

# Grundlagen Rechnernetze und Verteilte Systeme

IN0010, SoSe 2018

## Übungsblatt 1

16. April – 20. April 2018

**Hinweis:** Mit \* gekennzeichnete Teilaufgaben sind ohne Lösung vorhergehender Teilaufgaben lösbar.

### Aufgabe 1 Schichtenmodelle

In dieser Aufgabe soll ein Schichtenmodell aus insgesamt **3 Schichten** entwickelt werden, welches das Verfassen, Versenden, Empfangen, und Lesen einer Werbebroschüre beschreibt. Da die meisten Empfänger Werbung nicht lesen, nehmen wir an, dass es sich um die überlebenswichtige Speisekarte des nächstgelegenen Pizzaservice handelt, an der Empfänger großes Interesse hat.

a)\* Handelt es sich bei dem Versand von Werbeunterlagen um eine *bidirektionale* Kommunikation, d. h. wird der Empfänger auf dem Postweg antworten?

Nein. Zwar wird der Empfänger möglicherweise eine Pizza bestellen, diese wird er aber nicht auf dem Postweg ordern. Es handelt sich bei Werbung um eine *unidirektionale* Form der Kommunikation.

b)\* Die untenstehende Abbildung dient als Vorlage für das Schichtenmodell. Überlegen Sie sich für die fehlenden Schichten sowie den Übertragungskanal sinnvolle Bezeichnungen und ergänzen Sie diese in der Abbildung.



c) Beschreiben Sie, welche Dienste jede der drei Schichten erbringt.

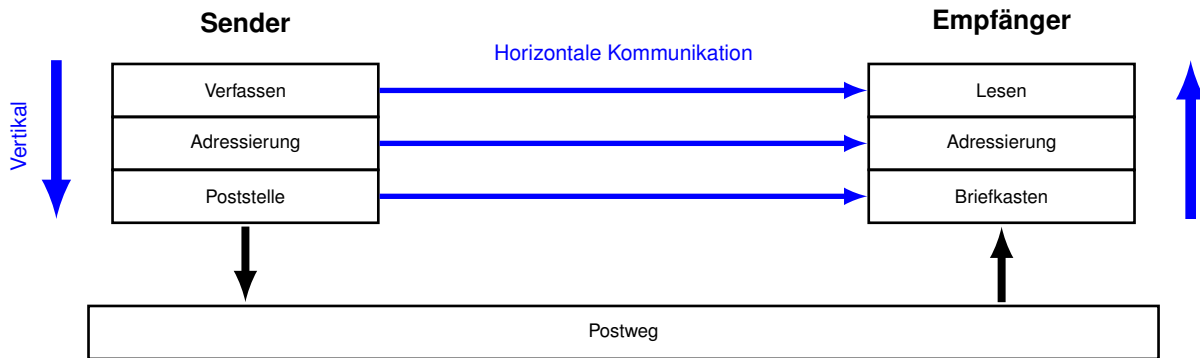
#### Sender:

- Verfassen: Werbetext wird zu Papier gebracht (Darstellung der Information in Schriftform)
- Adressierung: Die Broschüre wird in einen Umschlag verpackt, welcher mit der Absender- und Empfängeradresse versehen wird
- Poststelle: Der Brief wird (zusammen mit vielen weiteren) zur Poststelle gebracht und verschickt

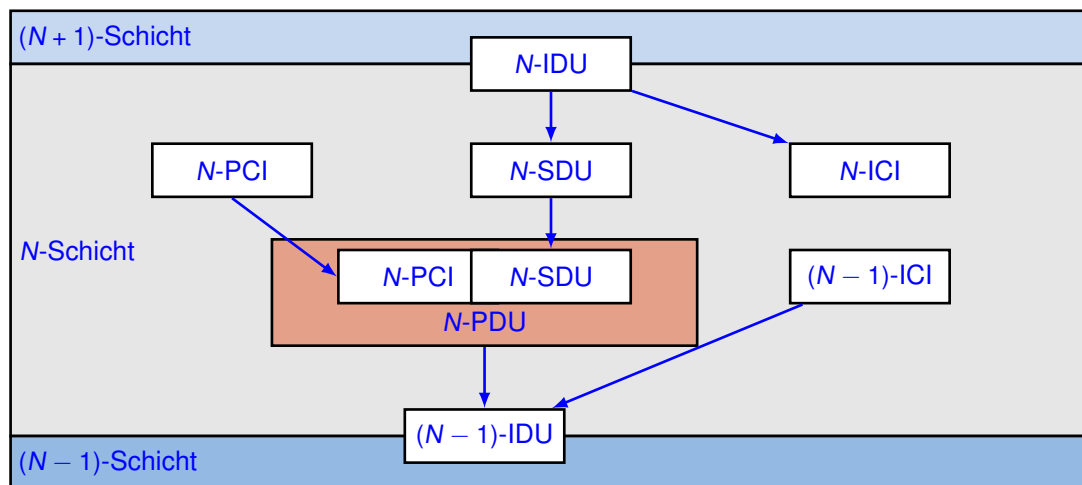
#### Empfänger:

