Blind Search

Ferienakademie im Sarntal — Kurs 1
Moderne Suchmethoden der Informatik: Trends und Potenzial

Sophia Koch

7. November 2014

Zusammenfassung

Die einfachste Herangehensweise an die Optimierung einer Funktion ist die sogenannte "Blind Search"-Methode. Sie beschränkt sich darauf, von einem aktuellen Punkt ausgehend, durch zufällige Sprünge einen besseren Punkt zu finden.

Hierbei ist vor allem die Sprunganzahl bis zum Erreichen des Optimums wichtig, da sie äquivalent zur Laufzeit dieser Methode ist. Deshalb wird im folgenden Paper eine obere und untere Schranke für den Erwartungswert der Sprunganzahl hergeleitet.

Die obere Schranke ist der leichtere Teil und benötigt nur die Betrachtung über Markovketten. Für die untere Schranke wird der Intervallprozess eingeführt und mit der Hilfe einer eigens definierten Potenzialfunktion untersucht.

Das Ergebnis wird bei einer größe des Suchraums von n eine Beschränkung der Laufzeit auf $\Theta\left((\log n)^2\right)$ sein.

Inhaltsverzeichnis

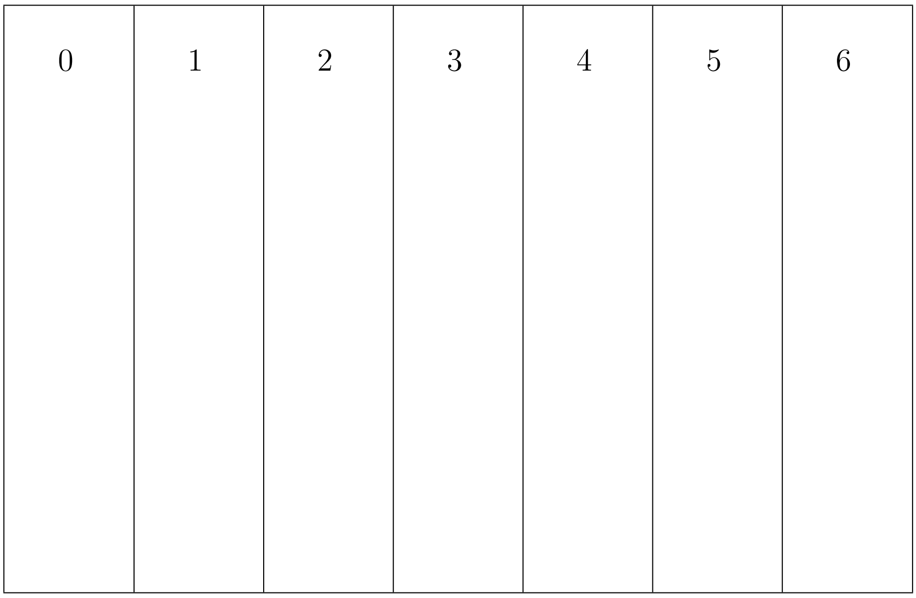

Abbildung 1: Blind Search - das Spiel

1 Einführung

Am einfachsten erklärt man "Blind Search" an einem kleinen Spiel, für das man einen Spielstein, ein Spielbrett (siehe 1) und einen Würfel braucht. Ziel des Spieles ist es mit möglichst wenigen Würfelwürfen die Null zu erreichen. Vor dem Spiel stellt man den Spielstein auf eine zufällige Position x mit $1 \leq x \leq 6$. In jedem Schritt wird dann gewürfelt und um die Augenzahl in Richtung der Null gezogen, falls sie hilfreich ist. Ist die Augenzahl zu groß, so verfällt sie ungenutzt.

Es ist wichtig sowohl nützliche als auch zu große, gewürfelte Augenzahlen als Würfelwürfe bis die Null erreicht ist zu zählen. Dieses Essay befasst sich mit einem Erwartungswert für genau diese Größe. Er wird in $\Theta\left((\log n)^2\right)$ liegen.

1.1 Diskreter Prozess

Dafür erfolgt zuerst eine mathematische Beschreibung des Spieles, dass als "Diskreter Prozess" bezeichnet wird. Das Spielbrett wird durch ein diskretes Intervall $A := \{0, \ldots, n\}$ beschrieben, wobei $n \in \mathbb{N}$ die Größe des Spielbretts und $x \in A$ die aktuelle Position des Spielsteines ist.

