommen  $R$  ist keine Singularität. Dann findet man einen Radius  $r$ , sodass  $f(z)$  auf  $U_r(R)$  analytisch ist. Wähle  $\tilde{r} = \frac{r}{3}$  und  $\tilde{z} = R - \tilde{r}$ . Dann lässt sich die Funktion  $f(z)$  um  $\tilde{z}$  in eine Potenzreihe entwickeln. D.h.  $f(z) = \sum_{n \geq 0} \tilde{a}_n (z - \tilde{z})^n$  für  $z \in \overline{U_{2\tilde{r}}(\tilde{z})}$ . Dann gilt

$$f(z + \tilde{z}) = \sum_{n \geq 0} \tilde{a}_n z^n \quad \text{für } z \in \overline{U_{2\tilde{r}}(\tilde{z})}$$

Somit gilt

$$\tilde{a}_n = [z^n]f(z + \tilde{z}) = \frac{f^{(n)}(\tilde{z})}{n!} = \frac{1}{n!} \sum_{k \geq 1} a_k \cdot k \cdot (k-1) \cdots (k-n+1) z^{k-n} = \sum_{n \geq m} \binom{n}{m} a_n \cdot \tilde{z}^{n-m}.$$

Daraus folgt

$$f(\tilde{z} + 2\tilde{r}) = f(R + \tilde{r}) = \sum_{n \geq 0} \left( \sum_{k \geq n} \binom{k}{n} a_n \cdot \tilde{z}^{k-n} \right) (2\tilde{r})^n.$$

Diese Doppelsumme konvergiert, da sie  $f(R + \tilde{r})$  darstellt,  $R + \tilde{r} \in U_r(R)$ . Sie ist sogar absolut konvergent, da die innere Summe immer nichtnegativ ist. Daraus folgt

$$f(R + \tilde{r}) = \sum_{k \geq 0} a_n \sum_{n=0}^k \binom{k}{n} \tilde{z}^{k-n} (2\tilde{r})^n = \sum_{k \geq 0} a_n (\tilde{z} + 2\tilde{r})^k = \sum_{k \geq 0} a_n (R + \tilde{r})^k.$$

Da die Potenzreihe  $f(z) = \sum_{n \geq 0} a_n z^n$  außerhalb des Konvergenzkreises divergiert, ist dies ein Widerspruch, da  $\tilde{r} > 0$ .  $\square$

**Definition 4.8.** *Sei  $\phi \in (0, \frac{\pi}{2})$  und  $r > \rho$ . Dann wird das offene Gebiet definiert durch*

$$\Delta_\rho(\phi, r) := \{z \in \mathcal{C} : z \neq \rho, |z| < r, \arg(z - \rho) \in (-\phi, \phi)\}.$$

*$\Delta$ -Gebiet genannt.*

#### 4 Boltzmann-Sampler

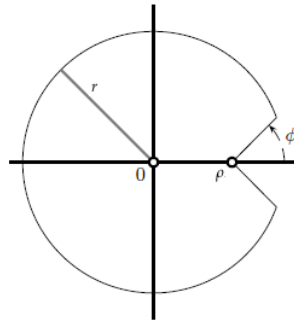


Abbildung 4.4:  $\Delta_\rho(\phi, R)$ -Gebiet

**Definition 4.9.** Die **Gammafunktion**  $\Gamma(s)$  ist durch die Eulersche Integraldarstellung

$$\Gamma(s) = \int_0^{\infty} e^{-t} \cdot t^{s-1} dt$$

für  $\operatorname{Re}(s) > 0$  definiert.

Nun kann man definieren, was es bedeutet, dass eine Boltzmann-Verteilung einer kombinatorischen Struktur flach ist.

**Definition 4.10.** Sei  $N$  die Größe eines zufälligen kombinatorischen Objekts unter der Boltzmann-Verteilung mit Parameter  $x$ . Sei  $\rho$  die dominante Singularität der erzeugenden Funktion. Dann nennt man die Boltzmann-Verteilung einer kombinatorischen Struktur **Aflach**, wenn die erzeugende Funktion  $A(z)$  analytisch an der Stelle 0 ist und die folgenden Bedingungen erfüllt:

1. Die einzige Singularität der Funktion  $A(z)$  auf dem Kreis  $\{z \in \mathbb{C} : |z| = \rho\}$  ist  $\rho$  und es gibt ein  $r > \rho$  und ein  $\phi \in (0, \frac{\pi}{2})$ , sodass  $A(z)$  auf dem Gebiet  $\Delta_\rho(r, \phi)$  analytisch fortsetzbar ist.
2. Es existiert ein  $\alpha \in \mathbb{R}^+$  sodass gilt:

$$A(z) \sim_{z \rightarrow \rho^-} a_0 \left(1 - \frac{z}{\rho}\right)^{-\alpha} + o\left(\left(1 - \frac{z}{\rho}\right)^{-\alpha}\right)$$

– $\alpha$  nennt man den **singulären Exponent** von  $A(z)$ .

Es kann nun eine Aussage über die Asymptotik des optimalen Parameter  $x_n$  und den Koeffizienten  $a_n$  der erzeugenden Funktion  $A(z)$  getroffen werden. Dazu benötigt man die aus der Analysis bekannte Gammafunktion.

#### 4 Boltzmann-Sampler

**Definition 4.11.** Die **Gammafunktion**  $\Gamma(s)$  ist durch die Eulersche Integraldarstellung

$$\Gamma(s) = \int_0^{\infty} e^{-t} \cdot t^{s-1} dt$$

für  $\operatorname{Re}(s) > 0$  definiert.

**Satz 4.7.** Sei  $\mathcal{A}$  eine kombinatorische Struktur mit erzeugender Funktion  $A(z) = \sum_{k \geq 0} a_k z^k$ . Sei  $\rho$  die dominante Singularität von  $A(z)$ . Sei  $x_n$  die Lösung der Gleichung  $x_n \frac{A'(x_n)}{A(x_n)} = n$ . Wenn die Boltzmann-Verteilung der Struktur flach mit singulärem Exponenten  $-\alpha$  ist und  $m \in [an, bn]$  für Konstanten  $a$  und  $b$ , dann ist  $\rho(1 - \frac{\alpha}{n})$  eine Approximation des Parameters  $x_n$ , und es gilt

$$a_m \sim \frac{a_0}{\Gamma(\alpha)} \rho^{-m} m^{\alpha-1};$$

$$A(x_n) \sim a_0 \left(\frac{n}{\alpha}\right)^\alpha;$$

$$\frac{a_m x_n^m}{A(x_n)} \sim \frac{\alpha^\alpha e^{-\alpha}}{\Gamma(\alpha)} \cdot \left(\frac{m}{n}\right)^\alpha \cdot \frac{1}{m} = \Theta\left(\frac{1}{m}\right).$$

*Beweis.* Die erste asymptotische Äquivalenz kann durch Anwendung der Singularitätsanalyse gewonnen werden. Hier wird auf eine umfassende Literatur [4] verwiesen.

Um eine Approximation für  $x_n$  zu erhalten betrachtet man die Definition von flachen Boltzmann-Verteilungen. Dadurch kann der Erwartungswert approximiert werden mit

$$\mathbb{E}_x(N) \sim_{x \rightarrow \rho} \frac{\alpha \frac{x}{\rho}}{1 - \frac{x}{\rho}}$$

Dadurch ergibt sich die Asymptotik von  $x_n$ . Und weiters folgt nun

$$\frac{a_m x_n^m}{A(x_n)} \sim \frac{a_0}{\Gamma(\alpha)} \rho^{-m} m^{\alpha-1} \cdot \left(\rho \left(1 - \frac{\alpha}{n}\right)\right)^m \cdot \frac{1}{a_0} \cdot \left(\frac{n}{\alpha}\right)^{-\alpha} = \frac{\alpha^\alpha}{\Gamma(\alpha)} \cdot \frac{1}{m} \cdot \left(\frac{m}{n}\right)^\alpha \cdot \left(1 - \frac{\alpha}{n}\right)^m.$$

Wenn nun angenommen wird, dass  $m \in [an, bn]$  folgt die Behauptung. □ □

Für die Definition einer unebenen Boltzmann-Verteilung benötigt die Eigenschaft der Hayman-Zulässigkeit.

**Definition 4.12.** Sei  $f(z)$  analytisch um 0 mit Konvergenzradius  $R > 0$ . Weiters gilt  $A(z) > 0$  für  $z \in (R_0, R)$  für ein  $R_0 \in (0, R)$ . Definiere

$$\mu_1(x) := x \frac{A'(x)}{A(x)}, \quad ,$$

#### 4 Boltzmann-Sampler

$$\mu_2(x) := x^2 \frac{A''(x)}{A(x)} + x \frac{A'(x)}{A(x)},$$

$$\sigma^2(x) := \mu_2(x) - \mu_1(x)^2.$$

$A(z)$  heißt **Hayman-zulässig**, falls folgende Bedingungen erfüllt sind:

1. **H1:**  $\lim_{x \rightarrow R} \mu_1(x) = \infty$  und  $\lim_{x \rightarrow R} \sigma^2(x) = \infty$
2. **H2:** Es gibt eine Funktion  $\delta(x)$  auf  $(R_0, R)$  mit  $0 < \delta(x) < \pi$ , sodass für  $|\phi| \leq \delta(x)$  gleichmäßig gilt:

$$A(xe^{i\phi}) = A(x) e^{i\phi \mu_1(x) - \frac{(\phi^2 \sigma^2(x))}{2} + o(1)} \quad \text{für } x \rightarrow R$$

3. **H3:** Es gilt gleichmäßig für  $\delta(x) \leq |\phi| < \pi$

$$A(xe^{i\phi}) = o\left(\frac{A(x)}{\sigma(x)}\right) \quad \text{für } x \rightarrow R$$

**Definition 4.13.** Sei  $N$  die Größe eines zufälligen kombinatorischen Objekts unter der Boltzmann-Verteilung mit Parameter  $x$ . Sei  $\rho$  die dominante Singularität der erzeugenden Funktion. Definiere  $\mu_1(x) := \mathbb{E}_x(N)$  und  $\mu_2(x) := \mathbb{E}_x(N^2)$ . Die Varianz ist daher  $\sigma^2(x) = \mu_2(x) - \mu_1(x)^2$ . Dann nennt man die Boltzmann-Verteilung einer kombinatorischen Struktur  $\mathcal{A}$  **uneben**, wenn folgende Bedingungen gelten

1. Varianzbedingung:  $\lim_{x \rightarrow \rho} \frac{\sigma(x)}{\mu_1(x)} = 0$ ;
2.  $A(x)$  ist Hayman-zulässig.

Wieder lassen sich Aussagen über die Asymptotik der Koeffizienten treffen.

**Satz 4.8.** Sei  $\mathcal{A}$  eine kombinatorische Struktur mit erzeugender Funktion  $A(z)$ . Sei  $x_n$  die Lösung der Gleichung  $x_n \frac{A'(x_n)}{A(x_n)} = n$  und sei  $\rho$  die dominante Singularität von  $A(z)$ . Wenn die Boltzmann-Verteilung der Struktur uneben ist, dann gilt  $x_n \xrightarrow{n \rightarrow \infty} \rho^-$  und

$$a_m \sim \frac{A(x_m)}{x_m^m \sigma(x_m) \sqrt{2\pi}}$$

sowie

$$\frac{a_m x_n^m}{A(x_n)} = \frac{1}{\sigma(x_n) \sqrt{2\pi}} \exp\left(-\frac{(m-n)^2}{2\sigma(x_n)^2} + o(1)\right)$$

gleichmäßig für alle  $m \in \mathbb{N}$ .

*Beweis.* Wegen **H1** gilt  $\lim_{x \rightarrow \rho^-} \mu_1(x) = \infty$  und daraus folgt direkt, dass  $x_n \xrightarrow{n \rightarrow \infty} \rho^-$ . Wir zeigen nun die stärkere Aussage

$$\frac{a_m x_n^m}{A(x_n)} = \frac{1}{\sigma(x_n) \sqrt{2\pi}} \exp\left(-\frac{(m-n)^2}{2\sigma(x_n)^2} + o(1)\right).$$

#### 4 Boltzmann-Sampler

Die Aussage über die Asymptotik der Koeffizienten  $a_m$  folgt dann durch Umformungen daraus. Man wendet nun Satz 4.4 an und erhält

$$a_m x^m = \frac{x^m}{2\pi i} \int_{\Gamma} \frac{A(z)}{z^{m+1}} dz.$$

Als Weg  $\Gamma$  wählt man nun den Weg der durch die Kurve  $x e^{i\phi}$  wobei  $\phi \in [0, 2\pi)$  gegeben ist. Dies bedeutet, dass der Rand des Konvergenzkreis einmal in positiver Richtung durchlaufen wird. Wendet man die Definition des Kurvenintegrals an, erhält man

$$a_m x^m = \frac{x^m}{2\pi i} \int_0^{2\pi} \frac{A(xe^{i\phi}) \cdot ix e^{i\phi}}{x^{m+1} e^{i\phi(m+1)}} d\phi = \frac{1}{2\pi} \int_0^{2\pi} \frac{A(xe^{i\phi})}{e^{i\phi m}} d\phi.$$

Dieses Integral wird nun folgendermaßen aufgespalten und transformiert:

$$a_m x^m = \frac{1}{2\pi} \int_{-\delta}^{\delta} \frac{A(xe^{i\phi})}{e^{i\phi m}} d\phi + \frac{1}{2\pi} \int_{\delta}^{2\pi-\delta} \frac{A(xe^{i\phi})}{e^{i\phi m}} d\phi,$$

wobei man  $\delta := \delta(x)$  aus Bedingung **H2** wählt. Aufgrund von Bedingung **H3** kann man das letzte Integral abschätzen durch

$$\left| \frac{1}{2\pi} \int_{\delta}^{2\pi-\delta} \frac{A(xe^{i\phi})}{e^{i\phi m}} d\phi \right| \leq \frac{1}{2\pi} 2\pi \cdot o\left(\frac{A(x)}{\sigma(x)}\right) = o\left(\frac{A(x)}{\sigma(x)}\right).$$

Für das erste Integral verwendet man **H2** um es umzuformen.

$$\begin{aligned} \frac{1}{2\pi} \int_{-\delta}^{\delta} \frac{A(xe^{i\phi})}{e^{i\phi m}} d\phi &= \frac{1}{2\pi} \int_{-\delta}^{\delta} A(x) \exp\left(i(\mu_1(x) - m)\phi - \frac{1}{2}\sigma^2(x)\phi^2 + o(1)\right) d\phi \\ &= \frac{A(x)}{2\pi} \int_{-\delta}^{\delta} \exp\left(i(\mu_1(x) - m)\phi - \frac{1}{2}\sigma^2(x)\phi^2\right) (1 + o(1)) d\phi \\ &= \frac{A(x)}{2\pi} \left( \int_{-\delta}^{\delta} \exp\left(i(\mu_1(x) - m)\phi - \frac{1}{2}\sigma^2(x)\phi^2\right) d\phi + o\left(\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\sigma(x)^2\phi^2\right) d\phi\right) \right) \\ &= \frac{A(x)}{2\pi} \left( \int_{-\delta}^{\delta} \exp\left(i(\mu_1(x) - m)\phi - \frac{1}{2}\sigma^2(x)\phi^2\right) d\phi + o\left(\frac{1}{\sigma(x)}\right) \right). \end{aligned}$$

Die erste Gleichheit gilt aufgrund von Bedingung **H2**. Die zweite folgt aus  $\exp(o(1)) = 1 + o(1)$ , die dritte Gleichheit aus  $|\exp(i\phi a)| = 1$  und die letzte Gleichheit folgt aus  $\int_{-\infty}^{\infty} \exp\left(-\frac{t^2 a}{2}\right) dt = \sqrt{\frac{2\pi}{a}}$ . Dies ist leicht nachrechenbar und wird an dieser Stelle als bekannt angesehen.

#### 4 Boltzmann-Sampler

Weiters gelten die Bedingung **H2** und **H3** für  $\phi = \delta(x)$  beide. Deshalb folgt für  $\phi = \delta(x)$

$$A(x)e^{i\phi\mu_1(x) - \frac{(\phi^2\sigma^2(x))}{2} + o(1)} = o\left(\frac{A(x)}{\sigma(x)}\right).$$

Daraus folgt

$$\sigma(x)e^{i\phi\mu_1(x) - \frac{(\phi^2\sigma^2(x))}{2}} \rightarrow 0$$

und weiters

$$\delta(x)^2\sigma(x)^2 \rightarrow \infty \quad \text{für } x \rightarrow \rho.$$

Der Ausdruck  $\frac{\sigma(x)\phi}{\sqrt{2}}$  wird durch  $u$  substituiert. Es gilt dann  $d\phi = \frac{\sqrt{2}}{\sigma(x)}du$ . Für  $x \rightarrow \rho$  gilt dann:

$$\begin{aligned} a_m x^m &= \frac{A(x)}{2\pi} \int_{-\delta\frac{\sigma(x)}{\sqrt{2}}}^{\delta\frac{\sigma(x)}{\sqrt{2}}} \left( \frac{\sqrt{2}}{\sigma(x)} \exp\left(-u^2 + iu\frac{\sqrt{2}}{\sigma(x)}(\mu_1(x) - m)\right) du + o\left(\frac{1}{\sigma(x)}\right) \right) du \\ &= \frac{A(x)}{\sqrt{2}\sigma(x)\pi} \int_{-\delta\frac{\sigma(x)}{\sqrt{2}}}^{\delta\frac{\sigma(x)}{\sqrt{2}}} \left( \exp\left(-u^2 + iu\frac{\sqrt{2}}{\sigma(x)}(\mu_1(x) - m)\right) du + o(1) \right) du \\ &= \frac{A(x)}{\sqrt{2}\sigma(x)\pi} \left( \int_{-\infty}^{\infty} \exp\left(-\left(u - i\frac{1}{\sqrt{2}\sigma(x)}(\mu_1(x) - m)\right)^2 - \frac{(\mu_1(x) - m)^2}{2\sigma^2(x)}\right) du + o(1) \right) \\ &= \frac{A(x)}{\sqrt{2}\sigma(x)\pi} \left( \exp\left(-\frac{(\mu_1(x) - m)^2}{2\sigma^2(x)}\right) \int_{-\infty - i\frac{1}{\sqrt{2}\sigma(x)}(\mu_1(x) - m)}^{\infty - i\frac{1}{\sqrt{2}\sigma(x)}(\mu_1(x) - m)} \exp(-s^2) ds + o(1) \right) \\ &= \frac{A(x)}{\sqrt{2}\sigma(x)\pi} \left( \exp\left(-\frac{(\mu_1(x) - m)^2}{2\sigma^2(x)}\right) \sqrt{\pi} + o(1) \right). \end{aligned}$$

Die zweite Behauptung folgt direkt daraus. Da  $\lim_{m \rightarrow \infty} x_m = \rho$ , ersetzt man in der Gleichung  $x$  durch  $x_m$  und erhält

$$a_m \sim \frac{A(x_m)}{x_m^m \sigma(x_m) \sqrt{2\pi}}. \quad \square$$

Als dritte Kategorie wird die spitze Boltzmann-Verteilung betrachtet. Sie unterscheidet sich von der flachen durch ihren singulären Exponenten.

## 4 Boltzmann-Sampler

**Definition 4.14.** Sei  $N$  die Größe eines zufälligen kombinatorischen Objekts unter der Boltzmann-Verteilung mit Parameter  $x$ . Sei  $\rho$  die dominante Singularität der erzeugenden Funktion. Dann nennt man die Boltzmann-Verteilung einer kombinatorischen Struktur  $\mathcal{A}$  **spitz**, wenn die erzeugende Funktion  $A(z)$  analytisch an der Stelle 0 ist und die folgenden Bedingungen erfüllt:

1. Die einzige Singularität der Funktion  $A(z)$  auf dem Kreis  $\{z \in \mathbb{C} : |z| = \rho\}$  ist  $\rho$ , und es gibt ein  $r > \rho$  und ein  $\phi \in (0, \frac{\pi}{2})$ , sodass  $A(z)$  auf dem Gebiet  $\Delta_\rho(r, \phi)$  analytisch fortsetzbar ist.
2. Es existiert ein  $\alpha \in \mathbb{R}^- \setminus \{0, -1, -2, -3, \dots\}$ , sodass gilt:

$$A(z) \underset{z \rightarrow \rho^-}{\sim} a_0 \left(1 - \frac{z}{\rho}\right)^{-\alpha} + o\left(\left(1 - \frac{z}{\rho}\right)^{-\alpha}\right).$$

Da bei den spitzen Boltzmann-Verteilungen Objekte von kleiner Größe sehr wahrscheinlich auftreten, ist es bei der Wahl des Parameter sehr schwierig, denn normalerweise möchte man Objekte sehr großer Größe erhalten. Um dies zu umgehen, können solche Verteilungen durch Anwendung des Auszeichnungsoperators zu flachen Boltzmann-Verteilungen transformiert werden.

