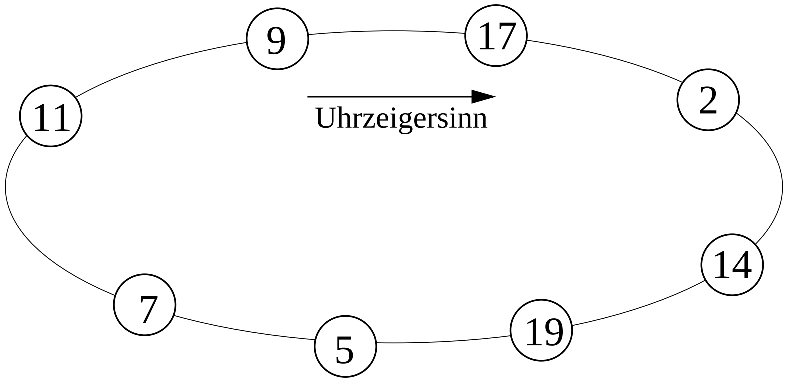
Election: Nachrichtenkomplexität

- Message-extinction-Prinzip von Chang und Roberts 1979
  - war einer der ersten verteilten Algorithmen



- *Worst-case Nachrichtenkomplexität* bei k Startern:
  n + (n-1) + (n-2) +...+ (n-k+1) = **nk - k(k-1)/2**

  --> $O(n^2)$ bei Ringgrösse n und k=n

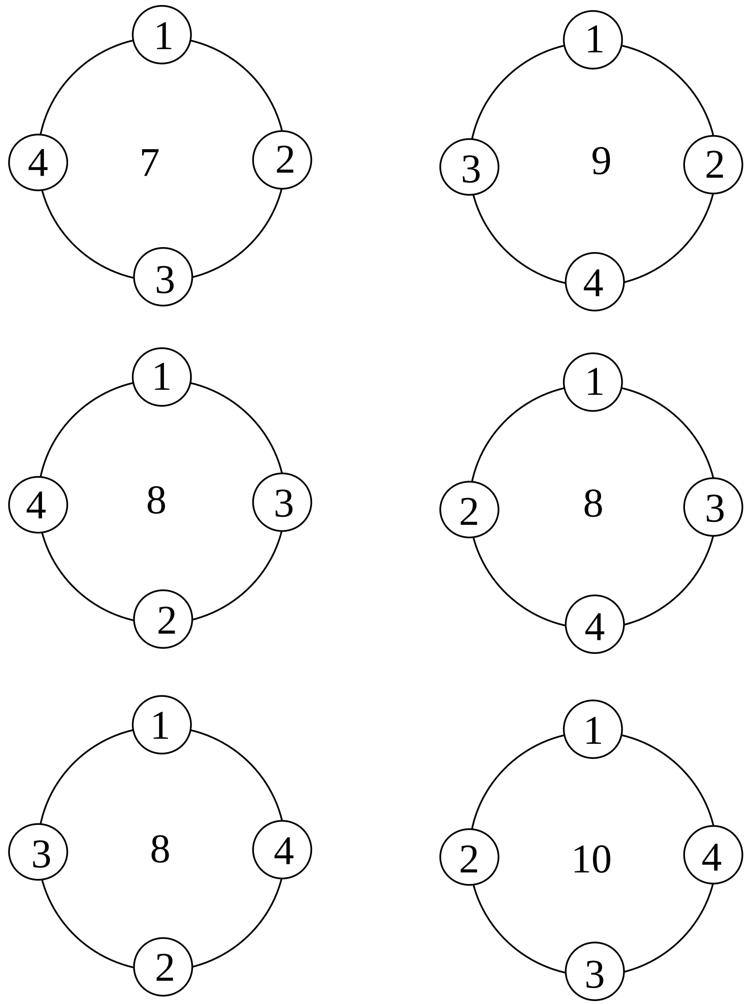
- Wie hoch ist die *mittlere* Nachrichtenkomplexität?
  - bei “zufällger” Permutation der Identitäten

- Wie hoch ist die *Zeitkomplexität*?
  - wenn alle gleichzeitig starten?
  - beim zeitversetzten Starten?

