Modal Logic & Kripke Semantics

Modal logic allows one to model possible truths – reasoning what is possible and what is simply not possible when we don't have complete knowledge. For example: it is possible that I don't have a car (maybe someone just stole my car 5 minutes ago and I don't know it yet). On the other hand, it is not possible that I am my own father.

1 Some Definitions

To begin setting up this idea we need to talk about *frames*. First, let $\mathcal{L}$ be a logical language (like that of predicate calculus). Let W be a set of **possible worlds** and let $S \subseteq W \times W$ a relation on W. S is the **successor relation**. For example, in temporal logic pSq (world p is succeeded by world q) if and only if q comes after p in time. Let $S(p) = \{q \in W \mid pSq\}$. This is the set of all worlds which succeed world p. A **frame** is a collection W, $S \subseteq W \times W$, $\mathcal{C} : W \to \{\text{interpretations of } \mathcal{L}\}$. So $\mathcal{C}$ is a function which takes a possible world as an input and outputs an interpretation (model) of our logical language $\mathcal{L}$.

Next, suppose that φ is a sentence in our language $\mathcal{L}$. We say a model M **forces** φ if φ is true when interpreted in the model M. If M forces φ, we write $M \Vdash \varphi$. Likewise, if Σ is a set of sentences and M forces each one, we say M forces Σ and write $M \Vdash \Sigma$.

Let $\mathcal{C}$ be a frame. We write $p \Vdash \varphi$ for $\mathcal{C}(p) \Vdash \varphi$ (φ is true in the model $\mathcal{C}(p)$). We say that a frame, $\mathcal{C}$, forces φ if for each world $p \in W$, $p \Vdash \varphi$. Finally, a sentence φ is a **logical consequence** of a set of sentences Σ if φ is forced by every frame which forces Σ.

In modal logic we have two new operators, $\Box$ and $\Diamond$. One way to interpret them is $\Box\varphi$ means φ is always true (true in all worlds). $\Diamond\varphi$ means that φ is possibly true (not false in all worlds), so $\Diamond = \neg\Box\neg$. In other words, $\Box$ is kind of like $\forall$ and $\Diamond$ is kind of like $\exists$. To be more formal, $p \Vdash \Box\varphi$ if and only if $q \Vdash \varphi$ for all $q \in S(p)$. And $p \Vdash \Diamond\varphi$ if there exists some $q \in S(p)$ such that $q \Vdash \varphi$.

A Proof We say that φ is a **local consequence** of Σ if, for every $\mathcal{L}$-frame $\mathcal{C} = (W, S, \mathcal{C}(p))$, $\forall p \in W[(\forall\psi \in \Sigma)(p \Vdash \psi) \to (p \Vdash \varphi)]$.

(i) Prove that if φ is a local consequence of Σ, then it is a logical consequence of Σ.

(ii) Prove that the converse of (i) fails, i.e., φ may be a logical consequence of Σ without being a local consequence.

(i) Assume that φ is a local consequence of Σ. Let $\mathcal{C} = (W, S, \mathcal{C}(p))$ be an $\mathcal{L}$-frame which forces every $\psi \in \Sigma$. Given a world $p \in W$, we have that $p \Vdash \psi$ for each $\psi \in \Sigma$. Stated another way, we have that $\forall p \in W(\forall\psi \in \Sigma)(p \Vdash \psi)$.

But we have assumed that φ is a local consequence of Σ, so for each $p \in W$ we have that $(\forall\psi \in \Sigma)(p \Vdash \psi) \to (p \Vdash \varphi)]$. Therefore, we conclude that for each $p \in W$ we must have that $p \Vdash \varphi$. Thus by definition, $\Vdash_{\mathcal{C}} \varphi$.