Der Operator des Auszeichnens wurde in Kapitel 2.2.6. behandelt. Im weiteren Verlauf wird die erzeugende Funktion der Struktur  $\Theta\mathcal{A}$  mit  $A^\bullet(x)$  bezeichnet. Es wurde gezeigt, dass für sie gilt:

$$A^\bullet(x) = xA'(x).$$

Die Erzeugung eines zufälligen Objektes von  $\mathcal{A}$  und  $\Theta\mathcal{A}$  unterscheidet sich durch Fallenlassen der Auszeichnung. Da für jedes Objekt von  $\mathcal{A}$  der Größe  $n$  aber  $n|\mathcal{A}_n|$  Objekte von  $\Theta\mathcal{A}$  der Größe  $n$  existieren sind die Boltzmann-Verteilungen aber nicht ident.

Erfüllt die erzeugende Funktion einer Struktur die Bedingungen einer spitzen Boltzmann-Verteilung mit Exponent  $-\alpha$ , dann erfüllt die erzeugende Funktion der ausgezeichneten Struktur ebenfalls die Bedingungen, aber mit singulärem Exponent  $-\alpha - 1$ .

Dies bedeutet, wenn man die Struktur auszeichnet, verringert sich der singuläre Exponent und die spitze Boltzmann-Verteilung kann zu einer flachen Boltzmann-Verteilung werden. Dadurch ergeben sich erhebliche Performanceverbesserungen. Um nun zur eigentlich gewünschten Struktur zurückzukommen, werden vom erzeugten Objekt einfach alle Auszeichnungen verworfen. Dass man tatsächlich für jede Struktur, deren Beschreibung den Auszeichnungs-Operator enthält, auch eine Beschreibung angeben kann, die ohne diesen Operator auskommt, zeigt folgender Satz.

**Satz 4.9.** Sei  $\mathcal{A}$  eine zerlegbare kombinatorische Struktur, dessen Beschreibung den Operator des Auszeichnens nicht enthält. Dann ist  $\Theta\mathcal{A}$  eine zerlegbare kombinatorische Struktur, die eine Beschreibung besitzt, die den Operator des Auszeichnens ebenfalls nicht enthält.

*Beweis.* Für den Auszeichnungsoperator gelten bei unmarkierten Strukturen offensichtlich folgende Bezie-



## 4 Boltzmann-Sampler

hungen zu den anderen vorgestellten Operatoren:

$$\begin{aligned}\Theta(\mathcal{A} + \mathcal{B}) &= \Theta\mathcal{A} + \Theta\mathcal{B}, \\ \Theta(\mathcal{A} \times \mathcal{B}) &= \Theta\mathcal{A} \times \mathcal{B} + \mathcal{A} \times \Theta\mathcal{B}, \\ \Theta(\text{SEQ}(\mathcal{A})) &= \text{SEQ}(\mathcal{A}) \times \Theta\mathcal{A} \times \text{SEQ}(\mathcal{A}), \\ \Theta(\text{SET}(\mathcal{A})) &= \Theta\mathcal{A} \times \text{SET}(\mathcal{A}), \\ \Theta(\text{MSET}(\mathcal{A})) &= \Theta\mathcal{A} \times \text{MSET}(\mathcal{A}).\end{aligned}$$

Für den Auszeichnungsoperator gelten bei markierten Strukturen offensichtlich folgende Beziehungen zu den anderen vorgestellten Operatoren:

$$\begin{aligned}\Theta(\mathcal{A} + \mathcal{B}) &= \Theta\mathcal{A} + \Theta\mathcal{B}, \\ \Theta(\mathcal{A} \star \mathcal{B}) &= \Theta\mathcal{A} \star \mathcal{B} + \mathcal{A} \star \Theta\mathcal{B}, \\ \Theta(\text{SEQ}(\mathcal{A})) &= \text{SEQ}(\mathcal{A}) \star \Theta\mathcal{A} \star \text{SEQ}(\mathcal{A}), \\ \Theta(\text{SET}(\mathcal{A})) &= \Theta\mathcal{A} \star \text{SET}(\mathcal{A}), \\ \Theta(\text{CYC}(\mathcal{A})) &= \Theta\mathcal{A} \star \text{CYC}(\mathcal{A}).\end{aligned}$$

Gegeben ist nun eine Beschreibung der kombinatorischen Struktur  $\mathcal{A} = (\mathcal{A}^{(1)}, \dots, \mathcal{A}^{(n)})$  mit

$$\mathcal{A}^{(i)} = \Psi_i(\square, \mathcal{A}^{(1)}, \dots, \mathcal{A}^{(n)}),$$

wobei  $\Psi_i$  ein Term ist, der aus Atomen  $\square$ , neutralen Objekten, den Strukturen  $\mathcal{A}^{(j)}$  und den in Kapitel 2 vorgestellten Konstruktionen ohne der Konstruktion des Auszeichnens aufgebaut ist. In Vektorschreibweise schreibt man kurz  $\mathcal{A} = \Psi(\square, \mathcal{A})$ .

Dann erhält man eine Beschreibung der kombinatorischen Struktur  $\Theta\mathcal{A}$ , indem man die Gleichungen  $\mathcal{A}^{(i)} = \Psi_i(\square, \mathcal{A}^{(1)}, \dots, \mathcal{A}^{(n)})$  übernimmt und zusätzlich folgende Gleichung in die Beschreibung aufnimmt:

$$\Theta\mathcal{A}^{(i)} = \Theta\Psi_i(\square, \mathcal{A}^{(1)}, \dots, \mathcal{A}^{(n)}, \Theta\mathcal{A}^{(1)}, \dots, \Theta\mathcal{A}^{(n)}).$$

Dabei werden  $\Theta\mathcal{A}^{(i)}$  als neue Strukturvariablen eingeführt. Um diese noch auszeichnungsfrei zu machen, werden die obigen Regeln solange angewandt, bis der Auszeichnungsoperator nur noch für Variablen angewandt wird. □

## 4.2 Auswertung der erzeugenden Funktion

Der Unterschied der Boltzmann-Sampler zur rekursiven Methode besteht darin, dass die Folge der Anzahlen der Objekte der Größe  $n$  nicht berechnet werden müssen. Stattdessen ist es aber notwendig die vorkommenden erzeugenden Funktionen an einer bestimmten Stelle  $x$  auszuwerten. Meist lassen sich die erzeugenden Funktion explizit angeben und somit auch leicht berechnen. Oft sind die erzeugenden Funktionen aber nur

durch eine Gleichung gegeben. In diesem Fall müssen die Funktionswerte approximiert werden. Hier werden zwei Methoden vorgestellt, wie dies realisiert werden kann.

### 4.2.1 Iterative Auswertung

Gegeben sei im folgenden immer eine Beschreibung einer kombinatorischen Struktur  $\mathcal{A} = \Psi(\square, \mathcal{A})$ .

Dabei sind einige Anforderungen an die Beschreibung gestellt, um diese Methode anwenden zu können. Dazu werden die nötigen Begriffe definiert. In [20] werden diese, zusammen mit dem Verfahren der iterativen Auswertung, vorgestellt.

**Definition 4.15.** Die Beschreibung  $\mathcal{A} = \Psi(\square, \mathcal{A})$  heißt *sinnvoll*, wenn sich für alle  $n \geq 0$  nur endlich viele kombinatorische Objekte der Größe  $n$  ableiten lassen.

**Definition 4.16.** Sei  $\Psi(\square, \mathcal{A}^{(1)}, \dots, \mathcal{A}^{(n)})$  eine Beschreibung. Dann ist die Ableitung  $\frac{\partial \Psi}{\partial \mathcal{A}^{(i)}}$  der kombinatorischen Struktur rekursiv definiert durch folgende Regeln

- $\frac{\partial \{\epsilon\}}{\partial \mathcal{A}^{(i)}} = \emptyset,$
- $\frac{\partial \{\square\}}{\partial \mathcal{A}^{(i)}} = \emptyset,$
- $\frac{\partial \mathcal{A}^{(i)}}{\partial \mathcal{A}^{(i)}} = \emptyset, \quad i \neq j,$
- $\frac{\partial \mathcal{A}^{(i)}}{\partial \mathcal{A}^{(i)}} = \{\mathcal{A}_*^{(i)}\},$  wobei  $\mathcal{A}_*^{(i)}$  ein neues Objekt der Größe 0 ist,
- $\frac{\partial (\mathcal{B}_1 + \mathcal{B}_2)}{\partial \mathcal{A}^{(i)}} = \frac{\partial \mathcal{B}_1}{\partial \mathcal{A}^{(i)}} + \frac{\partial \mathcal{B}_2}{\partial \mathcal{A}^{(i)}},$
- $\frac{\partial (\mathcal{B}_1 \times \mathcal{B}_2)}{\partial \mathcal{A}^{(i)}} = \frac{\partial \mathcal{B}_1}{\partial \mathcal{A}^{(i)}} \times \mathcal{B}_2 + \mathcal{B}_1 \times \frac{\partial \mathcal{B}_2}{\partial \mathcal{A}^{(i)}},$
- $\frac{\partial \text{SEQ}(\mathcal{B})}{\partial \mathcal{A}^{(i)}} = \text{SEQ}(\mathcal{B}) \times \frac{\partial \mathcal{B}}{\partial \mathcal{A}^{(i)}} \times \text{SEQ}(\mathcal{B}),$
- $\frac{\partial \text{SEQ}_{<k}(\mathcal{B})}{\partial \mathcal{A}^{(i)}} = \sum_{j=0}^{k-1} \text{SEQ}_{<j}(\mathcal{B}) \times \frac{\partial \mathcal{B}}{\partial \mathcal{A}^{(i)}} \times \text{SEQ}_{<k-1-j}(\mathcal{B}),$
- $\frac{\partial \text{CYC}(\mathcal{B})}{\partial \mathcal{A}^{(i)}} = \text{SEQ}(\mathcal{B}) \times \frac{\partial \mathcal{B}}{\partial \mathcal{A}^{(i)}},$
- $\frac{\partial \text{CYC}_{<k}(\mathcal{B})}{\partial \mathcal{A}^{(i)}} = \text{SEQ}_{<k-1}(\mathcal{B}) \times \frac{\partial \mathcal{B}}{\partial \mathcal{A}^{(i)}},$
- $\frac{\partial \text{SET}(\mathcal{B})}{\partial \mathcal{A}^{(i)}} = \text{SET}(\mathcal{B}) \times \frac{\partial \mathcal{B}}{\partial \mathcal{A}^{(i)}},$
- $\frac{\partial \text{SET}_{<k}(\mathcal{B})}{\partial \mathcal{A}^{(i)}} = \text{SET}_{<k-1}(\mathcal{B}) \times \frac{\partial \mathcal{B}}{\partial \mathcal{A}^{(i)}}.$

Falls  $n = 1$  wird die Ableitung mit  $\Psi'$  bezeichnet.

#### 4 Boltzmann-Sampler

Es kann durch Induktion gezeigt werden, dass die Ableitung einer kombinatorischen Struktur genau ein neues Objekt der Größe 0 enthält. Somit kann das kartesische Produkt von zwei kombinatorischen Struktur als Ableitung aufgefasst werden. Es gilt nämlich die Isomorphie

$$\frac{\partial \Psi}{\partial \mathcal{A}^{(i)}} \times \mathcal{B} \sim \frac{\partial \Psi}{\partial \mathcal{A}^{(i)}} \Big|_{\mathcal{A}^{(i)} = \mathcal{B}}.$$

**Definition 4.17.** Die **Jacobi-Matrix**  $\frac{\partial \Psi}{\partial \mathcal{A}}$  ist definiert als Matrix mit den Einträgen  $\frac{\partial \Psi_i}{\partial \mathcal{A}^{(i)}}$ .

Eine Matrix  $A$  von kombinatorischen Strukturen ist **nilpotent**, wenn eine natürliche Zahl  $n$  existiert, sodass  $A^n$  eine Matrix ist, deren Einträge alle  $\emptyset$  ist. Die kleinste solche Zahl definiert die **Ordnung** der Nilpotenz.

Mithilfe der Jacobi-Matrix lässt sich prüfen ob eine kombinatorische Struktur sinnvoll ist.

**Satz 4.10.** Sei  $\mathcal{A} = \Psi(\square, \mathcal{A})$  eine kombinatorische Beschreibung für die gilt, dass  $\Psi(\emptyset, \emptyset) = \emptyset$ . Dann ist  $\mathcal{A}$  genau dann sinnvoll, wenn die Jacobi-Matrix  $\frac{\partial \Psi}{\partial \mathcal{A}}(\emptyset, \emptyset)$  nilpotent ist.

*Beweis.* Sei angenommen es existiert mindestens eine natürliche Zahl  $n$ , sodass unendliche viele kombinatorische Objekte der Struktur mit Größe  $n$  existieren. Gibt es mehrere solcher Größen, dann wähle die kleinste. Wählt man nun ein beliebiges Objekt  $\alpha = (\alpha_1, \dots, \alpha_n)$  der Struktur  $\mathcal{A}$  der Größe  $n$ . Dies bedeutet, dass  $\alpha_i$  ein Objekt der Struktur  $\mathcal{A}_i$  ist und die Größe von  $\alpha_i$  maximal  $n$  sein kann. Gibt es ein  $\alpha_i$ , sodass  $|\alpha_i| = n$ , dann gilt für  $j \neq i$ , dass  $|\alpha_j| = 0$ . Dies bedeutet aber, dass  $\alpha_j = \emptyset$  und daher ist  $\alpha$  ein Objekt von  $\frac{\partial \Psi}{\partial \mathcal{A}}(\emptyset, \emptyset) \times \beta$ , wobei  $\beta$  ein Vektor von Objekten von  $\mathcal{A}$  der Größe  $n$  ist. Die Struktur  $\frac{\partial \Psi}{\partial \mathcal{A}}(\emptyset, \emptyset)$  kann nicht öfter iteriert werden als die Ordnung der Nilpotenz und deswegen können nur endlich viele Objekte der Größe  $n$  existieren.

Gibt es kein  $\alpha_i$ , sodass  $|\alpha_i| = n$ , dann gibt es nur eine endliche Anzahl von solchen Zerlegungen und damit nur eine endliche Anzahl von Objekten der Größen  $n$ .

Wenn  $\alpha$  ein Objekt der Struktur  $\mathcal{A}$  der Größe  $n$  ist, wobei  $n$  die kleinste natürliche Zahl ist, sodass es mindestens ein Objekt der Größe  $n$  gibt und die Matrix  $\frac{\partial \Psi}{\partial \mathcal{A}}(\emptyset, \emptyset)$  nicht nilpotent ist, dann gibt es für alle natürliche Zahlen  $p$  ein nichtleeres Objekt  $\beta$  der Struktur  $\left(\frac{\partial \Psi}{\partial \mathcal{A}}(\emptyset, \emptyset)\right)^p$  der Größe 0, sodass  $\beta \cdot \alpha$  ein Objekt von  $\mathcal{A}$  der Größe  $n$  ist. Dies ist jedoch ein Widerspruch zur endlichen Anzahl an Objekten der Größe  $n$ .  $\square$

Ist die kombinatorische Beschreibung  $\mathcal{A} = \Psi(\square, \mathcal{A})$  sinnvoll, dann konvergiert die Iteration  $\mathcal{A}^{[n+1]} = \Psi(\square, \mathcal{A}^{[n]})$  wobei  $\mathcal{A}^{[0]} = \emptyset$  gegen die Lösung  $\mathcal{A}$ . Die Definition von Konvergenz von kombinatorischen Strukturen wird im folgenden gegeben.

**Definition 4.18.** Zwei kombinatorische Strukturen  $\mathcal{A}$  und  $\mathcal{B}$  besitzen **Kontakt der Ordnung**  $k$ , wenn alle Objekte der Größe  $\leq k$  ident sind. Man schreibt dafür  $\mathcal{A} =_k \mathcal{B}$ .

**Definition 4.19.** Eine Folge von kombinatorischen Strukturen  $(\mathcal{A}^{[n]})_{n \in \mathbb{N}}$  **konvergiert** gegen eine kombinatorische Struktur  $\mathcal{A}$ , wenn

$$\forall k \geq 0 \exists N \geq 0 \forall n \geq N : \mathcal{A}^{[n]} =_k \mathcal{A}.$$

#### 4 Boltzmann-Sampler

**Satz 4.11.** Sei  $\mathcal{A} = \Psi(\square, \mathcal{A})$  eine sinnvolle kombinatorische Beschreibung mit  $\Psi(\emptyset, \emptyset) = \emptyset$ . Die Iteration  $\mathcal{A}^{[n+1]} = \Psi(\square, \mathcal{A}^{[n]})$ , wobei  $\mathcal{A}^{[0]} = \emptyset$ , konvergiert gegen die Lösung der Gleichung  $\mathcal{A} = \Psi(\square, \mathcal{A})$ .

Der Satz ist eine Folgerung des impliziten Spezientheorems. Dieses wird in [10] bewiesen.

**Satz 4.12.** Für die Iteration  $\mathcal{A}^{[n+1]} = \Psi(\square, \mathcal{A}^{[n]})$ , wobei  $\mathcal{A}^{[0]} = \emptyset$ , gilt  $\mathcal{A}^{[n]} \subset \mathcal{A}^{[n+1]}$ .

*Beweis.* Dies wird mit vollständiger Induktion bewiesen.

**Induktionsanfang:** Es gilt  $\mathcal{A}^{[0]} = \emptyset$  und  $\mathcal{A}^{[1]} = \Psi(\square, \emptyset)$ . Da  $\Psi(\square, \emptyset)$  ein Vektor von Termen ist, die aus Atomen, neutralen Objekten und den in Kapitel 2 vorgestellten Konstruktionen aufgebaut ist, kann die Struktur nicht leer sein.

**Induktionsschritt:**  $\mathcal{A}^{[n+1]} = \Psi(\square, \mathcal{A}^{[n]}) \subset \Psi(\square, \mathcal{A}^{[n+1]}) = \mathcal{A}^{[n+2]}$ , da mit Induktionsannahme  $\mathcal{A}^{[n]} \subset \mathcal{A}^{[n+1]}$  gilt. □

Die Aussage über die Konvergenz kombinatorischer Strukturen lässt sich nun überführen in eine Konvergenzaussage für die erzeugenden Funktionen.

Für die kombinatorische Struktur lässt sich mithilfe der Regeln für die erzeugenden Funktionen der Konstruktionen, die in Kapitel 2 vorgestellt wurden, eine Gleichung für die erzeugende Funktion der Struktur aufstellen:  $A(z) = \psi(z, A(z))$ .

**Satz 4.13.** Sei  $\mathcal{A} = \Psi(\square, \mathcal{A})$  eine sinnvolle kombinatorische Beschreibung mit  $\Psi(\emptyset, \emptyset) = \emptyset$ . Die Iteration  $A^{[n+1]}(z) = \psi(z, A^{[n]}(z))$  mit  $A^{[0]}(z) = 0$  konvergiert gegen die erzeugende Funktion  $A(z) = \sum_{n=0}^{\infty} a_n z^n$  der Struktur  $\mathcal{A}$ .

*Beweis.* Aufgrund der in Kapitel 2 vorgestellten Regeln für die erzeugenden Funktionen gilt, dass  $A^{[n]}(z)$  die erzeugende Funktion der kombinatorischen Struktur  $\mathcal{A}^{[n]}$  ist. Sei  $\text{ord}(A(z))$  der kleinste Index  $i$ , sodass  $a_i \neq 0$ .

