Sample Exam Question 1

December 14, 2017

1 Part 1

Prove *semantically* that: if an agent believes ϕ then she strongly believes that this belief is safe.

$$B_a\phi \Rightarrow Sb_a\Box_a\phi$$

Proof: Let $\mathbf{S} = (S, \leq_a, \sim_a, ||\cdot||)$ be an arbitrary plausibility model and $s \in S$ and arbitrary world. Assume $s \models B_a\phi$. We need to show that $s \models Sb_a\Box_a\phi$. By the semantics of belief, $best_a s(a) \subseteq ||\phi||$. We can assume that $s(a)$ is non-empty since at least $s \in s(a)$. We prove:

a) that there exist some $\Box_a\phi$-worlds, and,

b) that every $\Box_a\phi$-worlds are strictly more plausible than every $\neg\Box_a\phi$-world.

For (a) we know that $s(a) \neq \emptyset$. Therefore, $best_a s(a) \neq \emptyset$ by converse well-foundedness. Take $w \in best_a s(a)$. Then for every t such that $w \leq_a t$, $t \in best_a s(a) \subseteq ||\phi||$. Hence $t \models \phi$ and so by the semantics of safe belief, $w \models \Box_a$. This shows (a) that there exist some $\Box_a\phi$-worlds.

For (b), take any $v \in s(a)$ such that

$$v \models \Box_a\phi \tag{1}$$

and suppose that $u \models \neg\Box_a\phi$ is such that $v \leq_a u$. By the semantics of negation and safe belief, there exists a $u' \geq_a u$ such that:

$$u' \models \neg\phi \tag{2}$$

But $v \leq_a u \leq_a u'$ and so $v \leq_a u'$ by transitivity. Hence, by the semantics of $\Box_a$ and (1), we obtain that $u' \models \phi$. But this contradicts (2). Therefore $u <_a v$, and since u and v were arbitrary worlds in $s(a)$, every $\Box_a\phi$-world is strictly more plausible than every $\neg\Box_a\phi$-world. This shows (b).

Putting (a) and (b) together, we have shown that $s \models Sb_a\Box_a\phi$, as required.

2 Part 2

To prove

$$B_a\phi \not\Rightarrow \Box_a Sb_a\phi$$

We need to give a model $\mathbf{M}$ and a world w in that model in which w satisfies $B_a\phi$ and $\neg\Box_a Sb_a\phi$. Using the hint given in the question, we need that w satisfies $B_a\phi$ and $\neg Sb_a\phi$.

In the following model, we set $\phi = p$. Either of worlds s or w satisfy $B_a p$ (since $best_a s(a) = best_a w(a) = \{w\} \subseteq ||p||$. However, this is not a strong belief, since there are non-p worlds above p-worlds

$$\begin{tikzcd}
p \arrow[r] & \neg p \arrow[r] & p \\
s & t & w
\end{tikzcd}$$

Q2.

(1). $$(\neg f, \neg s) \to \overset{*}{(\neg f, s)} \to (f, \neg s) \to (f, s)$$

Real world is $(\neg f, s)$.

(2). $\Uparrow (\neg f \vee \neg s)$ gives us

$$(f, s) \to (\neg f, \neg s) \to \overset{*}{(\neg f, s)} \to (f, \neg s)$$

(3). (a) $B \neg s$ (believes service is bad)

(b). NOT safe (since in fact this belief is FALSE: in the real world $(\neg f, s)$, the service is good!)

(4). (a) We apply $\uparrow s$ to the model in part (2), obtaining

$$\begin{tikzcd} f, s \arrow[r] & \neg f, \neg s \arrow[r] & f, \neg s \arrow[r] & \neg f, s^{*} \end{tikzcd}$$

(b) Helen believes food is bad ($B\neg f$)

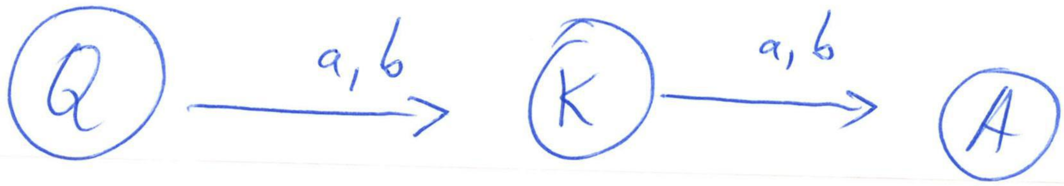
(5). (a) We apply $!(\neg f)$ to the model in part (4), obtaining

$$\begin{tikzcd} \neg f, \neg s \arrow[r] & \neg f, s^{*} \end{tikzcd}$$

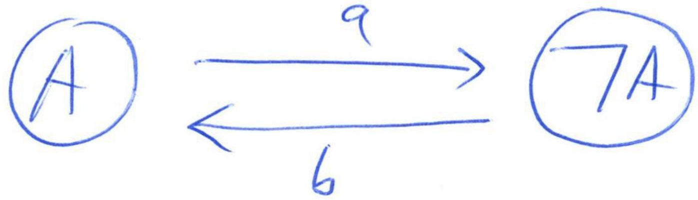
(b). NO: she believes the real world!

