Grading Guidelines
IN3070/IN4070
Autumn 2024

The maximum number of marks for the whole exam was 100.

The minimum number of marks required for each grade was as follows: For E, min. 30. For D, min. 43. For C, min. 50. For B, min. 66. For A, min. 82.

These boundaries are based on an alignment of the delivered work with the grade definitions given here:

`https://www.uio.no/studier/eksamen/karakterer/`

The final grade for each candidate is based on an evaluation of the delivered work as a whole, and may therefore in some cases deviate from the boundaries given above.

Question 1 – True or False

[1 mark for a correct answer, -1 for incorrect, 0 for missing answer. Min 0 for the whole question]

- Are the following terms unifiable? $f(x, a)$ and $f(g(y, b), y)$

  Answer: False. $\sigma = [x \backslash g(a, b), y \backslash a]$ is a mgu.

- Is validity of formulas in classical first-order logic decidable?

  Answer: False.

- When a set of formulas A is satisfiable, and a subset of some larger set $A \subseteq B$, then B also has to be satisfiable.

  Answer: False. Monotonicity goes the other way.

- A resolution proof has shown unsatisfiability of a clause set when the empty set of clauses has been derived.

  Answer: False. The empty clause has to be derived. An empty set of clauses is in fact satisfied by all interpretations.

- A β rule splits a branch of a derivation into two branches.

  Answer: True

Grading: Automatic grading

Question 2 – Sequent Calculi LK and LJ

Prove the validity of the following formulae using the given calculus. Note that the first two formulas are to be proven in the classical logic LK, while the last is to be proven in the intuitionistic logic LJ.

In this task you can submit a hand-written answer. Use the sketching paper handed to you in the exam room for this. See instructions in the link below the task bar.

A) $(p \to r) \to ((q \to r) \to ((p \vee q) \to r))$ using propositional LK [5 marks]

Answer:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{q \to r, p \;\Rightarrow\; p, r}\,\text{ax} \quad \dfrac{}{r, q \to r, p \;\Rightarrow\; r}\,\text{ax}}{p \to r, q \to r, p \;\Rightarrow\; r}\,{\to}\text{-l} \quad \dfrac{\dfrac{}{p \to r, q \;\Rightarrow\; q, r}\,\text{ax} \quad \dfrac{}{r, p \to r, q \;\Rightarrow\; r}\,\text{ax}}{p \to r, q \to r, q \;\Rightarrow\; r}\,{\to}\text{-l}}{p \to r, q \to r, p \vee q \;\Rightarrow\; r}\,{\vee}\text{-l}}{p \to r, q \to r \;\Rightarrow\; (p \vee q) \to r}\,{\to}\text{-r}}{p \to r \;\Rightarrow\; (q \to r) \to ((p \vee q) \to r)}\,{\to}\text{-r}}{\Rightarrow\; (p \to r) \to ((q \to r) \to ((p \vee q) \to r))}\,{\to}\text{-r}$$

Grading: The order of rule applications can differ. But there have to be at least four branches. 1 mark for the three $\to$-r rules, 1 mark for correct $\vee$-l and $\to$-l, 1 mark one mark for correct rule labels.

B) $(\forall x(p(x) \to p(fx))) \to (\forall x(p(x) \to p(ffx))$ using first-order LK. Here p is a unary predicate. [7 marks]

Answer:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{p(c), \forall \;\Rightarrow\; p(c), p(ffc)}\,\text{ax} \quad \dfrac{\dfrac{\dfrac{}{p(c), p(fc), \forall \;\Rightarrow\; p(fc), p(fffc)}\,\text{ax} \quad \dfrac{}{p(c), p(fc), p(ffc), \forall \;\Rightarrow\; p(ffc)}\,\text{ax}}{p(c), p(fc), p(fc) \to p(fffc), \forall \;\Rightarrow\; p(ffc)}\,{\to}\text{-l}}{p(c), p(fc), \forall x(p(x) \to p(fx)) \;\Rightarrow\; p(ffc)}\,\forall\text{-l},\ [x \backslash f(c)]}{p(c), p(c) \to p(fc), \forall x(p(x) \to p(fx)) \;\Rightarrow\; p(ffc)}\,{\to}\text{-l}}{p(c), \forall x(p(x) \to p(fx)) \;\Rightarrow\; p(ffc)}\,\forall\text{-l},\ [x \backslash c]}{\forall x(p(x) \to p(fx)) \;\Rightarrow\; p(c) \to p(ffc)}\,{\to}\text{-r}}{\forall x(p(x) \to p(fx)) \;\Rightarrow\; \forall x(p(x) \to p(ffx))}\,\forall\text{-r}}{\Rightarrow\; (\forall x(p(x) \to p(fx))) \to (\forall x(p(x) \to p(ffx)))}\,{\to}\text{-r}$$

Grading: 2 marks per correct quantifier rule, 1 for the correct propositional reasoning. Subtract 1 mark for forgetting to keep (or indicate) a copy of the γ formula.