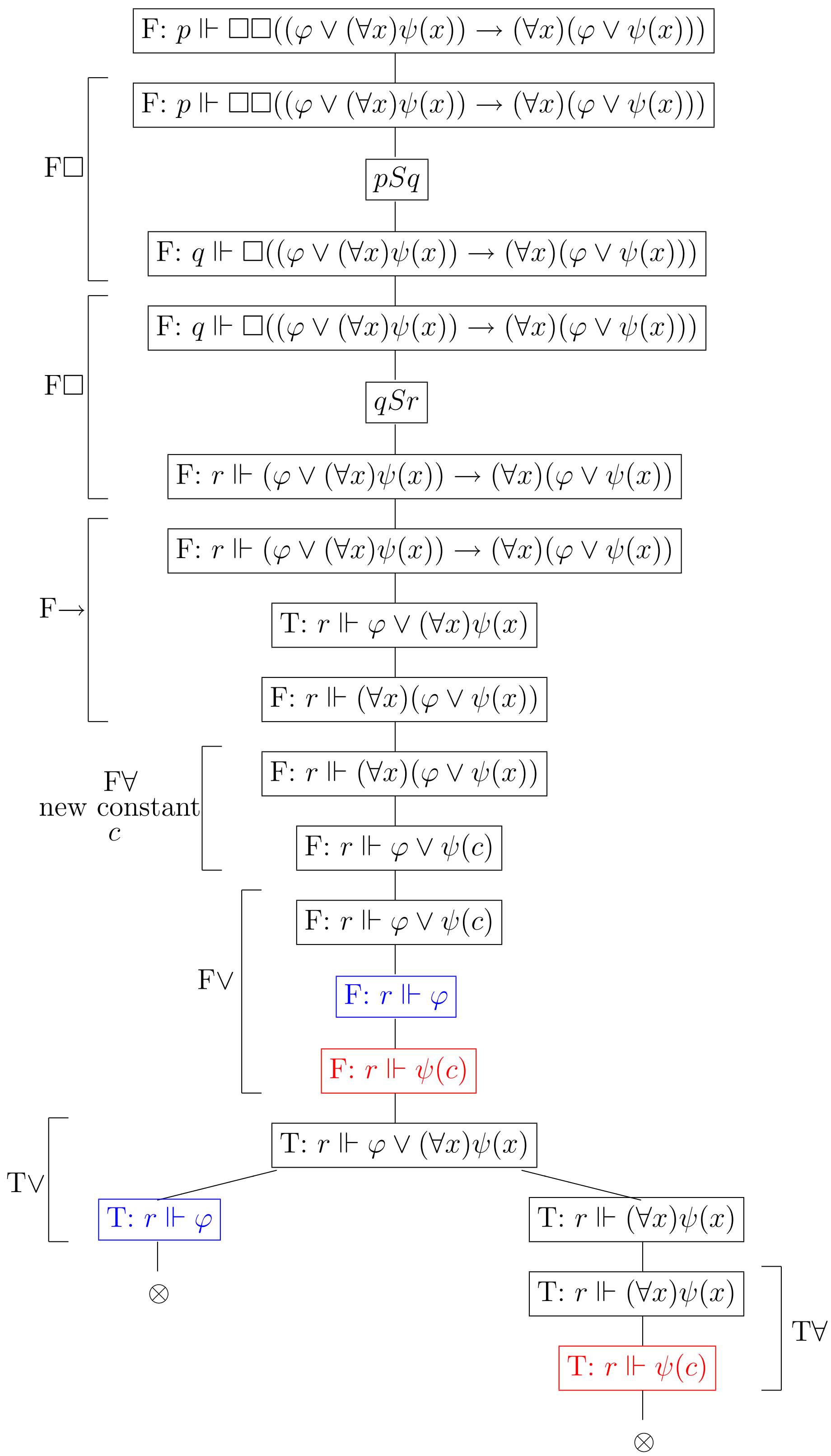
So we have shown that φ is forced by every $\mathcal{L}$-frame which forces every $\psi \in \Sigma$. Thus $\Sigma \vDash \varphi$ (i.e. φ is a logical consequence of Σ).

(ii) Consider the set $\Sigma = \{\varphi\}$ and the sentence $\Box\varphi$. Let $\mathcal{C} = (W, S, \mathcal{C}(p))$ be an $\mathcal{L}$-frame such that $\Vdash_{\mathcal{C}} \varphi$. This implies that for every world $p \in W$ we have that $p \Vdash \varphi$. Let $p \in W$ and $q \in W$ such that pSq. We know that $q \Vdash \varphi$ holds since the φ is forced at every world. Therefore, $p \Vdash \Box\varphi$. Since p was arbitrary, we have that $\Vdash_{\mathcal{C}} \Box\varphi$. Thus $\Sigma \vDash \varphi$.

