

Modalitäten

- Übung: Dienstag, 16:00 Uhr – 17:30 Uhr (wöchentlich), [BBB](#)
- Folien und Hausaufgaben per [GitHub](#)
- Kontakt:
 - Ben Lorenz: lorenzb@tu-freiberg.de
 - Jonas Treumer: treumer@tu-freiberg.de
 - (HUM-114 / 115)

Modalitäten

- Mündliche Prüfung bei Prof. Froitzheim + Beisitzer (online?)
- Anmeldung im Prüfungsamt
- Gruppen mit bis zu drei Leuten (?)
- Gruppenbonus: Erstes Prüfungsthema vorschlagen
- Individuelle Terminvereinbarung
- 9 Leistungspunkte

Übungsinhalte

- Wiederholung: Informationsbegriff, Quellcodierung
- ISO/OSI-Referenzmodell, Netzwerktopologien, Hardware
- Send-and-Wait, Sequenznummern, Sliding Window, HDLC
- Routing I (Warteschlangen, Distanzvektorverfahren)
- Routing II (Link-State-Verfahren, Pfadvektorverfahren)
- MAC- und IP-Adressen, ARP
- Ethernet
- IP
- TCP / UDP
- Praxisübung (Socket-Programmierung in C etc.)
- ... ?

Informationsbegriff

- Informationsquelle: Diskrete Zufallsvariable X
- X emittiert Zeichen aus Alphabet $S = \{s_1, s_2, \dots, s_n\}$
- Wahrscheinlichkeitsverteilung von X ist bekannt und es gelten:

$$- P(X = s_i) > 0$$

$$- \sum_{i=1}^n P(X = s_i) = 1$$

- Kurzschreibweise: $p(s_i)$ für $P(X = s_i)$
- X ist gedächtnislos, d. h., die Auftrittswahrscheinlichkeiten sind konstant und hängen nicht von zuvor emittierten Zeichen ab.

Informationsbegriff

- Der Informationsgehalt I eines Zeichens ergibt sich aus dessen Auftrittswahrscheinlichkeit:

$$I(s_i) = -\lg p(s_i)$$

- Einheit des Informationsgehalts ist das *Bit*.
- *Aufgaben:*
 - Skizzieren Sie den Graphen von $f(x) = -\lg x$ für $0 < x \leq 1$.
 - Wie viele Bit an Informationsgehalt kann ein einzelnes Zeichen maximal tragen?
 - Was passiert, wenn das Alphabet nur ein einzelnes Zeichen enthält?
 - Lässt sich eine Ampel als Informationsquelle über dem Alphabet $S = \{\text{rot}, \text{gelb}, \text{rotgelb}, \text{grün}\}$ modellieren?

Informationsbegriff

- Die Entropie H einer Informationsquelle ergibt sich als gewichtete Summe der Informationsgehalte ihrer Zeichen. Die Gewichte sind die jeweiligen Auftretswahrscheinlichkeiten:

$$\begin{aligned} H(X) &= \sum_{i=1}^n p(s_i) \cdot I(s_i) \\ &= \sum_{i=1}^n p(s_i) \cdot -\lg p(s_i) \\ &= -\sum_{i=1}^n p(s_i) \cdot \lg p(s_i) \end{aligned}$$

- Formal stellt die Entropie den *mittleren Informationsgehalt* dar.

Informationsbegriff

- *Aufgaben:*

- *Gegeben seien die Informationsquelle X sowie das zugehörige Alphabet S mit $n=2$ Zeichen. Bestimmen Sie die Wahrscheinlichkeitsverteilung über S , sodass die Entropie $H(X)$ den maximal möglichen Wert annimmt (Tipp: Funktion aufstellen, ableiten, Produkt- und Kettenregeln anwenden).*
- *Wie lässt sich das Ergebnis „vermutlich“ für alle n generalisieren? Welche maximale Entropie folgt abhängig von n ?*
- *Leiten Sie aus diesen Ergebnissen einen normierten Entropiebegriff her, der stets zwischen 0 und 1 liegt. Zeigen Sie, dass er für eine Informationsquelle mit n Zeichen äquivalent zu folgendem Ausdruck ist:*

$$- \sum_{i=1}^n p(s_i) \cdot \log_n p(s_i)$$

Informationsbegriff

- *Aufgaben:*

- *Gegeben sei eine Informationsquelle X über einem zweistelligen Alphabet:*

$$S = \{s_1, s_2\}$$

Zeichnen Sie den Graphen der Entropie $H(X)$ in Abhängigkeit von der Auftrittswahrscheinlichkeit $p(s_1)$.

