

**NAME DES DOZENTEN:** Prof. Dr.-Ing. Himmelspach



**Klausur: A105 – GESTALTUNG VON INFORMATIONSSYSTEMEN**

**QUARTAL II/2021**

**Zenturie:** A19a-c,A18

**Dauer:** 90 Minuten

**Datum:** 24.06.2021

**Seiten der Klausuraufgaben ohne Deckblatt: 7**

**Hilfsmittel:** NORDAKADEMIE-Taschenrechner, inkludierte Formelsammlung.

**Bemerkungen:**

- Bitte prüfen Sie zunächst die Klausur (alle Teile) auf Vollständigkeit.
- Bitte vermerken Sie auf Ihren Antwortbögen folgende Angaben:
  - Name
  - Matrikelnummer
  - Zenturie
  - Seitenzahl
  - Aufgabennummer
  - ModulNr. der Lehrveranstaltung/Klausur

Es sind 100 Punkte erreichbar.

Zum Bestehen der Klausur sind 50 Punkte ausreichend.

Aufgabe	Erreichbare Punkte
1	25
2	12
3	16
4	16
5	16
6	15
Summe	100

## Aufgabe 1 (25 Punkte)

### Unternehmensarchitektur

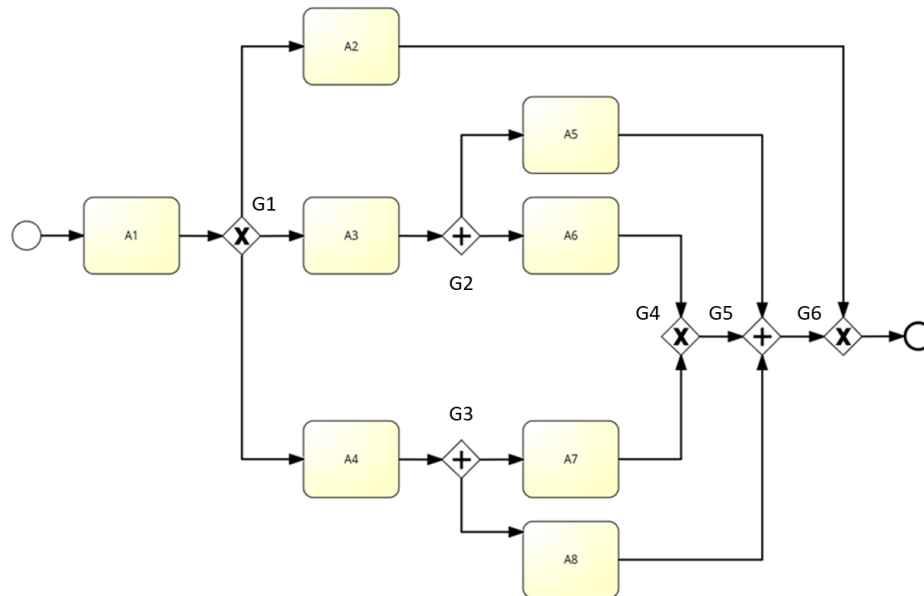
- (1.1) Schätzen Sie folgende Behauptung ein: Die unterschiedlichen Ansätze zur Unternehmensarchitektur müssen ohne Adaption in Unternehmen eingesetzt werden - nur so wird diesen hinreichend Rechnung getragen. (1)
- A. Trifft zu  
B. Trifft nicht zu
- (1.2) Was sind Ziele, die mittels einer Unternehmensanwendungsarchitektur erreicht werden sollen (nennen Sie mindestens 3)? (3)
- (1.3) Geben Sie eine kurze Antwort auf die Frage: „Wozu dient die ADM in TOGAF?“ (3)
- (1.4) Für jede der TOGAF ADM-Phasen definiert TOGAF „Eingaben“. Nennen Sie 2 dieser Eingaben. (2)
- (1.5) Erläutern Sie den Zweck einer GAP Analyse. (4)
- (1.6) Softwareprodukte A,B und C werden an unterschiedlichen Teilmengen der möglichen Orte ( $O_1 \dots O_5$ ) in einem Unternehmen von den Rollen  $R_1 \dots R_{10}$  eingesetzt. Dabei tauschen die Softwareprodukte Informationen ( $I_1 \dots I_4$ ) untereinander aus. Die Softwareprodukte benötigen dabei eine spezifische Infrastruktur ( $H_1 \dots H_3$ ) und dienen zur Unterstützung der Geschäftsprozesse ( $GP_1 \dots GP_7$ ).
- i. Skizzieren Sie zwei unterschiedliche Matrixdarstellungen, die jeweils eine Teilmenge der obigen Informationen wiedergibt. Erläutern Sie den Nutzen ihrer Darstellungen kurz. (12)

