COMPUTATIONAL LOGIC: Mock Final (sample 2017-18)

IMPORTANT NOTE: This exam consists of 3 questions, *worth in total 100 points*. As the mid-term, the exam is **open-book**.

Question 1. (*10 points*) Prove (via a *counterexample*) that

$$\Box_a \phi \not\Rightarrow B_a^{\psi} \phi$$

(where $\Box$ is safe belief).

HINT: You need to give an example of two specific sentences ϕ, ψ (e.g. atomic formulas p, q), a specific model $\mathbf{M}$ and a world w in that model, such that w satisfies both $\Box_a \phi$ and $\neg B_a^{\psi} \phi$.

Question 2. (*40 points*) There are two agents, Alexandru and Bob. A coin is on the table. **It is common knowledge that**: (a) *none of the agents knows if the other one saw* the upper face of the coin or not; (b) *if an agent didn't see the face of the coin, then he considers to be equally plausible* that the coin lies Heads up as that the coin lies Tails up; (c) *none of them believes that the other saw the face, nor he believes that the other one didn't see the face* (-in other words: each of them considers that it is *equally plausible* that the other one saw the face as that he didn't see it); (d) obviously, each of the two agents *knows whether he himself saw* the face of the coin.

1. (*10 points*) Represent all the agents' beliefs, knowledge and conditional beliefs, using a **plausibility model $\mathbf{M}_0$** with two agents (a and b).

   HINT: It's a cube!

2. (*5 points*) Now Alexandru announces: "*I* ***know*** *the coin lies Heads up*". It is **common knowledge that**: (a) Bob *strongly trusts* Alexandru (so that Bob performs a **radical upgrade** $\Uparrow$ with this information); (b) of course, Alexandru **knows** if he's lying or not (so he **doesn't change his beliefs** in any way after this announcement).

   Represent this action as an event plausibility model $\mathbf{\Sigma}_0$.

   HINT: There are two possible actions, one in which Alexandru tells the truth, the other in which Alexandru lies. Alexandru can distinguish between the two.

3. (*10 points*) **Represent (draw) a state plausibility model $\mathbf{M}_1$** for the situation *after* the action described in the previous part, by computing the Action-Priority update $\mathbf{M}_1 = \mathbf{M}_0 \otimes \mathbf{\Sigma}_0$.

   HINT: It's still a cube!

4. (*5 points*) **After** the previous action, Bob announces "*I don't know the face of the coin*". **It is common knowledge that**: (a) Alexandru *strongly distrusts* Bob (so that he performs a **negative radical upgrade** $\Uparrow^-$ with this information); (b) while of course Bob *doesn't change his beliefs* after this announcement (since he knows if he's lying or not).

   Represent this action as an event plausibility model $\mathbf{\Sigma}_1$.

5. (*10 points*) *Represent (draw) a model* $\mathbf{M}_2$ for the situation *after* this *second* action (as described in the previous part), by computing the Action-Priority update $\mathbf{M}_2 = \mathbf{M}_1 \otimes \mathbf{\Sigma}_1$.

   HINT: It's always a cube!

Question 3 (*50 points*): Consider the following scenario. There are three agents, Albert, Bernard and Cheryl. **The following facts are all common knowledge**: Cheryl has two secret "bits" x and y (i.e. numbers $x, y \in \{0, 1\}$), that are *known only to her* (Cheryl) but not to the other two; both Albert and Bernard *strongly believe* that $x = 1$; and they both also *believe* that $y = 1$ (though this is *not* a strong belief).

In reality (unknown to Alice and Bernard), their (strong) belief that $x = 1$ is *false*; but their belief that $y = 1$ is *true*.

It follows from all of the above that both Alice and Bernard *believe* that **at least one of the two numbers is equal to** 1. And in reality, **this last belief is safe**.

1. (*9 points*) *Draw a multi-agent plausibility model* $\mathbf{M}_0$ for the above situation (with three agents, a for Albert, b for Bernard and c for Cheryl).

2. (*5 points*) *What is the real world*? In other words: **what are the two numbers**?

3. (*9 points*) Now Cheryl says publicly: "*The sum of the two numbers is* 1". **The following facts are common knowledge**: (a) Albert and Bernard have *opposite attitudes* towards Cheryl, namely one of them *strongly trusts* her ($\Uparrow$) and the other one *strongly distrusts* her $\Uparrow^-$; (b) each of them *knows his own attitude* towards Cheryl, and (c) Cheryl *doesn't know* which one of the two trusts her and which doesn't, but considers the two possibilities to be *equally plausible.*

   Draw an **event plausibility model $\mathbf{\Sigma}$** for the above action.

4. (*9 points*) *Represent (draw) a model* $\mathbf{M}_1$ for the situation *after* the action described in the previous part, by computing the Action-Priority update $\mathbf{M}_1 = \mathbf{M}_0 \otimes \mathbf{\Sigma}$.

5. (*9 points*) We now change the previous scenario a bit. Cheryl makes exactly the same public announcement as above ("*The sum of the two numbers is* 1"). But now **it is common knowledge** that: (a) Albert and Bernard have the **same attitude** towards Cheryl, namely either they **both strongly trust** her ($\Uparrow$), or they are **both neutral** (*id*, neither trusting nor distrusting her), or they **both strongly distrust** her ($\Uparrow^-$); (b) each of them *knows his own attitude* towards Cheryl; (c) although Cheryl *doesn't know* the others' attitude, she *believes* that they **strongly trust** her; while IF this is **not** the case, then (**conditional on this information**) she *believes* that they are **neutral**.

   Represent this action as an **event plausibility model $\mathbf{\Sigma}'$**.

   HINT: This is an event model with 6 possible actions.

6. (*9 points*) Starting again from the original situation (in part 1), suppose the action that we described in the previous part happens.

   Represent (draw) a model $\mathbf{M}'_1$ for the situation *after* this action, by computing the Action-Priority update $\mathbf{M}'_1 = \mathbf{M}_0 \otimes \mathbf{\Sigma}'$.