Aus der Konvergenz der Strukturen folgt durch Übersetzung in erzeugende Funktionen  $\text{ord}(A^{[n]}(z) - A(z)) \rightarrow \infty$ . Daraus folgt dann die Konvergenz der Potenzreihen. □

Wählt man nun einen konkreten Parameter  $z$ , der im Konvergenzkreis liegt, dann konvergiert die Folge der ausgewerteten erzeugenden Funktion gegen den Wert der ausgewerteten erzeugenden Funktion der Struktur.

**Satz 4.14.** Sei  $\mathcal{A} = \Psi(\square, \mathcal{A})$  eine sinnvolle kombinatorische Beschreibung mit  $\Psi(\emptyset, \emptyset) = \emptyset$  und sodass die erzeugende Funktion  $A(z) = \sum_{n=0}^{\infty} a_n z^n$  von  $\mathcal{A}$  analytisch ist. Dann hat die erzeugende Funktion  $A(z)$  positiven Konvergenzradius  $R$ . Sei  $|\tilde{z}| < R$ , dann konvergiert die Iteration  $\alpha^{[n+1]} = \psi(\tilde{z}, \alpha^{[n]})$  mit  $\alpha^{[0]} = 0$  gegen den Wert  $A(\tilde{z})$ .

#### 4 Boltzmann-Sampler

Ein Beweis wird in [20] geführt.

Mithilfe dieser Sätze lässt sich durch Iteration die erzeugende Funktion an der Stelle  $x$  numerisch berechnen. Je mehr Iterationsschritte durchgeführt werden, desto genauer ist der errechnete Wert.

**Beispiel:** Anhand der ebenen Binärbäume wird diese iterative Auswertung demonstriert. Die Beschreibung dieser lautet  $\mathcal{A} = \{\square\} + \{\square\} \star \mathcal{A} \star \mathcal{A}$ . Zuerst wird die kombinatorische Iteration veranschaulicht. Dabei werden die ersten vier Iterationsschritte visualisiert.

$$\mathcal{A}^{[0]} = \emptyset$$


---

$$\mathcal{A}^{[1]} = \square$$


---

$$\mathcal{A}^{[2]} = \square \quad \begin{array}{c} \square \\ / \quad \backslash \\ \square \quad \square \end{array}$$


---

$$\mathcal{A}^{[3]} = \square \quad \begin{array}{c} \square \\ / \quad \backslash \\ \square \quad \square \end{array} \quad \begin{array}{c} \square \\ / \quad \backslash \\ \square \quad \square \\ / \quad \backslash \\ \square \quad \square \end{array} \quad \begin{array}{c} \square \\ / \quad \backslash \\ \square \quad \square \\ \backslash \quad / \\ \square \quad \square \end{array} \quad \begin{array}{c} \square \\ / \quad \backslash \\ \square \quad \square \\ / \quad \backslash \\ \square \quad \square \\ / \quad \backslash \\ \square \quad \square \end{array}$$


---

$$\mathcal{A}^{[4]} = \square \quad \begin{array}{c} \square \\ / \quad \backslash \\ \square \quad \square \end{array} \quad \begin{array}{c} \square \\ / \quad \backslash \\ \square \quad \square \\ / \quad \backslash \\ \square \quad \square \end{array} \quad \begin{array}{c} \square \\ / \quad \backslash \\ \square \quad \square \\ \backslash \quad / \\ \square \quad \square \end{array} \quad \begin{array}{c} \square \\ / \quad \backslash \\ \square \quad \square \\ / \quad \backslash \\ \square \quad \square \\ \backslash \quad / \\ \square \quad \square \end{array} \quad \begin{array}{c} \square \\ / \quad \backslash \\ \square \quad \square \\ / \quad \backslash \\ \square \quad \square \\ / \quad \backslash \\ \square \quad \square \end{array} \quad \begin{array}{c} \square \\ / \quad \backslash \\ \square \quad \square \\ \backslash \quad / \\ \square \quad \square \end{array} \quad \begin{array}{c} \square \\ / \quad \backslash \\ \square \quad \square \\ / \quad \backslash \\ \square \quad \square \\ / \quad \backslash \\ \square \quad \square \end{array}$$
  

$$\begin{array}{c} \square \\ / \quad \backslash \\ \square \quad \square \\ / \quad \backslash \\ \square \quad \square \\ / \quad \backslash \\ \square \quad \square \end{array} \quad \begin{array}{c} \square \\ / \quad \backslash \\ \square \quad \square \\ / \quad \backslash \\ \square \quad \square \\ \backslash \quad / \\ \square \quad \square \end{array} \quad \begin{array}{c} \square \\ / \quad \backslash \\ \square \quad \square \\ / \quad \backslash \\ \square \quad \square \\ / \quad \backslash \\ \square \quad \square \end{array} \quad \begin{array}{c} \square \\ / \quad \backslash \\ \square \quad \square \\ / \quad \backslash \\ \square \quad \square \\ \backslash \quad / \\ \square \quad \square \end{array} \quad \begin{array}{c} \square \\ / \quad \backslash \\ \square \quad \square \\ / \quad \backslash \\ \square \quad \square \\ \backslash \quad / \\ \square \quad \square \end{array} \quad \begin{array}{c} \square \\ / \quad \backslash \\ \square \quad \square \\ / \quad \backslash \\ \square \quad \square \\ \backslash \quad / \\ \square \quad \square \end{array}$$
  

$$\begin{array}{c} \square \\ / \quad \backslash \\ \square \quad \square \\ / \quad \backslash \\ \square \quad \square \\ / \quad \backslash \\ \square \quad \square \end{array} \quad \begin{array}{c} \square \\ / \quad \backslash \\ \square \quad \square \\ / \quad \backslash \\ \square \quad \square \\ \backslash \quad / \\ \square \quad \square \end{array} \quad \dots$$


---

#### 4 Boltzmann-Sampler

Aufgrund der in Kapitel 2 vorgestellten Sätze für erzeugende Funktion gilt folgende Gleichung für die erzeugende Funktion der Struktur der ebenen Binärbäume:

$$A(z) = z + z(A(z))^2.$$

Die Iteration der erzeugenden Funktion aus Satz 4.12 lautet für die ebenen Binärbäume also

$$A^{[n+1]}(z) = z + z(A^{[n]}(z))^2, \quad A^{[0]} = 0.$$

Im Folgenden werden die ersten fünf Iterationsschritte angegeben.

$$A^{[0]}(z) = 0,$$

$$A^{[1]}(z) = z,$$

$$A^{[2]}(z) = z^3 + z,$$

$$A^{[3]}(z) = z + z(z^3 + z)^2 = z^7 + 2z^5 + z^3 + z,$$

$$A^{[4]}(z) = z + z(z^7 + 2z^5 + z^3 + z)^2 = z^{15} + 4z^{13} + 6z^{11} + 6z^9 + 5z^7 + 2z^5 + z^3 + z,$$

$$\begin{aligned} A^{[5]}(z) &= z + z(z^{15} + 4z^{13} + 6z^{11} + 6z^9 + 5z^7 + 2z^5 + z^3 + z)^2 = \\ &= z^{31} + 8z^{29} + 28z^{27} + 60z^{25} + 94z^{23} + 116z^{21} + 114z^{19} + 94z^{17} + 69z^{15} + 44z^{13} + 26z^{11} + 14z^9 + \\ &\quad + 5z^7 + 2z^5 + z^3 + z. \end{aligned}$$

Wenn man die Anzahlen der ebenen Binärbäume mit  $n$  Knoten, die in der Struktur  $\mathcal{A}^{[n]}$  enthalten sind betrachtet, ist zu erkennen, dass die erzeugende Funktion der Struktur  $\mathcal{A}^{[n]}$  genau die Funktion  $A^{[n]}(z)$  ist. Als letzter Schritt um die erzeugende Funktion der Struktur auszuwerten wird die Iteration von Satz 4.13. angewendet. Als Beispiel möchte man die erzeugende Funktion an der Stelle  $\tilde{z} = 0.2$  auswerten. Dies ist möglich, da sich  $\tilde{z}$  im Konvergenzkreis befindet. Es werden wieder die ersten fünf Schritte der Iteration berechnet.

$$\alpha^{[0]} = 0,$$

$$\alpha^{[1]} = 0.2,$$

$$\alpha^{[2]} = 0.208,$$

$$\alpha^{[3]} = 0.2086528,$$

$$\alpha^{[4]} = 0.208707198,$$

$$\alpha^{[5]} = 0.208711739.$$

Hier zeigt sich bereits nach fünf Schritten eine gute Näherung des Funktionswertes. Je öfter man die Iteration laufen lässt, desto genauer wird dieser Wert, wenn man voraussetzt, dass sich der Wert  $\tilde{z}$  innerhalb des Konvergenzkreises befindet.

Es gibt aber auch noch eine andere Methode, die eine erzeugende Funktion an einer bestimmten Stelle numerisch auswertet. Diese konvergiert im Allgemeinen schneller als die eben vorgestellte Iteration.

#### 4.2.2 Newton-Methode

Die klassische Newton-Methode zur Näherung der Nullstellen einer Funktion wird in [20] auf kombinatorische Gleichungen umgelegt um eine Näherung der ausgewerteten erzeugenden Funktion zu erhalten. Mit der Newton-Methode lässt sich im Allgemeinen eine quadratische Konvergenz erreichen.

**Satz 4.15.** Sei  $\mathcal{A} = \Psi(\square, \mathcal{A})$  eine sinnvolle kombinatorische Beschreibung mit  $\Psi(\emptyset, \emptyset) = \emptyset$ . Man betrachte die im Folgenden definierte Funktion  $\mathcal{N}_\Psi(\square, \mathcal{A})$ .

$$\mathcal{N}_\Psi(\square, \mathcal{A}) := \mathcal{A} + \left( I - \frac{\partial \Psi}{\partial \mathcal{A}}(\square, \mathcal{A}) \right)^{-1} \times (\Psi(\square, \mathcal{A}) - \mathcal{A})$$

Dann konvergiert die Iteration  $\mathcal{A}^{[0]} = \emptyset, \mathcal{A}^{[n+1]} = \mathcal{N}_\Psi(\square, \mathcal{A}^{[n]})$  für  $n \geq 1$  gegen  $\mathcal{A}$ , und die Konvergenz ist quadratisch.

Um diesen Satz zu beweisen, wird zuerst gezeigt, dass die Newton-Iteration wohldefiniert und eindeutig ist. Dies geschieht mit den nächsten beiden Sätzen.

**Satz 4.16.** Die Iteration  $\mathcal{A}^{[0]} = \emptyset, \mathcal{A}^{[n+1]} = \mathcal{N}_\Psi(\square, \mathcal{A}^{[n]})$  für  $n \geq 1$  ist wohldefiniert für alle  $n \in \mathbb{N}$ .

*Beweis.* Um zu zeigen, dass die Iteration wohldefiniert ist, muss die Inklusion  $\mathcal{A}^{[n]} \subset \Psi(\square, \mathcal{A}^{[n]})$  für alle  $n \in \mathbb{N}$  gezeigt werden, denn nur so ist die Subtraktion möglich. Dies erfolgt mit vollständiger Induktion.

Induktionsanfang: Für  $n = 0$  ist  $\mathcal{A}^{[0]}$  die leere Struktur und somit Teilmenge von  $\Psi(\square, \emptyset)$ .

Induktionsschritt: Angenommen es gilt  $\mathcal{A}^{[n]} \subset \Psi(\square, \mathcal{A}^{[n]})$ . Aufgrund der Taylorformel in [13] gilt

$$\Psi(\mathcal{A} + \mathcal{B}) \supset \Psi(\mathcal{A}) + \frac{\partial \Psi}{\partial \mathcal{A}}(\mathcal{A}) \times \mathcal{B}.$$

Daraus folgt

$$\Psi(\square, \mathcal{A}^{[n+1]}) \supset \Psi(\square, \mathcal{A}^{[n]}) + \frac{\partial \Psi}{\partial \mathcal{A}}(\square, \mathcal{A}^{[n]}) \times (\mathcal{A}^{[n+1]} - \mathcal{A}^{[n]})$$

Laut Labelle [12] gilt folgende Identität über die Inverse der Jacobi-Matrix:

$$\left( I - \frac{\partial \Psi}{\partial \mathcal{A}}(\square, \mathcal{A}) \right)^{-1} = \sum_{k \geq 0} \left( \frac{\partial \Psi}{\partial \mathcal{A}}(\square, \mathcal{A}) \right)^k.$$

#### 4 Boltzmann-Sampler

Deshalb gilt weiters

$$\begin{aligned} \Psi(\square, \mathcal{A}^{[n+1]}) &\supset \mathcal{A}^{[n]} + \left( \Psi(\square, \mathcal{A}^{[n]}) - \mathcal{A}^{[n]} \right) + \frac{\partial \Psi}{\partial \mathcal{A}}(\square, \mathcal{A}^{[n]}) \times \sum_{k \geq 0} \left( \frac{\partial \Psi}{\partial \mathcal{A}}(\square, \mathcal{A}^{[n]}) \right)^k \times (\Psi(\square, \mathcal{A}^{[n]}) - \mathcal{A}^{[n]}) \\ &\supset \mathcal{A}^{[n]} + \sum_{k \geq 0} \left( \frac{\partial \Psi}{\partial \mathcal{A}}(\square, \mathcal{A}^{[n]}) \right)^k \times (\Psi(\square, \mathcal{A}^{[n]}) - \mathcal{A}^{[n]}) = \mathcal{A}^{[n+1]}. \quad \square \end{aligned}$$

**Satz 4.17.** Die Iteration  $\mathcal{A}^{[0]} = \emptyset, \mathcal{A}^{[n+1]} = \mathcal{N}_{\Psi}(\square, \mathcal{A}^{[n]})$  für  $n \geq 1$  ist eindeutig für alle  $n \in \mathbb{N}$ .

*Beweis.* Die Iteration ist eindeutig, wenn für alle Strukturen  $\mathcal{B} \subset \Psi(\square, \mathcal{B})$  gilt, dass  $\mathcal{B}$  und  $\left( \frac{\partial \Psi}{\partial \mathcal{A}}(\square, \mathcal{B}) \right)^k \times (\Psi(\square, \mathcal{B}) - \mathcal{B}), k \in \mathbb{N}$  paarweise disjunkt sind. Dies kann mit vollständiger Induktion gezeigt werden. Induktionsanfang: Für  $k = 0$  ist dies klar, denn  $\mathcal{B}$  und  $(\Psi(\square, \mathcal{B}) - \mathcal{B})$  sind definitionsgemäß disjunkt. Induktionsschritt: Angenommen die Aussage gilt für  $k$ . Dann muss dies auch für  $k + 1$  gelten. Denn jedes Objekt aus  $\left( \frac{\partial \Psi}{\partial \mathcal{A}}(\square, \mathcal{B}) \right)^{(k+1)} \times (\Psi(\square, \mathcal{B}) - \mathcal{B})$  lässt sich eindeutig auf zwei Objekte aufteilen, sodass eines der Objekte aus der Struktur  $\frac{\partial \Psi}{\partial \mathcal{A}}(\square, \mathcal{B})$  und eines aus der Struktur  $\left( \frac{\partial \Psi}{\partial \mathcal{A}}(\square, \mathcal{B}) \right)^k \times (\Psi(\square, \mathcal{B}) - \mathcal{B})$  ist. Da das Objekt aus  $\left( \frac{\partial \Psi}{\partial \mathcal{A}}(\square, \mathcal{B}) \right)^k \times (\Psi(\square, \mathcal{B}) - \mathcal{B})$  nicht in  $\Psi(\square, \mathcal{B})$  enthalten ist, kann das Objekt aus  $\left( \frac{\partial \Psi}{\partial \mathcal{A}}(\square, \mathcal{B}) \right)^{(k+1)} \times (\Psi(\square, \mathcal{B}) - \mathcal{B})$  somit auch nicht in der Struktur  $\Psi(\square, \mathcal{B})$  sein. Da  $\mathcal{B} \subset \Psi(\square, \mathcal{B})$  ist das Objekt somit nicht in der Struktur  $\mathcal{B}$  enthalten.  $\square$

Die Newton-Iteration hat den Vorteil gegenüber der iterativen Auswertung, dass sie quadratische Konvergenz ermöglicht. Dies bedeutet der Kontakt der kombinatorischen Strukturen verdoppelt sich mit jedem Schritt.

*Beweis. (von Satz 4.15)*

Angenommen zwei in der Iteration aufeinanderfolgende Strukturen besitzen Kontakt  $k$ , d.h.  $\mathcal{A}^{[n-1]} =_k \mathcal{A}^{[n]}$ . Weiters betrachtet man ein Objekt  $\alpha$  der Struktur  $\mathcal{A}^{[n+1]} - \mathcal{A}^{[n]}$ , dessen Größe höchstens  $2k + 1$  beträgt. Also muss  $\alpha$  ein Objekt der Struktur  $\left( \frac{\partial \Psi}{\partial \mathcal{A}}(\square, \mathcal{A}^{[n]}) \right)^l \times (\Psi(\square, \mathcal{A}^{[n]}) - \mathcal{A}^{[n]})$  für ein  $l \in \mathbb{N}$  sein. Genau eines der Objekte  $\beta$  aus  $\mathcal{A}^{[n]}$ , aus denen  $\alpha$  besteht besitzt eine Größe  $> k$ . Wenn dies auf mehr als nur ein Objekt zutreffen würde, dann wäre die Größe von  $\alpha > 2k + 1$ , und dies ist ein Widerspruch zur Annahme. Wenn kein Objekt größer als  $k$  wäre, dann wären alle diese Objekte, die  $\alpha$  zerlegen, Objekte von  $\mathcal{A}^{[n-1]}$  und dadurch müsste  $\alpha$  ein Objekt von  $\mathcal{A}^{[n]}$  sein. Laut Annahme ist aber  $\alpha \in \mathcal{A}^{[n+1]} - \mathcal{A}^{[n]}$ . Dieses Objekt  $\beta$  muss in der Struktur  $\Psi(\square, \mathcal{A}^{[n]})$  sein. Wäre dies nicht der Fall, dann müsste das letzte Objekt  $\gamma$ , aus dem  $\alpha$  besteht, aus der Struktur  $\Psi(\square, \mathcal{A}^{[n-1]}) \subset \mathcal{A}^{[n]}$  sein. Auch  $\beta$  muss aus einer Struktur  $\left( \frac{\partial \Psi}{\partial \mathcal{A}}(\square, \mathcal{A}^{[n-1]}) \right)^{l'} \times (\Psi(\square, \mathcal{A}^{[n-1]}) - \mathcal{A}^{[n-1]})$  für ein  $l' \in \mathbb{N}$  sein. Und deshalb ist  $\gamma \in \left( \frac{\partial \Psi}{\partial \mathcal{A}}(\square, \mathcal{A}^{[n-1]}) \right)^{l'+1} \times (\Psi(\square, \mathcal{A}^{[n-1]}) - \mathcal{A}^{[n-1]}) \subset \mathcal{A}^{[n]}$ . Dies ergibt einen Widerspruch,



#### 4 Boltzmann-Sampler

und deshalb existiert kein Objekt der Größe  $\leq 2k + 1$  aus der Struktur  $\mathcal{A}^{[n+1]} - \mathcal{A}^{[n]}$ . Dies ist äquivalent zu  $\mathcal{A}^{[n+1]} =_{2k+1} \mathcal{A}^{[n]}$ .

