Name	Vorname	Matrikel-Nr.

Institut für Informatik
Prof. Christian Cachin

Universität Bern

Algorithmen, Wahrscheinlichkeit und Information

Prüfung FS 2019

Vorgehen

- Beschriften Sie als erstes dieses Blatt mit Ihrem Namen, Vornamen und Ihrer Matrikel-Nr.
- Lösen Sie die Aufgaben auf dem jeweiligen Aufgabenblatt. Zusätzliche Blätter werden bei Bedarf ausgehändigt.
- Beschriften Sie zusätzliche Blätter mit Ihrem Namen und bearbeiten Sie pro Blatt nicht mehr als eine Aufgabe. Bitte geben Sie *alle* Aufgabenblätter ab.
- Benützen Sie *keine* rote Farbe.
- Dauer: 60 Minuten.

Hinweise

- Als *Hilfsmittel* zugelassen sind selbstverfasste, handgeschriebene Notizen auf 2 A4-Seiten (resp., 1 Blatt doppelseitig), jedoch keine Taschenrechner oder sonstige elektronischen Apparate.
- Wir müssen Ihre Lösungen verstehen können. Argumentieren Sie möglichst klar und präzise. Geben Sie an, welche Theoreme und andere Resultate aus der Vorlesung Sie verwenden.
- Numerische Ergebnisse sollen vereinfacht werden. Sie können als Bruch geschrieben werden, und in einem Ergebnis dürfen Ausdrücke wie $\log(3)$ oder $h(1/8)$ auftreten.
- Der natürliche Logarithmus von x wird geschrieben als $\ln x$; alle anderen Logarithmen sind zur Basis zwei.

Viel Erfolg!

Auswertung

1	2	3	4	5	6	Total	Note

Name	Vorname	pt

1 Finde das Minimum (8pt)

Der folgende Algorithmus erhält als Eingabe ein Array A aus n paarweise verschiedenen Zahlen.

```
Permutiere die Elemente von A zufällig
m ← +∞
for i ← 1, ..., n do
    if A[i] < m then
        m ← A[i]                    (*)
return m
```

Wie oft wird die Zeile (*) durchschnittlich ausgeführt, als Funktion von n? Begründen Sie das Ergebnis.

Name	Vorname	pt

2 Darts (8pt)

Auf eine Dartscheibe werfen Sie m Pfeile, welche sich zufällig und uniform auf den n Sektoren verteilen (auf dieser Scheibe gibt es sehr viele Sektoren).

a) Wie gross ist die Wahrscheinlichkeit dafür, dass ein bestimmter Sektor i leer bleibt?

b) Wie gross ist die *erwartete Anzahl* von Sektoren i, so dass Sektor i und sein Nachbarsektor $i+1$ beide leer sind?

Hinweis: Verwenden Sie die Annäherung $1 + x \approx e^x$ für kleine x.

Name	Vorname	pt

3 Suchalgorithmus (8pt)

Ein probabilistischer Algorithmus A findet die Lösung eines Suchproblems für beliebige Eingaben mit Wahrscheinlichkeit p. Falls die Suche nicht erfolgreich ist, wird A nochmals neu gestartet; dies wiederholt sich, bis A eine Lösung findet.

a) Wie viele Wiederholungen von A sind im Mittel nötig?

b) Geben Sie eine obere Schranke für die Wahrscheinlichkeit, dass die Suche mindestens m Mal so lange braucht wie erwartet, unter Verwendung der Chebyshev-Ungleichung.

c) Angenommen A findet die Lösung mit Wahrscheinlichkeit $1/3$. Was ist die Schranke aus der letzten Teilaufgabe für die Wahrscheinlichkeit, dass die Suche mindestens 51 Durchgänge von A braucht?

Name	Vorname	pt

4 Passwort erraten (8pt)

Alice schützt ihr System durch ein Passwort X, modelliert als eine Zufallsvariable. Sie kann sich nur 10 verschiedene Passwörter merken und wählt eines mit uniformer Verteilung.

Der neugieriege Bruder Bob hat Alice ausspioniert und schon Information Y über X gewonnen. Konkret ergibt die Beobachtung $Y = y$ von Bob die folgende Verteilung über die möglichen Passwörter:

Passwort X	a	b	c	d	e	f	g	h	i	j
$P_{X\|Y=y}$	$1/4$	$1/4$	$1/4$	$1/16$	$1/16$	$1/16$	$1/64$	$1/64$	$1/64$	$1/64$

a) Wie geht Bob vor, um mit der besten Strategie Alices Passwort zu erraten und sich in ihr System einzuloggen? Viele Versuche benötigt Bob dabei im Durchschnitt?

b) Was ist die Entropie der Zufallsvariable X gegeben Bobs Beobachtung $Y = y$?

c) Vergleichen Sie die beiden Ergebnisse und geben Sie eine kurze, intuitive Erklärung dafür den Unterschied.

Name	Vorname

pt

5 Entropie (8pt)

Die Zufallsvariablen $X, Y \in \{0, 1, 2\}$ sind unabhängig voneinander und wie folgt verteilt:

X	0	1	2
P_X	$1/3$	$1/3$	$1/3$

Y	0	1	2
P_Y	$1/2$	$1/4$	$1/4$

Sei $S = X + Y$ die Summe und $P = X \cdot Y$ das Produkt der Zufallsvariablen. Darüber hinaus sei $Z = 10X + Y + 1$.

Berechnen Sie folgende Grössen:

a) $H(X)$; $H(Y)$;

b) $H(XY)$; $H(Z)$;

c) $H(S)$; $H(Y|P = 0)$.

Name	Vorname	pt

6 Codes (8pt)

a) Eine Quelle X nimmt vier verschiedene Werte an mit den Wahrscheinlichkeiten:

$$[0.6, 0.2, 0.15, 0.05]$$

i. Konstruieren Sie einen optimalen binären präfixfreien Code C_1 für X und berechnen Sie die erwartete Codewort-Länge.

ii. Konstruieren Sie einen binären Shannon-Code C_2 für X und berechnen Sie die erwartete Codewort-Länge.

b) Für allgemeine Quellen ist die durchschnittliche Codewortlänge eines Huffman-Codes höchstens so gross wie jene eines Shannon-Codes. Seien y und z die zwei wenigst wahrscheinlichen Quellsymbole mit Wahrscheinlichkeiten $p_y \geq p_z$. Beschreiben Sie eine möglichst allgemeine Bedingung für p_y und p_z, unter welcher ein resultierender Huffman-Code im Mittel kürzer ist als ein Shannon-Code.