## Aufgabe 2 (12 Punkte)

## (2.1) Korrekte BPMN Modellierung

(6)

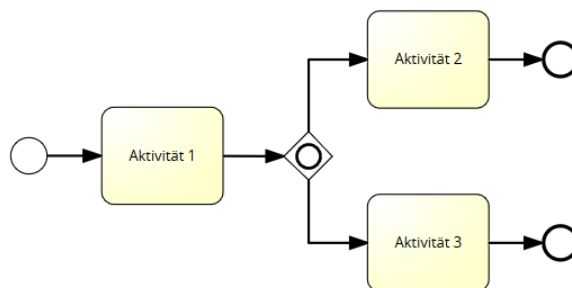
In dem folgenden Prozess kann es zu einem „Deadlock“ kommen, d.h. zu einer Situation, an der eine Prozessinstanz nicht weiterläuft, weil ein Token „stecken bleibt“. Benennen Sie die entsprechende Stelle, an der dieser Deadlock auftritt und erläutern Sie kurz, wie dieser zustande kommen kann.



## (2.2) Korrekte BPMN Modellierung

(6)

Der nachfolgende Prozess wurde in der BPMN mit einem OR-Gateways (inklusive Oder) dargestellt. Erstellen Sie ein Prozessmodell in der BPMN, das diesen Prozess mit derselben Ablauflogik ohne Verwendung des OR-Gateways darstellt.



### Aufgabe 3 (16 Punkte)

#### BPMN Modellierung

Modellieren Sie den folgenden Geschäftsprozess in der BPMN unter Beachtung der in der Vorlesung besprochenen Modellierungsregeln: Wenn der Bewerber eine interessante Stellenausschreibung entdeckt, erstellt er eine Bewerbung und schickt diese an das Unternehmen. Dort wird nach Erhalt der Bewerbung eine Eingangsbestätigung an den Bewerber gesendet. Im Unternehmen wird danach die Bewerbung geprüft. Wenn die Bewerbung interessant ist, wird eine Intervieweinladung an den Bewerber geschickt, andernfalls eine Absage. Wenn der Bewerber eine Eingangsbestätigung erhalten hat, wartet er auf die Einladung oder Absage. Im Falle einer Einladung bestätigt er die Teilnahme am Interview. Im Falle einer Absage ergeben sich keine weiteren Aktionen. Wenn der Bewerber das Interview bestätigt hat und der Interviewtermin erreicht ist, wird das Interview zwischen Bewerber und Unternehmen durchgeführt.

Verwenden Sie jeweils eigene Pools für den Bewerber und für das Unternehmen.

### Aufgabe 4 (16 Punkte)

#### Entscheidungsbaum

Sie sollen einen Entscheidungsbaum erstellen, der es erlaubt vorherzusagen, welchen Wert Gewinner annehmen wird, wenn die anderen Werte gegeben sind.

- (4.1) Erstellen Sie dazu unter Verwendung des Hunt-Algorithmus und des GINI-Index einen Entscheidungsbaum. Geben Sie alle benötigten Rechenschritte und Entscheidungen an. Als Stopp-Kriterium verwenden Sie bitte: die maximale Baumtiefe beträgt 2 (Wurzel + 1 Ebene mit Blättern; AlgoDat: Pfadlänge = 1)! (10)

Sport	Mittagessen	Sportart 1	Gewinner
Profi	Pizza Margarita	Tischtennis	ja
Profi	Pizza Margarita	Tischtennis	nein
Amateur	Frikadellenbrötchen	Tennis	nein
Amateur	Pizza Margarita	Tischtennis	ja
Hobby	Canneloni	Squash	nein
Hobby	Pizza Margarita	Tennis	nein
Amateur	Pizza Margarita	Tischtennis	nein
Hobby	Canneloni	Tischtennis	nein
Profi	Pizza Margarita	Squash	ja
Amateur	Canneloni	Squash	ja

- (4.2) Zeichnen Sie den resultierenden Entscheidungsbaum. (4)
- (4.3) Beschreiben Sie, wie der Algorithmus ohne das Stopp-Kriterium fortfahren würde. (2)