Nun muss noch gezeigt werden, dass die Iteration gegen die Struktur  $\mathcal{C}$  strebt, die die Gleichung  $\mathcal{A} = \Psi(\square, \mathcal{A})$  löst. Durch Umformen der Iteration ergibt sich für  $\mathcal{A}^{[n+1]} - \mathcal{A}^{[n]}$

$$\mathcal{A}^{[n+1]} - \mathcal{A}^{[n]} = \Psi(\square, \mathcal{A}^{[n]}) - \mathcal{A}^{[n]} + \frac{\partial \Psi}{\partial \mathcal{A}}(\square, \mathcal{A}^{[n]}) \times (\mathcal{A}^{[n+1]} - \mathcal{A}^{[n]}).$$

Weil sich aber für zwei aufeinanderfolgende Strukturen  $\mathcal{A}^{[n+1]}$  und  $\mathcal{A}^{[n]}$  der Kontakt immer verdoppelt muss  $\mathcal{A}^{[n+1]} - \mathcal{A}^{[n]}$  gegen  $\emptyset$  streben. Und wegen obiger Gleichheit muss auch  $\Psi(\square, \mathcal{A}^{[n]}) - \mathcal{A}^{[n]}$  gegen  $\emptyset$  streben.  $\square$

Nun wurde gezeigt, dass die angegebene Iteration eine Folge von kombinatorischen Strukturen erzeugt, die tatsächlich quadratisch gegen die Lösung der Gleichung  $\mathcal{A} = \Psi(\square, \mathcal{A})$  konvergiert. Was für den Fall der Auswertung der erzeugenden Funktion einer kombinatorischen Struktur aber wichtig ist, ist die Konvergenz der erzeugenden Funktionen der Strukturen.

**Satz 4.18.** *Sei  $\mathcal{A} = \Psi(\square, \mathcal{A})$  eine sinnvolle, analytische Beschreibung mit  $\Psi(\emptyset, \emptyset) = \emptyset$ . Sei  $\tilde{z}$  innerhalb der Konvergenzkreisscheibe der erzeugenden Funktion  $A(z)$  von  $\mathcal{A}$ . Dann konvergiert folgende Iteration gegen  $A(\tilde{z})$ :*

$$\alpha^{[n+1]} = \alpha^{[n]} + \left( I - \frac{\partial \psi}{\partial A}(\tilde{z}, \alpha^{[n]}) \right)^{-1} \cdot (\psi(\tilde{z}, \alpha^{[n]}) - \alpha^{[n]}), \quad \alpha^{[0]} = 0.$$

*Beweis.* Aufgrund von Satz 4.15 ist bekannt, dass der Grenzwert der Folge  $\mathcal{A}^{[n]}$ , die durch die Iteration  $\mathcal{A}^{[0]} = \emptyset, \mathcal{A}^{[n+1]} = \mathcal{N}_{\Psi}(\square, \mathcal{A}^{[n]})$  für  $n \geq 1$  gegeben ist, gleich der Lösung der Gleichung  $\mathcal{A} = \Psi(\square, \mathcal{A})$  ist. Da die Beschreibung  $\mathcal{A} = \Psi(\square, \mathcal{A})$  sinnvoll ist, gibt es nur endlich viele Objekte in der Struktur  $\mathcal{A}^{[n]}$ , für alle  $n \in \mathbb{N}$ . Deshalb ist auch die Beschreibung  $\mathcal{A} = \mathcal{N}_{\Psi}(\square, \mathcal{A})$  sinnvoll. Die Analytizität überträgt sich von  $\Psi$  auf  $\mathcal{N}_{\Psi}$ . Nun kann man Satz 4.14 auf  $\mathcal{N}_{\Psi}$  anwenden und erhält, dass die gegebene Iteration für  $\alpha^{[n]}$  gegen die an der Stelle  $\tilde{z}$  ausgewertete erzeugende Funktion von  $\mathcal{A}$  konvergiert.  $\square$

**Beispiel:** Die Newton-Methode wird wieder anhand der ebenen Binärbäume demonstriert. Die Iteration der kombinatorischen Strukturen, wie in Satz 4.15 angegeben, wird für die ersten drei Schritte visualisiert. Die Iteration ist im Fall der ebenen Binärbäume gegeben durch

$$\mathcal{A}^{[n+1]} = \mathcal{A}^{[n]} + \text{SEQ}(\{\square\} \times \mathcal{A}^{[n]}) \times (\{\square\} + \{\square\} \times \mathcal{A}^{[n]} - \mathcal{A}^{[n]}), \quad \mathcal{A}^{[0]} = \emptyset.$$

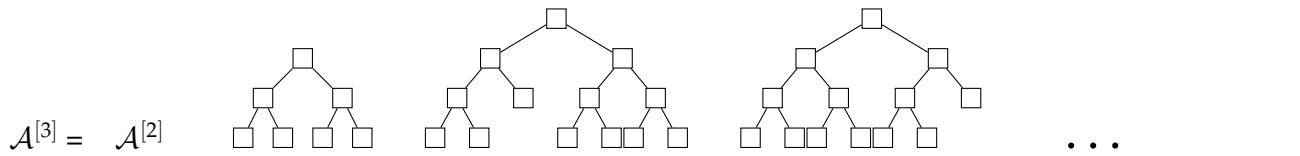
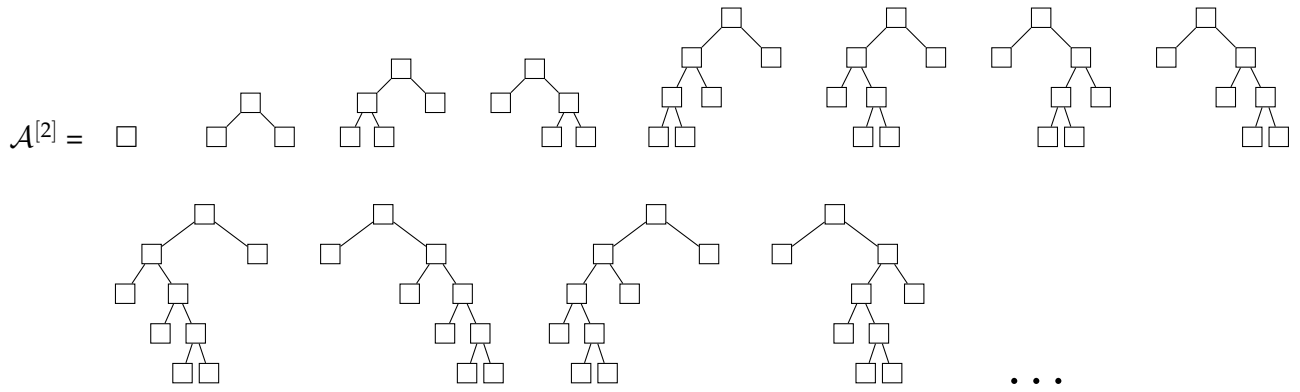
#### 4 Boltzmann-Sampler

$$\mathcal{A}^{[0]} = \emptyset$$


---

$$\mathcal{A}^{[1]} = \square$$


---



Um die erzeugende Funktion der ebenen Binärbäume zu generieren werden die ersten 3 Iterationsschritte der folgenden Iteration berechnet.

$$A^{[n+1]}(z) = A^{[n]}(z) + \left( I - \frac{\partial \psi}{\partial A}(z, A^{[n]}(z)) \right)^{-1} \left( \psi(z, A^{[n]}(z)) - A^{[n]}(z) \right), \quad A^{[0]}(z) = 0.$$

Nach Einsetzen der entsprechenden Funktionen für ebene Binärbäume erhält man dann folgende Rekursion.

$$A^{[n+1]}(z) = A^{[n]}(z) + \frac{1}{1 - 2zA^{[n]}(z)} \left( z + z(A^{[n]}(z))^2 - A^{[n]}(z) \right), \quad A^{[0]}(z) = 0.$$

Nun erhält man für die ersten vier Schritte folgende erzeugende Funktionen:

$$A^{[0]}(z) = 0,$$

$$A^{[1]}(z) = z,$$

$$A^{[2]}(z) = z + \frac{z^3}{1-2z^2} = z + z^3 \sum_{k=0}^{\infty} (2z^2)^k = z + z^3 + 2z^5 + 4z^7 + 8z^9 + \dots,$$

$$A^{[3]}(z) = z + z^3 + 2z^5 + 5z^7 + 14z^9 + 42z^{11} + \dots$$

Für die Auswertung der erzeugenden Funktion ist diese Iteration wie auch bei der iterativen Methode nicht

## 4 Boltzmann-Sampler

notwendig. Es soll damit aber demonstriert werden, dass hier die erzeugenden Funktion keine endlichen Polynome sind, wie es bei der iterativen Methode der Fall ist, sondern schon ab dem zweiten Iterationsschritt eine unendliche Reihe entsteht. Die numerische Auswertung der erzeugenden Funktion mit den Wert  $\tilde{z}$  lässt sich nun mit der Iteration aus Satz 4.18 realisieren. Die Iteration angewendet auf die ebenen Binärbäume ist dann definiert durch

$$\alpha^{[n+1]} = \alpha^{[n]} + \frac{\tilde{z} + \tilde{z}(\alpha^{[n]})^2 - \alpha^{[n]}}{1 - 2\tilde{z}\alpha^{[n]}}, \quad \alpha^{[0]} = 0.$$

Dadurch ergeben sich die folgenden ersten vier Iterationsschritte:

$$\begin{aligned}\alpha^{[0]} &= 0, \\ \alpha^{[1]} &= 0.2, \\ \alpha^{[2]} &= 0.208695652, \\ \alpha^{[3]} &= 0.208712152, \\ \alpha^{[4]} &= 0.208712153.\end{aligned}$$

Hier ist bereits zu erkennen, dass bei der Newton-Methode die Iteration deutlich schneller konvergiert. Der vierte Schritt unterscheidet sich erst in der neunten Nachkommastelle.

### 4.3 Gewöhnliche Boltzmann-Sampler

Es ist nun noch nicht gezeigt, dass solche Boltzmann-Sampler tatsächlich konstruiert werden können. Mit Hilfe der in Kapitel 2 vorgestellten Konstruktionen ist dies möglich, indem man die Basiskonstruktionen in Boltzmann-Sampler transformiert. Diese Prozeduren wurden erstmals in [16] eingeführt. Man muss nun zwischen unmarkierten und markierten Strukturen unterscheiden. Zuerst werden die unmarkierten Strukturen behandelt, bei denen die gewöhnlichen erzeugenden Funktionen angewendet werden.

Bei Strukturen, die nur eine endliche Menge an Objekten beinhalten, bedarf es keiner aufwendigen Theorie, um Zufallsobjekte zu erzeugen. Dennoch benötigt man diese öfter, um daraus komplexere Konstrukte aufzubauen.

Die einfachsten dieser endlichen Strukturen sind  $\{\epsilon\}$  und  $\{\square\}$ . Hier ist es ganz einfach, einen Boltzmann-Sampler dafür anzugeben.

$$\bullet \mathcal{A} = \{\epsilon\} \quad \Rightarrow \quad A(x) = 1$$

---

```
1: procedure  $\Gamma A(x)$ 
2:   return  $\epsilon$ 
3: end procedure
```

## 4 Boltzmann-Sampler

**Satz 4.19.** *Der obige Algorithmus ist ein Boltzmann-Sampler der Struktur, die ausschließlich das leere Objekt enthält.*

*Beweis.*

$$\mathbb{P}_x(\epsilon) = \frac{x^{|\epsilon|}}{A(x)} = \frac{x^0}{1} = 1$$

Der Algorithmus liefert mit Wahrscheinlichkeit 1 das leere Objekt zurück. □

•  $\mathcal{A} = \{\square\} \quad \Rightarrow \quad A(x) = x$

```

1: procedure  $\Gamma A(x)$ 
2:   return  $\square$ 
3: end procedure

```

**Satz 4.20.** *Der obige Algorithmus ist ein Boltzmann-Sampler der Struktur, die ausschließlich ein Atom enthält.*

*Beweis.*

$$\mathbb{P}_x(\square) = \frac{x^{|\square|}}{A(x)} = \frac{x^1}{x} = 1$$

Der Algorithmus liefert mit Wahrscheinlichkeit 1 das Atom zurück. □

Andere unmarkierte kombinatorische Strukturen sind aus den Basiskonstruktionen aufgebaut. Für gegebene kombinatorische Strukturen  $\mathcal{B}$  und  $\mathcal{C}$  mit deren Boltzmann-Sampler  $\Gamma B$  und  $\Gamma C$  können Boltzmann-Sampler  $\Gamma(B + C)$ ,  $\Gamma(B \times C)$ ,  $\Gamma\text{SEQ}(B)$ ,  $\Gamma\text{SET}(B)$  und  $\Gamma\text{MSET}(B)$  konstruiert werden.

- **Disjunkte Vereinigung:**  $\mathcal{A} = \mathcal{B} + \mathcal{C}$

$\text{BERNOULLI}(p)^2$  stellt eine Funktion dar, die eine Bernoulli-verteilte Zufallszahl mit Parameter  $p$  generiert.

```

1: procedure  $\Gamma A(x)$ 
2:   if  $\text{BERNOULLI}(\frac{B(x)}{A(x)}) = 1$  then
3:     return  $\Gamma B(x)$ 
4:   else
5:     return  $\Gamma C(x)$ 

```

<sup>2</sup> Im Anhang findet man Algorithmen zur Erzeugung von bernoulli-, geometrisch-, poisson- und logarithmisch-verteilten Zufallszahlen

## 4 Boltzmann-Sampler

6: **end if**  
7: **end procedure**

---

**Satz 4.21.** Seien  $\mathcal{B}$  und  $\mathcal{C}$  unmarkierte kombinatorische Strukturen, und seien Boltzmann-Sampler für  $\mathcal{B}$  und  $\mathcal{C}$  gegeben. Dann ist der obige Algorithmus ein Boltzmann-Sampler von  $\mathcal{B} + \mathcal{C}$ .

*Beweis.* Gemäß Satz 2.1 gilt  $A(x) = B(x) + C(x)$ . Betrachte ein zufälliges Objekt von  $\mathcal{A}$  der Boltzmann-Verteilung mit Parameter  $x$ . Die Wahrscheinlichkeit, dass ein Objekt aus  $\mathcal{B}$  stammt ist

$$\sum_{\beta \in \mathcal{B}} \frac{x^{|\beta|}}{A(x)} = \sum_{n \geq 0} \frac{B_n x^n}{A(x)} = \frac{B(x)}{A(x)}$$

Da der Algorithmus mithilfe einer Bernoulli-verteilten Zufallszahl mit Parameter  $\frac{B(x)}{A(x)}$  realisiert wird entspricht dies der gewünschten Wahrscheinlichkeit.  $\square$

- **Kartesisches Produkt:**  $\mathcal{A} = \mathcal{B} \times \mathcal{C}$

---

1: **procedure**  $\Gamma A(x)$   
2:     **return**  $(\Gamma B(x), \Gamma C(x))$   
3: **end procedure**

---

**Satz 4.22.** Seien  $\mathcal{B}$  und  $\mathcal{C}$  unmarkierte kombinatorische Strukturen, und seien Boltzmann-Sampler für  $\mathcal{B}$  und  $\mathcal{C}$  gegeben. Dann ist der obige Algorithmus ein Boltzmann-Sampler von  $\mathcal{B} \times \mathcal{C}$ .

*Beweis.* Im Kapitel 2 wurde bewiesen, dass  $A(x) = B(x) \cdot C(x)$ . Die Wahrscheinlichkeit von  $\alpha \in \mathcal{A}$  im Boltzmann-Modell ist nun

$$\mathbb{P}_x(\alpha) = \frac{x^{|\alpha|}}{A(x)} = \frac{x^{|\beta|}}{B(x)} \frac{x^{|\gamma|}}{C(x)} = \mathbb{P}_x(\beta) \cdot \mathbb{P}_x(\gamma)$$

Dies zeigt nun, dass die unabhängige Ausführung der Boltzmann-Sampler von  $\mathcal{B}$  und  $\mathcal{C}$  einen Boltzmann-Sampler für  $\mathcal{A}$  ergibt.  $\square$

- **Folge:**  $\mathcal{A} = \text{SEQ}(\mathcal{B})$

Der Boltzmann-Sampler der Folgen lässt sich auf zwei Arten lösen. Die erste Methode ist eine Rekursive.

#### 4 Boltzmann-Sampler

```

1: procedure  $\Gamma A(x)$ 
2:   if  $\text{BERNOULLI}(B(x)) = 1$  then
3:     return  $(\Gamma B(x), \Gamma A(x))$ 
4:   else
5:     return  $\epsilon$ 
6:   end if
7: end procedure

```

---

**Satz 4.23.** *Sei  $\mathcal{B}$  eine unmarkierte kombinatorische Struktur, und sei ein Boltzmann-Sampler für  $\mathcal{B}$  gegeben. Dann ist der obige Algorithmus ein Boltzmann-Sampler von  $\text{SEQ}(\mathcal{B})$ .*

*Beweis.* Laut Definition ist die Struktur  $\text{SEQ}(\mathcal{B})$  gegeben durch  $\{\epsilon\} + \mathcal{B} + \mathcal{B} \times \mathcal{B} + \dots$ . Dies entspricht der rekursiven Struktur  $\mathcal{A} = \{\epsilon\} + \mathcal{B} \times \mathcal{A}$ . Da Boltzmann-Sampler für  $\{\epsilon\}$ , für  $\mathcal{B}$ , für das kartesische Produkt und für die Summe bereits bekannt sind, kann man diese zusammenfügen, um einen Boltzmann-Sampler der Struktur der Folge zu erhalten. Dabei ist der Parameter der Bernoulli-Verteilung aufgrund von Satz 4.21  $\frac{A(x) \cdot B(x)}{A(x)} = B(x)$ .  $\square$

Die andere Methode benutzt die geometrische Verteilung um die Folgenlänge zu bestimmen. Dabei sei  $\text{GEOM}(p)$  eine Funktion die eine geometrisch-verteilte Zufallszahl mit Parameter  $p$  erzeugt.

---

```

1: procedure  $\Gamma A(x)$ 
2:    $k := \text{GEOM}(B(x))$ 
3:   return  $\underbrace{(\Gamma B(x), \dots, \Gamma B(x))}_{k\text{-Tupel}}$ 
4: end procedure

```

---

**Satz 4.24.** *Sei  $\mathcal{B}$  eine unmarkierte kombinatorische Struktur und sei ein Boltzmann-Sampler für  $\mathcal{B}$  gegeben. Dann ist der obige Algorithmus ein Boltzmann-Sampler von  $\text{SEQ}(\mathcal{B})$ .*

*Beweis.* Für die erzeugende Funktion von  $A(x)$  gilt  $A(x) = \frac{1}{1-B(x)}$ . Die Wahrscheinlichkeit eine Folge  $\alpha = (\beta_1, \beta_2, \dots, \beta_n)$  mit  $n$  Objekten aus  $\mathcal{B}$  zu erhalten ist deshalb

$$\begin{aligned}
 \mathbb{P}_x(\alpha) &= \frac{x^{|\alpha|}}{A(x)} = (1 - B(x)) \cdot x^{|\beta_1| + \dots + |\beta_n|} = B(x)^n \cdot (1 - B(x)) \cdot \frac{x^{|\beta_1|}}{B(x)} \cdots \frac{x^{|\beta_n|}}{B(x)} = \\
 &= B(x)^n \cdot (1 - B(x)) \cdot \mathbb{P}_x(\beta_1) \cdots \mathbb{P}_x(\beta_n)
 \end{aligned}$$

#### 4 Boltzmann-Sampler

Damit lässt sich der Boltzmann-Sampler auch mithilfe einer geometrisch-verteilten Zufallszahl erzeugen. □

- **Multimenge:**  $\text{MSET}(A)$

Der Boltzmann Sampler verwendet eine Funktion  $\text{MAXINDEX}(x)$ , die einen zufälligen maximalen Index unter folgender diskreten Verteilung liefert.

$$\mathbb{P}_x(K \leq k) = \frac{1}{A(x)} \prod_{j \leq k} \exp\left(\frac{B(x^j)}{j}\right)$$

Wie solch eine Zufallszahl generiert werden kann, findet man unter anderem in [3]. Weiters sei  $\text{POISSON}(p)$  eine Funktion, die eine poisson-verteilte Zufallszahl mit Parameter  $p$  erzeugt.  $\text{POISSON}_{\geq 1}(p)$  liefert eine poisson-verteilte Zufallszahl, die nicht Null ist.