Informationsbegriff

- Weitere Größen:

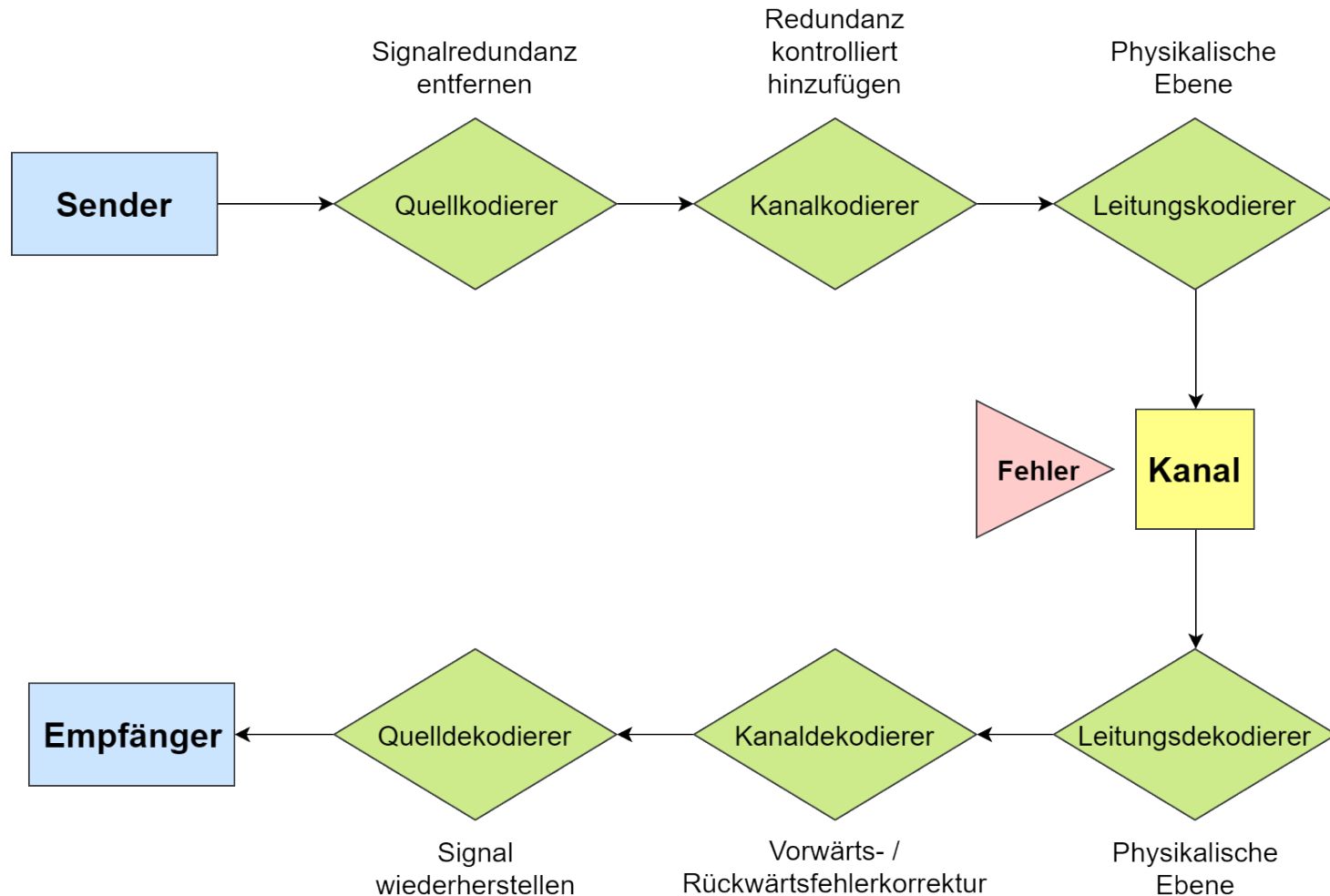
- Mittlere Codewortlänge: $\tilde{L} = \sum_{i=1}^n p(s_i) \cdot l(s_i)$

- Effizienz: $\frac{H}{\tilde{L}}$

- Redundanz: $\tilde{L} - H$

- *Aufgabe: Berechnen Sie die drei oben beschriebenen Größen für das Alphabet, das von der Zeichenkette ABRAKADABRA induziert wird (Buchstaben zählen). Codieren Sie die Buchstaben dazu mit aufsteigenden Binärzahlen (→ Codewortlänge konstant).*