Der Würfel wird durch ein ebenfalls diskretes Intervall $D := \{1, \ldots, n\}$ von entsprechenden Sprunggrößen ersetzt. Für den diskreten Prozess müssen nicht alle Sprunggrößen gleich wahrscheinlich sein, deshalb gibt es eine Wahrscheinlichkeitsverteilung μ auf D anhand der in jedem Schritt eine Sprunggröße $d \in D$ ausgewählt wird.

Der Ablauf des Prozesses wird durch eine Folge $X := (X_0, X_1, \ldots)$ beschrieben, für die gilt:

- der Startzustand ist $X_0 \in A \backslash \{0\}$ und
- der Zustand nach dem t-ten Sprung wird ausgedrückt durch

$$X_t = \begin{cases} X_{t-1} - d & \text{falls } X_{t-1} \geq d \\ X_{t-1} & \text{falls } X_{t-1} < d \end{cases} \quad \text{für alle } t \in \mathbb{N}. \tag{1}$$

Die Anzahl der Würfelwürfe bis zum Erreichen der Null erhält eine eigene Variable $T_x := \min\{t \,|\, X_t = 0\}$ deren Erwartungswert $E_\mu(T_x)$ die Aufmerksamkeit dieses Essays gilt.

1.2 Motivation

Doch bevor wir uns näher mit dem besagten Erwartungswert beschäftigen, soll zuerst einmal die Frage geklärt werden, wie man auf diese Fragestellung kommt.

Der Ursprungsgedanke war die Suche nach einem Minimum einer Funktion $f : [0, 1] \to \mathbb{R}$ deren Definition nicht bekannt ist. Die einzige Möglichkeit etwas über diese Funktion herauszufinden ist Beispielpunkte auszuprobieren. So versucht man durch eine iterative Herangehensweise bei einem zufälligen Punkt zu beginnen und dann durch zufällige Störungen in positive oder negative Richtung den Funktionswert zu verbessern.

Diese Art der Suche wird als "blind" bezeichnet, weil nicht versucht wird herauszufinden, wie nahe man dem Minimum bereits ist um die Größenklasse der Störungen dann entsprechend anzupassen. Stattdessen steht die Wahrscheinlichkeit der einzelnen Störungsgrößen von vornherein fest, weshalb μ auch als Suchstrategie bezeichnet wird.

Um das Minimum aller möglicherweise auftretenden Funktionen zu finden ist es notwendig eine skalierungsunabhängige Wahrscheinlichkeitsverteilung zu haben, also eine bei der Störungen aller Größenordnungen ähnlich wahrscheinlich sind.

Diese erhält man, indem man zuerst eine minimale Störungsgröße ϵ festlegt. Über diese wird die Genauigkeit des Algorithmus $p := \ln\left(\frac{1}{\epsilon}\right)$ berechnet.

Mit diesen Werten kann man sich eine Wahrscheinlichkeitsdichtefunktion auf ganz $\mathbb{R}$ definieren:

$$h(t) := \begin{cases} \frac{1}{pt} & \text{falls } \epsilon \leq t \leq 1 \\ 0 & \text{sonst} \end{cases} \tag{2}$$

Das dies tatsächlich eine Wahrscheinlichkeitsdichtefunktion ist, kann in einer einfachen Rechnung nachgewiesen werden.

Eine entsprechende Sprunggröße aus dem Intervall $(0, 1]$ kann nun wie folgt ausgewählt werden: $d := \exp(-pu)$, wobei $u \in [0, 1]$.

Dies ergibt jedoch sehr komplizierte Berechnungen für den Erwartungswert der Schritte bis man die Null entsprechend gut angenähert hat, deshalb soll die Funktion jetzt vereinfach als unimodal angenommen werden. Das heißt, dass sie im Intervall $[0, x^*]$ streng fallend ist und im Intervall $[x^*, 1]$ streng steigend, wobei x^* das gesuchte Minimum der Funktion ist.

Für diese Funktion kann bereits eine erste Schranke für den Erwartungswert hergeleitet werden, indem man die Entfernung τ mit $\tau \geq 2\epsilon$ vom aktuellen Punkt x zum Minimum x^* betrachtet. Eine Störung d mit $\frac{\tau}{2} \leq d \leq \tau$ führt, angewendet auf die richtige Richtung, zu einer Halbierung der Entfernung zum Minimum. Die Wahrscheinlichkeit für eine solche Halbierung in einem Schritt ergibt sich damit aus dem Integral

$$\frac{1}{2} \int_{\frac{\tau}{2}}^{\tau} \frac{dt}{pt} = \frac{\ln 2}{2p} \tag{3}$$

Die Berechnung dieses Integrals ergibt, dass die Wahrscheinlichkeit einer Halbierung unabhängig von der aktuellen Position ist. Außerdem erhält man nach $\frac{2p}{\ln 2}$ Schritten sicherlich eine Halbierung. Da die Anzahl der nötigen Halbierungen logarithmisch im Kehrwert der minimalen Störungsgröße ist, folgt für die maximale Schrittanzahl: $\log\left(\frac{1}{\epsilon}\right) \cdot \frac{2p}{\ln 2} = O\left(\left(\log\left(\frac{1}{\epsilon}\right)\right)^2\right)$.