Folgender Algorithmus liefert einen Boltzmann Sampler für Multimengen:

```

1: procedure  $\Gamma A(x)$ 
2:    $\alpha := \emptyset$ 
3:    $k := \text{MAXINDEX}(x)$ 
4:   for  $j$  from 1 to  $k-1$  do
5:      $k_j := \text{POISSON}\left(\frac{B(x^j)}{j}\right)$ 
6:      $\beta_j := \Gamma B(x^j)$ 
7:     for  $i$  from 1 to  $k_j$  do
8:        $\alpha := \alpha \cup \{\beta_j\}$ 
9:     end for
10:  end for
11:   $k_k := \text{POISSON}_{\geq 1}\left(\frac{B(x^k)}{k}\right)$ 
12:   $\beta_k := \Gamma B(x^k)$ 
13:  for  $i$  from 1 to  $k_k$  do
14:     $\alpha := \alpha \cup \{\beta_k\}$ 
15:  end for
16:  return  $\alpha$ 
17: end procedure

```

**Satz 4.25.** Sei  $\mathcal{B}$  eine unmarkierte kombinatorische Struktur, und sei ein Boltzmann-Sampler für  $\mathcal{B}$  gegeben. Dann ist der obige Algorithmus ein Boltzmann-Sampler von  $\text{MSET}(\mathcal{B})$ .

*Beweis.* Wie zuvor bewiesen, lässt sich die erzeugende Funktion einer kombinatorischen Struktur

#### 4 Boltzmann-Sampler

$\mathcal{A} = \text{MSET}(\mathcal{B})$ , die durch die Konstruktion einer Multimenge definiert ist, ausdrücken durch

$$A(x) = \exp\left(\sum_{k \geq 1} \frac{B(x^k)}{k}\right).$$

Für alle Elemente  $\beta$  von  $\mathcal{B}$  kann eine zufällige Anzahl an Kopien erstellt werden, die geometrisch verteilt mit Parameter  $x^{|\beta|}$  sind. Dies bedeutet, der Boltzmann-Sampler würde lauten:

$$\Gamma \mathcal{A}(x) = \prod_{\beta \in \mathcal{B}} \beta^{\text{Geom}(x^{|\beta|})}$$

Geometrisch verteilte Zufallsvariablen lassen sich erzeugen durch Generierung von poisson-verteilten Zufallsvariablen mit  $\text{GEOM}(x^{|\beta|}) = \sum_{i \geq 1} i \text{POISSON}\left(\frac{x^{|\beta|}}{i}\right)$ . Deshalb gilt weiters

$$\Gamma \mathcal{A}(x) = \prod_{\beta \in \mathcal{B}} \beta^{\sum_{i \geq 1} i \text{POISSON}\left(\frac{x^{|\beta|}}{i}\right)} = \prod_{i \geq 1} \prod_{\beta \in \mathcal{B}} \beta^{i \text{POISSON}\left(\frac{x^{|\beta|}}{i}\right)} = \prod_{i \geq 1} \underbrace{\prod_{\gamma \in \{(\beta, \beta, \dots, \beta)\}}}_{i\text{-mal}} \gamma^{\text{POISSON}\left(\frac{x^{|\beta|}}{i}\right)}.$$

Der Algorithmus für  $\prod_{\beta \in \mathcal{B}} \beta^{\text{POISSON}(cx^{|\beta|})}$  wird durch Erzeugung einer poisson-verteilten Zufallszahl mit Parameter  $cB(x)$  realisiert.

Die Wahrscheinlichkeit, dass  $\prod_{\beta \in \mathcal{B}} \beta^{\text{POISSON}(cx^{|\beta|})}$  eine Multimenge  $\{\beta_1^{r_1}, \beta_2^{r_2}, \dots, \beta_s^{r_s}\}$  generiert, ist

$$\prod_{i=1}^s \frac{(cx^{|\beta_i|})^{r_i}}{r_i!} \prod_{\beta \in \mathcal{B}} e^{-cx^{|\beta|}} = c^{\sum_{i=1}^s r_i} x^{\sum_{i=1}^s r_i |\beta_i|} e^{-cB(x)} \prod_{i=1}^s \frac{1}{r_i!}.$$

Erzeugt man eine poisson-verteilte Zufallszahl mit Parameter  $cB(x)$ , dann wird die Wahrscheinlichkeit, dass die Multimenge  $\{\beta_1^{r_1}, \beta_2^{r_2}, \dots, \beta_s^{r_s}\}$  generiert wird, errechnet, indem man eine Folge  $\beta_1, \dots, \beta_{\sum_{i=1}^s r_i}$

erzeugt und diese auf  $\binom{\sum_{i=1}^s r_i}{r_1, \dots, r_s}$  Arten umordnet. Die Folge  $\beta_1, \dots, \beta_{\sum_{i=1}^s r_i}$  hat die Wahrscheinlichkeit

$$e^{-cB(x)} \frac{(cB(x))^{\sum_{i=1}^s r_i}}{(\sum_{i=1}^s r_i)!} \frac{x^{|\beta_1|}}{B(x)} \dots \frac{x^{|\beta_{\sum_{i=1}^s r_i}|}}{B(x)}.$$

Damit der Algorithmus nicht endlos läuft, wird vorher der maximale Index berechnet. □

Für jede zerlegbare Struktur findet sich nun ein Boltzmann-Sampler. Dieser besteht dann aus möglicherweise



## 4 Boltzmann-Sampler

rekursiven Aufrufen der Boltzmann-Sampler der Basiskonstruktionen.

### Beispiel: ebene Wurzelbäume

Als Beispiel wird die Struktur der unmarkierten ebenen Wurzelbäume angeführt. Dieser wird definiert durch die Beschreibung  $\{\mathcal{A} = \{\square\} \times \text{SEQ}(\mathcal{A})\}$ . Sei  $A(x)$  die erzeugende Funktion der Struktur. Dann lässt sich mithilfe der Beschreibung eine Gleichung dafür aufstellen.

$$A(x) = \frac{x}{1 - A(x)}.$$

Daraus lässt sich für die erzeugende Funktion und die Anzahl der ebenen Wurzelbäume eine explizite Darstellung errechnen.

$$A(x) = \frac{1 - \sqrt{1 - 4x}}{2},$$
$$A_n = \frac{1}{n} \binom{2n - 2}{n - 1}.$$

Zuerst werden zwei unabhängige Sampler aufgerufen, nämlich der Boltzmann-Sampler für  $\{\square\}$  und für  $\text{SEQ}(\mathcal{A})$ . Da  $\{\square\}$  eine Menge mit nur einem Objekt ist, wird vom Sampler  $\square$  zurückgegeben. Für den Sampler von  $\text{SEQ}(\mathcal{A})$  wird eine geometrisch-verteilte Zufallszahl  $k$  generiert und dann ein  $k$ -Tupel erzeugt, das den Sampler von  $\mathcal{A}$  rekursiv aufruft.

---

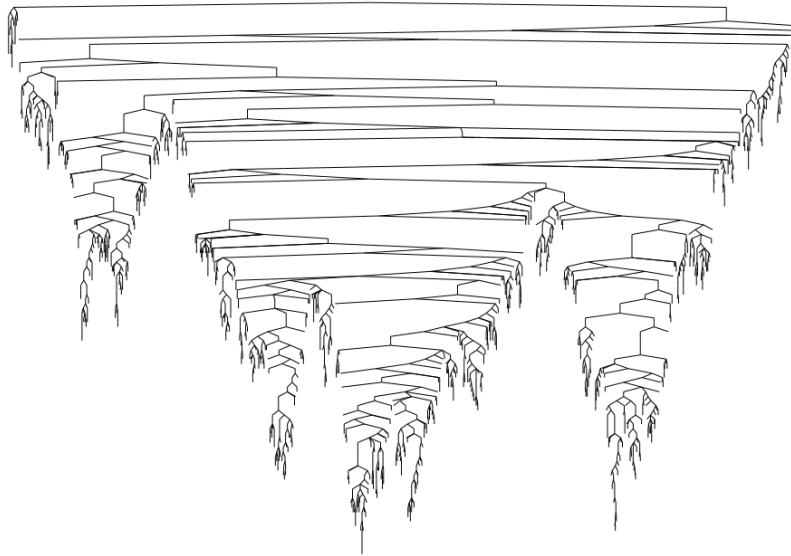
```
1: procedure  $\Gamma A(x)$ 
2:    $k := \text{GEOM}(\frac{1 - \sqrt{1 - 4x}}{2})$ 
3:   return  $(\square, \Gamma A(x), \dots, \Gamma A(x))$ 
            $\underbrace{\hspace{10em}}_{(k+1)\text{-Tupel}}$ 
4: end procedure
```

---

## 4.4 Exponentielle Boltzmann-Sampler

Für markierte kombinatorische Strukturen, die mithilfe der Konstruktionen der Vereinigung, des Produkts, der Folge, der Menge und der Zyklen aufgebaut sind, lassen sich auf die gleiche Weise Boltzmann-Sampler konstruieren. Hierbei werden exponentielle erzeugende Funktionen anstatt der gewöhnlichen verwendet. Die Sampler für die Strukturen  $\{\epsilon\}$ ,  $\{\square\}$  und die der disjunkten Vereinigung bleiben unverändert.

## 4 Boltzmann-Sampler



**Abbildung 4.5:** Beispiel eines durch den Boltzmann-Sampler erzeugten Wurzelbaum

- **Produkt:**  $\mathcal{A} = \mathcal{B} \star \mathcal{C}$

Das Produkt von markierten Strukturen ist analog zum kartesischen Produkt von unmarkierten Strukturen definiert. Dabei ersetzt lediglich das Produkt das kartesische Produkt. Da die erzeugende Funktion von  $\mathcal{A}$  auch das Produkt der erzeugenden Funktionen von  $\mathcal{B}$  und  $\mathcal{C}$  ist, besteht der Boltzmann-Sampler von  $\Gamma A(x)$  aus zwei unabhängigen Aufrufen von  $\Gamma B(x)$  und  $\Gamma C(x)$ .

Zu beachten sind dabei die Markierungen, die danach noch zufällig und konsistent auf die Atome verteilt werden müssen.

- **Folge:**  $\mathcal{A} = \text{SEQ}(\mathcal{B})$

Da Folgen aus Vereinigung und Produkt aufgebaut sind, bleibt auch der Boltzmann-Sampler für Folgen zu dem für unmarkierte Strukturen unverändert.

- **Menge:**  $\mathcal{A} = \text{SET}(\mathcal{B})$

Der Boltzmann-Sampler der Menge einer markierten kombinatorischen Struktur wird durch folgenden Algorithmus realisiert.

---

```
1: procedure  $\Gamma A(x)$ 
2:    $k := \text{POISSON}(B(x))$ 
3:   return  $(\underbrace{\Gamma B(x), \dots, \Gamma B(x)}_{k\text{-Tupel}})$ 
4: end procedure
```

---

## 4 Boltzmann-Sampler

**Satz 4.26.** Sei  $\mathcal{B}$  eine markierte kombinatorische Struktur und sei ein Boltzmann-Sampler für  $\mathcal{B}$  gegeben. Dann ist der obige Algorithmus ein Boltzmann-Sampler von  $\text{SET}(\mathcal{B})$ .

*Beweis.* In Satz 2.9 wurde gezeigt, dass für die erzeugende Funktion von  $\mathcal{A}$  gilt  $A(x) = \exp(B(x))$ . Deshalb ist die Wahrscheinlichkeit, dass ein Element  $\alpha$  von  $\mathcal{A}$  der Boltzmann-Verteilung gewählt wird

$$\mathbb{P}_x(\alpha) = \frac{1}{A(x)} \frac{x^{|\alpha|}}{|\alpha|!} = e^{-B(x)} \frac{x^{|\alpha|}}{|\alpha|!},$$

und die Wahrscheinlichkeit, dass die gewählte Menge aus  $k$  Elementen besteht daher

$$\frac{1}{A(x)} \frac{B(x)^k}{k!} = e^{-B(x)} \frac{B(x)^k}{k!}.$$

Damit lässt sich der Boltzmann-Sampler für Mengen mithilfe der Poisson-Verteilung lösen. Da  $e^{-B(x)} \frac{B(x)^k}{k!}$  die Wahrscheinlichkeit ist, dass eine poisson-verteilte Zufallszahl mit Parameter  $B(x)$  den Wert  $k$  annimmt. □

• **Zyklen:**  $\mathcal{A} = \text{CYC}(\mathcal{B})$

Der Boltzmann-Sampler wird mithilfe der Logarithmus-Verteilung realisiert. Dabei bezeichnet  $\text{LOGARITH}(p)$  eine Funktion, die eine logarithmisch verteilte Zufallszahl mit Parameter  $p$  erzeugt.

---

```

1: procedure  $\Gamma A(x)$ 
2:    $k := \text{LOGARITH}(B(x))$ 
3:   return  $(\underbrace{\Gamma B(x), \dots, \Gamma B(x)}_{k\text{-Tupel}})$ 
4: end procedure

```

---

**Satz 4.27.** Sei  $\mathcal{B}$  eine markierte kombinatorische Struktur, und sei ein Boltzmann-Sampler für  $\mathcal{B}$  gegeben. Dann ist der obige Algorithmus ein Boltzmann-Sampler von  $\text{CYC}(\mathcal{B})$ .

*Beweis.* Für die erzeugende Funktion von  $\mathcal{A}$  gilt  $A(x) = \log(\frac{1}{1-B(x)})$ . Die Wahrscheinlichkeit, dass ein Element von  $\mathcal{A}$  aus  $k$  Elementen von  $\mathcal{B}$  besteht ist

$$\frac{1}{A(x)} \frac{B(x)^k}{k} = \frac{1}{\log(\frac{1}{1-B(x)})} \frac{B(x)^k}{k}.$$

Dies ist die Wahrscheinlichkeit, dass eine logarithmisch-verteilte Zufallszahl mit Parameter  $B(x)$  den Wert  $k$  annimmt. □

## 4 Boltzmann-Sampler

### Beispiel: Mengenpartitionen

Wie bereits in Kapitel 2 vorgestellt sind die Mengenpartitionen definiert durch

$$\mathcal{P} = \text{SET}(\text{SET}_{\geq 1}(\{\square\}))$$

Die exponentielle erzeugende Funktion der Mengenpartitionen ist die erzeugende Funktion der bekannten Bell-Zahlen. Deshalb gilt:

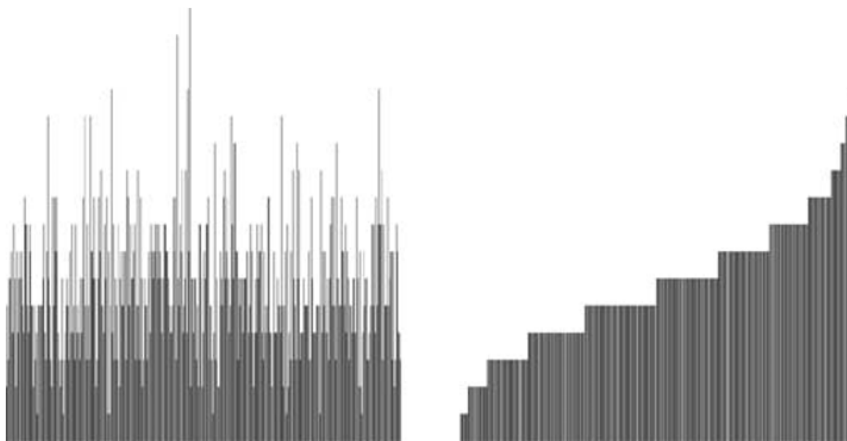
$$P(x) = e^{e^x - 1}$$

Der Boltzmann-Sampler der Mengenpartitionen lässt sich realisieren, indem man zuerst eine poisson-verteilte Zufallszahl  $k$  mit Parameter  $e^x - 1$  erzeugt, die die Anzahl der Blöcke angibt. Danach werden  $k$  unabhängige poisson-verteilte Zufallszahlen mit Parameter  $x$  erzeugt, die die Unterteilungen festlegen.

---

```
1: procedure  $\Gamma P(x)$ 
2:    $k := \text{POISSON}(e^x - 1)$ 
3:   return  $(\underbrace{\text{POISSON}(x), \text{POISSON}(x), \dots, \text{POISSON}(x)}_{k\text{-Tupel}})$ 
4: end procedure
```

---



**Abbildung 4.6:** Abbildung aus [16]: Beispiel einer durch den Boltzmann-Sampler erzeugten Mengenpartition mit einer Größe von 2356 in 409 Blöcken

In Abbildung 4.6 ist ein Beispiel einer zufällig erzeugten Mengenpartition mit dem gewählten Parameter  $x = 6$  gegeben. Links sind die Größen der Blöcke visualisiert. Rechts sind die Blöcke nach Größen geordnet dargestellt.

## 4.5 Toleranzbereich

Der Nachteil der Boltzmann-Sampler im Gegensatz zur rekursiven Methode ist, dass die Größe des kombinatorischen Objekts einer gewissen Wahrscheinlichkeitsverteilung entspricht und meist nicht der exakten gewünschten Größe, denn die Wahrscheinlichkeit, ein Objekt der Größe  $n$  zu generieren, beträgt  $\mathbb{P}_x(|\alpha| = n) = \frac{A_n x^n}{\omega_n A(x)}$ .

Der Anwender muss nun wählen, wie genau der Sampler die gewünschte Größe treffen soll. Eine höhere Genauigkeit bewirkt jedoch auch eine höhere Laufzeit. In [16] wird dabei zwischen **Approximate-Size-Sampler** und **Exact-Size-Sampler** unterschieden.

Nun ist die Effizienz des Boltzmann-Samplers stark von der Form der Boltzmann-Verteilung abhängig.

### 4.5.1 Approximate-Size-Sampler

Sei eine Größe  $n$  und eine Toleranz  $\epsilon > 0$  gegeben. In Kapitel 4.1 wurde beschrieben wie man den Parameter  $x_n$  der Boltzmann-Verteilung wählt, sodass die erwartete Größe gleich der gewünschten Größe  $n$  ist.

Der Approximate-Size-Sampler liefert nun als Ergebnis ein zufälliges Objekt der kombinatorischen Struktur, wobei die Größe im Toleranzbereich  $[(1 - \epsilon)n, (1 + \epsilon)n]$  liegt. Dies gelingt mit einer Verwerfungsmethode. Es wird ein zufälliges Objekt mit dem Boltzmann-Sampler  $\Gamma\mathcal{A}(x_n)$  mit Parameter  $x_n$  generiert. Liegt dieses Objekt nicht im Toleranzbereich, dann wird es verworfen und erneut ein Objekt generiert. Dies geschieht solange, bis die Größe des Objekts akzeptiert wird.

Man muss sich nun aber die Frage stellen, wie effizient diese Verwerfungsmethode ist. Wählt man den Toleranzbereich sehr klein und hat die Boltzmann-Verteilung eine sehr ungünstige Form, so kann der Sampler für diese Struktur und diese Genauigkeitsanforderung unbrauchbar werden.

Man kann die Typen der Boltzmann-Verteilungen unterscheiden, um Komplexitätsaussagen treffen zu können. In [17] werden Komplexitätsfragen für die verschiedenen Typen gesondert abgehandelt. Für unebene Verteilungen lässt sich mithilfe der Tschebyscheff-Ungleichung (siehe [1]) folgender Satz beweisen.

**Satz 4.28.** *Sei  $\mathcal{A}$  eine kombinatorische Struktur, und sei  $\epsilon > 0$  eine Toleranzgröße. Sei  $x_n$  jener Parameter der Boltzmann-Verteilung, sodass  $n = \mathbb{E}(n)$ . Erfüllt die Boltzmann-Verteilung folgende beiden Bedingungen:*

- *Varianzbedingung:*  $\lim_{x \rightarrow \rho} \frac{\sigma(x)}{\mu_1(x)} = 0$
- *Mittelwertbedingung:*  $\lim_{x \rightarrow \rho} \mu_1(x) = \infty$

*Dann strebt die Wahrscheinlichkeit, dass der Approximate-Size-Sampler mit Toleranzbereich  $[(1 - \epsilon)n, (1 + \epsilon)n]$  in einem Versuch die gewünschte Genauigkeit erreicht, für  $n \rightarrow \infty$  gegen 1. Die Gesamtkosten betragen im Durchschnitt  $\mathcal{O}(n)$ .*

#### 4 Boltzmann-Sampler

Da unebene Verteilungen diese Bedingungen erfüllen, führt ein Approximate-Size-Sampler in diesem Fall in linearer Zeit zum Ergebnis. Wie es sich bei Exact-Size-Sampler verhält, zeigt sich im nächsten Abschnitt.