Quellcodierung



Quellcodierung

- Ziel der Quellcodierung: Redundanz entfernen
- Beliebte Verfahren (datenabhängig!):
 - Lauflängencodierung (RLE: *run-length encoding*)
 - Lempel-Ziv-Welch-Codierung (LZW): dynamisches Wörterbuch
 - Entropiecodierung: Codewörter haben variable Länge
 - Arithmetische Codierung
 - Shannon-Fano-Codierung
 - Huffman-Codierung

Quellcodierung

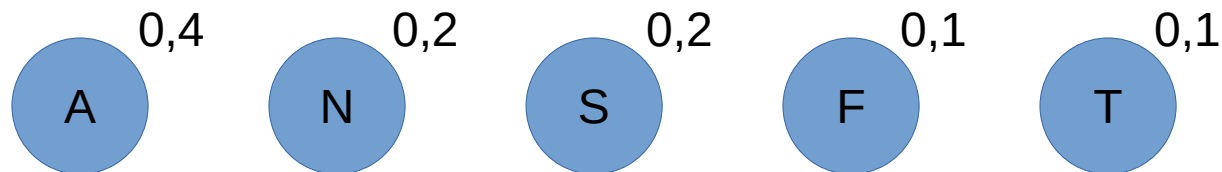
- Beispiel: ANANASSAFT
- Naiver Ansatz:
 - Sortiere nach Symbolwahrscheinlichkeit:

A	0,4
N	0,2
S	0,2
F	0,1
T	0,1

- Codiere mit aufsteigenden Binärzahlen: 0, 1, 10, 11, 100
- Problem?

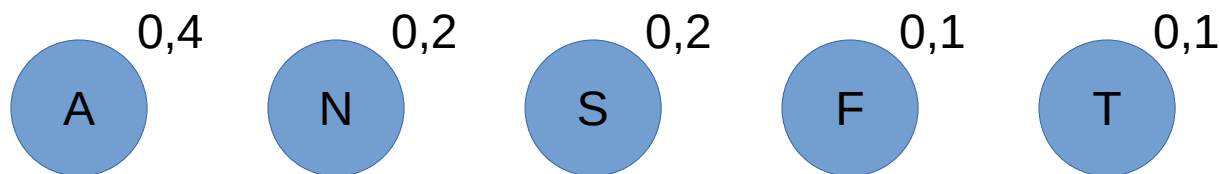
Quellcodierung

- „In einem Präfix-Code ist kein Codewort Präfix eines anderen.“
- Huffman-Codierung: Erzeugung eines (in Bezug auf die Länge der Codewörter) optimalen Präfix-Codes für eine Nachrichtenquelle mit bekannter Symbolwahrscheinlichkeit
- Algorithmus:
 - Knoten für jedes Symbol erstellen, Wahrscheinlichkeit an der Wurzel annotieren

