Die Collatz-Vermutung und nicht-triviale Collatz-Zyklen

Christian Hercher

Europa-Universität Flensburg

13.09.2017

Überblick

Definitionen

Sei $C : \mathbb{N}_{>0} \to \mathbb{N}_{>0}$ gegeben durch

$$n \mapsto \begin{cases} \frac{n}{2} & \text{für } n \text{ gerade} \\ 3n+1 & \text{für } n \text{ ungerade.} \end{cases}$$

Dann heißt die Folge C_n, die gegeben ist durch

$$n, \; C(n), \; C(C(n)), \; \dots$$

Collatz-Folge zur Startzahl n.

Definitionen

Sei $T : \mathbb{N}_{>0} \to \mathbb{N}_{>0}$ gegeben durch

$$n \mapsto \begin{cases} \frac{n}{2} & \text{für } n \text{ gerade} \\ \frac{3n+1}{2} & \text{für } n \text{ ungerade.} \end{cases}$$

Dann heißt die Folge T_n, die gegeben ist durch

$$n, \ T(n), \ T(T(n)), \ \ldots$$

reduzierte Collatz-Folge zur Startzahl n.

Beispiele

Die Collatz-Folge C_{13} zur Startzahl 13 lautet wie folgt:

13 , 40 , 20 , 10, 5 , 16 , 8, 4, 2, 1 , 4 , . . .

Die Collatz-Folge C_7 zur Startzahl 7 lautet wie folgt:

7 , 22 , 11 , 34 , 17 , 52 , 26 , 13 , 40 , 20, 10, 5, 16, 8, 4, 2, 1, 4, . . .

Beispiele

Die Collatz-Folge C_{27} zur Startzahl 27 lautet wie folgt:

27 , 82, 41, 124, 62, 31 , 94, 47, 142, 71, 214, 107, 322, 161, 484, 242, 121, 364, 182, 91, 274, 137, 412, 206, 103, 310, 155, 466, 233, 700, 350, 175, 526, 263, 790, 395, 1186 , 593, 1780, 890, 445, 1336, 668, 334, 167, 502, 251, 754, 377, 1132, 566, 283, 850, 425, 1276, 638, 319, 958, 479, 1438, 719, 2158, 1079, 3238,1619, 4858, 2429, 7288, 3644, 1822, 911, 2734, 1367, 4102, 2051, 6154, 3077, 9232, 4616, 2308, 1154, 577, 1732, 866, 433, 1300, 650, 325, 976, 488, 244, 122, 61, 184, 92, 46, 23, 70, 35, 106, 53, 160, 80, 40, 20, 10, 5, 16, 8, 4, 2, 1, 4, . . .

Collatz-Vermutung

Vermutung (Collatz-Vermutung)

Für jede Startzahl $n \in \mathbb{N}_{>0}$ geht die Collatz-Folge C_n mit Startzahl n irgendwann in den trivialen Zyklus $4 \to 2 \to 1 \to 4 \to \dots$ über.

Für die reduzierte Collatz-Folge T_n zur Startzahl n gilt analog, dass sie in den trivialen Zyklus $2 \to 1 \to 2 \to \dots$ übergeht.

Heuristik

Beobachtung:

- Nächster Schritt hängt nur von letzter Binärziffer (gerade/ ungerade) ab.
- Nach einer T-Iteration kein „a priori“-Wissen über Parität mehr bekannt.

$\Rightarrow$ nehmen zufällige Folge von „Halbierungs“- u. „$\frac{3n+1}{2}$“-Schritten an.

Heuristik

Für $n \geq 2$ ist $\frac{3n+1}{2} \leq \frac{3n+\frac{1}{2}n}{2} = \frac{7}{4}n$.

$\Rightarrow$ Betrachten als „obere Abschätzung“ für T folgenden Operator $T' : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$:

$$T'(n) := \begin{cases} \frac{1}{2}n & \text{, wenn Münze Zahl zeigt} \\ \frac{7}{4}n & \text{, wenn Münze Kopf zeigt.} \end{cases}$$

Wie verhält sich eine mit T' gebildete Zahlenfolge? Ist linear in n, d. h., es kommt nur auf das Produkt der Vorfaktoren $\left(\frac{1}{2}\right)^k \cdot \left(\frac{7}{4}\right)^\ell$ an.