Jede Nachricht benötigt eine Zeiteinheit ("Einheitszeitmodell")

Mittlere Nachrichtenkomplexität (1)

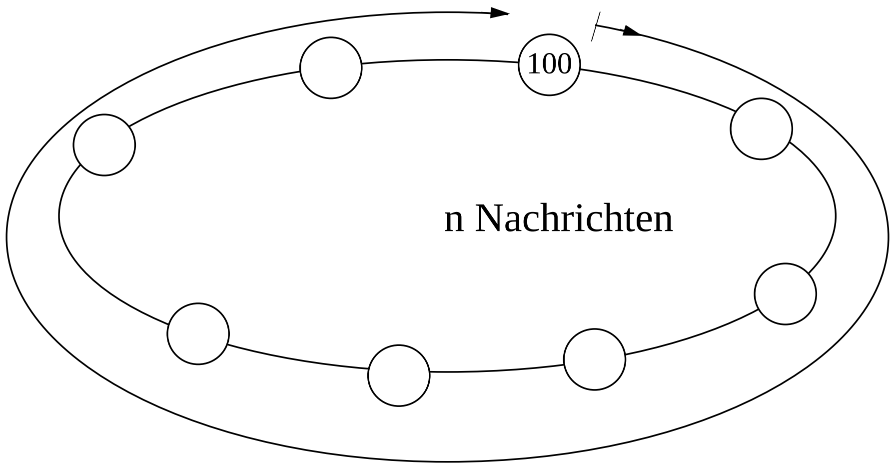
- Beispiel: Sei k = n = 4
- Über alle Permutationen mitteln (wieviele?)

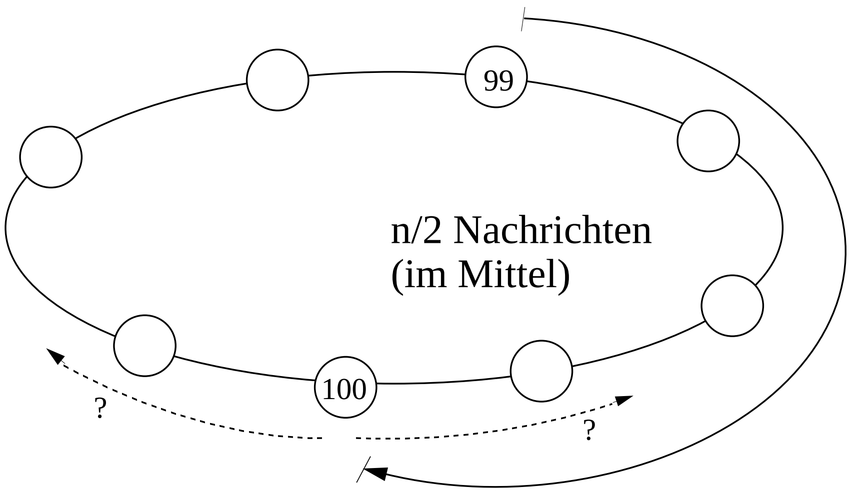


50/6 = 8.333... Nachrichten im Mittel

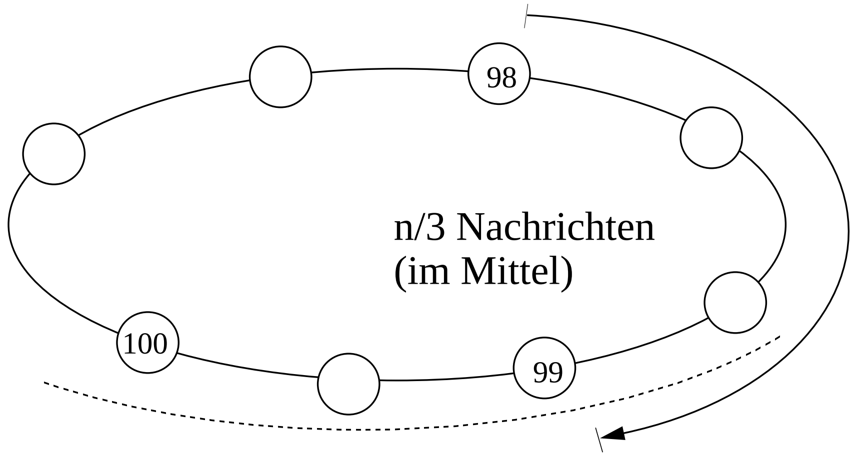
Mittlere Nachrichtenkomplexität (2)

Sei n = 100; id = 1, 2, ..., 100





Mittlere Nachrichtenkomplexität (3)



Vermutung: Drittgrösster sollte im Mittel nach n/3 Schritten auf einen grösseren Prozess (99 oder 100) treffen

- Lässt sich durch explizites Nachrechnen aller ca. n^2 Fälle auch beweisen...

- Allgemeiner Beweis für die Vermutung “n/i Schritte beim i-t grössten”?

Mittlere Nachrichtenkomplexität (4)

- Grösste Identität: n Nachrichten (immer)
- Zweitgrösste: im Mittel n/2
- Drittgrösste: im Mittel n/3 ??

  ...

- i-t grösster: im Mittel n/i ← stimmt das?