- Briefkasten: Der Brief wird ausgetragen und in den Briefkasten des Empfängers eingeworfen
- Adressierung: Der Empfänger prüft für gewöhnlich nochmals, ob der zugestellte Brief wirklich an ihn adressiert war, und wird anschließend aus dem Umschlag genommen
- Lesen: Die im Brief enthaltene Broschüre wird gelesen

d) Was versteht man unter *horizontaler* und *vertikaler Kommunikation* im Kontext von Schichtenmodellen? Zeichnen Sie beide Kommunikationstypen in die Abbildung aus Teilaufgabe b) ein.



Wir betrachten nun die Schicht 2 etwas näher. Aus der Vorlesung kennen Sie die folgende Abbildung:



e)\* Welche Teile des Briefs entsprechen der PCI (Protocol Control Information), SDU (Service Data Unit) und PDU (Protocol Data Unit) aus Sicht von Schicht 2?

- PCI: Die auf dem Briefumschlag befindliche Adressinformation
- SDU: Die Werbebroschüre selbst, also der Inhalt des Briefs
- PDU: Der verschlossene und beschriftete Brief

## Aufgabe 2 Quellenentropie

Gegeben sei eine binäre, gedächtnislose Nachrichtenquelle  $Q$ , welche voneinander statistisch unabhängige Zeichen aus dem Alphabet  $\mathcal{X} = \{a, b\}$  emittiert. Wir modellieren diese Nachrichtenquelle als diskrete Zufallsvariable  $X$ . Die Wahrscheinlichkeit, dass die Quelle das Zeichen  $X = a$  emittiert, betrage  $p_a = \Pr[X = a] = 0.25$ .

a)\* Bestimmen Sie die Wahrscheinlichkeit  $p_b$ , dass das Zeichen  $X = b$  emittiert wird.

Da  $p_a + p_b = 1$  folgt  $p_b = 0.75$ .

b) Bestimmen Sie den Informationsgehalt  $I(a)$  und  $I(b)$  beider Zeichen.

$$I(a) = -\log_2 p_a = 2,00 \text{ bit}$$

$$I(b) = -\log_2 p_b \approx 0,42 \text{ bit}$$

c) Bestimmen Sie die Entropie  $H$  der Quelle.

$$H(X) = \sum_{x \in \mathcal{X}} p_x I(x) = 0,81 \text{ bit}$$

**d)** Bestimmen Sie die Auftretswahrscheinlichkeiten  $p_0$  und  $p_1$  einer anderen binären Nachrichtenquelle  $Q'$ , so dass deren Entropie  $H$  maximal ist.

Zunächst drücken wir  $p_1$  durch  $p_0$  aus und schreiben  $p_1 = 1 - p_0$ . Zur Vereinfachung schreiben wir  $p_0 = p$ . Anschließend lässt sich die Entropie  $H$  als Funktion in Abhängigkeit von  $p$  ausdrücken und die gesuchte Wahrscheinlichkeit mittels Ableitung bestimmen:

$$\begin{aligned} H(p) &= -p \log_2(p) - (1-p) \log_2(1-p) \\ \frac{dH}{dp} &= -\log_2(p) - \frac{p}{p \ln(2)} + \log_2(1-p) + \frac{1-p}{(1-p) \ln(2)} \\ &\Rightarrow \log_2(p) + \frac{p}{p \ln(2)} \stackrel{!}{=} \log_2(1-p) + \frac{1-p}{(1-p) \ln(2)} \end{aligned}$$

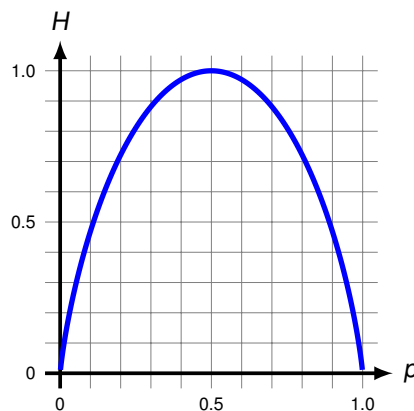
Vergleich beider Seiten liefert  $p = 1 - p = 1/2$ .