**Aufgabe 5** (16 Punkte)**Datenauswertung (Wahrheiten)**

(5.1) Gegeben sind folgende Daten:

(16)

**Fakten**

Pro- dukt	Kunde	Datum	Um- satz
P02	K30	2019-02	100
P02	K20	2019-02	200
P03	K30	2019-02	250
P01	K10	2019-02	150
P01	K10	2019-03	100
P03	K30	2019-03	300
P02	K20	2019-03	150
P02	K30	2019-03	50
P01	K10	2019-03	200
P01	K10	2019-04	250
P01	K10	2019-04	200
P01	K30	2019-04	100
P03	K30	2019-04	250
P01	K20	2019-05	50
P03	K30	2019-05	300
P02	K20	2019-05	100
P01	K30	2019-05	50
P01	K30	2019-05	200

**Dimension: Produkt**

Produkt	Produkt- gruppe	gültig von	gültig bis
P01	PG Y	01.01.2000	31.03.2019
P01	PG Z	01.04.2019	31.12.2050
P02	PG Z	01.02.2019	31.12.2019
P03	PG Z	01.01.2019	30.04.2019
P03	PG Y	01.05.2019	31.12.2050

**Dimension: Kunde**

Kunde	Kunden- gruppe	gültig von	gültig bis
K10	KG A	01.01.2000	31.01.2019
K10	KG B	01.02.2019	31.12.2019
K20	KG A	01.01.2000	31.03.2019
K20	KG B	01.04.2019	31.12.2019
K30	KG B	01.01.2000	31.12.2050

Stellen Sie die monatlichen Umsätze nach aktueller Struktur (Juni 2019) und nach gültiger Struktur im März 2019 gleichzeitig gruppiert nach Kundengruppen und Produktgruppe dar.

Aktuelle Struktur (Juni 2019)

Kundengruppe	Produktgruppe	2019-02	2019-03	2019-04	2019-05

Gültige Struktur im März 2019

Kundengruppe	Produktgruppe	2019-02	2019-03	2019-04	2019-05

## Aufgabe 6 (15 Punkte)

### Assoziationsanalyse

- (6.1) Führen Sie auf Grundlagen der Datenbasis und einem  $\text{minSupport} \geq 0.6$  den ersten Schritt des Apriori-Algorithmus (Finden häufiger Item-Mengen) durch. (9)

Gegeben sind folgende Transaktionen von Einkäufen:

TID	Items
1	Buschbohnen, Orangen, Rosen, Saubohne
2	Milch, Orangen, Rosen, Saubohne
3	Buschbohnen, Milch, Pommes, Saubohne
4	Orangen, Pommes, Rosen, Saubohne
5	Buschbohnen, Milch, Orangen, Pommes
6	Buschbohnen, Orangen, Rosen, Saubohne
7	Milch, Rosen
8	Orangen, Pommes, Rosen, Saubohne
9	Buschbohnen, Milch, Orangen, Pommes
10	Buschbohnen, Orangen, Rosen, Saubohne

- (6.2) 2. Schritt des Algorithmus (6)

Sie haben aus dem ersten Schritt des A-Priori Algorithmus folgende Item-Mengen, jeweils gegeben mit ihren absoluten Häufigkeiten, als relevante Mengen ermittelt. Erzeugen Sie alle möglichen Assoziationsregeln und bewerten Sie diese auf Basis der  $\text{minKonfidenz} = 0.6$ ,  $|D| = 10$  (Anzahl der Datensätze)

$|\{\text{Ketchup}\}| = 6$   $|\{\text{Wiener}\}| = 6$   $|\{\text{Scholle}\}| = 7$   $|\{\text{Forelle}\}| = 7$   $|\{\text{Forelle, Wiener}\}| = 6$   
 $|\{\text{Forelle, Scholle}\}| = 6$   $|\{\text{Scholle, Wiener}\}| = 6$   $|\{\text{Forelle, Scholle, Wiener}\}| = 6$

Viel Erfolg!