Wenn das stimmt -->

einfach aufaddieren?

$$\frac{n}{1} + \frac{n}{2} + \frac{n}{3} + \ldots + \frac{n}{k} = n \underbrace{\sum_{i=1}^{k} \frac{1}{i}} = n \underbrace{H_k}$$

Def. der sogenannten *harmonischen Reihe*

Was ergibt sich für n = k = 4 ?

$$4 \; \frac{12+6+4+3}{12} = 25/3 = 8.333\ldots$$

Stimmt mit früherem Wert überein!

==> *Vermutung*: n H_k ist korrekt

- Man kann tatsächlich exakt zeigen:

  Die mittlere Nachrichtenkomplexität beträgt

$$n \sum_{i=1}^{k} \frac{1}{i} = n\,H_k \approx \mathbf{n \ln k}$$

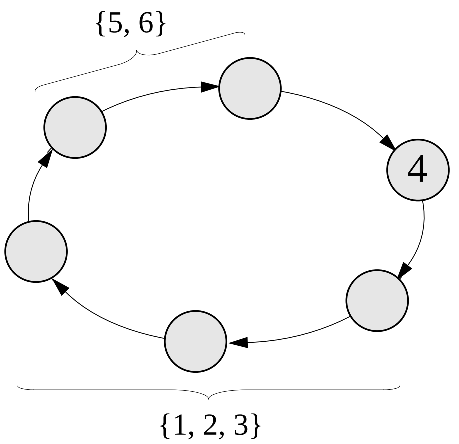
Wir wollen dies gleich genauer herleiten...

Unabhängige Fälle?

Im allgemeinen sind die "Zufallsvariablen", die die Länge der von Prozess i initiierten Nachrichtenkette repräsentieren, nicht unabhängig voneinander!

Beispiel: Gegeben 6 Prozesse 1, 2, ... 6 im Ring. Sei *A* das Zufallsereignis "Nachrichtenkette von Prozess 4 hat die Länge 4" und *B* das Zufallsereignis "Nachrichten-kette von Prozess 5 hat die Länge 2"

Aus *A* folgt nebenstehende Situation; dann kann die von 5 initiierte Nachrichtenkette jedoch nur die Länge 1 oder 5 haben, nicht jedoch 2, wie von *B* gefordert



==> prob(A $\cap$ B) = 0, obwohl prob(A) $\times$ prob(B) $\neq$ 0

==> A und B sind *nicht unabhängig* voneinander!

Allerdings (hier ohne Rechtfertigung):

- die Zufallsvariablen in obigem Beispiel sind *paarweise unkorreliert*
- *Erwartungswerte* (d.h. die mittleren Längen der Nachrichtenketten) dürfen *aufaddiert* werden, um den Mittelwert bzgl. der Gesamtnachrichtenzahl zu erhalten (in "Formeln": E(X+Y) = E(X) + E(Y))

Die harmonische Reihe H_n ...

Beh.: H_n divergiert

Bew.: Fasse 4 Reihenglieder ab 1/4 zusammen,
(jew. grösser als 1/8), 8 Reihenglieder ab 1/8...

Beh. (o. Bew.) $\sum_i \frac{1}{i^r}$ konvergiert gdw. r > 1

(r = 2 --> $\pi^2/6$ ... Riemann'sche Zeta-Funktion)

==> Harmonische Reihe "divergiert gerade noch"

Beh. (o.Bew.) $\gamma = \lim_{n \to \infty} (H_n - \ln n) = 0.5772156649...$

Euler'sche Konstante

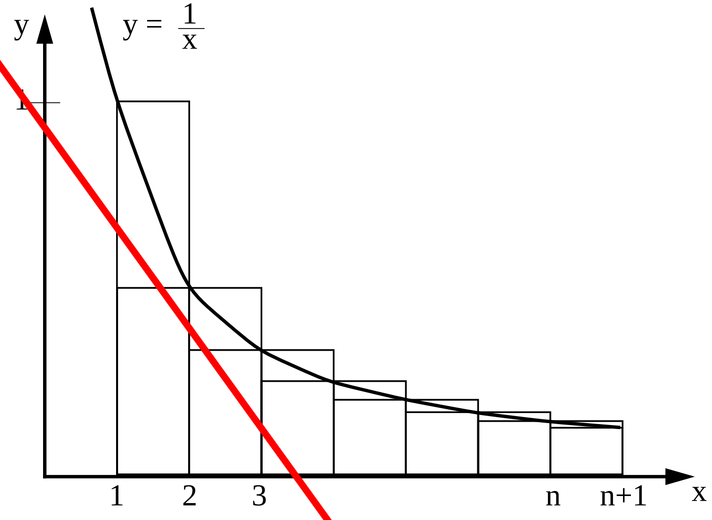
==> H_n lässt sich zumindest für grosse n durch ln n abschätzen

- diese Abschätzung lässt sich sogar noch präzisieren...

