

Unterrichtsentwurf zum Thema

Vergleich von Morse- und ASCII-Code

Lernziele Die SchülerInnen wenden die Begriffe der mittleren Codewortlänge, Präfixfreiheit und binären Kodierung in der Beschreibung des Morse- und ASCII-Codes an.

- FZ 1: Die SchülerInnen ermitteln die mittlere Codewortlänge der Codes.
- FZ 2: Die SchülerInnen erkennen, dass der Morsecode ohne Pausen nicht eindeutig dekodierbar ist und begründen dies mit der fehlenden Präfixfreiheit.
- FZ 3: Die SchülerInnen leiten die Forderung ab, dass ein Binärcode mit variabler Codewortlänge präfixfrei sein muss.

Angestrebter Kompetenzzuwachs

Analyse und Vergleich formaler Merkmale einer Kodierung

Lernvoraussetzungen

Genereller Lernstand und Lernverhalten

Die 8 Jungen und 3 Mädchen des Leistungskurses Informatik in der 1. Qualifikationsphase befinden sich im zweiten Lernjahr Informatik. Einige SchülerInnen bringen darüber hinaus weitere Erfahrungen aus dem Wahlpflichtfachunterricht ein. Einen Teil der SchülerInnen habe ich im vergangenen Jahr im Basiskurs Informatik unterrichtet, die verbleibenden TeilnehmerInnen sind mir aus Hospitationen bekannt. Die recht leistungsstarke Lerngruppe zeigt sich sehr interessiert und arbeitet konzentriert und engagiert.

Spezielle Lernvoraussetzungen

Der Leistungskurs hat sich in den vergangenen Stunden intensiv mit der Datenstruktur Baum auseinandergesetzt und dabei unter anderem balancierte binäre Suchbäume (AVL-Bäume) implementiert. Als eine weitere Anwendung wurde im Vorfeld dieser Stunde mit der Sortierung mittels einer Halde das effiziente Sortierverfahren *Heapsort* erarbeitet. Die in dieser Stunde in Auszügen vorgestellten Codes dürften den SchülerInnen oberflächlich bekannt sein, so dass nach einer kurzen Einstimmung zügig zur Erarbeitung ihrer Merkmale übergegangen werden kann.

Sachauseinandersetzung und Begründung der Auswahl

Die Konstruktion eines Huffman-Codes mit Hilfe der rekursiven Datenstruktur Binärbaum verbindet die als Schwerpunkte der Grundlagen großer Programmsysteme für das 2. Unterrichtsjahr Informatik vorgesehenen Aspekte der Rekursion und der Betrachtung der Komplexität von Laufzeitverhalten und Speicherbedarf¹. Der Algorithmus ermittelt für eine spezifische Texteingabe einen präfixfreien Binärcode mit variabler Codewort-

¹ vgl. Senatsverwaltung für Bildung, Jugend und Sport: Curriculare Vorgaben für die gymnasiale Oberstufe – Informatik., S. 17

länge, der eine für die Häufigkeit der im Text auftretenden Zeichen optimale mittlere Codewortlänge aufweist. Exemplarisch wird hiermit ein Optimierungsproblem vorgestellt, das mit der Strategie eines *greedy algorithm* gelöst werden kann.

Durch das Verfahren lässt sich eine effiziente Kompression von Daten realisieren, es wird z.B. im Fax-Protokoll oder in der Kompression von Bilddateien im JPEG-Format eingesetzt. Dabei handelt es sich um einen Vorgang, den die SchülerInnen im Alltag vor allem durch Programme wie ZIP aktiv initiieren. Um das Verfahren nachvollziehen zu können, müssen die SchülerInnen jedoch zunächst relevante Begriffe wie die mittlere Codewortlänge, Präfixfreiheit und Binärcodierung sicher verwenden. Hier scheint der Vergleich des Morse- und ASCII-Codes im Besonderen geeignet, weil beide einen Teil der zu erarbeitenden Merkmale aufweisen. Mit der Gegenüberstellung der jeweiligen Vorteile von kurzem Code einerseits und präfixfreier Binärcodierung andererseits sollen die SchülerInnen für eine Lösung, die beide Vorteile in einem vereint, motiviert werden.

Begründung des didaktischen und methodischen Vorgehens

In einem kurzen **Einstieg (Phase I)** wird an einem Cartoon die Telegrafie mittels Morsecode thematisiert. Der Cartoon kontrastiert diese Technologie mit der älteren Kommunikation durch Rauchzeichen. Die SchülerInnen sollen nun spekulieren, was man heute der Telegrafie als moderne Form des Nachrichtenaustauschs entgegenstellen könnte, um so zur binären Repräsentation von Textnachrichten in Computersystemen, z.B. im ASCII-Code, überzuleiten.

