

## Hinweise für den Prüfling

**Auswahlzeit:** 30 Minuten

**Bearbeitungszeit (insgesamt):** 180 Minuten

### Auswahlverfahren

Es gibt zwei Aufgabengruppen A und B, aus denen jeweils ein Vorschlag zu bearbeiten ist. Der vorliegende Vorschlag aus der Gruppe A (objektorientierte Modellierung) ist ein Pflichtvorschlag.

Wählen Sie von den zwei vorliegenden Vorschlägen der Gruppe B (Konzepte und Anwendungen der theoretischen Informatik) einen zur Bearbeitung aus. Der nicht ausgewählte Vorschlag muss am Ende der Auswahlzeit der Aufsicht führenden Lehrkraft zurückgegeben werden.

### Erlaubte Hilfsmittel

1. ein Wörterbuch der deutschen Rechtschreibung
2. eine Liste der fachspezifischen Operatoren

### Sonstige Hinweise

ohne PC-Nutzung


### In jedem Fall vom Prüfling auszufüllen

Name: _____	Vorname: _____
Prüferin/Prüfer: _____	Datum: _____

## UTF-8-Codierung

Unicode ist ein internationaler Standard, in dem für jedes Schriftzeichen aller bekannten Schriftkulturen und Zeichensysteme ein digitaler Code festgelegt wird. UTF-8 ist die am weitesten verbreitete Codierung von Unicode-Zeichen. Dabei wird jedem Unicode-Zeichen eine kodierte Bytefolge zugeordnet, die bis zu vier Byte lang sein kann. Ein Byte besteht aus acht Bit, wobei jedes Bit den Wert 0 oder 1 haben kann.

Beispiele:

Unicode-Zeichen	UTF-8-Codierung
Buchstabe y	01111001
Buchstabe ä	11000011 10100100
eingetragene Marke ®	11000010 10101110
Eurozeichen €	11100010 10000010 10101100
Violinschlüssel 	11110000 10011101 10000100 10011110

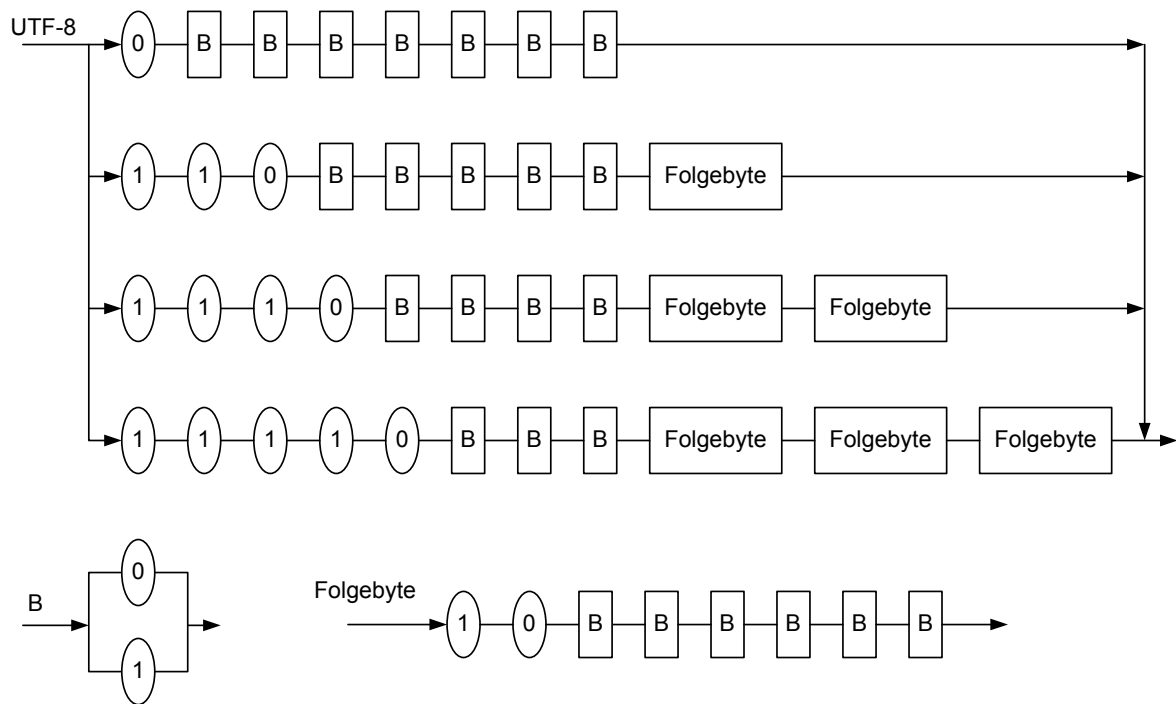
## Aufgaben

- 1.1 Beschreiben Sie die Syntaxdiagramme für die UTF-8-Codierung in Material 1. **(6 BE)**
  
- 1.2 Im Folgenden ist ein Ausschnitt aus einer Webseite als Bytefolge dargestellt. Geben Sie direkt in diesem Ausschnitt durch entsprechende Markierungen die darin enthaltenen vollständigen UTF-8-Codierungen an.  
... 10011011 01101101 11100010 10000000 10100010 01101101 11010100 10011101 11000110 ... **(4 BE)**
  
- 1.3 Bestimmen Sie, wie viele verschiedene Zeichen sich mit einem Byte und wie viele sich mit zwei Bytes gemäß der UTF-8-Codierung codieren lassen. **(2 BE)**
  
2. Modellieren Sie das Zustandsdiagramm eines endlichen Automaten, der prüft, ob eine Bitfolge eine korrekte UTF-8-Codierung der Länge ein oder zwei Byte ist. **(7 BE)**

- 3.1 Entwerfen Sie zu den Syntaxdiagrammen aus Material 1 eine kontextfreie Grammatik  $G$  und geben Sie diese in der üblichen formalen Form an.  
**(8 BE)**
- 3.2 Beschreiben Sie, was man unter einer regulären und einer kontextfreien Grammatik versteht, und bestimmen Sie, ob Ihre Grammatik  $G$  regulär oder nur kontextfrei ist.  
**(6 BE)**
- 3.3 Begründen Sie, ob die Sprache  $L(G)$  regulär oder kontextfrei ist.  
**(3 BE)**
4. Gegeben ist ein korrekt codierter Unicode-Text mit zehn Unicode-Zeichen.  
Begründen Sie, ob ein endlicher Automat feststellen kann, dass das €-Zeichen, dessen UTF-8-Codierung in der Tabelle als Beispiel angegeben ist, im Unicode-Text enthalten ist.  
**(4 BE)**

Material 1

Die Syntaxdiagramme für die UTF-8-Codierung



**Hinweis:** B steht abkürzend für Bit.