Diese Abschätzung gibt uns schon einen ersten Einblick, wir möchten aber eine genauere und mathematisch fundiertere Analyse durchführen, weshalb wir den betrachteten Prozess erneut vereinfachen, indem wir das Minimum fest auf die Null legen und dann die Schrittanzahl bis zum Erreichen dieser betrachten. Der wichtigste Vorteil dieser Vereinfachung ist, dass die Richtung, in die man Gehen muss um das Minimum zu erreichen nun eindeutig ist.

Dieser Prozess wurde bereits als “Diskreter Prozess” (siehe 1.1) eingeführt und soll im Folgenden untersucht werden.

2 Obere Schranke

Zuerst soll eine obere Schranke für den diskreten Prozess hergeleitet werden. Dies erfolgt über die Einführung von Markovketten.

2.1 Markovkette

Zur Vereinfachung wird folgende Notation für die Schreibweise der diskreten Intervalle eingeführt: $\langle a, b\rangle := \{a, \ldots, b\}$ für alle $a, b \in \mathbb{Z}$. μ soll weiterhin eine Wahrscheinlichkeitsverteilung über $\langle 1, n\rangle$ sein.

Der Prozess $X = (X_0, X_1, \ldots)$ über $A = \langle 0, n\rangle$ soll ab sofort als Markovkette verstanden werden, da für jeden Zustand gilt, dass er ausschließlich von dem vorangehenden Zustand abhängt. Für die Zustände folgt also:

- X_0 ist gleichmäßig verteilt in $A \setminus \{0\} = \langle 1, n\rangle$.
- Die Übergangswahrscheinlichkeit in den t-ten Zustand ist

$$Pr\left(X_t = x' \mid X_{t-1} = x\right) = \begin{cases} \mu\left(x - x'\right) & \text{falls } x' < x \\ 1 - \sum_{d=1}^{x} \mu\left(d\right) & \text{falls } x' = x \\ 0 & \text{falls } x' > x \end{cases} \tag{4}$$

Weiterhin wollen wir in dieser Analyse den Erwartungswert $E_\mu\left(T_x\right)$ in Abhängigkeit der Wahrscheinlichkeitsverteilung für die Schrittanzahl bis zum Erreichen der Null
$T_x = \min\{t \mid X_t = 0\}$ betrachten.

Aus diesen Informationen können bereits erste Folgerungen gezogen werde.

Bemerkung 1. $\mu\left(1\right) = 0 \Rightarrow E_\mu\left(T_x\right) = \infty$

Dies gilt, weil der Startzustand mit Wahrscheinlichkeit $\frac{1}{n}$ die Eins ist und im Falls $\mu\left(1\right) = 0$ für diesen Startzustand nie die Null erreicht werden kann. Deshalb soll für die weitere Analyse gelten:

Voraussetzung 1. $\forall\mu$ Wahrscheinlichkeitsverteilung: $\mu\left(1\right) > 0$

Ab sofort soll außerdem folgende vereinfachende Schreibweise benutzt werden:
$F\left(x\right) := \mu\left(\langle 1, x\rangle\right) = \sum\limits_{d=1}^{x} \mu\left(d\right)$ für $x \in A \setminus \{0\}$.

Damit kann bereits folgende Formel für den Erwartungswert hergeleitet werden.

Proposition 1. $E_\mu\left(T_x\right) = \frac{1}{n} \cdot \sum\limits_{1 \leq x_1 < \ldots < x_l \leq n} \frac{\mu(x_2 - x_1) \cdot \ldots \cdot \mu(x_l - x_{l-1})}{F(x_1) \cdot \ldots \cdot F(x_l)}$

Beweis. Bemerkung 1 (siehe 1) ergibt: $\forall x \in A \setminus \{0\}\colon\ F\left(x\right) > 0$ mit der Definition von $F\left(x\right)$. Weiter gilt für die Schrittanzahl bei Start in x:

$$T_x = \underbrace{1}_{\text{aktueller Schritt}} + \underbrace{\left(1 - F\left(x\right)\right) \cdot T_x}_{\text{W'keit: } d \in D \text{ zu große },T_x \text{ entsprechend}} + \underbrace{\sum_{d=1}^{x} \mu\left(d\right) \cdot T_{x-d}}_{\text{W'keiten: } d \in D \text{ passend },\ T_{x-d} \text{ “übrigbleibend”}} \tag{5}$$