In der **Hauptlernaktion** dieser Stunde in der darauf folgenden **Erarbeitung (Phase II)** untersuchen die SchülerInnen in Partnerarbeit Morse- und ASCII-Code hinsichtlich der mittleren Codewortlänge und der binären Kodierbarkeit. Diese entdeckende Schüleraktivität wird durch den Arbeitsbogen *Morse- und ASCII-Code im Vergleich* (siehe Anhang) gesteuert. Um die Konzentration auf die für das Stundenziel funktionalen Begriffe zu lenken, wird dabei die Betrachtung der Codes auf die symbolische Ebene reduziert und von der technischen Signalebene (konkrete Zeiteinheiten, Trennpausen zwischen den Signalen) abstrahiert. Darüber hinaus wird der Zeichenvorrat des Klartexts auf Großbuchstaben und eine Auswahl von Satzzeichen beschränkt. Ein affektiver Zugang durch das Kodieren eigener Texte könnte die SchülerInnen zusätzlich motivieren. Dies ist wie Erfahrungen aus Unterricht zur Kryptologie zeigen jedoch fehleranfällig und zeitaufwendig, und für das gewählte Stundenziel nicht funktional.

In der **Ergebnissicherung (Phase III)** stellen die SchülerInnen die gewonnenen Erkenntnisse vor und diskutieren ihre Lösungen. In einem vorstrukturierten Tafelbild werden die jeweiligen Vorteile der beiden Codes fixiert, um eine optimierte Lösung zu motivieren, die beide Vorteile vereint. Im Diskurs wird festgestellt, dass hierfür die Präfixfreiheit des Binärcodes sicherzustellen ist. Ggf. können weitere Aspekte wie der Performanz auf technischer Ebene (Vergleich der Übertragungsdauer mehrerer langer Signale (O = - - - = 15 Zeiteinheiten) gegenüber einem ASCII-Zeichen (O = 1001111 = 7 Zeiteinheiten)) thematisiert werden. In der **Fortführung des Unterrichts** werden die SchülerInnen nun erfahren, wie diese Anforderung durch die rekursive Konstruktion eines optimalen Codebaums zu erfüllen ist.

Operationsobjekte und antizipierte Ergebnisse

Cartoon Morsecode und Rauchzeichen



Antizipiertes Tafelbild

	Morsecode	ASCII-Code
Alphabet	kurzes / langes Signal, kurze / lange Pause	0, 1
Codewortlänge	variabel, je nach Häufigkeit der Zeichen	fest: 7 Bit
Vorteil	kurzer Code	binär kodierbar

gesucht: ein präfixfreier Binärcode, der für einen Text eine optimale mittlere Codewortlänge hat

Morse- und ASCII-Code im Vergleich

1) Dekodiert die folgenden Nachrichten:

-- • - - • - • • • • - - - - - • - • • • • • -
 • • • • • - - • • - • • • • • • • - - • •

10010011001000101001001000001001100100010110100111010
 1000100000100000110100111000011100100110010010101110



Samuel Morse

2) Ermittelt für jede Nachricht die mittlere (= durchschnittliche) Codewortlänge! Zählt dabei • und - als je ein Zeichen. Welcher Code ist bezüglich der mittleren Codewortlänge effizienter? Wie ist die Differenz zu erklären?

3) Während die Fernschreiber gelesene Zeichen sofort ausgeben, erwarten wir heute, dass der Computer empfangene Texte (z.B. Emails) für eine spätere Verwendung speichert. Die Signale müssten also in eine binäre Darstellung übertragen werden. Das Alphabet des ASCII-Codes besteht aus den zwei Symbolen 0 und 1. Könnte man eine Nachricht im Morsecode zur Bearbeitung durch Computer mit • = 0 und - = 1 codieren? Versucht das folgende nach diesem Verfahren codierte Wort zum Dekodieren in den Morsecode zurückzuführen:

000100010

- Welches Problem tritt dabei auf?
- Aus welchen Symbolen besteht das Alphabet des Morsecodes?
- Warum spielt das Problem beim ASCII-Code keine Rolle?

4) Vergleicht nun abschließend die Vorteile von Morse- und ASCII-Code! Was müsste für einen Code gelten, der die Vorteile in einem vereint?