**e)** Wie hoch ist demnach die maximale Entropie einer binären Quelle?

Die Entropie wird maximiert, wenn  $\Pr[X = a] = \Pr[X = b] = 0.5$  gilt. Die maximale Entropie beträgt daher

$$H_{\max} = -2 \cdot 0.5 \cdot \log_2(0.5) = 1 \text{ bit}.$$

**f)** Skizzieren Sie die Quellenentropie  $H$  einer binären Quelle allgemein in Abhängigkeit der Auftretswahrscheinlichkeit  $p$ .



**g)** Offensichtlich ist die Entropie  $H(X) < 1$  nicht maximal. Welche Schlussfolgerung lässt sich aus dieser Tatsache für den von der Quelle  $Q$  emittierten Datenstrom hinsichtlich Redundanz ableiten?

Die von  $Q$  emittierte Zeichenkette, welche nichts anderes ist als verschiedene Realisierungen der Zufallsvariable  $X$ , beinhaltet Redundanz. Der von  $Q$  erzeugte Datenstrom ist durchschnittlich mit weniger als 1 bit/ Symbol darstellbar.

**h)** Verallgemeinern Sie die Ergebnisse der Teilaufgaben d) und e) auf eine  $N$ -äre Quelle, d. h. auf eine Quelle, die  $N$  unterschiedliche Zeichen emittiert.

Allgemein gilt für die Entropie

$$H(X) = \sum_{x \in \mathcal{X}} I(x) p_i.$$

Mit der Forderung  $p_i = p$ , d. h. alle Zeichen treten mit derselben Wahrscheinlichkeit auf, folgt sofort  $p = 1/N$  und damit

$$H = \sum_{x \in \mathcal{X}} l(x)p = - \sum_{i=1}^N \log_2 \left( \frac{1}{N} \right) \frac{1}{N} = \log_2(N).$$

### Aufgabe 3 Binärpräfixe

Der Unterschied zwischen Binärpräfixen und SI-Präfixen sorgt immer wieder für Verwirrung. Das Problem besteht in widersprüchlichen Angaben insbesondere auf Seiten der Betriebssysteme: Häufig wird die Speicherbelegung von Massenspeichern in Binärpräfixen angegeben, obwohl die angegebenen Einheiten SI-Präfixe enthalten.

Ein Beispiel: Sie kaufen eine Festplatte mit einer vom Hersteller ausgewiesenen Kapazität von 3 TB. Im Kleingedruckten auf der Verpackung finden Sie den Hinweis „1 TB =  $10^{12}$  B“. Es handelt sich also klar um SI-Präfixe. Nehmen wir an, das verwendete Betriebssystem rechnet mit Binärpräfixen.

SI-Präfix		Wert	Binärpräfix		Wert
k	(kilo)	$10^3$	Ki	(Kibi)	$2^{10}$
M	(Mega)	$10^6$	Mi	(Mebi)	$2^{20}$
G	(Giga)	$10^9$	Gi	(Gibi)	$2^{30}$
T	(Tera)	$10^{12}$	Ti	(Tebi)	$2^{40}$
P	(Peta)	$10^{15}$	Pi	(Pebi)	$2^{50}$

**Tabelle 1:** SI-Präfixe und Binärpräfixe im Vergleich

a)\* Geben Sie die Kapazität der Festplatte in TiB an.

$$3 \text{ TB} = 3 \cdot 10^{12} \text{ B} = \frac{3 \cdot 10^{12}}{2^{40}} \text{ TiB} \approx 2,73 \text{ TiB}$$

b)\* Bestimmen Sie für die in Tabelle 1 angegebenen Präfixe den prozentualen Unterschied zwischen SI- und Binär-präfixen.