Bei den flachen Boltzmann-Verteilungen kann ebenfalls eine Aussage über die Effizienz der Approximate-Size-Sampler getroffen werden.

**Satz 4.29.** *Sei  $\mathcal{A}$  eine kombinatorische Struktur und  $\epsilon > 0$  eine Toleranzgröße, und sei die Boltzmann-Verteilung der Struktur  $\mathcal{A}$  flach mit einem singulären Exponenten  $-\alpha$ . Sei  $x_n$  jener Parameter der Boltzmann-Verteilung, sodass  $n = \mathbb{E}(n)$ . Dann verhält sich der Erwartungswert der Anzahl an Versuchen, die der Sampler mit der Verwerfungsmethode und Toleranzbereich  $[(1 - \epsilon)n, (1 + \epsilon)n]$  benötigt, asymptotisch wie*

$$\frac{1}{\zeta_\alpha(\epsilon)}, \quad \text{mit } \zeta_\alpha(\epsilon) = \frac{\alpha^\alpha}{\Gamma(\alpha)} \int_{-\epsilon}^{\epsilon} (1+s)^{\alpha-1} e^{-\alpha(1+s)} ds.$$

*Beweis.* Für einen umfassenden Beweis sei auf [16] verwiesen. □

Deshalb lassen sich kombinatorische Strukturen mit flacher Boltzmann-Verteilung in linearer Zeit generieren.

In Kapitel 4.1 wurde bereits erwähnt, wie man die Erzeugung eines Objektes einer kombinatorischen Struktur mit einer spitzen Boltzmann-Verteilung durch die Erzeugung eines Objektes einer kombinatorischen Struktur mit einer flachen Boltzmann-Verteilung realisieren kann. Angenommen man möchte ein Objekt der Struktur  $\mathcal{A}$ , wobei deren Boltzmann-Verteilung spitz ist mit singulärem Exponent  $-\alpha$ . Ist  $-\alpha > 1$ , dann ist die Boltzmann-Verteilung der Struktur  $\Theta\mathcal{A}$  ebenfalls spitz mit singulärem Exponenten  $-\alpha - 1$ . Ist  $-\alpha \in (0, 1)$ , dann ist die Boltzmann-Verteilung der Struktur  $\Theta\mathcal{A}$  flach mit singulärem Exponenten  $-\alpha - 1$ .

Die gleichverteilte Zufallsgenerierung eines Objekts in  $\mathcal{A}_n$  entspricht einer gleichverteilte Zufallsgenerierung eines Objekts in  $\Theta\mathcal{A}_n$ . Deshalb kann zur Erzeugung eines zufälligen Objekts aus  $\mathcal{A}$  ein zufälliges Objekt aus  $\Theta\mathcal{A}$  erzeugt und die Markierung fallen gelassen werden. Oft genug angewendet wird also aus der Erzeugung eines Objektes einer Struktur, die eine spitze Boltzmann-Verteilung besitzt, die Erzeugung eines Objektes einer Struktur, die eine flache Boltzmann-Verteilung besitzt. Dadurch ergeben sich erhebliche Performance-Vorteile.

**Satz 4.30.** *Sei  $\mathcal{A}$  eine kombinatorische Struktur und  $\epsilon > 0$  eine Toleranzgröße, und sei die Boltzmann-Verteilung der Struktur  $\mathcal{A}$  spitz mit einem singulären Exponenten  $-\alpha$ . Sei  $x_n$  jener Parameter der Boltzmann-Verteilung, sodass  $n = \mathbb{E}(n)$ . Sei  $\Theta^k\mathcal{A}$  jene Struktur, die entsteht, wenn man auf die Struktur  $\mathcal{A}$  den Auszeichnungsoperator  $k$  mal anwendet. Dann verhält sich der Erwartungswert der Anzahl an Versuchen, die der Sampler der Struktur  $\Theta^{\max(0, \lceil -\alpha \rceil)}\mathcal{A}$  mit der Verwerfungsmethode und Toleranzbereich  $[(1 - \epsilon)n, (1 + \epsilon)n]$  benötigt, asymptotisch wie*

$$\frac{1}{\zeta_{\alpha + \max(0, \lceil -\alpha \rceil)}(\epsilon)}.$$

## 4 Boltzmann-Sampler

*Beweis.* Dies folgt direkt aus Satz 4.29. □

Dadurch ergibt sich eine durchschnittliche lineare Gesamtlaufzeit für Objekte der kombinatorischen Strukturen mit spitzen Boltzmann-Verteilungen.

### 4.5.2 Exact-Size-Sampler

Bei den Exact-Size-Sampler, die ebenfalls in [16] eingeführt wurden, werden nur Objekte akzeptiert, die von der gewünschten Größe  $n$  sind. Dies würde einem Approximate-Size-Sampler mit Toleranz  $\epsilon := \frac{1}{n}$  entsprechen. Für flache Verteilungen lässt sich nun folgendes Resultat über die Laufzeit eines Exact-Size-Samplers zeigen.

**Satz 4.31.** *Sei  $A$  eine kombinatorische Struktur, und sei die Boltzmann-Verteilung der Struktur  $A$  uneben. Sei  $x_n$  jener Parameter der Boltzmann-Verteilung, sodass  $n = \mathbb{E}(n)$ . Dann verhält sich die durchschnittliche Anzahl der Versuche, die der Exact-Size-Sampler benötigt, um ein Objekt der Größe  $n$  zu generieren asymptotisch wie*

$$\sqrt{2\pi}\sigma(x_n).$$

*Die Gesamtkosten betragen im Mittel also  $\mathcal{O}(n\sigma(x_n))$ .*

*Beweis.* Aufgrund von Satz 4.8 gilt

$$\frac{a_m x_n^m}{A(x_n)} = \frac{1}{\sigma(x_n)\sqrt{2\pi}} \exp\left(-\frac{(m-n)^2}{2\sigma(x_n)^2} + o(1)\right)$$

Für  $m = n$  erhält man sofort das gewünschte Resultat. □

Für Strukturen mit flachen Boltzmann-Verteilungen ergeben sich aufgrund von Satz 4.29 durchschnittliche Gesamtkosten von  $\mathcal{O}(n^2)$ .

### 4.5.3 Singuläre Sampler

Man kann die Boltzmann-Verteilung verwenden, die als Parameter  $x$  die dominante Singularität  $w$  verwendet, vorausgesetzt die erzeugende Funktion ist an der Stelle  $w$  endlich. Solche Boltzmann-Sampler nennt man dann **singuläre Sampler**. In den meisten Fällen trifft dies auf kombinatorische Strukturen zu, die durch eine rekursive Beschreibung gegeben sind. Wenn  $C(x) = C(w) + a(1 - \frac{x}{w})^{\frac{1}{2}} + o((1 - \frac{x}{w})^{\frac{1}{2}})$  für  $x \rightarrow w$ , dann ist der Erwartungswert der Größe des singulären Samplers unendlich. Obwohl daraus resultiert, dass die Wahrscheinlichkeit ein Objekt der gewünschten Größe zu erhalten, in diesem Fall groß sind, heißt dies auch,

#### 4 Boltzmann-Sampler

dass die erwartete Laufzeit des Samplers unendlich ist.

Die erwartete unendliche Laufzeit ergibt sich durch die Durchläufe, bei denen das zurückgelieferte Objekt weit größer ist als die gewünschte Größe  $n$ . Dies lässt sich dadurch verbessern, dass man den Durchlauf unterbricht, wenn das generierte Objekt bereits zu groß ist. Wendet man diese Modifizierung an, dann ergibt sich für eine Approximate-Size-Sampler eine Laufzeit von  $\Theta(n)$  und für einen Exact-Size-Sampler eine Laufzeit von  $\Theta(n^{\frac{3}{2}})$ .



# 5 Anwendungen der Boltzmann-Sampler

## 5.1 Reguläre Sprachen mit linearer Mischung

In [19] widmen sich Darrasse, Panagiotou, Roussel und Soria der Zufallserzeugung von Wörtern regulärer Sprachen, die durch die gewöhnlichen regulären Operatoren und dem Operator des linearen Mischens entstanden sind. Das Erzeugen von Wörtern ohne die Verwendung des Mischens lässt sich durch die in Kapitel 4 vorgestellten Boltzmann-Sampler von unmarkierten kombinatorischen Strukturen realisieren. Das Problem des Mischens von regulären Sprachen besteht in der Anpassung der Wahrscheinlichkeitsverteilungen, da zwischen unmarkierten und markierten Strukturen gewechselt werden muss. In diesem Kapitel wird darauf eingegangen, wie sich solch eine Anpassung umsetzen lässt.

Zu Beginn werden einige grundlegende Definitionen und Aussagen über reguläre Sprachen behandelt.

**Definition 5.1.** Ein **Alphabet**  $\Sigma$  ist eine endliche Menge von **Zeichen**.

Ein **Wort** ist eine endliche Folge von Zeichen. Die Menge aller Wörter wird mit  $\Sigma^*$  bezeichnet.

Eine **Sprache** ist eine Menge von Wörtern über einem Alphabet  $\Sigma$ . Sei  $\Sigma$  ein Alphabet und sei  $\mathcal{A}^1$  die Menge aller Wörter der Länge 1 über  $\Sigma$ . Eine Teilmenge von  $\mathcal{A}^1$  nennt man dann **Elementarsprache**.

**Definition 5.2.** Seien  $\mathcal{A}$  und  $\mathcal{B}$  Sprachen.

- Man bezeichnet  $\mathcal{C} = \mathcal{A} + \mathcal{B}$  als **disjunkte Vereinigung**, wenn  $\mathcal{A}$  und  $\mathcal{B}$  disjunkt sind und  $\mathcal{C} = \mathcal{A} \cup \mathcal{B}$ .
- Man bezeichnet  $\mathcal{C} = \mathcal{A} \times \mathcal{B}$  als **Produkt**, wenn  $\mathcal{C} = \{vw : w \in \mathcal{A}, v \in \mathcal{B}\}$ , wobei  $vw$  die Verkettung zweier Folgen bezeichnet.
- Man bezeichnet  $\mathcal{C} = \mathcal{A}^*$  als **Folge** von  $\mathcal{A}$ 's, wenn  $\mathcal{C} = \{w_1w_2 \cdot \dots \cdot w_n : w_i \in \mathcal{A}\}$ , wobei  $\mathcal{A} \neq \emptyset$ .

**Definition 5.3.** Eine Sprache  $\mathcal{A}$  über einem Alphabet  $\Sigma$  heißt **regulär**, wenn sie sich durch endlich viele Anwendungen der Operationen Vereinigung, Produkt und Folge auf Elementarsprachen über  $\Sigma$  erzeugen lässt.

**Definition 5.4.** Die **lineare Mischung**  $u \sqcup\sqcup v$  von zwei Wörtern  $u, v \in \Sigma^*$  ist definiert als

$$u \sqcup\sqcup v = \{u_1v_1u_2v_2 \dots u_nv_n : u = u_1 \dots u_n, v = v_1 \dots v_n \in \Sigma^*, n \in \mathbb{N}, \forall 1 \leq i \leq n, u_i, v_i \in \Sigma^*\}$$

## 5 Anwendungen der Boltzmann-Sampler

Die **lineare Mischung**  $\mathcal{A} \sqcup \mathcal{B}$  von zwei Sprachen  $\mathcal{A}$  und  $\mathcal{B}$  ist definiert als

$$\mathcal{A} \sqcup \mathcal{B} = \{u \sqcup v : u \in \mathcal{A}, v \in \mathcal{B}\}$$

Als reguläre Sprache mit linearer Mischung bezeichnen wir nun Sprachen, die durch endlich viele Anwendungen der Operationen Vereinigung, Produkt, Folge und linearer Mischung auf Elementarsprachen erzeugt werden können.

Reguläre Sprachen werden hier als kombinatorische Strukturen betrachtet, deren man erzeugende Funktionen zuordnet. Diese lassen sich durch die Regeln für erzeugende Funktionen aus Kapitel 2 übersetzen.

Im Falle des linearen Mischens müssen exponentielle erzeugende Funktionen verwendet werden. Dabei gilt für die exponentielle erzeugende Funktion Folgendes:

**Satz 5.1.** *Seien  $\mathcal{A}$  und  $\mathcal{B}$  reguläre Sprachen und  $\widehat{A}(z)$  und  $\widehat{B}(z)$  deren exponentielle erzeugende Funktionen. Weiters sei  $\widehat{C}(z)$  die exponentielle erzeugende Funktion von  $\mathcal{C} = \mathcal{A} \sqcup \mathcal{B}$ . Dann gilt*

$$\widehat{C}(z) = \widehat{A}(z)\widehat{B}(z).$$

*Beweis.* Sei  $\widehat{A}(z) = \sum_{n \geq 0} a_n \frac{z^n}{n!}$ ,  $\widehat{B}(z) = \sum_{n \geq 0} b_n \frac{z^n}{n!}$  und  $\widehat{C}(z) = \sum_{n \geq 0} c_n \frac{z^n}{n!}$ . Dann gilt für die Anzahl  $c_n$  der Wörter der Länge  $n$  aus  $\mathcal{C}$

$$c_n = \sum_{k=0}^n \binom{n}{k} a_k b_{n-k}.$$

Daraus folgt mit dem Cauchyproduktsatz

$$\widehat{C}(z) = \sum_{n \geq 0} \sum_{k=0}^n \frac{n!}{k!(n-k)!} a_k b_{n-k} \frac{z^n}{n!} = \sum_{n \geq 0} \sum_{k=0}^n \frac{1}{k!(n-k)!} a_k b_{n-k} z^k z^{n-k} = \widehat{A}(z)\widehat{B}(z). \quad \square$$

Es ist nun nicht so einfach möglich einen gewöhnlichen Boltzmann-Sampler für reguläre Sprache mit linearem Mischen anzugeben. Aufgrund des letzten Satzes ist dies jedoch für exponentielle Boltzmann-Sampler simpler. Nun stellt sich aber die Frage, wie man zwischen den gewöhnlichen und den exponentiellen Boltzmann-Sampler wechseln kann, damit die Eigenschaft erhalten bleibt, dass zwei kombinatorische Objekte (Wörter) einer Länge  $n$  gleichwahrscheinlich sind.

Um ein Wort einer regulären Sprache mit linearem Mischen zu generieren, werden exponentielle Boltzmann-Sampler verwendet. Danach wird der exponentielle Boltzmann-Sampler in einen entsprechenden gewöhnlichen Boltzmann-Sampler transformiert.

Diese Transformation kann mithilfe einer Laplace-Transformation<sup>3</sup> realisiert werden.

---

<sup>3</sup> Die Definition der Laplace-Transformierten und der Faltung ist im Anhang A.3 angeführt.

## 5 Anwendungen der Boltzmann-Sampler

**Satz 5.2.** Sei  $\mathcal{A}$  eine kombinatorische Struktur, und seien  $A(z)$  und  $\widehat{A}(z)$  die gewöhnliche und die exponentielle erzeugende Funktion von  $\mathcal{A}$ . Sei  $x$  innerhalb des Konvergenzradius von  $A(z)$ . Dann definiert  $\delta_x^A(u) = \frac{\widehat{A}(xu)}{A(x)} e^{-u}$  eine Wahrscheinlichkeitsverteilung.

*Beweis.* Die erzeugende Funktion  $A(z)$  ist die Laplace-Transformierte von  $\widehat{A}(z)$ . Das heißt es gilt  $A(z) = \int_0^\infty \widehat{A}(zu) e^{-u} du$ . Und daraus folgt direkt die Behauptung.  $\square$

**Satz 5.3.** Ist ein exponentieller Boltzmann-Sampler  $\widehat{\Gamma}\mathcal{A}$  für eine kombinatorische Struktur  $\mathcal{A}$  gegeben, so erhält man einen exponentiellen Boltzmann-Sampler durch Erzeugung einer Zufallszahl  $u$ , deren Wahrscheinlichkeit  $\delta_x^A(u)$  beträgt. Dann ist der dazugehörige gewöhnliche Boltzmann-Sampler der Struktur  $\mathcal{A}$  gegeben durch einen Aufruf von  $\widehat{\Gamma}\mathcal{A}(xu)$ . Die Wahrscheinlichkeit, ein Objekt  $\alpha \in \mathcal{A}$  zu erhalten, beträgt dann  $\frac{x^{|\alpha|}}{A(x)}$ .

*Beweis.* Sei  $\alpha$  ein Objekt aus  $\mathcal{A}$  der Größe  $n$ . Sei  $U$  nach der Wahrscheinlichkeitsverteilung aus Satz 5.2 verteilt. Dann ist die erwartete Wahrscheinlichkeit, dass der gewöhnliche Boltzmann-Sampler das Objekt zurückliefert

$$\begin{aligned} \mathbb{E}(\mathbb{P}(\widehat{\Gamma}\mathcal{A}(xU) = \alpha)) &= \int_0^\infty \mathbb{P}(\widehat{\Gamma}\mathcal{A}(xu) = \alpha) \delta_x^A(u) du = \\ &= \int_0^\infty \frac{(xu)^n}{n! \widehat{A}(xu)} \delta_x^A(u) du = \frac{x^n}{A(x)} \frac{\int_0^\infty u^n e^{-u} du}{n!} = \frac{x^n}{A(x)}. \end{aligned} \quad \square$$

Wenn man nun exponentielle Boltzmann-Sampler für Beschreibungen von regulären Sprachen mit linearen Mischen kennt, können diese mithilfe des letzten Satzes in gewöhnliche Boltzmann-Sampler transformiert werden.

Zu beachten ist nun, dass die Folge der Anzahlen der Wörter einer bestimmten Größe sowohl bei der gewöhnlichen erzeugenden Funktion als auch bei der exponentiellen erzeugenden Funktion gleich ist. Bei der Vereinigung von zwei Sprachen hat dies keine Auswirkungen, und es gilt  $\widehat{A}(z) = \widehat{B}(z) + \widehat{C}(z)$ . Für die Anzahl der Wörter eines Produkts von zwei Sprachen gilt  $a_n = \sum_{i=0}^n b_i c_{n-i}$ . Beim Produkt von exponentiellen erzeugenden Funktionen werden auch die Anzahlen  $\binom{n}{i}$  der verschiedenen Markierungen berücksichtigt. Dies ist bei Sprachen jedoch nicht der Fall. Deshalb lässt sich die exponentielle erzeugende Funktion in diesem Fall nicht einfach durch das Produkt der beiden exponentiellen erzeugenden Funktionen ausdrücken.

**Satz 5.4.** Seien  $\mathcal{B}$  und  $\mathcal{C}$  zwei reguläre Sprachen und  $\widehat{B}(z)$  und  $\widehat{C}(z)$  deren exponentielle erzeugende

## 5 Anwendungen der Boltzmann-Sampler

Funktionen. Weiters sei  $\mathcal{A} = \mathcal{B} \times \mathcal{C}$ . Dann gilt für die exponentielle erzeugende Funktion  $\widehat{A}(z)$  von  $\mathcal{A}$

$$\widehat{A}(z) = b_0 \widehat{C}(z) + \int_0^z \widehat{B}'(z-t) \widehat{C}(t) dt.$$

*Beweis.*

$$\begin{aligned} \int_0^z \widehat{B}(z-t) \widehat{C}(t) dt &= \int_0^z \sum_{n \geq 0} \frac{b_n}{n!} (z-t)^n \sum_{m \geq 0} \frac{c_m}{m!} t^m dt = \int_0^z \sum_{n \geq 0} \sum_{i=0}^n \frac{b_i}{i!} \frac{c_{n-i}}{(n-i)!} (z-t)^i t^{n-i} dt = \\ &= \sum_{n \geq 0} \sum_{i=0}^n \frac{b_i}{i!} \frac{c_{n-i}}{(n-i)!} \int_0^z (z-t)^i t^{n-i} dt. \end{aligned}$$

Mithilfe der Laplace-Transformation lässt sich der Ausdruck  $\int_0^z (z-t)^i t^{n-i} dt$  vereinfachen. Das Faltungstheorem (Satz A.3) besagt, dass die Laplace-Transformierte einer Faltung die Multiplikation der Laplace-Transformierten der gefalteten Funktionen ist.