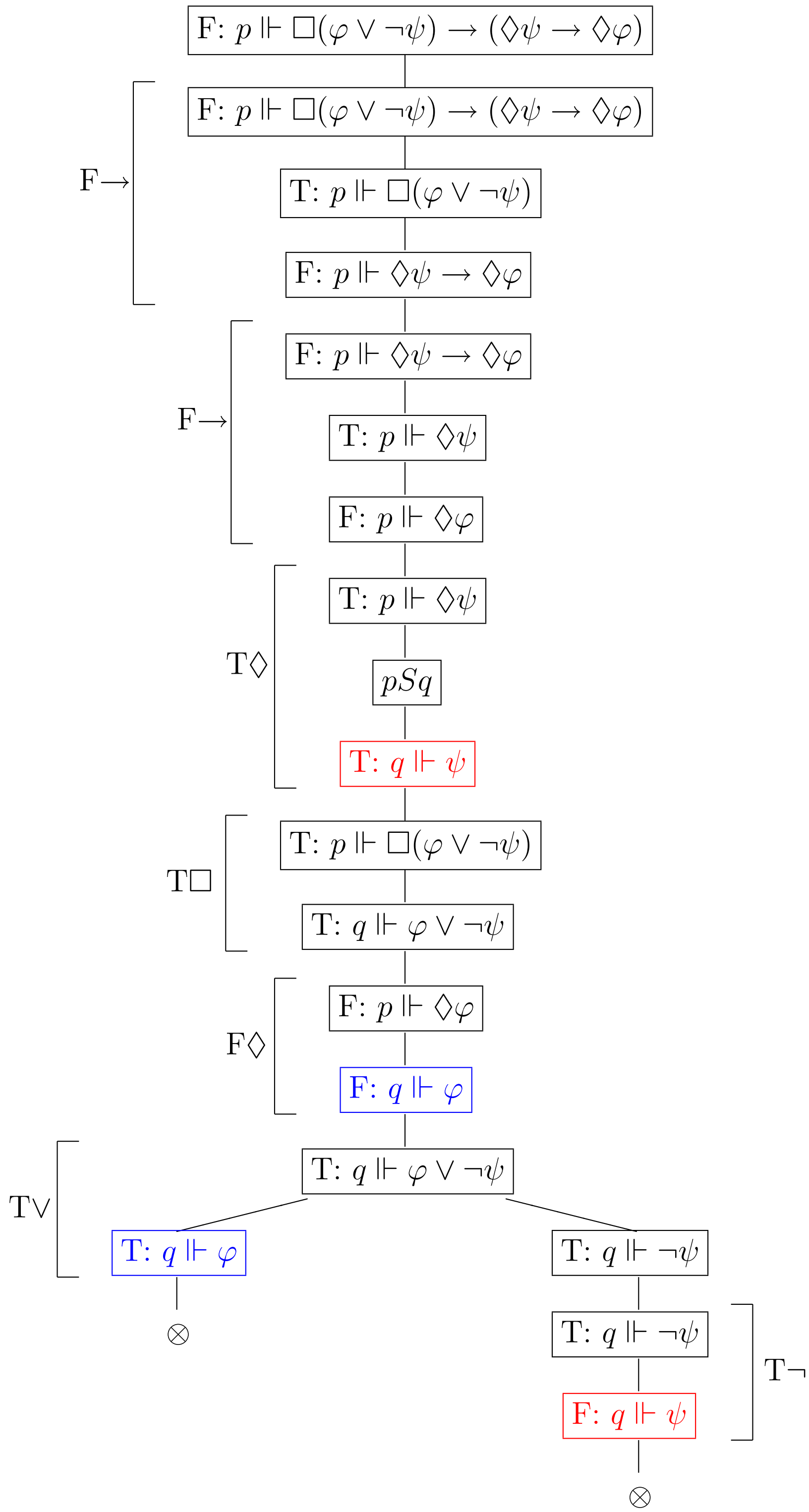
Now consider a frame $\mathcal{C} = (W, S, \mathcal{C}(p))$ where $W = \{p, q\}$, $S = \{(p, q)\}$, $p \Vdash \varphi$, and $q \Vdash \neg\varphi$. We have that $p \Vdash \varphi$, but pSq and $q \nVdash \varphi$. Therefore, $p \nVdash \Box\varphi$. Put another way, we have that $p \Vdash \varphi \nrightarrow p \Vdash \Box\varphi$. Thus $\Box\varphi$ is a logical but not a local consequence of Σ.

2 A Couple of Tableau Proofs

Prove: $\Box\Box((\varphi \vee (\forall x)\psi(x)) \to (\forall x)(\varphi \vee \psi(x)))$, x not free in φ.



Prove: $\Box(\varphi \vee \neg\psi) \to (\Diamond\psi \to \Diamond\varphi)$.



3 E-Mail in Modal Logic

Building on our previous example, we wish to construct a set of sentences in a modal first-order logic that characterize the essential operations of an email system, this time using temporal modalities to express the temporal aspects.