## Assoziationsanalyse

Items $G$	Grundgesamtheit $G$ von Bezeichnern
Item-Menge $X$	Nicht-leere Teilmenge $X \subseteq G$
$k$ -Item-Menge	Item-Menge mit $k$ Elementen ( $k \geq 1$ )
Datenbank/-basis $D$	Menge $D$ von Item-Mengen (Transaktionen); die Anzahl der Transaktionen ist $ D $
Absolute Häufigkeit $n(X)$	Anzahl der Item-Menge $X$ in der Datenbank $D$
Assoziationsregel	Implikation der Form $X \rightarrow Y$ mit $X \cap Y = \emptyset$
Support einer Item-Menge $X$	$Support(X) = \frac{n(X)}{ D }$
Support einer Assoziationsregel $X \rightarrow Y$	$Support(X \rightarrow Y) = Support(X \cup Y)$
Konfidenz einer Assoziationsregel $X \rightarrow Y$	$Konfidenz(X \rightarrow Y) = \frac{n(X \cup Y)}{n(X)}$
Lift einer Assoziationsregel $X \rightarrow Y$	$Lift(X \rightarrow Y) = \frac{Konfidenz(X \rightarrow Y)}{Support(Y)} = \frac{\hat{P}(X Y)}{\hat{P}(Y)}$

## Entscheidungsbäume

Accuracy (Treffgenauigkeit)

$$Accuracy = \frac{\text{korrekte Klassifikationen}}{\text{alle Klassifikationen}} = \frac{r_p + r_n}{r_p + f_p + f_n + r_n}$$

Error rate (Klassifikationsfehler)

$$Errorrate = \frac{\text{falsche Klassifikationen}}{\text{alle Klassifikationen}} = \frac{f_p + f_n}{r_p + f_p + f_n + r_n}$$

### Homogenitätsmaße

Gegeben: Knoten  $T$  mit  $|T|$  Datensätzen in  $k$  Klassen (Partitionen von  $T$ )

$p_i$  = relative Anzahl der Datensätze der Klasse  $i$ , mit ( $i = 1, \dots, k$ )

Entropy des Knotens  $T$

$$Entropy(T) = -\sum_{i=1}^k p_i \log_2 p_i$$

Gini-Index des Knotens  $T$

$$Gini(T) = 1 - \sum_{i=1}^k p_i^2$$

Klassifikationsfehler des Knotens  $T$

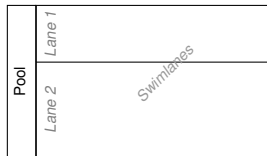
$$Klassifizierungsfehler(T) = 1 - \max\{p_i\}$$

Information Gain bei Aufspaltung des Knotens  $T$

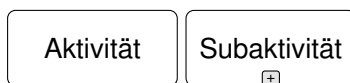
$$Information\ Gain = H(T) - \sum_{i=1}^m \frac{|T_i|}{|T|} H(T_i)$$

## BPMN

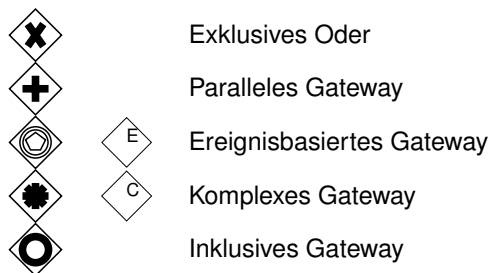
### Pools und Schwimmbahnen



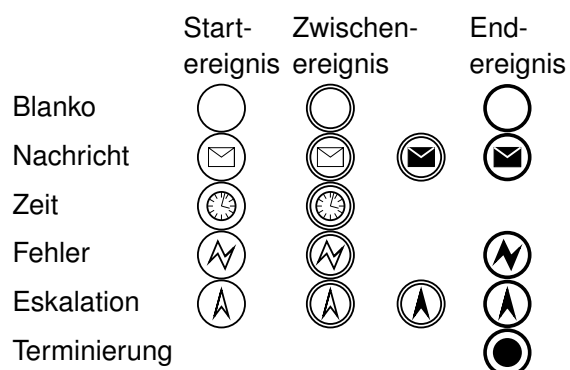
### Aktivitäten



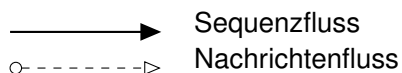
### Gateways



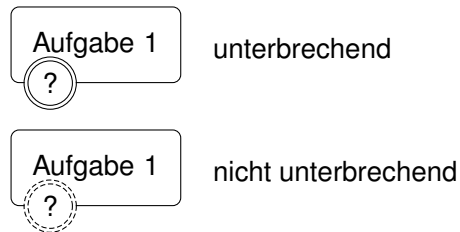
### Ereignisse



### Verbindungen



### Angeheftete Ereignisse



## Archimate

### Assoziationen