Sei  $f(t) := t^i$  und  $g(t) := t^{n-i}$ . Dann gilt

$$\int_0^z (z-t)^i t^{n-i} dt = (f \star g)(z) = \mathcal{L}^{-1}(\mathcal{L}(f) \cdot \mathcal{L}(g))(z).$$

Weiters gilt für die Laplace-Transformierten der Funktionen  $f$  und  $g$  und für die Rücktransformation der Multiplikation folgendes:

$$\mathcal{L}(f)(z) = \int_0^\infty t^i \exp(-zt) dt = \frac{i!}{z^{i+1}},$$

$$\mathcal{L}(g)(z) = \int_0^\infty t^{n-i} \exp(-zt) dt = \frac{(n-i)!}{z^{n-i+1}},$$

$$\int_0^z (z-t)^i t^{n-i} dt = \mathcal{L}^{-1}\left(\frac{i!(n-i)!}{z^{n+2}}\right) = \frac{i!(n-i)!}{(n+1)!} z^{n+1}.$$

Eingesetzt in obige Gleichung folgt daraus

$$\int_0^z \widehat{B}(z-t) \widehat{C}(t) dt = \sum_{n \geq 0} \sum_{i=0}^n \frac{b_i}{i!} \frac{c_{n-i}}{(n-i)!} \frac{i!(n-i)!}{(n+1)!} z^{n+1} = \sum_{n \geq 0} \sum_{i=0}^n \frac{b_i c_{n-i}}{(n+1)!} z^{n+1}.$$

## 5 Anwendungen der Boltzmann-Sampler

Durch Differentiation beider Seiten erhält man

$$\widehat{B}(0)\widehat{C}(z) + \int_0^z \widehat{B}'(z-t)\widehat{C}(t)dt = \sum_{n \geq 0} \sum_{i=0}^n \frac{b_i c_{n-i}}{n!} z^n.$$

Da die Anzahl der Wörter der Größe  $n$  aus  $\mathcal{A}$   $\sum_{i=0}^n b_i c_{n-i}$  beträgt, ergibt sich

$$b_0 \widehat{C}(z) + \int_0^z \widehat{B}'(z-t)\widehat{C}(t)dt = \widehat{C}(z). \quad \square$$

Die  $k$ -te Ableitung der exponentiellen erzeugenden Funktion liefert eine Folge, die die Anzahl der Wörter der Länge  $n - k$  angibt. Der letzte Satz begründet nun auch die erweiterte Definition einer exponentiellen Boltzmann-Verteilung.

**Definition 5.5.** Sei  $\mathcal{A}$  eine reguläre Sprache,  $k \in \mathbb{N}$  und  $\alpha \in \mathcal{A}$ . Dann nennt man die Wahrscheinlichkeitsverteilung mit

$$\mathbb{P}_{x,k} = \begin{cases} 0 & , |\alpha| < k \\ \frac{x^{|\alpha|-k}}{(|\alpha|-k)\widehat{A}^{(k)}(x)} & , |\alpha| \geq k \end{cases}$$

die **exponentielle Boltzmann-Verteilung der Ordnung  $k$** . Man bezeichnet mit  $\widehat{\Gamma}\mathcal{A}(x,k)$  den Boltzmann-Sampler, dessen Verteilung der exponentiellen Boltzmann-Verteilung der Ordnung  $k$  entspricht.

In Folge werden exponentielle Boltzmann-Sampler der Ordnung  $k$  für die Basis-Konstruktionen Vereinigung, Produkt, Folge und lineares Mischen entwickelt. Für die Sprache, die ausschließlich das leere Wort enthält, wird nur im Falle  $k = 0$  das leere Wort zurückgeliefert. Für die Sprache, die ausschließlich ein Wort der Länge 1 enthält, wird im Falle  $k = 0$  oder  $k = 1$  dieses Wort zurückgeliefert. In den anderen Fällen kann nichts ausgegeben werden, und der Algorithmus liefert hier einen Fehler.

- **Vereinigung:**  $\mathcal{A} = \mathcal{B} + \mathcal{C}$

Für die exponentielle erzeugende Funktion von  $\mathcal{A}$  der Ordnung  $k$  gilt aufgrund der Summenregel des Differenzierens  $\widehat{A}^{(k)}(z) = \widehat{B}^{(k)}(z) + \widehat{C}^{(k)}(z)$ . Deshalb lässt sich der Boltzmann-Sampler der Ordnung  $k$  realisieren durch eine bernoulliverteilte Zufallsvariable mit Parameter  $\frac{\widehat{B}^{(k)}(z)}{\widehat{A}^{(k)}(z)}$ . Falls die Zufallszahl 1 beträgt, wird der Boltzmann-Sampler der Ordnung  $k$  von  $\mathcal{B}$  ausgeführt, im anderen Fall der Boltzmann-Sampler der Ordnung  $k$  von  $\mathcal{C}$ .

- **Produkt:**  $\mathcal{A} = \mathcal{B} \times \mathcal{C}$

Für die exponentielle erzeugende Funktion von  $\mathcal{A}$  der Ordnung  $k$  gilt durch Differentiation der

## 5 Anwendungen der Boltzmann-Sampler

Gleichung aus Satz 5.4

$$\widehat{A}^{(k)}(z) = \sum_{i=0}^k b_i \widehat{C}^{(k-i)}(z) + \int_0^z \widehat{B}'(z-t) \widehat{C}(t) dt.$$

Um einen Boltzmann-Sampler für das Produkt von zwei Sprachen zu entwickeln werden zwei Wahrscheinlichkeitsverteilungen verwendet. Man bezeichnet mit  $P_{k,x,\mathcal{A}}$  eine Zufallsvariable die folgender Wahrscheinlichkeitsverteilung genügt:

$$\mathbb{P}(P_{k,x,\mathcal{A}} = l) = \frac{1}{\widehat{A}^{(k)}} \cdot \begin{cases} b_l \widehat{C}^{(k-l)}(x), l \in \{0, \dots, k\} \\ \int_0^x \widehat{B}^{(k+1)}(x-t) \widehat{C}(t) dt, l = k+1 \end{cases}$$

Weiters wird mit  $B_{k,x,\mathcal{A}}$  eine Zufallsvariable der Wahrscheinlichkeitsverteilung mit der Dichtefunktion

$$f(t) = \frac{\widehat{B}^{(k+1)}(x-t) \widehat{C}(t)}{\int_0^x \widehat{B}^{(k+1)}(x-u) \widehat{C}(u) du}$$

bezeichnet. Nun sei  $P(x, k, \mathcal{A})$  eine Funktion, die eine solche Zufallszahl generiere. Und sei  $\text{UNIFORM}(\mathcal{A})$  eine Funktion die ein gleichverteiltes Zufallswort der Länge  $l$  aus  $\mathcal{A}_l$  erzeugt. Damit erhält man einen exponentiellen Boltzmann-Sampler für das Produkt durch folgenden Algorithmus.

---

```

1: procedure  $\widehat{\Gamma} \mathcal{A}(x,k)$ 
2:    $l := P(k, x, \mathcal{A})$ 
3:   if  $0 \leq l \leq k$  then
4:      $\alpha_{\mathcal{B}} := \text{UNIFORM}(\mathcal{B}_l)$ 
5:      $\alpha_{\mathcal{C}} := \Gamma \mathcal{C}(x, k-l)$ 
6:   else
7:      $t := B(k, x, \mathcal{A})$ 
8:      $\alpha_{\mathcal{B}} := \Gamma \mathcal{B}(x-t, k+1)$ 
9:      $\alpha_{\mathcal{C}} := \Gamma \mathcal{C}(t, 0)$ 
10:
11:   end if
12:   return  $\alpha_{\mathcal{B}} \alpha_{\mathcal{C}}$ 
13: end procedure

```

---

- **Folge:**  $\mathcal{A} = \text{SEQ}(\mathcal{B})$

## 5 Anwendungen der Boltzmann-Sampler

Für die Konstruktion eines Boltzmann-Samplers für Folgen wird folgende Wahrscheinlichkeitsverteilung definiert

$$\mathbb{P}(S_{k,x,\mathcal{A}} = l) = \frac{1}{\widehat{A}^{(k)}(x)} \cdot \begin{cases} 0 & , l = 0 \\ b_l \widehat{C}^{(k-l)}(x) & , l \in \{1, \dots, k\} \\ \int_0^x \widehat{B}^{(k+1)}(x-t) \widehat{C}(t) dt & , l = k+1 \end{cases}$$

Sei wieder  $S(x, k, \mathcal{A})$  eine Funktion, die eine Zufallszahl mit dieser Wahrscheinlichkeitsverteilung erzeugt. Dann ist der Boltzmann-Sampler der Ordnung  $k$  gegeben durch nachfolgenden Algorithmus.

---

```

1: procedure  $\widehat{\Gamma}\mathcal{A}(x,k)$ 
2:    $l := S(k, x, \mathcal{A})$ 
3:   if  $l = 0$  then
4:     return  $\epsilon$ 
5:   else
6:     if  $1 \leq l \leq k$  then
7:        $\alpha_{\mathcal{B}} := \text{UNIFORM}(\mathcal{B}_l)$ 
8:        $\alpha_{\mathcal{A}} := \Gamma\mathcal{A}(x, k-l)$ 
9:     else
10:       $t := B(k, x, \mathcal{A})$ 
11:       $\alpha_{\mathcal{B}} := \Gamma\mathcal{B}(x-t, k+1)$ 
12:       $\alpha_{\mathcal{A}} := \Gamma\mathcal{A}(t, 0)$ 
13:     end if
14:   end if
15:   return  $\alpha_{\mathcal{B}}\alpha_{\mathcal{A}}$ 
16: end procedure

```

---

- **lineares Mischen:**  $\mathcal{A} = \mathcal{B} \sqcup \sqcup \mathcal{C}$

Auch hier muss wieder eine Wahrscheinlichkeitsverteilung definiert werden. Diese ergibt sich aus der Formel von Satz 5.1. Sei also  $D_{k,x,\mathcal{A}}$  eine Zufallsvariable, die folgender Verteilung folgt:

$$\mathbb{P}(D_{k,x,\mathcal{A}} = l) = \frac{1}{\widehat{A}^{(k)}(x)} \cdot \binom{k}{l} \widehat{B}^{(l)}(x) \widehat{B}^{(k-l)}(x)$$

Dann lässt sich die Zufallserzeugung eines Wortes aus  $\mathcal{A}$  durch folgenden Algorithmus realisieren.

---

```

1: procedure  $\widehat{\Gamma}\mathcal{A}(x,k)$ 
2:    $l := D(k, x, \mathcal{A})$     $\alpha_{\mathcal{B}} := \Gamma\mathcal{B}(x, l)$ 
3:    $\alpha_{\mathcal{C}} := \Gamma\mathcal{C}(x, k - l)$ 
4:   Sei  $b_{\mathcal{B}}e_{\mathcal{B}} = \alpha_{\mathcal{B}}$ , wobei  $|b_{\mathcal{B}}| = l$ 
5:   Sei  $b_{\mathcal{C}}e_{\mathcal{C}} = \alpha_{\mathcal{C}}$ , wobei  $|b_{\mathcal{C}}| = k - l$ 
6:    $b = \text{LINEARMISCHEN}(b_{\mathcal{B}}, e_{\mathcal{B}})$ 
7:    $e = \text{LINEARMISCHEN}(b_{\mathcal{C}}, e_{\mathcal{C}})$ 
8:   return  $be$ 
9: end procedure

```

---

Die Funktion LINEARMISCHEN führt eine zufällige lineare Mischung der beiden Wörter durch.

Durch die letzten Algorithmen für die exponentiellen Boltzmann-Sampler der Ordnung  $k$  der einfachen Konstruktionen inklusive des linearen Mischens lässt sich für die betrachteten Wörter ein exponentieller Boltzmann-Sampler erzeugen. Im Falle, dass  $\mathcal{A} = \mathcal{B} \sqcup \mathcal{C}$  wird der Boltzmann-Sampler  $\widehat{\Gamma}(\mathcal{B} \sqcup \mathcal{C})(ux, 0)$  zurückgeliefert, wobei  $u$  eine Zufallsvariable ist, die der Dichtefunktion  $\gamma_x^{\mathcal{A}}(u) = \frac{\widehat{A}(xu)}{\widehat{C}(x)} e^{-u}$  genügt. Diese exponentiellen Sampler lassen sich dann mit dem beschriebenen Verfahren in einen gewöhnlichen Boltzmann-Sampler überführen. In [19] wird ein Beweis geführt der belegt, dass der Boltzmann-Sampler lineare Laufzeit in der Größe des zurückgelieferten Wortes besitzt.

## 5.2 Apollonische Netzwerkstrukturen

In diesem Abschnitt werden sogenannte spezielle ungerichtete Graphen betrachte, sogenannte zufällige apollonische Netzwerkstrukturen. Zur Erzeugung und zur Herleitung von Eigenschaften wird eine Bijektion zwischen diesen Graphen und ternären Bäumen hergestellt. So lassen sich apollonische Netzwerkstrukturen erzeugen, indem man einen Boltzmann-Sampler für ternäre Graphen verwendet und diese dann überführt. Für die Wahrscheinlichkeitsverteilung des Grades des Zentrums einer apollonischen Netzwerkstruktur wird ein bivariates Boltzmann-Modell für ternäre Graphen verwendet, mithilfe dessen man auf die Verteilung des Grades schließen kann. Dieses und wie sich diese Graphen auch verallgemeinern lassen, findet man in einer Arbeit von Darrasse und Soria in [2].

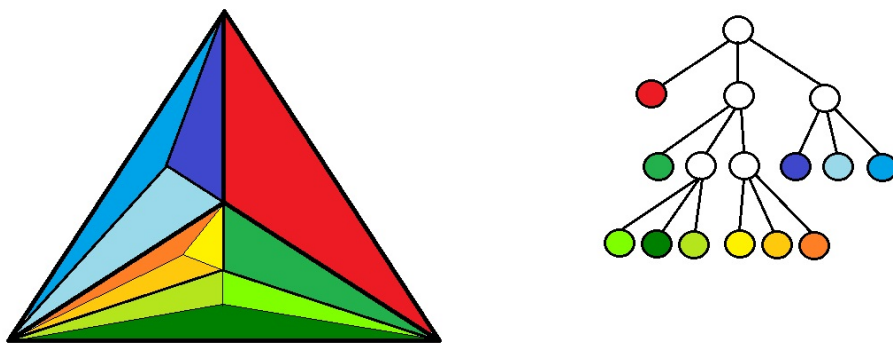
**Definition 5.6.** *Eine zufällige apollonische Netzwerkstruktur (RANS) ist entweder*

- *ein leeres Dreieck, oder*
- *ein Dreieck, das aus drei Dreiecken besteht, die wiederum zufällige apollonische Netzwerkstrukturen sind. Den Knoten in der Mitte des äußeren Dreiecks nennt man dabei **Zentrum** der zufälligen apollonischen Netzwerkstruktur.*



Die Ordnung der RANS ist ebenfalls rekursiv definiert.

**Definition 5.7.** Die **Ordnung** des RANS ist 0, falls es das leere Dreieck ist. Die **Ordnung** des RANS im Falle eines nicht-leeren Dreiecks ist die Summe der Ordnung der enthaltenen 3 Unterdreiecke um eins erhöht.



**Abbildung 5.1:** RANS der Ordnung 5 mit zugehörigem ternären Baum

Es lässt sich nun eine Bijektion zwischen zufälligen apollonischen Netzwerkstrukturen der Ordnung  $N$  und ternären Bäumen mit  $N$  internen Knoten herstellen. Diese wird in Abbildung 5.1 veranschaulicht.

Für weitere Zusammenhänge zwischen dem RANS und dem entsprechenden Baum werden einige Begriffe definiert.

Die **Nachbarschaft eines äußeren Knotens** einer RANS ist die Summe aller Nachbarn des Knotens in der RANS außer den beiden anderen äußeren Knoten.

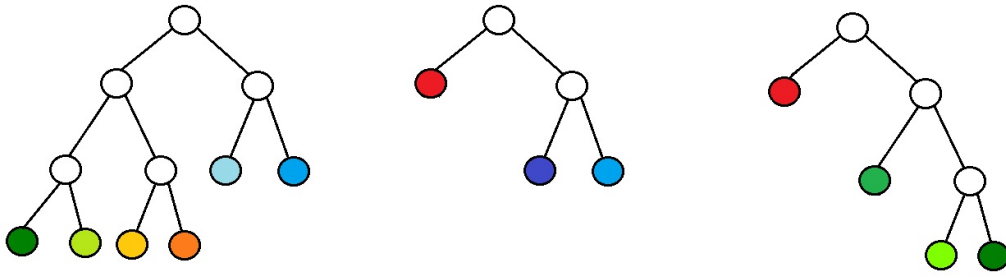
Die **Nachbarschaft des Zentrums** einer RANS ist die Summe aller Nachbarn des Knoten in der RANS außer den äußeren Knoten. Nun ist die Größe der Nachbarschaft eines äußeren Knotens in einer RANS gleich der Anzahl der internen Knoten des Binärbaums, der entsteht, wenn man vom zur RANS gehörigen ternären Baum alle linken Nachfolger streicht. Man nennt diesen Binärbaum auch **1-Unterbaum**. Streicht man aus dem ternären Baum jeweils die mittleren bzw. rechten Nachfolger, nennt man den entstandenen Binärbaum **2-Unterbaum** bzw. **3-Unterbaum**.

Weiters ist der Grad des Zentrums der RANS gleich der um drei erhöhten Anzahl der Nachbarn des Zentrums. Man kann Boltzmann-Sampler dazu verwenden, um Aussagen über die Verteilung des Grads der Knoten in einer RANS zu treffen. Der Grad des Zentrums einer RANS ist gleich drei plus der Summe der Anzahlen der Knoten des 1-, 2- und 3-Unterbaums des zur RANS gehörigen ternären Baums. Im Weiteren wird dieser Wert **R-Grad** genannt. Aufgrund dieser Aussage wird die Verteilung des R-Grades untersucht.

Jeder Knoten in der RANS ist Zentrum einer Unter-RANS und deshalb ist der Grad des Knoten der R-Grad des zugehörigen Unterbaums des ternären Baums.

Für die Analyse der Verteilung des Grads werden nun eine Menge an zufälligen ternären Bäumen erzeugt und

## 5 Anwendungen der Boltzmann-Sampler



**Abbildung 5.2:** 1-Unterbaum, 2-Unterbaum und 3-Unterbaum des ternären Baum aus Abbildung 5.1

die Tatsache verwendet, dass die Menge aller Unterbäume eines zufälligen Baums die gleichen statistischen Parameter besitzen wie eine Menge an zufällig erzeugten Bäumen. Dies liegt an den unabhängigen Aufrufen zur Erzeugung der Unterbäume beim Boltzmann-Sampling eines Baums. Sei also  $T = (T_1, T_2, T_3)$  ein ternärer Baum, wobei  $T_1, T_2$  und  $T_3$  der linke, mittlere und rechte Nachfolger ist und diese ebenfalls ternäre Bäume sind. Im Weiteren wird der R-Grad des Baumes mit  $D(T)$  bezeichnet. Die Anzahl der Knoten eines Baumes  $T$  wird mit  $|T|$  bezeichnet und  $U_i(T)$  bezeichnet den  $i$ -Unterbaum von  $T$ . Es gilt also für  $D(T)$  folgendes:

$$D(T) = |U_1(T_1)| + |U_2(T_2)| + |U_3(T_3)| + 3.$$

Mithilfe einer bivariaten erzeugenden Funktion  $D(z, u)$ , die außer der Größe eines Baums auch dessen R-Grad berücksichtigt, wird auf die Verteilung des R-Grades geschlossen.