Remark: For the following discussion $\mathcal{C} = (T, S, \mathcal{C}(p))$ will denote an $\mathcal{L}$-frame.

3.1 Modalities and S-Relations

I will use two modalities to formalize my email system in modal logic: $\Box$ = Forever and $\Diamond$ = Eventually. Let $t \in T$ (t is a *moment in time*). By $t \in W$ $t \Vdash \text{Forever}(\varphi)$, we that for all $t' \in T$ such that tSt' we have that $t' \Vdash \varphi$. Also, $t \in W$ $t \Vdash \text{Eventually}(\varphi)$, means that there exists some $t' \in T$ such that tSt' and $t' \Vdash \varphi$. Notice that Forever is equivalent to $\neg$Eventually$\neg$.

Roughly in English, Forever(φ) is intended to mean that φ will hold at all times in the future (not necessarily including the present). Eventually(φ) is intended to mean that φ will hold at some future moment in time.

Before going further, we need to describe the S relation between moments in time. To capture the necessary properties let's adapt the previous assignment's the axioms for Time.

- S is transitive. $(\forall t_1, t_2, t_3 \in T)\ t_1St_2 \wedge t_2St_3 \rightarrow t_1St_3$. This says that if time t_2 comes after time t_1 and time t_3 comes after t_2, then certainly t_3 comes after t_1.
- S is anti-symmetric. $(\forall t_1, t_2 \in T)\ t_1St_2 \rightarrow \neg(t_2St_1)$. This says that if time t_2 comes after time t_1, then t_1 cannot possibly come after t_2.
- S is totally ordered. $(\forall t_1, t_2 \in T)\ t_1St_2 \vee t_2St_1$. This says that given two moments in time, one must be an earlier time and the other a later time.

3.2 Adapting Axioms

The first few sections of my axiomization don't involve time explicitly. I will assume that network connections, email account owners, and hosts don't change over time. So let's skip to the section entitled "Messages".

First, we must modify the functions sent and received. Before we defined sent(m) to be the moment in time when the message m was sent and received(m) to be the moment when the message was received. Instead let Sent and Received be unary predicates. For a message m, Sent(m) is true at all moments in time after the message has been sent. Also, Received(m) is true at every moment in time after the message has been received.

Next, we must change the function location. Before, location$(m, t) = h$ meant that m was residing on host h at time t. Instead let's turn Location into a binary predicate, so that Location(m, h) means that message m is on host h. Note the change: location$(m, t) = h$ has been replaced by $t \Vdash \text{Location}(m, h)$. Let's guarantee that Location only holds for messages paired with hosts.

$$\forall m, h\ \text{Location}(m, h) \rightarrow \text{Message}(m) \wedge \text{Host}(h)$$

Now let's modify some of the old axioms.

- $\forall m$ Message$(m) \rightarrow$ Eventually(Sent$(m)) \wedge$ Eventually(Received(m)). All messages are eventually sent and received.

- $\forall m$ Message$(m) \wedge$ Sent$(m) \rightarrow$ Forever(Sent(m)) Once a message has been sent, it retains the property of being sent.

- $\forall m$ Message$(m) \wedge$ Received$(m) \rightarrow$ Forever(Received(m)) Once a message has been received, it retains the property of being received.

- $\forall m$ Message$(m) \wedge \neg$Sent$(m) \rightarrow$ Location$(m$, hostedBy(sender$(m))) \vee$ Eventually(Location$(m$, hostedBy(sender(m))))). Every message that hasn't been sent is either currently on the sender's host or (if it doesn't exist on a host yet) will eventually be on the sender's host.

- $\forall m$ Message$(m) \wedge$ Received$(m) \rightarrow$ Location$(m$, hostedBy(recipient(m)))). Every message that has been delivered resides on the host of the recipient's email account.

- $\forall m$ $(\exists h$ Location$(m, h)) \leftrightarrow$ Sent$(m) \vee$ Location$(m$, hostedBy(sender(m)))). Every message that is sitting on some host is either sitting on the sender's host or it has been sent. Conversely, if a message is sitting on the sender's host or it has been sent, then it resides on some host.

- $\forall m, h_1$ Location$(m, h_1)) \wedge \neg$Received$(m) \rightarrow \exists h_2$ Connection$(h_1, h_2) \wedge$ Eventually(Location(m, h_2))). Every undelivered message sitting on a host will eventually be transferred to a host which is directly connected to the current host.