Außerdem gilt $T_0 = 0$. Dieser Ausdruck wird nach T_x umgeformt:

$$T_x = \frac{1}{F(x)} \cdot \left(1 + \sum_{d=1}^{x} \mu\,(d) \cdot T_{x-d}\right) \text{ für alle } 1 \le x \le n \tag{6}$$

und $x = 1, \cdots, 3$ eingesetzt:

$$\begin{aligned} T_1 &= \tfrac{1}{\mu(1)} = \tfrac{1}{F(1)}, \\ T_2 &= \tfrac{1}{F(2)} + \tfrac{\mu(1)}{F(1)\cdot F(2)}, \\ T_3 &= \tfrac{1}{F(3)} + \tfrac{\mu(2)}{F(1)\cdot F(3)} + \tfrac{\mu(1)}{F(2)\cdot F(3)} + \tfrac{(\mu(1))^2}{F(1)\cdot F(2)\cdot F(3)} \end{aligned} \tag{7}$$

Durch Induktion erhält man damit:

$$T_x = \sum_{1 \le x_1 < \ldots < x_l = x} \frac{\mu\,(x_2 - x_1) \cdot \ldots \cdot \mu\,(x_l - x_{l-1})}{F\,(x_1) \cdot \ldots \cdot F\,(x_l)} \tag{8}$$

Zur Berechnung des Erwartungswert für T_x muss nun noch (8) über alle möglichen Startpositionen aufsummiert werden:

$$E_\mu(T_x) = \frac{1}{n} \cdot \sum_{1 \le x_1 < \ldots < x_l \le n} \frac{\mu\,(x_2 - x_1) \cdot \ldots \cdot \mu\,(x_l - x_{l-1})}{F\,(x_1) \cdot \ldots \cdot F\,(x_l)} \tag{9}$$

□

Diese Herleitung ist zwar richtig, hilft aber für die obere Schranke leider nicht weiter.

2.2 Herleitung

Seien $L := \lfloor \log_2 n \rfloor$, $l_i := \langle 2^i, 2^{i+1} - 1 \rangle$ für alle $0 \le i < L$, $I_L := \langle 2^L, n \rangle$ und $p_i := \sum_{d \in I_i} \mu\,(d)$, wobei $0 \le i \le L$. Die Wahrscheinlichkeit beim Start im i-ten oder einem größeren Intervall eine Sprunggröße $d \ge 2^{i-1}$ zu bekommen ist

$$Pr\left(X_t \le X_{t-1} - 2^{i-1} \mid X_{t-1} \ge 2^i\right) \ge p_{i-1} \text{ wobei } t \ge 1,\ i \ge 1 \tag{10}$$

Da man maximal zwei solches Sprunggrößen braucht um das i-te Intervall sicher wieder verlassen zu haben, kann man die Schrittanzahl bis diese sicher eingetreten sind für alle Intervalle abwärts aufsummieren. Zusätzlich gilt: nach dem Verlassen des ersten Intervalls befinden wir uns in Null oder Eins und brauchen maximal noch einmal die Sprunggröße eins. Damit erhalten wir folgende obere Schranke:

$$T_x \le \frac{2}{p_{j-1}} + \frac{2}{p_{j-2}} + \ldots + \frac{2}{p_1} + \frac{3}{p_0} \text{ für } x \in I_j \tag{11}$$

Die p_i können nun mit $\alpha > 0$ konstant abgeschätzt werden durch: $p_0, \ldots, p_{L-1} \ge \frac{\alpha}{L}$. Da über $2j - 1$ Elemente aufsummiert wird, erhalten wir: $T_x \le (2j+1)\,\frac{L}{\alpha}$. Da $x \in I_j$ also $j \le \log x$ und mit der Definition von L folgt der Erwartungswert:

$$E_\mu\,(T_x) = O\,((\log x)\,(\log n)) = O\left((\log n)^2\right) \tag{12}$$

3 Untere Schranke

Da jetzt die obere Schranke vollständig bewiesen ist, wenden wir uns der unteren zu. Im Laufe dieses Kapitels soll für sie folgendes gezeigt werden:

Satz 1. $\forall\,\mu\colon\ E_\mu\,(T_x) = \Omega\left((\log n)^2\right)$

Beweis: später

Weiter werden wir die folgende Kenntnisse aus den vorangehenden Kapiteln benutzen:

Voraussetzung 2. μ fest, $\mu\,(1) > 0$, $I_0, \ldots, I_L$, $\forall 0 \le i \le L:\ p_i$.