C) $(p \to r) \to (p \to (q \vee r))$ using propositional LJ (the intuitionistic calculus) [8 marks]

Answer:

$$\dfrac{\dfrac{\dfrac{}{p \to r, p \;\Rightarrow\; p}\,\text{ax} \qquad \dfrac{\dfrac{}{r, p \;\Rightarrow\; r}\,\text{ax}}{r, p \;\Rightarrow\; q \vee r}\,\vee\text{-r}_2}{p \to r, p \;\Rightarrow\; q \vee r}\,\to\text{-l}}{p \to r \;\Rightarrow\; p \to (q \vee r)}\,\to\text{-r}}{(p \to r) \to (p \to (q \vee r))}\,\to\text{-r}$$

Grading:

- 8 marks for a correct proof
- minus 2 for forgetting to keep a copy in $\to$-left
- minus 5 for more than one formula in a succedent

Question 3 – Soundness and Completeness

In the context of a sequent calculus, which of these statements are equivalent to soundness of the calculus, or completeness, or neither.

[1 mark for a correct answer, -1 for incorrect, 0 for missing answer. Min 0 for the whole question]

(For each question, one of three options could be chosen, soundness, completeness, neither. The statements were presented to students in a randomized order.)

- Every derivable formula is valid **Answer:** Soundness
- Every satisfiable formula is derivable **Answer:** Neither
- Every formula that has no proof is falsifiable **Answer:** Completeness
- There is a closed derivation for every valid formula **Answer:** Completeness
- A falsifiable formula cannot be derived **Answer:** Soundness

Question 4 – Classical first-order semantics

A) Show that the formula $\forall x(p(x) \leftrightarrow \neg p(f(x)))$ is satisfiable by constructing satisfying interpretations, namely

A1) an interpretation with a finite domain [6 marks], and

A2) an interpretation with an infinite domain (take e.g. the natural numbers) [6 marks]

Answer: If the formula is satisfied in an interpretation $\mathcal{I} = (D, \iota)$ then f^ι has to take every domain element $d \in p^\iota$ to one with $f^\iota(d) \notin p^\iota$ and vice versa.

For A1, we can build a finite interpretation with two domain elements $D = \{a, b\}$, $p^i = \{a\}$, $f^\iota(a) = b$ and $f^\iota(b) = a$. For any variable assignment α, either $\alpha(x) \in p^\iota$ or $f^\iota(\alpha(x)) \in p^\iota$, but not both, so $p(x) \leftrightarrow \neg p(f(x))$ always holds, and therefore also the universally quantified formula.

For A2, we can take the natural numbers as domain, $D = \mathbb{N}$. One way of making sure that f^ι will take every domain element in p^ι to one outside, and vice versa, is to take $f^\iota(x) := x + 1$, the successor function, and to interpret $p^\iota = \{2n \mid n \in \mathbb{N}\}$, the set of even numbers. The successor of every even number is odd, and the successor of every odd number is even.

So again, for any variable assignment α, either $\alpha(x) \in p^\iota$ or $f^\iota(\alpha(x)) = \alpha(x) + 1 \in p^\iota$, but not both, so $p(x) \leftrightarrow \neg p(f(x))$ always holds, and therefore also the universally quantified formula.

Grading: For each part:

- 2 marks for intuition
- 3 marks for correctly described interpretations
- 1 mark for explanation using model semantics.

B) Show that the same formula is not valid by giving a falsifying interpretation. [6 marks] An additional [2 marks] if the domain of your interpretation has only one element.

In this task you can submit a hand-written answer. Use the sketching paper handed to you in the exam room for this. See instructions in the link below the task bar.

Answer: You could start to build a sequent proof. This will actually have many open branches. The Herbrand universe of these branches will be $H = \{o, f(o), f(f(o)), f(f(f(o), \ldots)\}$ starting from some dummy constant o.

One possible interpretation (from one of the open branches) would be (H, ι) with $f\iota(x) = f(x)$ and $p^\iota = \emptyset$, i.e. p is not true for any domain element. For any α, e.g. with $\alpha(x) = o$, neither $p(x)$ nor $p(f(x))$ are true in this interpretation, and therefore the universally quantified formula is falsified.

Following the hint with an interpretation with only element will not only give more points, it is also less work. But the interpretation cannot be read off from an open branch.