Abschätzung der harmonischen Reihe H_n

Beh.: $\underline{\underline{\ln n < H_n < 1 + \ln n}}$

Bew.:



- *Fläche der Obertreppe* von 1 bis n+1: H_n
- *Fläche unter Kurve* von 1 bis n+1: $\int_1^{n+1} \frac{1}{x}\,dx = \ln(n+1)$

(wg. ln(1) = 0)

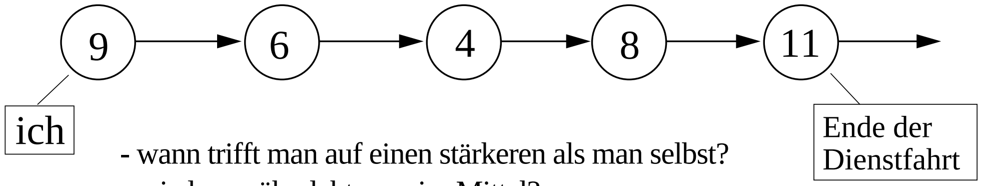
==> $H_n > \ln(n+1)$ ==> $\underline{\underline{H_n > \ln n}}$

- *Fläche der Untertreppe* von 1 bis n: H_n-1
- *Fläche unter Kurve* von 1 bis n: $\int_1^{n} \frac{1}{x}\,dx = \ln(n)$

==> $H_n - 1 < \ln(n)$ ==> $\underline{\underline{H_n < 1 + \ln n}}$

Wartezeit bis zum ersten Rekord

- Wie weit kommt die von einem "x-beliebigen" Knoten initiierte Nachrichtenkette im Mittel?



- wann trifft man auf einen stärkeren als man selbst?
- wie lange überlebt man im Mittel?

In einem Teich schwimmen n Fische verschiedener Grösse - wieviele Fische muss man im Mittel noch fangen, bis man einen fängt, der grösser als der erste gefangene Fisch ist? (Oder bis der Teich leer ist)

Lösungsansatz (aber darf man wirklich so vorgehen?):

Fängt man den grössten zuerst --> n Fänge ("Pech", Teich wird leer)
Fängt man den zweitgrössten zuerst --> n/2 Fänge im Mittel
Fängt man den drittgrössten zuerst --> ...

Zurückführung auf ein bek. math. Modell: *Urnenmodell ohne Zurücklegen*

Z.B. n=100, 7.grösster Fisch initial:

--> 6 schwarze und 93 weisse Kugeln

--> Sei p die Wahrscheinlichkeit, eine schwarze Kugel zu ziehen ("Treffer")

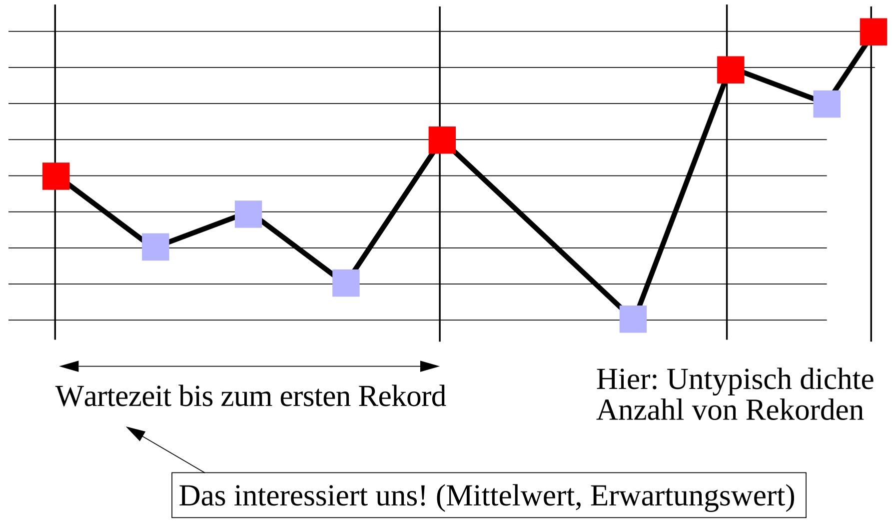
--> Beispiel: Wahrscheinlichkeit, eine weisse und dann eine schwarze zu ziehen (d.h. genau die Entfernung 2 zu schaffen): (1-p) mal Wahrscheinlichkeit für "Treffer bei 6 schwarzen und <u>92</u> weissen"

--> ... alles gewichtet aufsummieren... (aufwendig!)

Warten auf einen neuen Rekord...

- Rekord = grösserer Wert als alle vorangehenden
- "left to right maxima" (einer Zahlenfolge)

- Z.B. "heissester August seit Anfang des Jahrhunderts"



- Anwendung: z.B. Verlobungshäufigkeit bei deletion-sort

Festhalten und weitersuchen