3.1 Intuition

Zuerst soll mit einer intuitiven Herangehensweise an die untere Schranke begonnen werden.

Der Erwartungswert für die Schrittanzahl zum Durchlaufen des i-ten Intervalls bei Start im (i+1)-ten ist abhängig von der Schrittanzahl die nötig ist, bis eine Sprunggröße kommt, die wirklich hilfreich ist, also $d \in \{I_{i-1}, \cdots, I_{i-2}\}$. Damit ist dieser Erwartungswert von unten beschränkt durch $\frac{1}{p_{i+1}+p_i+p_{i-1}+p_{i-2}}$.

Angenommen $\exists \beta > 0$ konstant mit $\forall\, 0 \le i < L-1$ gilt I_{i+1} wird mit Wahrscheinlichkeit β besucht, dann gilt:

$$E\left(T_x\right) \geq \sum_{j=1}^{\frac{L}{2}-1} \frac{\beta}{p_{2j+1} + p_{2j} + p_{2j-1} + p_{2j-2}} \tag{13}$$

Es wird statt p_i über p_{2j} aufsummiert um die Vierfachzählung der Wahrscheinlichkeit einer Sprunggröße aus dem jeweiligen Intervall zu verhindern.

Nun ist der aufsummierte Quotient kleinstmöglich, falls alle p_i den selben Wert haben. Die p_i's werden dopellt aufsummiert, damit muss für alle $i < L$ gelten: $p_i = \frac{2}{L}$. Eingesetzt ergibt sich damit die untere Schranke:

$$E\left(T_x\right) \geq \beta \cdot \frac{L}{2} \cdot \frac{L}{2} \cdot \frac{1}{2} = \Omega\left(L^2\right) = \Omega\left((\log n)^2\right) \tag{14}$$

Leider kann man nicht für alle Wahrscheinlichkeitsverteilungen ein solches β finden, sodass dies nicht in einem formellen Beweis umgesetzt werden kann.

3.2 Intervallprozess

Um den stetigen Prozess weiter analysieren zu können, soll zuerst eine andere Sichtweise auf diesen, der sogenannte Intervallprozess eingeführt werden.

Der Ablauf dieses Prozesses wird mit $Y := (Y_0, Y_1, \ldots)$ bezeichnet. Hierbei soll für Y_0 nicht ein konkreter Wert festgelegt werden, sondern nur festgestellt werden, dass x zufällig gleichverteilt in $A \setminus \{0\}$ ist. Weiter soll gelten:

$$Y_t \begin{cases} > 0 & \Rightarrow \; X_t \text{ gleichmäßig in } \langle 1, Y_t \rangle \text{ verteilt} \\ = 0 & \Rightarrow \; X_t = 0 \end{cases} \tag{15}$$

also $Y_0 = n$.

Außerdem kann $Y_t \in \langle 1, n \rangle$ in Abhängigkeit zu seinem Vorgänger beschrieben werden durch

$$Y_t = \begin{cases} 0 & \text{falls } x = d \quad \text{mit Wahrscheinlichkeit } \frac{1}{y} \\ Y_{t-1} - d & \text{falls } x > d \quad \text{mit Wahrscheinlichkeit } \frac{y-d}{y} \\ d-1 & \text{falls } x < d \quad \text{mit Wahrscheinlichkeit } \frac{d-1}{y} \end{cases} \tag{16}$$

wobei $d \in D$ bzgl. μ und $d \leq Y_{t-1}$ gelten soll, weil für $d > Y_{t-1}$ offensichtlich gilt, dass es nicht hilfreich ist.

Des weiteren ist auch der Intervallprozess eine Markovkette weshalb die Übergangswahrscheinlichkeiten $Pr\left(Y_t = y' \mid Y_{t-1} = y\right)$ betrachtet werden können (siehe 2)

Die Schrittanzahl bis zum Erreichen der Null soll mit $T_y := \min\{t \mid Y_t = 0\}$ bezeichnet werden. Für den Erwartungswert soll ohne Beweis gelten: $E_\mu\left(T_x\right) = E_\mu\left(T_y\right)$.