With only one domain element, $D = \{a\}$, there is only one possible interpretation for f, namely $f^\iota(a) = a$. There are only two such interpretations, one where $p^\iota = D$, the other where $p^\iota = \emptyset$. In both cases, $p(x) \leftrightarrow \neg p(f(x))$ is not going to hold for the one domain element.
Grading: Like for part A. The additional 2 points are only given if the answer really contains a sensible interpretation.

Question 5 – Deontic Logic

We call a binary relation $R \subseteq W \times W$ serial if for all $u \in W$ there is at least one $v \in W$ with uRv. In other words, every world has an R-successor, a world that is reachable from it.

A Kripke frame (W, R) is called serial if R is a serial relation on W.

A) Show that the formula $\Box p \to \Diamond p$ is satisfied in every serial Kripke Frame, i.e. for every $w \in W$, $w \Vdash \Box p \to \Diamond p$ [8 marks]
Answer: Let (W, R) be a serial Kripke frame, and $w \in W$ an abritrary but fixed world. Since the frame is serial, w has an R-successor; let u be one such successor. To show that $w \Vdash \Box p \to \Diamond p$, we have to show that *if* $w \Vdash \Box p$, *then* $w \Vdash \Box p \to \Diamond p$. So assume $w \Vdash \Box p$. Then $v \Vdash p$ for all R-successors of w. Since u is an R-successor of w, we have $u \Vdash p$. This again implies that $w \Vdash \Diamond p$.

B) Suppose that (W, R) is a Kripke frame, and $\Box p \to \Diamond p$ is forced by every world of every Kripke structure $((W, R), \{I(w)\}_{w \in W}\}$ with that frame. Show that the frame is serial.

Hint. You will need to show that an arbitrary world w has an R-successor. Use an interpretation where p is true exactly in the R-successors of w, and investigate what $\Box p \to \Diamond p$ then tells you about the world w.

[12 marks]
Answer: To show that the frame is serial, we need to show that all worlds w have a successor. Let w be an arbitratry but fixed world. we know that $w \Vdash \Box p \to \Diamond p$ in any Kripke structure based on (W, R). In particular, we choose the specific structure that makes p true in exactly the worlds reachable from w.

Since p is true in *all* v reachable from w, we have that $w \Vdash \Box p$ in that structure. But $w \Vdash \Box p \to \Diamond p$, so we can conclude that $w \Vdash \Diamond p$, i.e. there is at least one successor of w that makes p true. The existence of this successor for every w proves that the frame is serial.

Alternative proof: The interpretation in the hint is not really needed. Let w be an arbitratry but fixed world of some Kripke interpretation. We know that $w \Vdash \Box p \to \Diamond p$. This means that either $w \nVdash \Box p$ or $w \Vdash \Diamond p$. In the first case, there has to be a successor world v of w with $v \nVdash p$. In the other case, there has to be a successor world v of w with $v \Vdash p$. So in any case, w has a successor, and therefore the frame is serial.

Question 6 – Resolution

Show that the following formula is not valid, by using the resolution calculus.

$$(\forall x \exists y.p(x,y)) \rightarrow (\exists y \forall x.p(x,y))$$

Remember that resolution is a refutation calculus, i.e. you can derive that a set of clauses is unsatisfiable. Also remember that variables should be made disjoint before applying resolution.

Transform the problem to clause form, then explain why a resolution proof is not possible from the clauses you get, and finally why this means that the formula is not valid.

- Arriving at a correct set of clauses: 5 marks
- Correct reasoning about invalidity: 5 marks

Answer:

Negation:
$\neg((\forall x \exists y.p(x,y)) \rightarrow (\exists y \forall x.p(x,y)))$

NNF:
$(\forall x \exists y.p(x,y)) \wedge (\forall y \exists x.\neg p(x,y)))$

Rename variables:
$(\forall x \exists y.p(x,y)) \wedge (\forall v \exists u.\neg p(u,v)))$

Prenex (pull out x, y, v, u in that order):
$\forall x \exists y \forall v \exists u(p(x,y) \wedge \neg p(u,v))$

Skolemisation:
$\forall x \forall v(p(x,f(x)) \wedge \neg p(g(x,v),v))$

Clauses (renaming x to z in the second clause to make varibles disjoint between clauses):
(1) $p(x,f(x))$
(2) $\neg p(g(z,v),v)$

No resolution step is possible between these two unit clauses, since one is not unifiable with the complement of the other. To unify the left subterms, we need to apply the substitution $[x \backslash g(z,v)]$, so it remains to unify $p(g(x,v),f(g(z,v))$ with $p(g(z,v),v)$. But unification of the second subterms fails because of the occurs check, v occurs in $f(g(z,v))$.