Heuristik

Auf lange Sicht (Gesetz der großen Zahlen) nähert sich Verhältnis von k u. ℓ (Anzahl gerader bzw. ungerader Zahlen in der Folge) an:

$\Rightarrow$ Vorfaktor verhält sich wie $\left(\frac{1}{2} \cdot \frac{7}{4}\right)^k = \left(\frac{7}{8}\right)^k \to 0$.

$\Rightarrow$ Die mittels T' gebildete Folge fällt mit hoher Wahrscheinlichkeit unter jede pos. Schranke

$\Rightarrow$ T_n wird irgendwann kleiner als 2, geht also in trivialen Zyklus $2 \to 1 \to 2 \to \dots$ über.

$\Rightarrow$ Collatz-Vermutung

Definition

Ist die Collatz-Vermutung falsch, so muss (min.) eines der beiden folgenden Ereignisse eintreten:

- Es gibt eine Startzahl n, deren Collatz-Folge unbeschränkt wächst.
- Es gibt eine Startzahl n, deren Collatz-Folge in einen vom trivialen Zyklus $4 \to 2 \to 1 \to 4 \to \ldots$ verschiedenen Zyklus übergeht.

Uns interessiert der zweite Fall. O. B. d. A sei n schon Teil des betrachteten nicht-trivialen Zyklus'.

Eigenschaften

Es sei $\Omega := \{n, C(n), C(C(n)), \dots\}$ die Menge der Folgenglieder eines nicht-trivialen Collatz-Zyklus', d. h. $|\Omega|$ ist dessen Länge.

Weiterhin seien $\Omega_g := \Omega \cap 2\mathbb{Z}$ und $\Omega_u := \Omega \cap (2\mathbb{Z}+1)$ die Mengen der beteiligten geraden bzw. ungeraden Folgenglieder mit $p := |\Omega_g|$ und $q := |\Omega_u|$. Dann ist

$$\prod_{n\in\Omega} n = \prod_{n\in\Omega} C(n) = \prod_{n\in\Omega_g} C(n) \cdot \prod_{n\in\Omega_u} C(n) = \prod_{n\in\Omega_g} \frac{n}{2} \cdot \prod_{n\in\Omega_u} (3n+1)$$

$$= \left(\prod_{n\in\Omega_g} n \cdot \prod_{n\in\Omega_g} \frac{1}{2}\right) \cdot \left(\prod_{n\in\Omega_u} n \cdot \prod_{n\in\Omega_u} \left(3 + \frac{1}{n}\right)\right).$$

Eigenschaften

Damit ist $2^p = \prod_{n\in\Omega_g} 2 = \prod_{n\in\Omega_u}\left(3+\frac{1}{n}\right)$, also

$$3^q = \prod_{n\in\Omega_u} 3 < \prod_{n\in\Omega_u}\left(3+\frac{1}{n}\right) = \left(\sqrt[q]{\prod_{n\in\Omega_u}\left(3+\frac{1}{n}\right)}\right)^q$$
$$\leq \left(\frac{1}{q}\cdot\sum_{n\in\Omega_u}\left(3+\frac{1}{n}\right)\right)^q = (3+\mu)^q\,, \text{ mit } \mu := \frac{1}{q}\cdot\sum_{n\in\Omega_u}\frac{1}{n}.$$

Logarithmieren liefert dann die Abschätzung

$$q\cdot\log(3) < p\cdot\log(2) \leq q\cdot\log\left(3+\mu\right) \text{ bzw.}$$
$$\log_2 3 < \frac{p}{q} \leq \log_2(3+\mu).$$

Kettenbrüche

Gesucht: rationale Näherung $\frac{p}{q}$ der Zahl $\log_2 3 \in \mathbb{R} \setminus \mathbb{Q}$:

$$\log_2 3 = 1.5849\dots = 1 + \cfrac{1}{1 + \cfrac{1}{1+\cfrac{1}{2+\cfrac{1}{2+\cfrac{1}{3+\cfrac{1}{1+\dots}}}}}}$$

Bricht man den Kettenbruch in Tiefe n ab, erhält man n-ten Näherungsbruch $\frac{p_n}{q_n}$.
Fakt: Ist $\frac{p_n}{q_n}$ Näherungsbruch der irrationalen Zahl $\alpha > 0$, so gilt für alle $p, q \in \mathbb{N}_{>0}$:

$$\left|\alpha - \frac{p}{q}\right| < \left|\alpha - \frac{p_n}{q_n}\right| \Rightarrow q > q_n.$$

erste Abschätzung

Wir wissen $\log_2 3 < \frac{p}{q} < \log_2(3+\mu)$.