Q3.

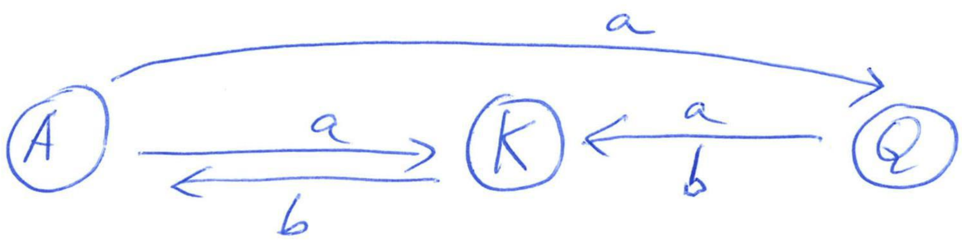
(1).



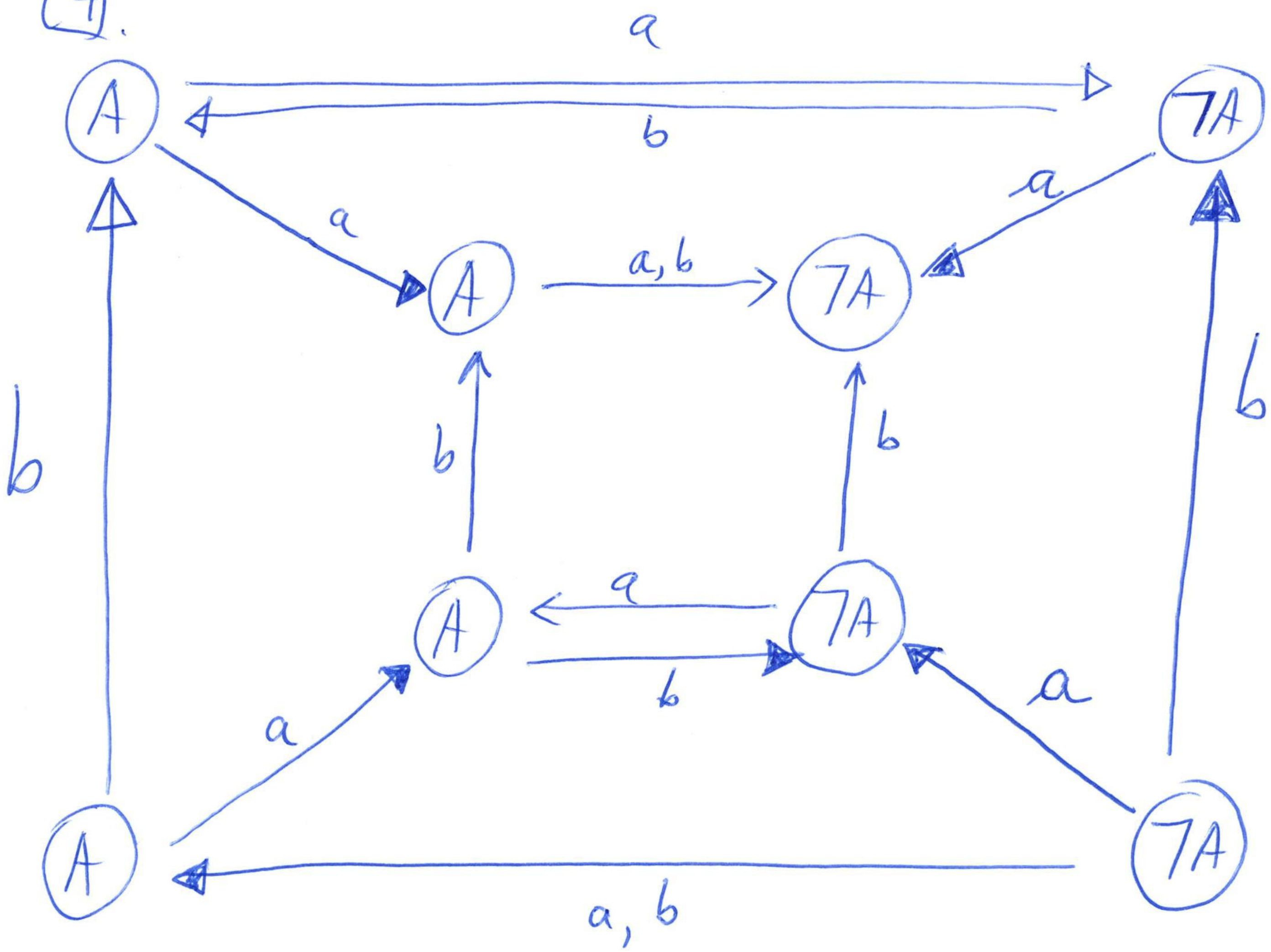
(2).