Sei  $\mathcal{T}$  die Menge aller ternären Bäumen. Dann definiert man die bivariaten erzeugenden Funktionen  $D(z, u)$  und  $T(z, u)$ :

$$D(z, u) = \sum_{T \in \mathcal{T}} z^{|T|} u^{D(T)},$$

$$T(z, u) = \sum_{T \in \mathcal{T}} z^{|T|} u^{|N_1(T)|}.$$

Da die Verteilung der Größe des 1-Unterbaums auch der Verteilung der Größe des 2- und 3-Unterbaums entspricht, gilt für  $D(z, u)$  folgendes:

$$D(z, u) = zu^3 T^3(z, u).$$

Markiert man jeweils nur einen der 3 Unterbäume, dann lässt sich für  $B(z, u)$  eine rekursive Gleichung aufstellen:

$$T(z, u) = 1 + uzT(z)T^2(z, u).$$

Über die Verteilung des Grades einer RANS lässt sich nun folgende Aussage treffen.

**Satz 5.5.** *Der erwartete Grad einer RANS beträgt 6 und die Wahrscheinlichkeit, dass der Grad der RANS*

## 5 Anwendungen der Boltzmann-Sampler

$n + 3$  beträgt, ist

$$\mathbb{P}(D = n + 3) = \frac{8}{9} \frac{1}{n+3} \binom{2n+2}{k} \frac{2^n}{9^n}.$$

*Beweis.* Sei  $B(t)$  die erzeugende Funktion der Binärbäume, d.h.  $B_n = \frac{1}{n+1} \binom{2n}{n}$ . Dann gilt  $T(z, u) = B(zuT(z))$ , da  $B(t) = 1 + tB^2(t)$ . Der Konvergenzradius von  $B(t)$  beträgt  $\frac{1}{4}$ .

Sei  $w$  die Singularität der erzeugenden Funktion  $T(z)$ . Dann ist die wahrscheinlichkeitserzeugende Funktion der Anzahl der Knoten des 1-Unterbaums eines ternären Baum, wenn dieser mit einem singulären Boltzmann-Sampler erzeugt wurde,

$$\begin{aligned} p(u) &= \frac{T(w, u)}{T(w, 1)} = \frac{1}{T(w)} \sum_{n \in \mathbb{N}} B_n T(w)^n w^n u^n = \\ &= \frac{1}{T(w)} \sum_{n \in \mathbb{N}} \frac{1}{n+1} \binom{2n}{n} T(w)^n w^n u^n. \end{aligned}$$

Der Erwartungswert für die Anzahl der Knoten des 1-Unterbaums beträgt  $\mu = p'(1) = 1$ .

In [16] wird ein singulärer Boltzmann-Sampler für Bäume vorgestellt. Die Singularität beträgt  $w = \frac{4}{27}$  und der Wert der erzeugenden Funktion an dieser Stelle  $T(w) = \frac{3}{2}$ . Und deshalb erhält man folgende Wahrscheinlichkeit für die Anzahl der Knoten des 1-Unterbaums:

$$\mathbb{P}(|1 - \text{Unterbaum}| = n) = \frac{2}{3(n+1)} \binom{2n}{n} \left(\frac{2}{9}\right)^n.$$

Weiters gilt für die bivariate erzeugende Funktion  $D(z, u) = zu^3 T^3(z, u)$  und somit für die Wahrscheinlichkeit, dass der R-Grad des ternären Baum  $n$  beträgt:

$$\mathbb{P}(D = n) = [u^{n+3}] \frac{u^3}{T(w)^3} B^3(uwT(w)).$$

Der  $n$ -te Koeffizient von  $B(t)$  lässt sich mit der Lagrange-Bürmannschen Formel (siehe [9]) bestimmen.

$$\begin{aligned} \mathbb{P}(D = n + 3) &= \frac{1}{T(w)^k} w^{3k} T(w)^{3k} \frac{3}{n+3} \binom{2n+2}{n} = \\ &= \frac{2^{n+3}}{9^{n+1}} \frac{1}{n+3} \binom{2n+2}{n}. \end{aligned}$$

Da die erwartete Anzahl an Knoten des 1-Unterbaums eines ternären Baums 1 beträgt, ist der erwartete Grad des Zentrums 6. Dies ergibt sich aus den drei Kanten des Zentrums zu den äußeren Eckpunkten und den drei Kanten in jedem der Unter-RANS. □

# A Anhang

## A.1 Landau-Symbole

Man betrachtet zwei Funktion  $f(x), g(x)$  mit  $x > 0$  und Werten in den reellen Zahlen. Man schreibt

$$f(x) = \mathcal{O}(g(x)) \text{ für } x \rightarrow \infty$$

falls  $f(x)$  für große Werte von  $x$  von der Größenordnung her  $g(x)$  nicht übersteigt. Dies bedeutet  $f$  wächst nicht schneller oder fällt nicht langsamer als  $g$ . Formal ausgedrückt bedeutet dies, dass für die beiden Funktionen eine Konstante  $c > 0$  und ein  $x_0 > 0$  existiert, so dass

$$|f(x)| \leq c \cdot |g(x)|$$

gilt, wenn  $x > x_0$ . Wenn  $\frac{f(x)}{g(x)} \rightarrow 0$ , dann schreibt man

$$f(x) = o(g(x)) \text{ für } x \rightarrow \infty.$$

Dies bedeutet, für jedes  $c > 0$  existiert ein  $x_0 > 0$ , so dass

$$|f(x)| \leq c \cdot |g(x)|$$

gilt, wenn  $x > x_0$ .

Wenn die betrachteten Funktion gleich schnell wachsen bzw. fallen, dann wird dies notiert mit

$$f(x) = \Theta(g(x)) \text{ für } x \rightarrow \infty$$

Dies ist der Fall, wenn  $f(x) = \mathcal{O}(g(x))$  und  $g(x) = \mathcal{O}(f(x))$  gilt.

Weiters wird hier folgende Notation für **asymptotische Äquivalenz** verwendet:

$$f(z) \sim_{z \rightarrow \zeta} g(z) \rightarrow \zeta : \iff \lim_{z \rightarrow \zeta} \frac{f(z)}{g(z)} = 1.$$

Ein nützlicher Satz, den man oft bei Asymptotik-Untersuchungen anwenden kann ist die Stirlingsche Formel, die unter anderem in [15] bewiesen ist.

**Satz A.1.** *Es gilt*

$$n! \sim_{n \rightarrow \infty} \sqrt{2\pi n} \left(\frac{n}{e}\right)^n.$$

## A.2 Zufallsgeneratoren

Für die Erzeugung einer Zufallszahl einer beliebigen Verteilung wird für gewöhnlich zunächst eine gleichverteilte Zufallszahl im Intervall  $[0, 1]$  erzeugt, bei der jede reelle Zahlen zwischen 0 und 1 mit der gleichen Wahrscheinlichkeit auftritt. Diese wird dann auf die entsprechende Wahrscheinlichkeitsverteilung transformiert. Für die Generierung gleichverteilter Zufallszahlen gibt es verschiedenste Algorithmen, von denen einige in [7] zu finden sind. Im Weiteren sei vorausgesetzt, dass eine Funktion  $\text{UNIFORM}()$  existiert, die eine gleichverteilte Zufallszahl im Intervall  $[0, 1]$  zurückliefert. In den nächsten Unterkapiteln wird kurz darauf eingegangen, wie man daraus eine entsprechende Zufallszahl einer anderen Verteilung errechnet. Dabei kann eine Methode verwendet werden, die mithilfe der Verteilungsfunktion eine Zufallsgröße einer beliebigen Verteilung erzeugt. Die folgenden Abschnitte der Zufallserzeugung befinden sich in ausführlicher Weise auch in [3].

### A.2.1 Inversionsmethode

Sei  $F : \mathbb{R} \rightarrow \mathbb{R}$  die Verteilungsfunktion einer Zufallsvariablen  $X$ . Dies bedeutet, dass

$$F(x) = \mathbb{P}(\{X \leq x\}), x \in \mathbb{R}.$$

Die **verallgemeinerte Umkehrfunktion** wird definiert als

$$F^{-1}(y) := \inf\{u \in \mathbb{R} | F(u) \geq y\}, y \in \mathbb{R};$$

wobei  $F^{-1}(1) = \infty$ , falls  $\{u \in \mathbb{R} | F(u) \geq 1\} = \emptyset$ .

Der folgende Satz findet sich in [3],[18],[6], in denen die Inversenmethode näher beschrieben wird, und lässt sich dafür anwenden, um die gewünschten Zufallszahlen mithilfe der verallgemeinerten Umkehrfunktion zu erhalten.

**Satz A.2.** *Sei  $F$  eine Verteilungsfunktion einer Zufallsvariablen  $X$ . Weiters sei  $U$  eine gleichverteilte Zufallsvariable im Intervall  $[0, 1]$ . Dann ist  $Y := F^{-1} \circ U$  eine Zufallsvariable einer Wahrscheinlichkeitsverteilung mit Verteilungsfunktion  $F$ .*

### A.2.2 Bernoulli-Verteilung

Hierbei handelt es sich um Experimente mit nur zwei möglichen Resultaten, die als 0 und 1 codiert werden. Die Wahrscheinlichkeit, mit der bei einem Experiment "1" eintritt, wird mit dem Parameter  $p$  bezeichnet. Die Dichtefunktion der Bernoulli-Verteilung lautet

$$f(x) = \begin{cases} p & \text{für } x = 1 \\ 1 - p & \text{für } x = 0 \\ 0 & \text{sonst} \end{cases} \quad (\text{A.1})$$

Es ist nun keine schwere Aufgabe aus einer gleichverteilten Zufallszahl im Intervall  $[0, 1]$  eine Bernoulli-verteilte Zufallszahl mit Parameter  $p$  zu transformieren. Folgende Funktion  $\text{BERNOULLI}(p)$  realisiert dies.

---

```

1: procedure BERNOULLI( $p$ )
2:   if UNIFORM()  $\leq p$  then
3:     return 1
4:   else
5:     return 0
6:   end if
7: end procedure

```

---

### A.2.3 Geometrische Verteilung

Der geometrischen Verteilung liegt ein Bernoulli-Experiment mit Parameter  $p$  zugrunde. Die Frage des geometrischen Experiments ist nun, wie oft das Bernoulli-Experiment ausgeführt werden muss, bis als Ergebnis "1" erscheint. Die geometrische Verteilung übernimmt den Parameter  $p$ , nämlich die "1"-Wahrscheinlichkeit bei jedem einzelnen Versuch. Die Dichtefunktion der geometrischen Verteilung lautet

$$f(x) = (1 - p)^{x-1} p.$$

Die Verteilungsfunktion der geometrischen Verteilung ist

$$F(x) = 1 - (1 - p)^x$$

Damit lässt sich die verallgemeinerte Umkehrfunktion ableiten.

## A Anhang

$$F^{-1}(x) = \lfloor \frac{\log(1-x)}{\log(1-p)} \rfloor$$

Um nun zu einer geometrisch verteilten Zufallszahl zu kommen, wird nun die Inversionsmethode angewendet. Sei  $U$  eine gleichverteilte Zufallszahl im Intervall  $[0, 1]$ . Dann lässt sich daraus eine geometrisch verteilte Zufallszahl  $X$  transformieren durch folgende Prozedur.

---

```
1: procedure GEOM( $p$ )
2:   return  $\lfloor \frac{\log(U)}{\log(1-p)} \rfloor$ 
3: end procedure
```

---

### A.2.4 Poissonverteilung

Die Poissonverteilung ist eine diskrete Wahrscheinlichkeitsverteilung, mit der die Anzahl von Ereignissen in einem festen Zeitintervall oder räumlichen Gebiet eintreten. Der Parameter  $\lambda$  gibt den Erwartungswert an. Die Dichtefunktion der Poissonverteilung ist

$$f(x) = \frac{\lambda^x}{x!} \exp(-\lambda)$$

Die Verteilungsfunktion der Poissonverteilung ist

$$F(x) = \exp(-\lambda) \sum_{k=0}^n \frac{\lambda^k}{k!}$$

Für die Erzeugung einer poissonverteilten Zufallszahl wird folgender Algorithmus angewendet.

---

```
1: procedure POISSON( $\lambda$ )
2:    $I := 0$ 
3:    $P := \text{UNIFORM}()$ 
4:   while  $P \geq e^{-\lambda}$  do
5:      $I := I + 1$ 
6:      $P := P \cdot \text{UNIFORM}()$ 
7:   end while
8:   return  $I$ 
9: end procedure
```

---

### A.2.5 Logarithmische Verteilung

Eine diskrete Zufallsgröße  $X$  ist logarithmisch verteilt mit dem Parameter  $\lambda$ , wenn sie die Dichtefunktion

$$f(x) = \frac{\lambda^x}{x} \cdot \frac{1}{-\log(1-\lambda)}$$

besitzt. Folgende Prozedur liefert eine logarithmisch verteilte Zufallszahl.

---

```

1: procedure LOGARITH( $\lambda$ )
2:    $I := 0$ 
3:    $U := \text{UNIFORM}()$ 
4:    $S := 0$ 
5:   while  $U < S$  do
6:      $I := I + 1$ 
7:      $S := S + \frac{\lambda^I}{I} \cdot \frac{1}{-\log(1-\lambda)}$ 
8:   end while
9:   return  $I$ 
10: end procedure

```

---

### A.2.6 Komplexitätsanalyse

Vorerst sei angenommen, die Erzeugung einer gleichverteilten Zufallszahl im Intervall  $[0, 1]$  habe eine konstante Anzahl an arithmetischen Operationen, d.h. befindet sich in  $\mathcal{O}(1)$ .

Durch einen einzigen Rechenschritt lässt sich daraus eine Bernoulli- und eine geometrischverteilte Zufallszahl transformieren. Beim Generator der Poisson- und der logarithmisch verteilten Zufallszahl sind  $\mathcal{O}(i + 1)$  arithmetische Operationen notwendig, wenn die Prozedur  $i$  als Ausgabe zurückgibt.

## A.3 Laplace-Transformation und Faltung

Der Anwendungsbereich der sogenannten Laplace-Transformation ist die Analyse von linearen und zeitinvarianten Systemen wie lineare Differentialgleichungen mit konstanten Koeffizienten. Die Theorie der Laplace-Transformation und der Faltung findet man unter anderem in [14]. Sie ist folgendermaßen definiert.



## A Anhang

**Definition A.1.** Sei eine Funktion  $f : \mathbb{R}^+ \rightarrow \mathbb{C}$  gegeben. Dann nennt man

$$\mathcal{L}(f)(s) := \int_0^{\infty} f(t) \exp(-st) dt, \quad s \in \mathbb{C},$$

die **Laplace-Transformierte** von  $f(t)$ , wenn dieses Integral existiert.

Sei  $F$  die Laplacetransformierte von  $f$ . Die Funktion  $f$  kann durch die Umkehrformel

$$\mathcal{L}^{-1}(F(s)) = \frac{1}{2\pi i} \int_{\gamma-i\infty}^{\gamma+i\infty} \exp(st) F(s) ds$$

bestimmt werden. Dabei ist  $\gamma$  größer als die Konvergenzabszisse von  $F$ .

Die Faltung ist eine sehr wichtige Operation in der Signalverarbeitung und ist wie folgt definiert:

**Definition A.2.** Seien zwei Funktionen  $f, g : \mathbb{R}^n \rightarrow \mathbb{C}$  gegeben. Dann nennt man

$$(f \star g)(x) := \int_{\mathbb{R}^n} f(t) g(x-t) dt$$

die **Faltung** von  $f$  und  $g$ .

**Satz A.3.** Seien zwei Funktionen  $f, g : \mathbb{R}^n \rightarrow \mathbb{C}$  gegeben. Dann [22] gilt

$$\mathcal{L}(f \star g) = \mathcal{L}(f) \star \mathcal{L}(g)$$

*Beweis.*

$$\begin{aligned} \mathcal{L}(f \star g)(s) &= \int_0^{\infty} (f \star g)(x) \exp(-sx) dx = \int_0^{\infty} \int_0^x f(t) g(x-t) dt \exp(-sx) dx = \\ &= \int_0^{\infty} \int_t^{\infty} f(t) g(x-t) \exp(-sx) dx dt = \int_0^{\infty} f(t) \exp(-sx) \int_t^{\infty} g(x-t) \exp(-s(x-t)) dx dt = \\ &= \int_0^{\infty} f(t) \exp(-sx) \int_0^{\infty} g(p) \exp(-sp) dp dt = \int_0^{\infty} f(t) \exp(-sx) dt \int_0^{\infty} g(p) \exp(-sp) dp = \mathcal{L}(f)(s) \cdot \mathcal{L}(g)(s) \end{aligned}$$

□

# Literaturverzeichnis

- [1] Robert B. Ash. *Real analysis and probability*. Academic Press, New York-London, 1972. Probability and Mathematical Statistics, No. 11.
- [2] Michèle Darrasse, Alexis und Soria. Degree distribution of random Apollonian network structures and Boltzmann sampling. In *2007 Conference on Analysis of Algorithms, AofA 07*, Discrete Math. Theor. Comput. Sci. Proc., AH, pages 313–324. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2007.
- [3] Luc Devroye. *Non-uniform random variate generation*. Springer-Verlag, New York, 1986.
- [4] Robert Flajolet, Philippe und Sedgewick. *Analytic combinatorics*. Cambridge University Press, Cambridge, 2009.
- [5] Rolf Freitag, Eberhard und Busam. *Funktionentheorie*. Springer-Lehrbuch. [Springer Textbook]. Springer-Verlag, Berlin, 1993.
- [6] Paul Glasserman. *Monte Carlo methods in financial engineering*, volume 53 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 2004. Stochastic Modelling and Applied Probability.
- [7] Ulrich Hedtstück. *Simulation diskreter Prozesse: Methoden und Anwendungen*. Springer Verlag, 2013.
- [8] Maurice Heins. *Complex function theory*. Pure and Applied Mathematics, Vol. 28. Academic Press, New York-London, 1968.
- [9] Peter Henrici. An algebraic proof of the Lagrange-Bürmann formula. *J. Math. Anal. Appl.*, 8:218–224, 1964.
- [10] André Joyal. Une théorie combinatoire des séries formelles. *Adv. in Math.*, 42(1):1–82, 1981.
- [11] Donald E. Knuth. Mathematical analysis of algorithms. pages 19–27, 1972.
- [12] Gilbert Labelle. éclosions combinatoires appliquées à l'inversion multidimensionnelle des séries formelles. *J. Combin. Theory Ser. A*, 39(1):52–82, 1985.

## Literaturverzeichnis

- [13] Gilbert Labelle. Dérivées directionnelles et développements de Taylor combinatoires. *Discrete Math.*, 79(3):279–297, 1990.
- [14] Joel L. Schiff. *The Laplace transform*. Undergraduate Texts in Mathematics. Springer-Verlag, New York, 1999. Theory and applications.
- [15] Alan Slomson. *An introduction to combinatorics*. Chapman and Hall Mathematics Series. Chapman and Hall, Ltd., London, 1991.
- [16] Duchon P. und Flajolet P. und Louchard G. und Schaeffer G. Boltzmann samplers for the random generation of combinatorial structures. *Combin. Probab. Comput.*, 13(4-5):577–625, 2004.
- [17] Bodini O. und Genitrini A. und Rolin N. Pointed versus singular boltzmann samplers: a comparative analysis. *Pure Math. Appl. (P.U.M.A.)*, 25(2):115–131, 2015.
- [18] Becker T. und Herrmann R. und Sandor V. und Schäfer D. und Wellisch U. *Stochastische Risikomodelierung und statistische Methoden*. Statistik und ihre Anwendungen. Chapman and Hall, Ltd., London, 2016.
- [19] Darrasse A. und Panagiotou K. und Roussel O. und Soria M. Boltzmann generation for regular languages with shuffle. In *GASCOM 2010 - Conference on random generation of combinatorial structures*, Montréal, Canada, September 2010. 12 pages.
- [20] Pivoteau C. und Salvy B. und Soria M. Boltzmann oracle for combinatorial systems. In *Fifth Colloquium on Mathematics and Computer Science*, Discrete Math. Theor. Comput. Sci. Proc., AI, pages 475–488. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2008.
- [21] Flajolet P. und Salvy B. und Zimmermann P. Automatic average-case analysis of algorithms. *Theoret. Comput. Sci.*, 79(1, (Part A)):37–109, 1991. Algebraic and computing treatment of noncommutative power series (Lille, 1988).
- [22] Flajolet P. und Zimmerman P. und Van Cutsem B. A calculus for the random generation of labelled combinatorial structures. *Theoret. Comput. Sci.*, 132(1-2):1–35, 1994.