3.3 Potenzialfunktion

Um eine untere Schranke für diesen Intervallprozess herleiten zu können benötigen wir außerdem eine Potenzialfunktion. Mit ihrer Hilfe werden wir über das Potenzial des Anfangszustandes $\Phi(Y_0) = \Omega\left((\log n)^2\right)$, die Beschränkung $E\left(\Phi\left(Y_{t-1}\right) - \Phi\left(Y_t\right) \mid Y_{t-1} = y\right) = O\left(1\right)$ und

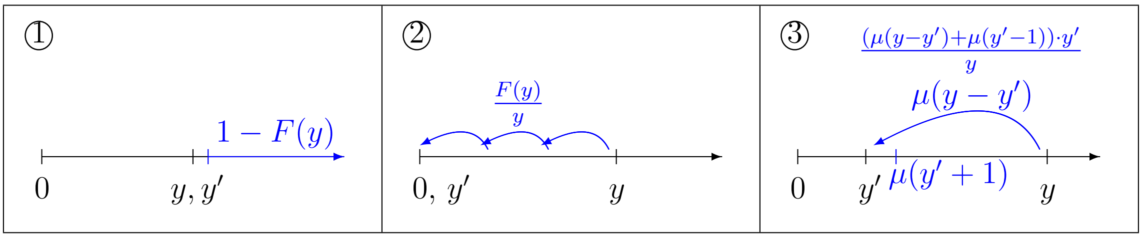

Abbildung 2: Übergangswahrscheinlichkeiten des Intervallprozesses

die daraus resultierende Folgerung für den Erwartungswert $E(T_y) = \Omega(\Phi(Y_0))$ die untere Schranke zeigen können.

Weiter werden wir $\psi_i := \frac{1}{\sum_{j=0}^{L} p_j \cdot c^{|j-i|}}$ für c konstant mit $\frac{1}{2} < c < 1$ brauchen. Es gilt:

$$\sum_{i=1}^{L-1} \psi_i^{-1} = \sum_{i=1}^{L-1} \sum_{j=0}^{L} p_j \cdot c^{|j-i|} = \sum_{j=0}^{L} p_j \sum_{i=1}^{L-1} c^{|j-i|} = O(1) \tag{17}$$

Die Summen können wegen der Kommutativität vertauscht werden. Die innere Summe kann mit der geometrischen Reihe abgeschätzt werden und die p_j's sind so definiert, dass sie aufsummiert eins ergeben.

Nach diesen Vorüberlegungen folgt nun die eigentliche Definition der Potenzialfunktion.

Definition 1. Für die Potenzialfunktion Φ werden σ und φ wie folgt definiert:

$$\begin{aligned}
\sigma_a &:= \sum_{d=1}^{n} \mu(d) \cdot 2^{\frac{-|\log a - \log d|}{2}} = \sum_{d=1}^{a} \mu(d) \sqrt{\tfrac{d}{a}} + \sum_{d=a+1}^{n} \mu(d) \sqrt{\tfrac{a}{d}} \text{ für } 1 \le a \le n, \\
\varphi_a &:= \frac{1}{a\sigma_a} \text{ wobei } 1 \le a \le n, \\
\Phi(y) &:= \sum_{a=1}^{y} \varphi_a \text{ für alle } 0 \le y \le n \text{ und} \\
\Phi_t &:= \Phi(Y_t) \text{ wobei } t = 0, 1, \ldots
\end{aligned}$$

Lemma 1. *Für die Potenzialfunktion gilt:*

1. $\Phi_t,\ t \ge 0$ *steigt nicht, wenn* t *steigt,*
2. $\Phi_t = 0 \Leftrightarrow Y_t = 0$ *und*
3. $\Phi_0 = \Omega\left((\log n)^2\right)$.

Beweis. zu 1. Folgt mit $\forall n \in \mathbb{N} : Y_{t+n} \le Y_t$ und der Definition der Potenzialfunktion.

zu 2. $\Phi_t = 0$ genau dann, wenn $\Phi(Y_t) = \sum_{a=1}^{Y_t} \varphi_a$ leere Summe (da φ_1 immer $\neq 0$) genau dann, wenn $Y_t = 0$.

zu 3. Sei $i(a) := \lfloor \log_2 a \rfloor = \max\{i \,|\, 2^i \le a\}$ für $1 \le a \le n$. Dann gilt:

$$\begin{aligned}
\sigma_a &\overset{\text{Def.}}{=} \sum_{d=1}^{n} \mu(d) \cdot 2^{\frac{-|\log a - \log d|}{2}} \\
&\overset{\text{Summe aufspalten}}{\le} \sum_{j \le i(a)} \sum_{d \in I_j} \mu(d) \cdot 2^{\frac{j+1-i(a)}{2}} + \sum_{j > i(a)} \sum_{d \in I_j} \mu(d) \cdot 2^{\frac{i(a)+1-j}{2}} \\
&\overset{\text{Def. } p_j}{=} \sum_{j \le i(a)} p_j \cdot 2^{\frac{j+1-i(a)}{2}} + \sum_{j > i(a)} p_j \cdot 2^{\frac{i(a)+1-j}{2}} \\
&\overset{2^{\frac{i(a)+1-j}{2}} = 2^{\frac{i(a)-j}{2}} \cdot 2^{\frac{1}{2}} = c^{|j-i(a)|} \cdot 2c}{=} 2c \cdot \left(\sum_{j=0}^{L} p_j \cdot c^{|j-i(a)|} \right) \text{ für } c := \frac{1}{\sqrt{2}}
\end{aligned} \tag{18}$$

Damit erhält man:

$$\sum_{a\in l_i}\varphi_a \overset{\text{Def.}}{=} \sum_{a\in I_i}\frac{1}{a\sigma_a} \overset{(18)}{\geq} \frac{\overbrace{2^i}^{\text{Anzahl Elemente in } I_i}}{2c\cdot \underbrace{2^{i+1}}_{\text{größtes Element in } I_i} \cdot \left(\sum_{j=0}^{L} p_j \cdot c^{|j-i|}\right)} \overset{\text{Def. } \psi_i}{=} \frac{\psi_i}{4c} \tag{19}$$

und

$$\sum_{i=0}^{L-1}\frac{\psi_i}{4c} \overset{(19)}{\leq} \sum_{i=0}^{L-1}\sum_{a\in I_i}\varphi_a \overset{\substack{\text{Summen}\\ \text{zusammenfassen}}}{=} \sum_{i=1}^{2^L-1}\varphi_i \overset{\text{Def. }\Phi}{=} \Phi(2^L-1) \overset{\text{Def. }\Phi}{\leq} \Phi(n) \overset{\Phi(Y_0)}{=} \Phi_0 \tag{20}$$

Setze $u_i := \frac{4c}{\psi_i}$ für $0 \leq i < L$. Es gilt $\sum\limits_{i=0}^{L-1} u_i \leq k$ wobei k konstant mit (17). Weiter gilt $\sum\limits_{i=0}^{L-1}\frac{1}{u_i}$ minimal, falls $\forall u_i : u_i = \frac{k}{L}$.
Damit folgt die Behauptung:

$$\Phi_0 \geq L\cdot\frac{L}{k} = \frac{L^2}{k} = \Omega\left((\log n)^2\right) \tag{21}$$

□

Somit ist der erste Teil der bereits angekündigten Resultate gezeigt.

3.4 Hauptlemma

Um den Beweis für die untere Schranke zu vervollständigen, muss noch der zweite Teil gefolgert werden.

Lemma 2 (Hauptlemma)**.** $\exists C$ *konst.* $\forall 0\leq y\leq n : E\left(\Phi_{t-1}-\Phi_t \mid Y_{t-1}=y\right) \leq C$

Beweis: später

Lemma 3. *Seien $R_1, R_2, \ldots$ beschränkte Zufallsvariablen, $g > 0$ konstant und $S := \min\{t \mid R_1+\ldots+R_t \geq g\}$. Falls $E(S) < \infty$ und $\forall t\in\mathbb{N} : E(R_t \mid S\geq t) \leq C$, dann gilt: $E(S) \geq \frac{g}{C}$*

Beweis: $\left(\forall\, \mu :\ E_\mu(T_x) = \Omega\left((\log n)^2\right)\right)$. Es gilt: $T_\Phi = \min\{t \mid \Phi_t = 0\} \overset{1}{=} \min\{t \mid Y_t = 0\} = T_y$.
Damit bleibt zu zeigen: $E(T_\Phi) = \Omega\left((\log n)^2\right)$

Setze $R_t := \Phi_{t-1}-\Phi_t$. Dann gilt mit Lemma 2 $\forall y\geq 1 :\ E(R_t \mid Y_{t-1}=y) \leq C$. Das ist äquivalent zu $E(R_t \mid T_y \geq t) = E(R_t \mid Y_{t-1} > 0) \leq C$, da $Y_{t-1} = y \geq 1 \Leftrightarrow T_y > t \Leftrightarrow Y_{t-1} > 0$.

Weiter gilt mit der Teleskopreihe $R_1+\ldots+R_t = \Phi_0 - \Phi_t$. Damit erhalten wir $T_\Phi = min\{t \mid R_1+\ldots+R_t \geq \Phi_0\}$.

