3.5 Lineare Komplexität und TURING-Komplexität

Eine **universelle** TURING-**Maschine** kann jede andere TURING-Maschine durch ein geeignetes Programm simulieren. Sei **M** eine solche, und sei $u \in \mathbb{F}_2^n$ eine Bitfolge der Länge n. Dann ist die TURING-KOLMOGOROV-CHAITIN-**Komplexität** $\chi(u)$ gleich der Länge des kürzesten Programms von **M**, das u als Output produziert. Ein Programm der ungefähren Länge n gibt es immer: Es nimmt einfach u als Input-Folge und gibt diese wieder aus.

Anmerkung. Die Funktion $\chi : \mathbb{F}_2^* \longrightarrow \mathbb{N}$ ist selbst nicht berechenbar; d. h., es gibt keine TURING-Maschine, die χ berechnet. Daher ist die TURING-KOLMOGOROV-CHAITIN-Komplexität als Komplexitätsmaß kaum praktisch verwendbar. Sie ist allerdings in den letzten Jahren durch die Arbeiten von VITANYI und anderen in präzisierter Form wieder in Mode gekommen; siehe dazu etwa:

- MING LI, PAUL VITANYI: *An Introduction to Kolmogorov Complexity and Its Applications.* Springer-Verlag, New York 1993, 1997.

Eine wichtige Aussage der Theorie ist:

$$\frac{1}{2^n} \cdot \#\{u \in \mathbb{F}_2^n \mid \chi(u) > n \cdot (1-\varepsilon)\} > 1 - \frac{1}{2^{n\varepsilon - 1}},$$

d. h., fast alle Folgen haben eine TKC-Komplexität nahe am maximalen Wert, also keine wesentlich kürzere Beschreibung als durch vollständiges Hinschreiben. Eine gängige Interpretation dieses Sachverhaltes ist: „Fast alle Folgen sind zufällig." Dies entspricht der intuitiven Vorstellung von Zufall sehr gut. Eine Folge mit einer kurzen Beschreibung wie „eine Million Mal abwechselnd 0 und 1" wird nämlich niemand als auch nur im geringsten zufällig ansehen.

Das Komplexitätsmaß „lineare Komplexität" λ, das auf dem sehr speziellen Maschinenmodell des linearen Schieberegisters beruht, hat dagegen auf den ersten Blick gravierende Mängel. Die Folge „999999 Mal die 0, dann eine 1" hat eine sehr geringe TKC-Komplexität – und eine sehr geringe intuitive Zufälligkeit –, aber die lineare Komplexität 1 Million. Der Vorteil der linearen Komplexität ist, wie gesehen, ihre leichte explizite Bestimmbarkeit, und sie beschreibt „im allgemeinen" die Zufälligkeit ener Bitfolge doch recht gut. Diese Aussage lässt sich überraschend präzise fassen (ohne Beweis):

Satz 2 (BETH/DAI, EUROCRYPT 89)

$$\begin{aligned}
\frac{1}{2^n} \cdot \#\{u \in \mathbb{F}_2^n \mid (1-\varepsilon)\lambda(u) \le \chi(u)\} &\ge 1 - \frac{8}{3 \cdot 2^{\frac{n\varepsilon}{2-\varepsilon}}}, \\
\frac{1}{2^n} \cdot \#\{u \in \mathbb{F}_2^n \mid (1-\varepsilon)\chi(u) \le \lambda(u)\} &\ge 1 - \frac{1}{3} \cdot \frac{1}{2^{n\varepsilon - (1-\varepsilon)(1+\log n)+1}} - \frac{1}{3} \cdot \frac{1}{2^n}.
\end{aligned}$$

Interpretation: „Für fast alle Bitfolgen stimmen die lineare Komplexität und die TKC-Komplexität mit vernachlässigbarer Abweichung überein."

Damit ist klar, dass die lineare Komplexität trotz ihrer Einfachheit ein gutes Komplexitätsmaß ist, und dass Bitfolgen hoher linearer Komplexität im allgemeinen auch mit anderen Ansätzen nicht kürzer erklärbar sind. Sie sind also kryptographisch brauchbar. (Ein anderer Ansatz wäre etwa, die lineare Komplexität auf andere endliche Körper oder Restklassenringe zu verallgemeinern – das würde also dem Kryptoanalytiker kaum nützen.) Jedes effiziente Vorhersageverfahren im Sinne der Kryptoanalyse von Bitstromchiffren wäre auch eine Kurzbeschreibung im Sinne der TKC-Komplexität, so dass man als Fazit festhalten kann: *Bitfolgen hoher linearer Komplexität sind im allgemeinen nicht vorhersagbar.*

Natürlich gibt es auch Ansätze, nichtlineare Schieberegister zur Beurteilung der Komplexität heranzuziehen, siehe etwa:

- Agnes Hui Chan, Richard A. Games: On the quadratic span of periodic sequences. Crypto 89, 82–89.
- Cees J. A. Jansen, Dick E. Boekee: The shortest feedback shift register that can generate a given sequence. Crypto 89, 90–96.