Ist $\log_2(3+\mu) < \frac{p_n}{q_n}$, so also $q > q_n$.

Es ist $\mu = \frac{1}{q}\sum_{n\in\Omega_u}\frac{1}{n} < \frac{1}{q}\sum_{n\in\Omega_u}\frac{1}{n_{\min}} = \frac{1}{n_{\min}} < \frac{1}{S}$.

Satz

Ist für alle Startzahlen unterhalb von $S = 2.17 \cdot 10^{20}$ bekannt, dass die Collatz-Vermutung erfüllen, so besteht ein nicht-trivialer Collatz-Zyklus aus mindestens 17,026,679,261 Folgengliedern.

Beweis.

Ist $S = 2.17 \cdot 10^{20}$, so $\frac{p}{q} < \log_2(3+\mu) < \frac{p_{21}}{q_{21}} = \frac{10439860591}{6586818670}$. $\square$

Verbesserungen

Bisher nur für $S = 5 \cdot 2^{60} = 5.76 \ldots \cdot 10^{18}$ bekannt, dass bis dorthin Collatz-Vermutung gilt.

Möglichkeiten der Verbesserung:

- Schranke S durch weitere Berechnungen erhöhen.
- Abschätzung $\mu < \frac{1}{S}$ verbessern.

Beide Varianten werden hier beschritten:

- Durch verbessertes algorithmisches Vorgehen für alle Startzahlen $< S = 87 \cdot 2^{60} = 1.003 \ldots \cdot 10^{20}$ Collatz-Vermutung nachweisen.
- Durch Zusatzüberlegungen und -berechnungen Abschätzung so verbessern, dass $\mu < \frac{1}{2.17 \cdot 10^{20}}$ folgt, wenn erste Teilaussage erfolgt ist, sodass der Satz dann weiterhin folgt.

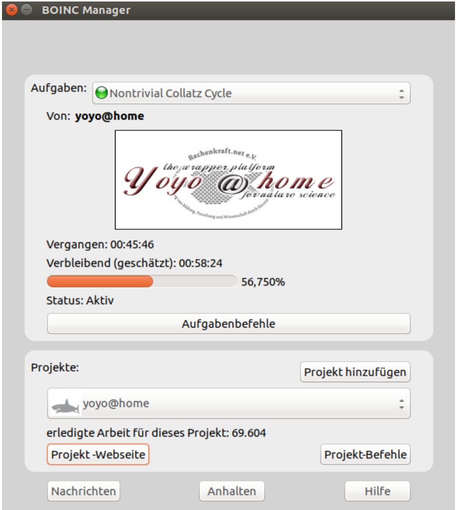
„Theoretische“ Methoden für Testbeschleunigung

- Betrachten reduzierte Collatz-Folgen.
- Brauchen nur prüfen, ob irgendwann Startzahl unterschritten wird. (Dann fällt Folge auf 1; Induktionsargument.)
- Genügt, nur bestimmte Restklasen mod Zweierpotenzen zu betrachten, da z. B. $4k+1 \to 6k+2 \to 3k+1 \Rightarrow$ Sieb
- Brauchen Startzahlen nicht betrachten, die in Folgen kleinerer Startzahlen münden, z. B. $64k+15 \to \cdots \to 81k+20$ und auch $32k+7 \to \cdots \to 81k+20$. $\Rightarrow$ verbessertes Sieb.
- Können mehrere Schritte auf einmal berechnen: $4k+1 \to 3k+1$ und $4k+3 \to 9k+8$. $\Rightarrow$ Mehrfachschritt