Morsegerät

Auszug aus dem Morsecode:

Zeichen	Codewort
A	• -
B	- • • •
C	- • - •
D	- • •
E	•
F	• • - •
G	- - •
H	• • • •
I	• •
J	• - - -

Zeichen	Codewort
K	- • -
L	• - • •
M	- -
N	- •
O	- - -
P	• - - •
Q	- - • -
R	• - •
S	• • •
T	-

Zeichen	Codewort
U	• • -
V	• • • -
W	• - -
X	- • • -
Y	- • - -
Z	- - • •
.	• - • - • -
,	- - • • - -
?	• • - - • •

Codewörter werden von einer kurzen Pause getrennt,
 Wörter des Klartexts werden von einer langen Pause (=Leerzeichen) getrennt.

Auszug aus dem ASCII-Code:

Zeichen	Codewort
A	1000001
B	1000010
C	1000011
D	1000100
E	1000101
F	1000110
G	1000111
H	1001000
I	1001001
J	1001010

Zeichen	Codewort
K	1001011
L	1001100
M	1001101
N	1001110
O	1001111
P	1010000
Q	1010001
R	1010010
S	1010011
T	1010100

Zeichen	Codewort
U	1010101
V	1010110
W	1010111
X	1011000
Y	1011001
Z	1011010
.	0101110
,	0101100
?	0111111
Leerzeichen	0100000

AB Morse- und ASCII-Code im Vergleich - Erwartungshorizont

- 1) Dekodiert die folgenden Nachrichten:

.. - . - . - - . - . - - .
 - . - - . -
M A C H T M O R S E N S P A S S ?
10010011001000101001001000001001100100010110100111010
1000100000100000110100111000011100100110010010101110
I H R L E S T A S C I I .

- 2) Ermittelt für jede Nachricht die mittlere (= durchschnittliche) Codewortlänge!
Zählt dabei • und - als je ein Zeichen.

Morsecode: ohne Pausen, ohne Leerzeichen: $48/17 = 2,8$
ohne Pausen, mit Leerzeichen: $50 / 19 = 2,6$
mit Pausen zwischen Zeichen: $65/17 = 3,8$

ASCII-Code: $112/16 = 7$

Welcher Code ist bezüglich der mittleren Codewortlänge effizienter?

Der Morsecode ist kürzer.

Wie ist die Differenz zu erklären?

Der Morsecode hat eine variable Codewortlänge in Abhängigkeit der Häufigkeit der Zeichen. Beim ASCII-Code ist jedes Zeichen 7 Bit lang.

- 3) Während die Fernschreiber gelesene Zeichen sofort ausgeben, erwarten wir heute, dass der Computer empfangene Texte (z.B. Emails) für eine spätere Verwendung speichert. Die Signale müssten also in eine binäre Darstellung übertragen werden. Das Alphabet des ASCII-Codes besteht aus den zwei Symbolen 0 und 1. Könnte man eine Nachricht im Morsecode zur Bearbeitung durch Computer mit • = 0 und - = 1 codieren? Versucht das folgende nach diesem Verfahren codierte Wort zum Dekodieren in den Morsecode zurückzuführen:

000100010 = . . . - . . . -. (STEIN) oder . . . -. . . -. (EINER)

- Welches Problem tritt dabei auf?
Die Codierung ist nicht eindeutig, da Pausen nicht mit kodiert sind.
010 / -. kann sowohl für AE als auch für R stehen kann.
- Aus welchen Symbolen besteht das Alphabet des Morsecodes?
Das Morsealphabet hat vier Zeichen:
{ kurzes Signal, langes Signal, kurze Pause, lange Pause }
Warum spielt dieses Problem beim ASCII-Code keine Rolle?
Aufgrund der festen Codewortlänge weiß man, wann das nächste Codewort beginnt.

- 4) Vergleicht nun abschließend die Vorteile von Morse- und ASCII-Code!

Vorteil Morsecode: kurzer Code durch variable Codewortlänge

Vorteil ASCII-Code: [präfixfrei] binär codierbar

Was müsste für einen Code gelten, der die Vorteile in einem vereint?

Ein Codewort darf nicht als Anfang eines weiteren Codeworts auftreten.

Er muss präfixfrei sein.