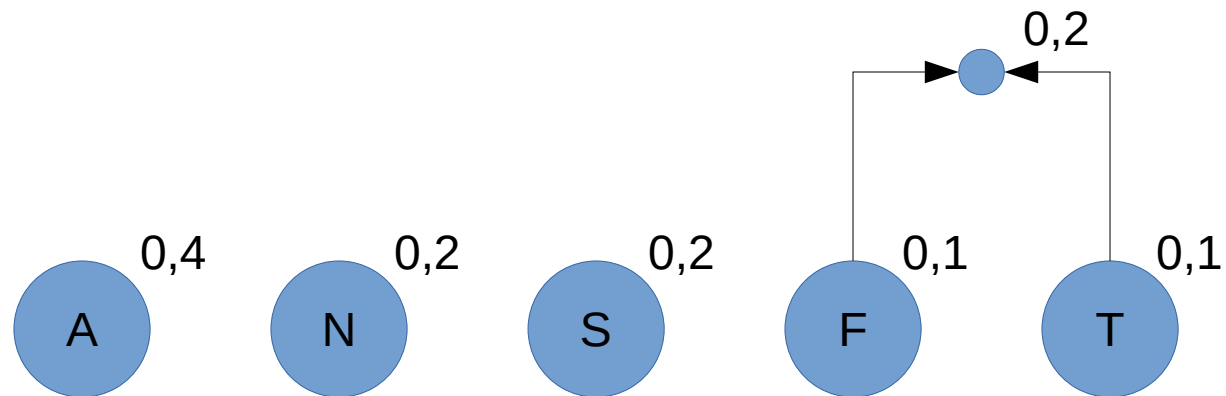


Quellcodierung

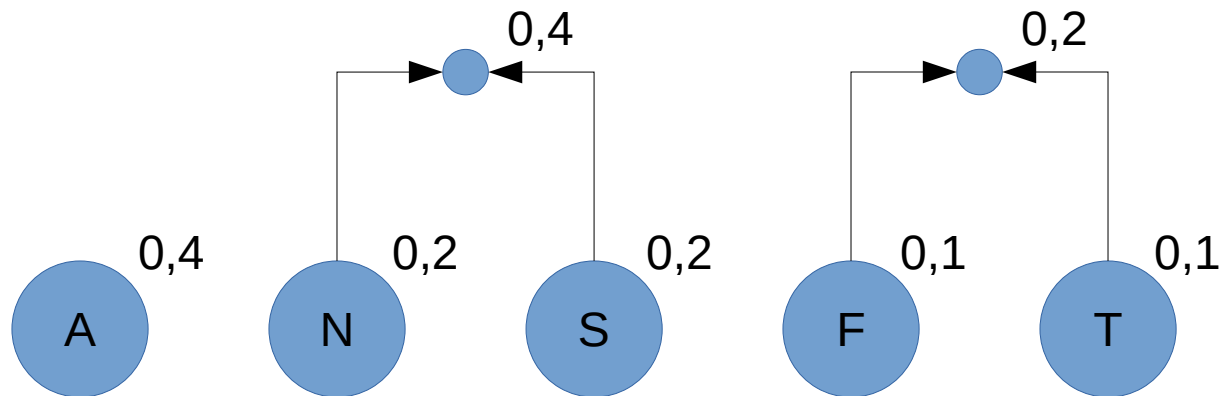
- „In einem Präfix-Code ist kein Codewort Präfix eines anderen.“
- Huffman-Codierung: Erzeugung eines (in Bezug auf die Länge der Codewörter) optimalen Präfix-Codes für eine Nachrichtenquelle mit bekannter Symbolwahrscheinlichkeit
- Algorithmus:
 - Knoten für jedes Symbol erstellen, Wahrscheinlichkeit an der Wurzel annotieren
 - Jeweils zwei Teilbäume mit minimaler summierter Wahrscheinlichkeit (bei Gleichstand: Tiefe) zusammenfassen



