CHALMERS
EXAMINATION / TENTAMEN

Course code/kurskod	Course name/kursnamn			
DIT023	Mathematical Foundations for Software Engineering			
Anonymous code Anonym kod		Examination date Tentamensdatum	Number of pages Antal blad	Grade Betyg
435		27/10/2022	9	5

* I confirm that I've no mobile or other similar electronic equipment available during the examination.
Jag intygar att jag inte har mobiltelefon eller annan liknande elektronisk utrustning tillgänglig under eximinationen.

Solved task Behandlade uppgifter No/nr		Points per task Poäng på uppgiften	Observe: Areas with bold contour are to completed by the teacher. Anmärkning: Rutor inom bred kontur ifylles av lärare.
1	X	6	
2	X	22	
3	X	19	
4	X	8	
5	X	17	
6	X	22	
7			
8		Σ=94	
9			
10			
11			
12			
13			
14			
15			
16			
17			
Bonus poäng		9,6	
Total examination points Summa poäng		104,0	

CHALMERS	Anonymous code / Anonym kod: DITO 23 435	Points for question (to be filled in by teacher) / Poäng på uppgiften (ifylles av lärare)	Consecutive page no. / Löpande sid nr: 1 — Question no. / Uppgift nr: 1

4

Question 1

1.1 d 1 pt

1.2 d 1 pt

1.3 d 1 pt

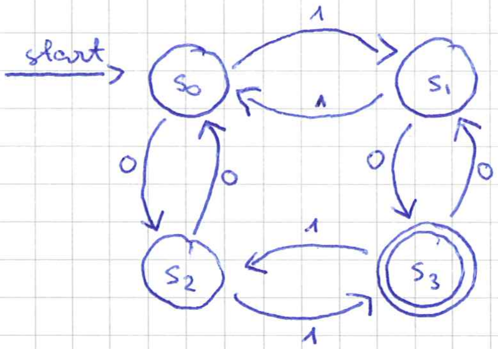
1.4 Word 1: AaAaa 3 pt
Word 2: AbBab
Word 3: KuKus

Question 2

2.1 b ✓ 1pt

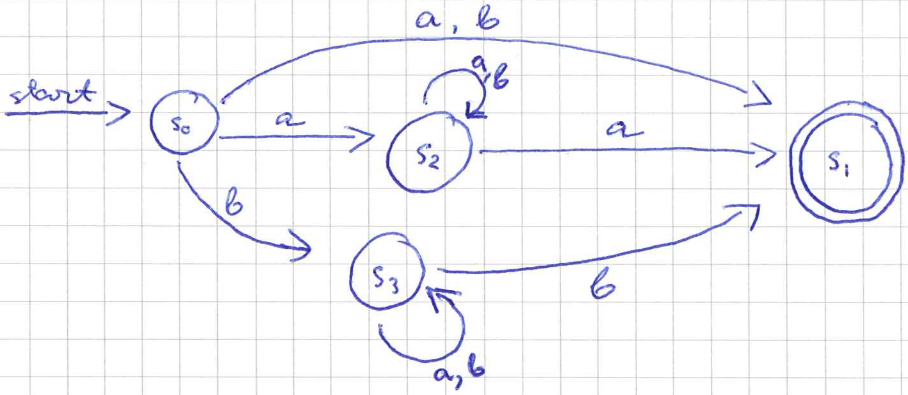
2.2 b ✓ 1pt

2.3



6pt

2.4 Because it is not specified in the task, my finate-state machine is nondeterministic.



14pt

$a(a \cup b)^* a \cup b(a \cup b)^* b \cup a \cup b$

4

CHALMERS	Anonymous code / Anonym kod: 435	Points for question (to be filled in by teacher) / Poäng på uppgiften (ifylles av lärare)	Consecutive page no. / Löpande sid nr: 3
			Question no. / Uppgift nr: 3

Question 3

3.1 $(((r \vee q) \wedge ((r \to q) \wedge p)) \to (p \wedge q))$

p	q	r	r->q	(r->q)∧p	r∨q	(r∨q)∧((r->q)∧p)	p∧q	(((r∨q)∧((r->q)∧p))->(p∧q))
T	T	T	T	T	T	T	T	T
T	T	F	T	T	T	T	T	T
T	F	T	F	F	T	F	F	T
T	F	F	T	T	F	F	F	T
F	T	T	T	F	T	F	F	T
F	T	F	T	F	T	F	F	T
F	F	T	F	F	T	F	F	T
F	F	F	T	F	F	F	F	T

8pt

3.2 The compound proposition from 3.1 is logically equivalent to

T which means, the compound proposition is a tautology. It is a logical statement that is always true.