$$\begin{aligned} \frac{k}{Ki} &= \frac{10^3}{2^{10}} \approx 97.66 \% \Rightarrow e = 2.34 \% \\ \frac{M}{Mi} &= \frac{10^6}{2^{20}} \approx 95.37 \% \Rightarrow e = 4.63 \% \\ \frac{G}{Gi} &= \frac{10^9}{2^{30}} \approx 93.13 \% \Rightarrow e = 6.87 \% \\ \frac{T}{Ti} &= \frac{10^{12}}{2^{40}} \approx 90.95 \% \Rightarrow e = 9.05 \% \\ \frac{P}{Pi} &= \frac{10^{15}}{2^{50}} \approx 88.82 \% \Rightarrow e = 11.18 \% \end{aligned}$$

**Übrigens:** Die Angabe von Binärpräfixen ist nur für Byte-Werte üblich. Bitwerte, z. B. kbit oder Mbit, werden ausschließlich mit SI-Präfixen angegeben.

## Aufgabe 4 Daten per LKW (Hausaufgabe)

Um Animationsfilme in München zu fördern wird eine Kooperation zwischen dem Hochleistungsrechenzentrum Garching und den Bavaria-Filmstudios geschlossen. Statt einer Datenleitung sollen LKWs einer Spedition die Daten vom Rechenzentrum in Garching zu den Filmstudios in Grünwald bringen. Um die Stadt nicht zu sehr zu belasten, fahren die LKWs den Weg zwischen Garching und Grünwald über A9 und A99, was einer Distanz von  $d = 52$  km entspricht. Im Mittel kann ein LKW die Strecke mit  $v = 55$  km/h befahren.

Der LKW werde mit einer Rate von  $r_{in} = 12$  Festplatten/min beladen und mit einer Rate von  $r_{out} = 15$  Festplatten/min entladen. Die Kapazität des LKWs betrage  $N = 512$  Festplatten. Zur Anwendung kommen Festplatten mit einer Kapazität von  $C = 12$  TB.

a)\* Wie lange dauert das Beladen des LKWs?

$$T_{in} = \frac{N}{r_{in}} = \frac{512 \text{ Festplatten}}{12 \text{ Festplatten/min}} \approx 42,67 \text{ min}$$

b) Wie lange dauert es, bis die Daten beim Filmstudio angekommen und entladen sind?

$$\begin{aligned} T &= T_{in} + T_{trans} + T_{out} = T_{in} + \frac{d}{v} + \frac{N}{r_{out}} \\ &= T_{in} + \frac{52 \text{ km}}{55 \text{ km/h}} \cdot \frac{60}{1} \text{ min/h} + \frac{512 \text{ Festplatten}}{15 \text{ Festplatten/min}} \\ &\approx 42,67 \text{ min} + 56,73 \text{ min} + 34,13 \text{ min} \approx 133,53 \text{ min} \approx 2 \text{ h } 14 \text{ min} \end{aligned}$$

c) Welcher Datenrate  $r$  in Gbit/s und GiB/s entspricht dies?

$$\begin{aligned} T &\approx 133,53 \text{ min} \approx 8012 \text{ s} \\ C_{ges} &= 512 \text{ Festplatten} \cdot 12 \frac{\text{TB}}{\text{Festplatten}} = 6144 \text{ TB} \\ r &= \frac{C_{ges}}{T} = \frac{512 \cdot 12 \cdot 8 \cdot 10^{12} \text{ bit}}{8012 \text{ s}} \approx 6,13 \text{ Tbit/s} \\ &= \frac{C_{ges}}{T} \text{ bit/s} \cdot \frac{1}{8} \text{ B/bit} \cdot \frac{1}{2^{30}} \text{ GiB/B} \approx 714,18 \text{ GiB/s} \end{aligned}$$

d) Angenommen es stehen genug LKWs zur Verfügung, so dass nach 2 min Pause bereits der nächste LKW beladen werden kann. Welche Datenrate  $r'$  ist jetzt zu erreichen?

Aus Teilaufgabe a) kennen wir bereits die Zeit  $T_{in}$  zum Beladen eines LKW. Wenn nun zwischen dem Beladen der einzelnen LKWs noch zusätzlich eine Pause von 2 min verstreicht, so kann alle  $T_{in} + 2$  min ein LKW Garching verlassen. Eben in diesen Zeitabständen erreichen LKWs auch Grünwald. Die Datenrate steigert sich damit auf

$$r' = \frac{T}{T_{in} + 2 \text{ min}} \cdot r \approx 18,34 \text{ Tbit/s.}$$