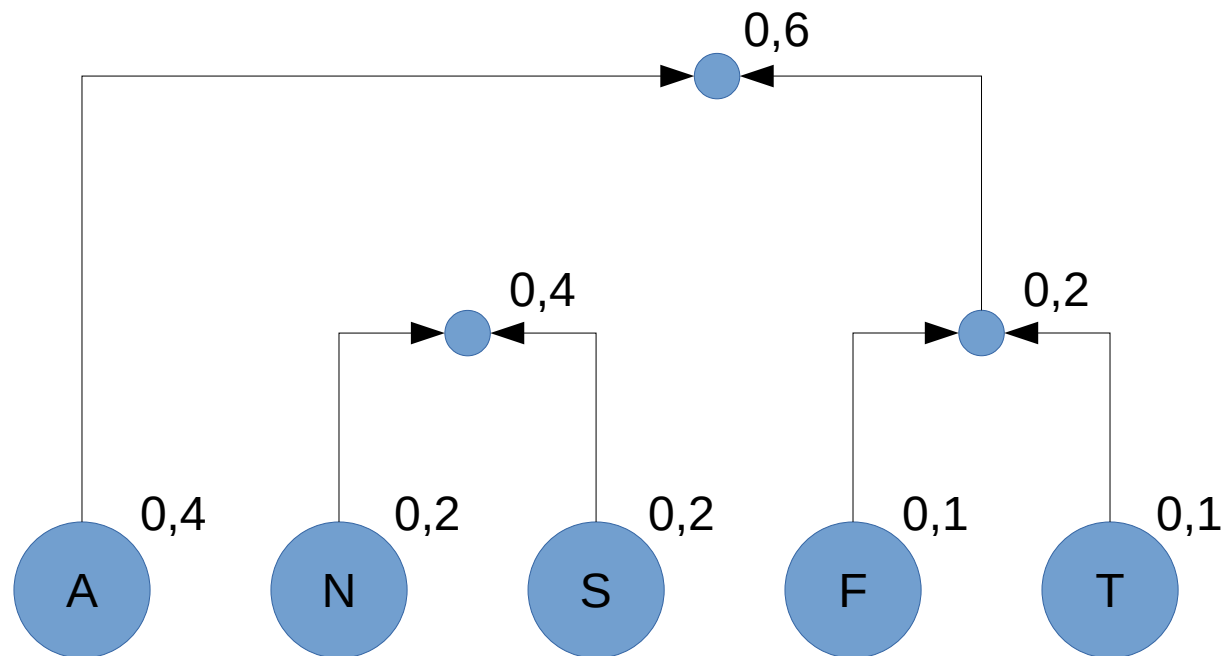
Quellcodierung



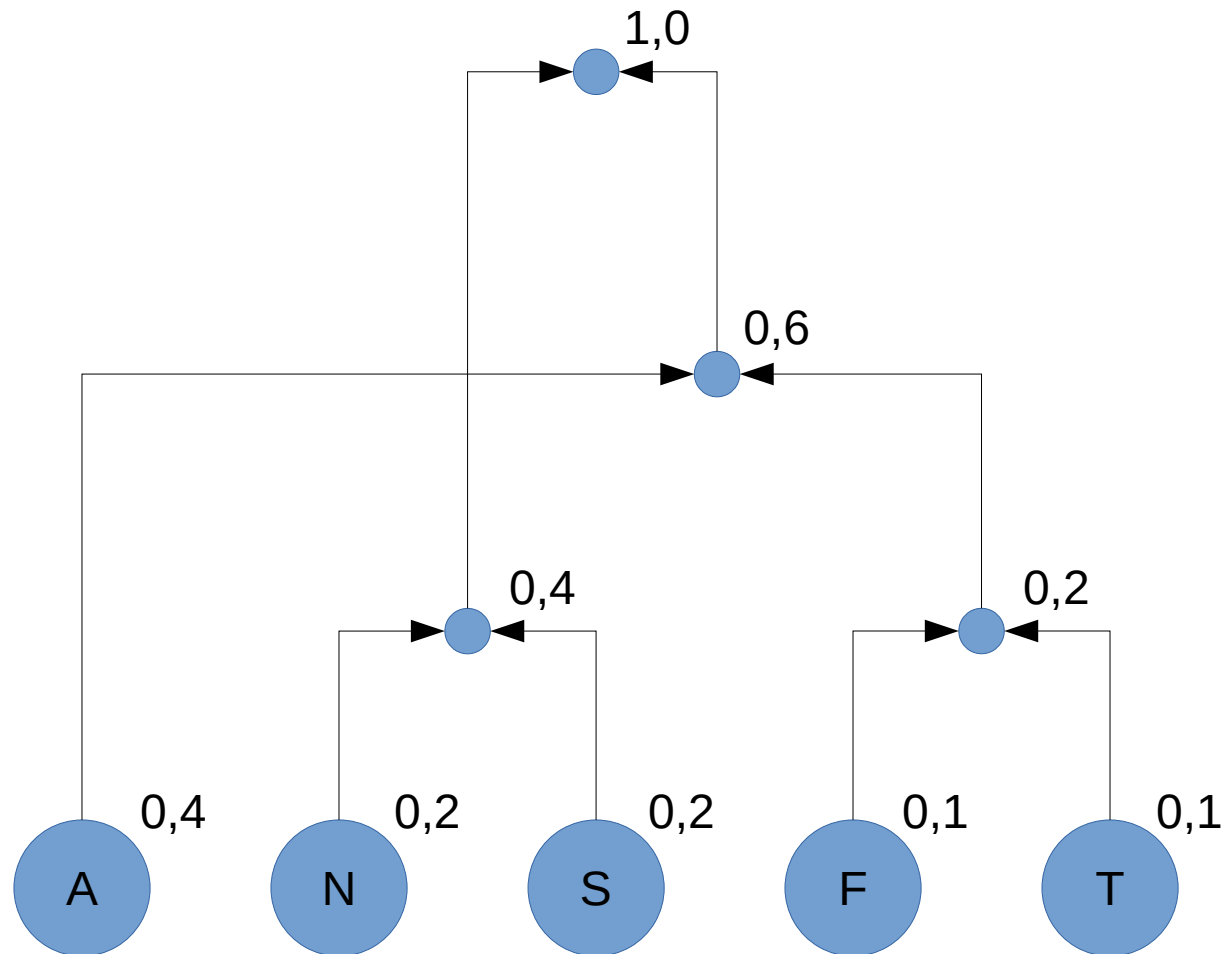
Quellcodierung



Quellcodierung



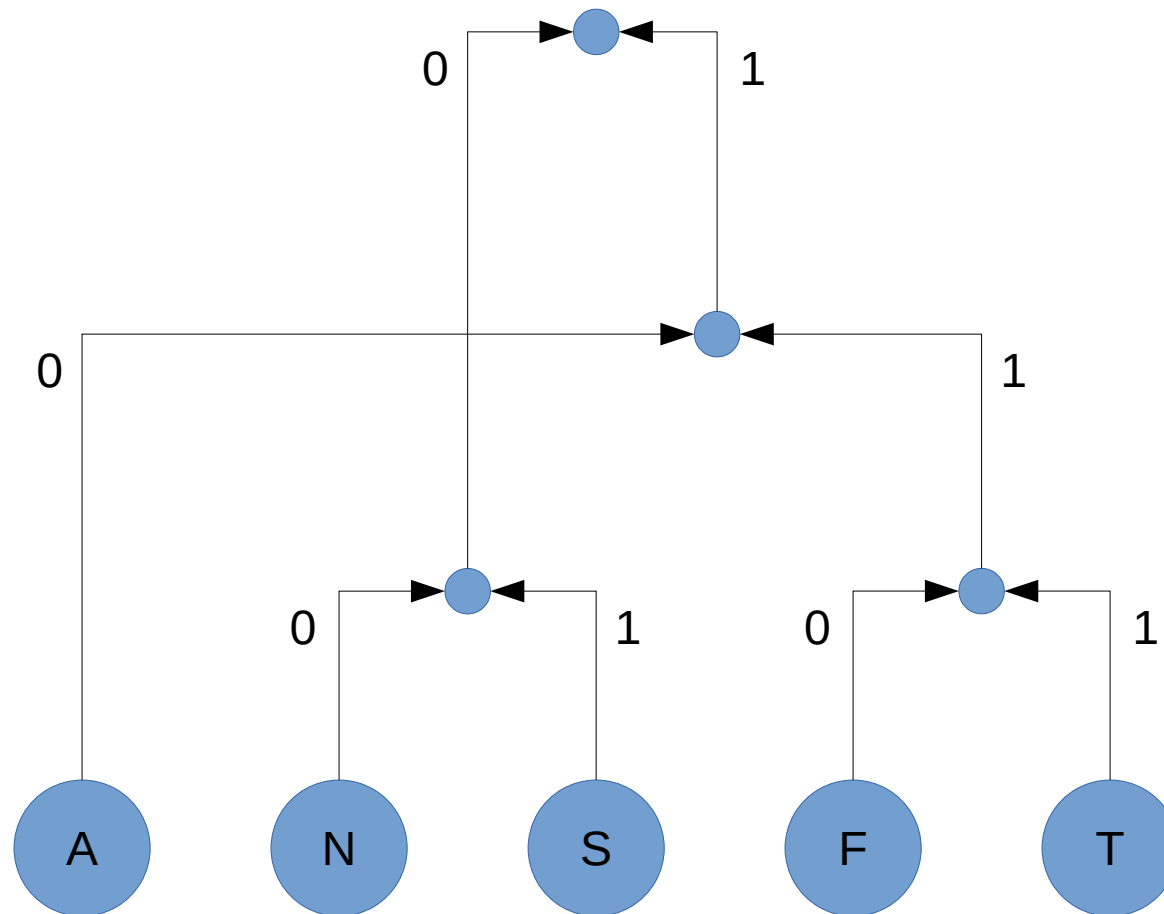
Quellcodierung



Quellcodierung

- Abschließend: Von der Wurzel aus rückwärts jede Verzweigung mit 0 bzw. 1 labeln und Codewörter auslesen

Quellcodierung



Quellcodierung

- Abschließend: Von der Wurzel aus rückwärts jede Kante mit 0 bzw. 1 beschriften und Codewörter auslesen
- Ergebnis:
 - A: 10
 - N: 00
 - S: 01
 - F: 110
 - T: 111
- *Aufgabe: Führen Sie die Huffman-Codierung für das ABRAKADABRA-Beispiel durch. Berechnen Sie anschließend die mittlere CWL und die Redundanz. Vergleichen Sie die Ergebnisse mit dem naiven Ansatz.*