2pt

3.3 b

1pt

3.4 i) A compound proposition is a contingency when it is not a tautology nor a contradiction. That means, it is a compound proposition, that is sometimes true and sometimes false. There exists a mix of truth values for the combined statements so that the compound proposition is true, and at least one so that it is false.

2pt

ii) A compound proposition is satisfiable when there exists an assignment of truth values so that the compound proposition is true. That means, it is at least once true.

1pt

iii) $(((r \to q) \wedge ((r \to q) \wedge p)) \to (p \wedge q))$

p	q	r	r->q	(r->q)∧p	(r->q)∧((r->q)∧p)	p∧q	((r->q)∧((r->q)∧p))->(p∧q)
T	T	T	T	T	T	T	T
T	T	F	T	T	T	T	T
T	F	T	F	F	F	F	T
T	F	F	T	T	T	F	F
F	T	T	T	F	F	F	T
F	T	F	T	F	F	F	T
F	F	T	F	F	F	F	T
F	F	F	T	F	F	F	T

5pt

The above compound proposition is not satisfiable when:

p = True
q = False
r = False

Answer only one question on this page. Do not write on the back of this paper
Behandla endast en uppgift på detta blad. Skriv ej på baksidan

CHALMERS	Anonymous code / Anonym kod: 435	Points for question (to be filled in by teacher) / Poäng på uppgiften (ifylles av lärare)	Consecutive page no. / Löpande sid nr: 4 / Question no. / Uppgift nr: 4

Question 4

4.1 Prove that $\sum_{i=1}^{n} i = 1 + \dots + n = \frac{5n + 5n^2}{10}$ for $n > 0$

Prove by mathematical induction

Basis step. Let's prove when $n = 1$.

$LHS_1 = 1$

$RHS_1 = \frac{5 \cdot 1 + 5 \cdot 1^2}{10} = 1$

$LHS_1 = RHS_1 = 1$

Basis step is true and holds.

1p

Inductive hypothesis: Let's assume that:

$$\underbrace{1 + \dots + n}_{LHS_0} = \underbrace{\frac{5n + 5n^2}{10}}_{RHS_0}$$

1p

Inductive step:

Prove that this statement holds for every $n > 0$.
$n = k$ for $k + 1$:

LHS_{k+1}: $\underbrace{1 + \dots + k}_{LHS_0 = RHS_0} + (k+1) = \frac{5k + 5k^2}{10} + (k+1)^{\cdot 10} =$

1p

$= \frac{5k + 5k^2 + 10k + 10}{10} = \frac{5k^2 + 15k + 10}{10} = \frac{\cancel{5}(k^2 + 3k + 2)}{\cancel{10}\,2} = \frac{k^2 + 3k + 2}{2}$

RHS_{k+1}: $\frac{5(k+1) + 5(k+1)^2}{10} = \frac{5k + 5 + 5(k^2 + 2k + 1)}{10} = \frac{5k + 5k^2 + 10k + 10}{10}$

$= \frac{5k^2 + 15k + 10}{10} = \frac{k^2 + 3k + 2}{2}$

$LHS_{k+1} = RHS_{k+1}$

We proved that this statement / equation holds for every $n > 0$.

1p

4

CHALMERS	Anonymous code / Anonym kod: 435	Points for question (to be filled in by teacher) / Poäng på uppgiften (ifylles av lärare)	Consecutive page no. / Löpande sid nr: 5
			Question no. / Uppgift nr: 4

4

Question 4

4.2

Prove that n^3+2n is divisible by 3 for $n \geq 0$

~~We can write the same statement~~

Proof by mathematical induction.

Basis step for $n=0$

$$0^3+2\cdot 0 = 0$$

0 is divisible by 3. 1p $\frac{0}{3}=0$

Inductive hypothesis

Let's assume that n^3+2n is divisible by 3. 1p

Inductive step

Prove that it holds for every $n>0$
$n=k$
So for $k+1$.

$$(k+1)^3+2(k+1) = (k+1)[(k+1)^2+2] = (k+1)(k^2+2k+3) =$$

$$= k^3+2k^2+3k+k^2+2k+3 =$$

$$= k^3+3k^2+3k+2k+3 =$$

$$= \underbrace{k^3+2k}_{\text{Inductive hypothesis}} + 3k^2+3k+3 = \underbrace{k^3+2k}_{\text{divisible by 3}} + \underbrace{3(k^2+k+1)}_{\text{divisible by 3}}$$

1p

Because of our inductive hypothesis, k^3+2k must be divisible by 3. The other half is multiplied by 3 which means that it is divisible by 3. 1p
So all together, the sum must also be divisible by 3.

We proved that n^3+2n is divisible for every $n \geq 0$.