Unterrichtsverlaufsplan zum Thema ‚Vergleich von Morse- und ASCII-Code‘

Zeit	Phasenfunktion / Sozialform	geplantes Lehrerverhalten	antizipiertes Schülerverhalten	Operationsobjekt	Kommentar
10:00 - 10:10 [10']	Phase I Einführung des Stundenthemas Sozialform: Geführter Diskurs	<i>Wir wollen uns zunächst einmal mit einer etwas älteren Erfindung beschäftigen. Was seht ihr?</i> <i>Wie funktioniert der Morsecode?</i> <i>Welche Kommunikationsform reicht weiter?</i> <i>Welche Beziehung könnten die beiden Indianer haben?</i> <i>Euer Großvater hat vielleicht noch telegraphiert. Was würdet ihr ihm heute als neue Technologie vorstellen, mit der man Nachrichten übermittelt?</i> <i>Wisst ihr, in welcher Codierung Computer heute Texte verarbeiten?</i> <i>Wir wollen heute den Morse- und den ASCII-Code genauer untersuchen.</i>	<i>Zwei Indianer. Einer macht Rauchzeichen mit einer Decke, einer mit einem Morsegerät.</i> <i>Übertragung einer Folge unterschiedlich langer Signale</i> <i>Die Kommunikation über Rauchzeichen ist auf Sichtweite beschränkt.</i> <i>Vater – Sohn / Großvater – Enkel</i> <i>Email / SMS / [Handy / Telefon / Internet]</i> <i>im ASCII-Code / [binär / als 0 und 1]</i>	Folie Rauch- und Morsezeichen	
<i>Feinziel 1: Die SchülerInnen ermitteln die mittlere Codewortlänge der Codes.</i> <i>Feinziel 2: Die SchülerInnen erkennen, dass der Morsecode ohne Pausen nicht eindeutig dekodierbar ist und begründen dies mit der fehlenden Präfixfreiheit.</i>					
10:10 - 10:30 [20']	Phase II Erarbeitung der Begriffe Sozialform: Gruppenarbeit	teilt AB aus, lässt Aufgaben vorlesen klärt Begriffe Zeichen, Wort, Codewort bietet Hilfestellung	dekodieren Code und ermitteln mittlere Codewortlänge, erarbeiten Fragen	AB Morse- und ASCII-Code im Vergleich	Alternative: Verbleibt weniger Zeit als vorgesehen, so wird die Analyse im Diskurs vervollständigt.

Zeit	Phasenfunktion / Sozialform	geplantes Lehrerverhalten	antizipiertes Schülerverhalten	Operationsobjekt	Kommentar
Feinziel 3: Die SchülerInnen leiten die Forderung ab, dass ein Binärcode mit variabler Codewortlänge präfixfrei sein muss.					
10:30 - 10:45 [15']	<p>Phase III Ergebnissicherung: Diskussion der Ergebnisse, Zusammenfassung und Forderung einer Optimierung</p> <p>Sozialform: Schülerpräsentation/ Geführter Diskurs</p>	<p><i>Bitte stellt eure Ergebnisse vor und füllt dabei die Tabelle an der Tafel aus. Diejenigen, die nicht vorstellen, ergänzen bitte wenn sie andere Lösungen oder Erklärungen haben.</i></p> <p>ergänzt an geeigneter Stelle den Begriff <i>präfixfrei</i></p> <p><i>Was müsste für einen binären Code mit variabler Codewortlänge gelten?</i></p> <p><i>Wie man so einen Code konstruieren kann, werden wir uns nach der Pause überlegen.</i></p> <p>ergänzt Tafelbild, fordert S auf, Tabelle und Forderung zu übernehmen</p> <p style="text-align: center;">▼ optional ▼</p> <p><i>Ist die Pause am Ende eines Codeworts beim Morsecode bei der mittleren Codewortlänge zu berücksichtigen?</i></p> <p><i>Warum ist kurzer Code besser?</i></p> <p><i>Wie wirkt sich ein kurzer Code auf die Speicherung von Nachrichten auf einem Computer aus?</i></p> <p><i>Deshalb wird dieses Verfahren auch zur Kompression von Daten z.B. in JPEG-Grafiken angewendet.</i></p> <p><i>Ist der Morsecode für jede Nachricht optimal?</i></p> <p><i>Für eine optimale Kompression sollte man individuelle Codes konstruieren.</i></p>	<p>stellen ihre Ergebnisse vor und begründen ggf. ihre Entscheidungen, beziehen begründet Stellung</p> <p><i>Ein Binärcode mit variabler Codewortlänge muss präfixfrei sein.</i></p> <p>übernehmen Tabelle und Forderung</p> <p style="text-align: center;">▼ optional ▼</p> <p><i>Ja, denn wenn man sie weg lässt, ist der Code nicht mehr lesbar.</i></p> <p><i>schnellere Kommunikation</i></p> <p><i>weniger Speicherbedarf</i></p> <p><i>Nein, je nach Sprache und Häufigkeitsverteilung der Zeichen.</i></p>	<p>AB Morse- und ASCII-Code im Vergleich, Tafel</p>	<p>siehe Erwartungshorizont für AB und antizipiertes Tafelbild</p> <p>Alternative: Verbleibt weniger Zeit, so erfolgt die Schlussfolgerung in der nächsten Stunde.</p> <p>Alternative: Verbleibt zusätzlich Zeit, so können die links aufgeführten weitergehenden Fragen erörtert werden.</p>