Mit Lemma 3 folgt die Behauptung: $E(T_\Phi) = \Omega\left((\log n)^2\right)$ □

Beweis: $\left(E\left(\Phi_{t-1}-\Phi_t \mid Y_{t-1}=y\right) = O(1)\right)$. Für diesen Beweis soll an einigen Stellen auf die genauen Abschätzungen verzichtet werden, da diese recht technisch sind. Der zu zeigende Erwartungswert soll zuerst als Summe über $\Delta(y,y') = (\Phi(y)-\Phi(y'))\cdot Pr(Y_t = y' \mid Y_{t-1}=y)$ geschrieben werden:

$$E\left(\Phi_{t-1}-\Phi_t \mid Y_{t-1}=y\right) = \sum_{y'=0}^{y}\Delta\left(y,y'\right) \tag{22}$$

Die $\Delta(y,y')$ werden einzeln abgeschätzt.

Beginnend mit $\Delta(y,y) = 0$, weil $\Phi(y) - \Phi(y) = 0$.

Weiter folgt: $\Delta(y,0) = \underbrace{\frac{1}{y} \cdot \underbrace{\left(\sum_{d\le y} \mu(d)\right)}_{F(y)}}_{\text{Übergangsw'keit}} \cdot \underbrace{\left(\sum_{a\le y} \varphi(a)\right)}_{\Phi(y)} \le \ldots < 2$

Fehlt nur noch:

$$\sum_{y'=1}^{y-1} \Delta(y,y') = \sum_{y'=1}^{y-1} \underbrace{\left(\sum_{a=y'+1}^{y} \varphi_a\right)}_{=\sum_{a=0}^{y}\varphi_a - \sum_{a=0}^{y'}\varphi_a} \cdot \underbrace{\frac{y'}{y}\left(\mu(y'+1) + \mu(y-y')\right)}_{\text{Übergangsw'keit}} = \frac{1}{y} \cdot \sum_{a=2}^{y} (\lambda_a + \gamma_a) \tag{23}$$

wobei $\lambda_a = \varphi_a \cdot \sum_{y'=1}^{a-1} \mu(y'+1)\, y'$ und $\gamma_a = \varphi_a \cdot \sum_{y'=1}^{a-1} \mu(y-y')\, y'$

λ_a und γ_a können wie folgt abgeschätzt werden: $\lambda_a < 1,\ \gamma_a \le \sqrt{\frac{y}{y-a+1}} + \sqrt{\frac{y}{a}}$. Eingesetzt ergibt sich damit: $\sum_{y'=1}^{y-1} \Delta(y,y') < \frac{1}{y} \cdot \sum_{a=2}^{y} \left(1 + \sqrt{\frac{y}{y-a+1}} + \sqrt{\frac{y}{a}}\right) < 5$

Damit erhalten wir die Behauptung: $\sum_{y'=0}^{y} \Delta(y,y') < 0 + 2 + 5 = 7$. □

So ist letztendlich auch die untere Schranke bewiesen.

4 Ausblick

In der Realität hat man aber in den wenigsten Fällen diskrete Funktionen zu optimieren. Deshalb kann man auch den so genannten stetigen Prozess betrachten. Dafür gilt: $A := [0, n+1)$, $n \in \mathbb{N}$, $x \in A$ und $D := (0, n+1)$, μ Wahrscheinlichkeitsverteilung, $d \in D$. Der Ablauf des Prozesses wird durch $X := (X_0, X_1, \ldots)$ beschrieben, wobei $X_0 \in [1, n+1)$. Für den t-ten Zustand in Abhängigkeit des vorherigen gilt:

$$X_t = \begin{cases} X_{t-1} & \text{falls } X_{t-1} < d \\ 0 & \text{falls } X_{t-1} \in [d, d+1) \\ X_{t-1} - d & \text{falls } X_{t-1} \ge d+1 \end{cases} \tag{24}$$

mit $t \in \mathbb{N}$. Weiter soll $T_x := \min\{t \mid X_t = 0\}$ die Schrittanzahl bis zum Erreichen der Null sein und wird der Erwartungswert $E_\mu(T_x)$ genauer untersucht.

Dies führt jedoch für dieses Paper zu weit und soll hier nur als Ausblick dienen.

5 Schluss

Abschließend lässt sich sagen, dass die Optimierungsmethode "Blind Search" nicht die schnellste ist, dafür aber relativ intuitiv und leicht zu implementieren.

Literatur

[1] Dietzfelbinger, Martin; Rowe, Jonathan E.; Wegener, Ingo; Woelfel, Philipp: "Tight Bounds for Blind Search on the Integers and the Reals". In: Combinatorics, Probability and Computing, Nr. 19, 2010, S. 711-728