4

4

CHALMERS	Anonymous code / Anonym kod: 435	Points for question (to be filled in by teacher) / Poäng på uppgiften (ifylles av lärare)	Consecutive page no. / Löpande sid nr: 6 · Question no. / Uppgift nr: 5

Question 5

5.1 $O(n^2)$

5.2 ~~$O(n)$~~ $O(\infty)$

5.3 $O(n^2)$

2pt

5.4 T I B E L P P A

8pt

5.5.1

A - $O(\log(n))$

B - $O(n^2)$

C - $O(n^3)$

Each algorithm spends 5s to process 10 000 items.

ok

A

$T_A(10\,000) = c \log 10000 = 5$

$c_A = \frac{5}{\log 10000} = 1{,}25$

For 20000 items:

$T_A(20\,000) = 1{,}25 \cdot \log 20000 \approx 5{,}38$ (s)

B

$T_B(10000) = c\, 10000^2 = 5$

$c = \frac{5}{10000^2} = \frac{1}{20000000}$

For 20000 items:

$T_B(20000) = \frac{1}{20000000} \cdot 20000^2 = 20$ (s)

3pt

C $T_C(10\,000) = c \cdot 10000^3 = 5$

$c = \frac{5}{10000^3} = 5 \cdot 10^{-12}$

For 20 000 items:

$T_C(20\,000) = 5 \cdot 10^{-12} \cdot 20\,000^3 = 40$ (s)

Answer: For 20000 items, the processing ~~the~~ time for ~~of~~ A is 5,38 s, for B 20 s and C 40 s.

1pt

5.5.2 It is better to choose algorithm B over C ~~for~~ $n > 10\,000$ items. If we assume that items are integers, that means for $n \geq 10001$.

Processing time for 10001 items for algorithm A is 5,00100 seconds and for algorithm B it is 5,00150 seconds.

1pt

CHALMERS	Anonymous code	Points for question (to be filled in by teacher)	Consecutive page no. Löpande sid nr 7
4	Anonym kod 435	Poäng på uppgiften (ifylles av lärare)	Question no. Uppgift nr 5

5.6

5.6.1 $$\frac{\sum_{v \in V} \deg^-(v)}{\sum_{v \in V} \deg^+(v)} = \frac{7}{7} = 1$$

1pt

5.6.2 ~~There does not exist any Euler paths in G5.~~
There exists ~~3~~ 6 possible Euler path.

5.6.3 C -> B -> D -> C -> C -> A -> B -> A

1pt

5.6.4 True

Question 6

6.1 a) I think we have heteroscedasity because all variance is not equal for every value. The plots do not seem to follow the assumptions for normality. 2p

b) We could but we should not. If we run t-tests, we would end up doing several of them which would end with alfa inflation. That means, that we would "allow" us to make more mistakes but we don't want that. In order for the alpha level to stay as low as possible, we can not run multiple t-tests. 2p

c) Null Hypothesis: Obsolete requirements do not have any impact on the estimates for the remaining requirements.

$\mu_1 = \mu_2 = \mu_3$

Alternative Hypothesis: Obsolete requirements have an impact on the estimates for the remaining requirements.

$\mu_1 \neq \mu_2 \neq \mu_3$ 2p

d) To see if the task group is explaining estimation variance I would calculate R^2. R^2 is the explained variance. In order to see where the differences exactly are, I would run a Post-Hoc test. F (ANOVA) 1p

6.2 a) Null Hypothesis is that $p < 0,05$ that means that if $p < 0,05$ then the data is not normally distributed. –

b) Group A and B are not normally distributed. 1p

c) Group C is normally distributed. 1p

d) If there are problems with normality
1) It is possible to modify our data with special tests / algorithms so it becomes normal.
2) Use ~~m~~ non parametric tests that do not assume / require normality. 2p

6.3 a) I believe they are, as the curves on the plot are symmetrical and bell-shaped. 1p

b) According to the Central Limit Theorem, whatever the population (if it is normally distributed or not), if we take large enough sample size, the sample means will be normally distributed. 1p

4

CHALMERS	Anonymous code / Anonym kod: 435	Points for question (to be filled in by teacher) / Poäng på uppgiften (ifylles av lärare)	Consecutive page no. / Löpande sid nr: 9 — Question no. / Uppgift nr: 6

6.3 6.4 ? Same number as the previous question...

a) There is a significant overall effect because the p-value is around 0,0000007 which is below 0,05.

2p

b) There are significant differences between the groups in terms of mean values.
If we did a Post-Hoc test, we would most likely find out that group A has a significant difference from the other groups. On the plot there is no overlapping between groups B&C and A.

3p

c) The effect size is ~ 60%

2p

d) The effect size shows the explained variance.
It means that ~ 60% of the variance is explained by obsolete requirements (for our research case).
In our research I would say that 60% is a relatively high percentage as there could be quite much noise in the research.

2p

No, by task groups