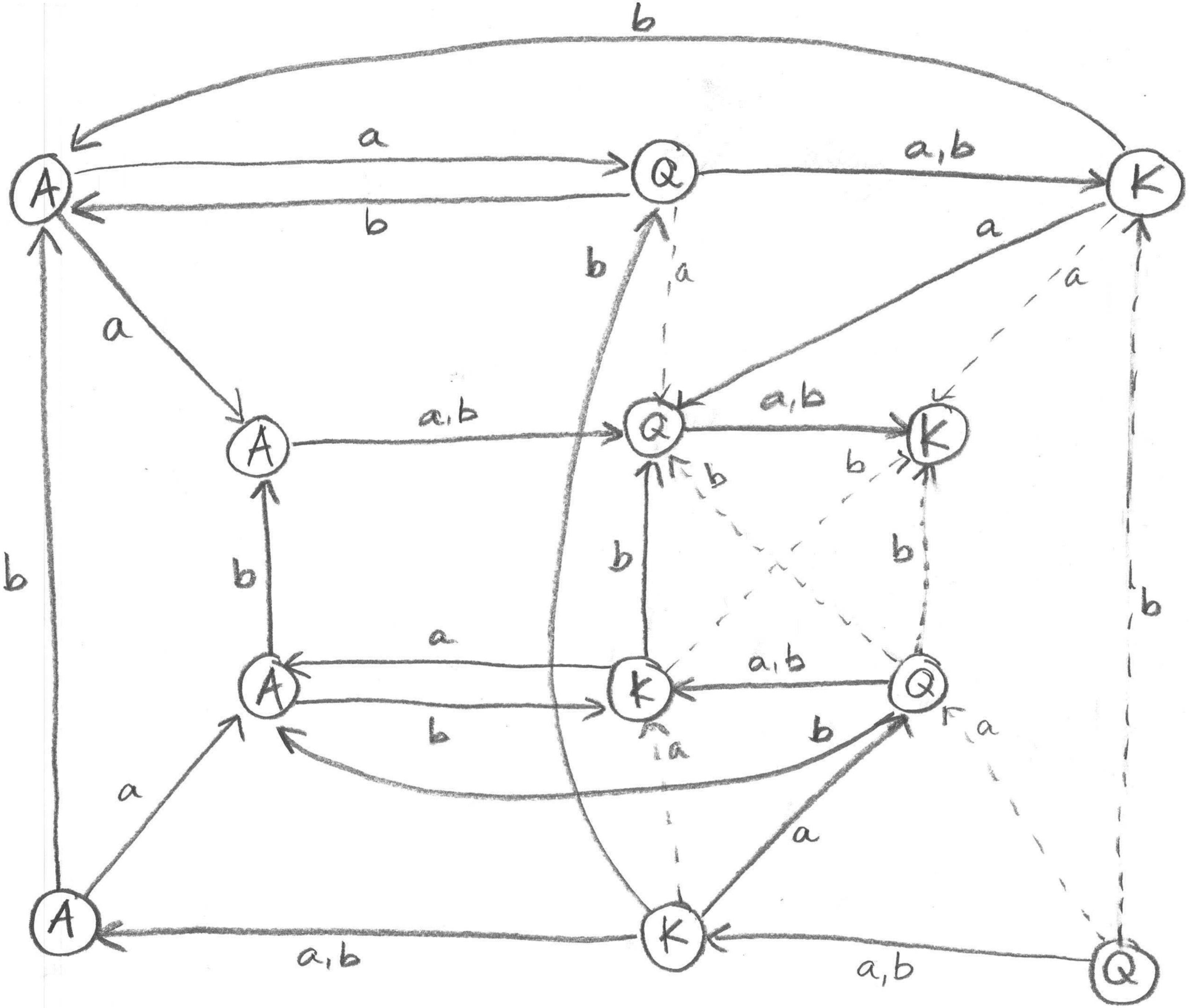
(3).



(4).



3.5



Dotted lines obtainable by transitivity. (Some not drawn).