„Praktische“ Methoden für Testbeschleunigung

- Umsetzung in maschinennaher Sprache: C
- Vortabellieren der Mehrfachschritte ⇒ weniger wiederholte Berechnungen, nur Lookups
- Ausnutzen von Speicher-Größen (L1-Cache) für schnellen Zugriff (10er-Mehrfachschritte)
- Schnelle Berechnungen mod 2^{64} für „End-Bits“ und Gleitkommazahlen für absolute Größe
- Bit-Schieben, bitweise log. Verknüpfungen statt $\cdot, :, \mod 2^k$
- (Parallelisierung auf einem CPU-Kern [MMX, SSE, ...])
- Parallelisierung über mehrere CPU-Kerne eines Rechners (Einzelne Arbeitspakete für jeden Kern: Restklassen mod 2^{32}, dann weiter bis 2^{58} sieben, dann Mehrfachschritte)
- Parallelisierung über viele Rechner ⇒ Distributed Computing

Nontrivial Collatz Cycle I



Nontrivial Collatz Cycle II

Yoyo @ home – the wrapper platform for natural science

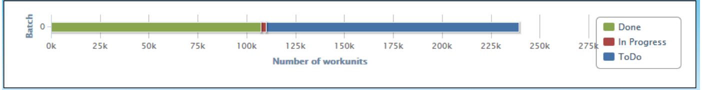
Forum | Chat | Project Description | Statistics | Rechenkraft.net

rechenKRAFT.net e.V.

Please visit donation page to help the project cover running costs.

Nontrivial Collatz Cycle wu status

5 Sep 2017 8:50:21 UTC



Found Candidates

Show 10 entries

Search:

Start	Path Maximum	Bit	ResCl	Date	by User
65031059431736043055	193985025440208840514448416160828209024	128	8831894	2017.08.31 22:10:16	Gonz, Daniel
87891032609386558311	103575467746546672653847387819815708976	127	7103817	2017.08.28 09:25:20	Danny, Daniel
55301489598103050351	82481055670255774134849743868820752544	126	9628625	2017.09.02 05:00:17	Eric_Kaiser, Michael H.W. Weber
79619863089615381531	73981282122345728534268703831645870996	126	286712	2017.08.11 01:00:00	Danny, scole of TSBT
73112120496948844455	64619408717738348059184819998594074856	126	5665994	2017.08.25 17:02:29	NeuralMiner, phoenicis
23314617123910465627	52159921375623872234470241595697335400	126	2901250	2017.08.20 13:35:20	zombie67 [MM], scole of TSBT

Ideen zur Verbesserung der Abschätzung

Gesucht: Bessere obere Abschätzung für $\mu = \frac{1}{q} \cdot \sum_{n \in \Omega_u} \frac{1}{n} < \frac{1}{S}$:

- Zerlegen Zyklus in Teilbereiche und zeigen für jeden, dass Durchschnitt der Reziproken darin $<$ Wunschwert.
- Fallunterscheidung nach Restklasse von erster ungerader Zahl u in Teilbereich mod Zweierpotenzen:
  Ist etwa $u = 4k + 1$, so ist nächste ungerade Zahl $\leq 3k + 1$, die aber auch $> S$ ist, also $u > \frac{4}{3}S$ bzw. $\frac{1}{u} < \frac{3}{4} \cdot \frac{1}{S}$.
- Ist kleinster pos. Vertreter der Restklasse „ $> \frac{1}{\text{Wunschwert}}$ “, so Bedingung automatisch erfüllt.
- Können kleinsten Vertreter der Restklasse $> S$ betrachten.

$\Rightarrow$ Sieb modulo immer größerer Zweierpotenzen (bis max. 2^{200}) zeigte: $S = 87 \cdot 2^{60}$ reicht aus.

Zusammenfassung

Das Collatzproblem ist nicht-trivial.

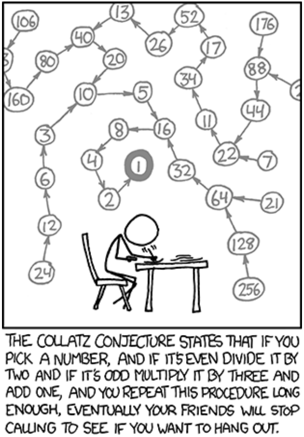
Durch theoretische und algorithmische Überlegungen sind Teilaspekte studierbar.

Quantitative Ergebnisse lassen sich zum Teil deutlich verbessern.

Ausblick: Verbesserung einer ähnlichen Aussage: Ein nicht-trivialer Zyklus einer reduzierten Collatz-Folge besitzt mindestens 80 (bisher 75) lokale Minima.

Abschluss

Vielen Dank für Ihre Aufmerksamkeit!



XKCD Collatz-Conjecture