It is therefore not possible to derive the empty clause from this clause set. The completeness theorem of the resolution calculus tells us that the clause set cannot be unsatisfiable, and there for the original formula is not valid.

Grading:

- 1 mark for each ingredient of the clause transformation (negation, NNF, renaming, prenex, skolemisation)
- 1 marks for confirming that no resolution step is possible
- 2 marks for stating that the occurs check is the problem
- 2 marks for correctly invoking the completeness theorem to arrive at the conclusion.

Question 7 – Propositional Substitution Lemma

We will explore a propositional logic variant of the substitution lemma. Instead of modifications to variable assignments, we will define modifications to propositional interpretations, defined as follows:

For an interpretation $\mathcal{I}$ and an atomic formula p, we define the modified interpretation $\mathcal{I}\{p\}$ resp. $\mathcal{I}\{\neg p\}$ to be identical to $\mathcal{I}$, except that p is interpreted as true in $\mathcal{I}\{p\}$ and false in $\mathcal{I}\{\neg p\}$. Formally, for two atomic formulas p and q,

$$\mathcal{I}\{p\}(q) := \begin{cases} T & \text{if } q = p \\ \mathcal{I}(q) & otherwise \end{cases}$$

and

$$\mathcal{I}\{\neg p\}(q) := \begin{cases} F & \text{if } q = p \\ \mathcal{I}(q) & otherwise \end{cases}$$

We can also define a substitution operator on formulas, that is very similar to the simplification operator defined in the DPLL lecture, but without the rewriting steps to reduce the formula. We want to define $A[p]$ as the result of replacing all occurrences of p in A by the formula $true$, and $A[\neg p]$ as the result of replacing all occurrences of p in A by the formula $false$.

A) Give a recursive definition of $A[p]$ (this will be recursive on the structure of A). You don't need to define the negative case, which is very similar. [10 marks]

B) prove the positive case of the substitution lemma: for all propositional formulae A, propositional atoms p, and interpretations $\mathcal{I}$,

$$\mathcal{I}\{p\} \models A \quad \text{iff} \quad \mathcal{I} \models A[p]$$

Answer: A) We recursively define:

- $p[p] := true$
- $q[p] := q$ for any atomic formula $q \neq p$
- $(\neg A)[p] := \neg(A[p])$

- $(A \wedge B)[p] := (A[p]) \wedge (B[p])$
- $(A \vee B)[p] := (A[p]) \vee (B[p])$
- $(A \rightarrow B)[p] := (A[p]) \rightarrow (B[p])$

B) We prove the statement by induction over the structure of A.
The base case is that A is an atomic formula. There are two sub-cases:

- $A = p$. Then $A[p] = p[p] = \mathit{true}$ according to the definition of $A[p]$ in part A), and therefore $\mathcal{I} \models p[p]$. Also $\mathcal{I}\{p\} \models p$ due to the definition of $\mathcal{I}\{p\}$.
- $A = q \neq p$. Then $A[p] = q[p] = q$ according to the definition of $A[p]$ in part A). On the other hand $\mathcal{I}\{p\} \models q$ iff $\mathcal{I} \models q$ because of the definition of $\mathcal{I}[p]$. So together, $\mathcal{I}\{p\} \models q$ iff $\mathcal{I} \models q$ iff $\mathcal{I} \models q[p]$.

For the induction step, the induction hypothesis is that the statement holds for two smaller formulas A and B, and we have to show it for a compound formulas.

- $\neg A$: $\mathcal{I}\{p\} \models \neg A$ iff $\mathcal{I}\{p\} \not\models A$ (by semantics of negation) iff $\mathcal{I} \not\models A[p]$ (by the induction hypothesis) iff $\mathcal{I} \models \neg(A[p])$ (by semantics of negation) iff $\mathcal{I} \models (\neg A)[p]$ (by the definition of $(\neg A)[p]$)
- $A \wedge B$: $\mathcal{I}\{p\} \models A \wedge B$ iff $\mathcal{I}\{p\} \models A$ and $\mathcal{I}\{p\} \models B$ (by semantics of conjunction) iff $\mathcal{I} \models A[p]$ and $\mathcal{I} \models B[p]$ (by the induction hypothesis for both A and B) iff $\mathcal{I} \models (A[p]) \wedge (B[p])$ (by semantics of conjunction) iff $\mathcal{I} \models (A \wedge B)[p]$ (by the definition of $(A \wedge B)[p]$)

The cases for $\vee$ and $\rightarrow$ are very similar.