Fara's Formula and the Supervaluational Thin Red Line*

Alex MALPASS

Received: 08.02.2012
Final version: 27.05.2012

BIBLID [0495-4548 (2013) 28: 77; pp. 267-282]

ABSTRACT: The focus of this paper is an argument presented by Fara (2010), which is against supervaluationism in the context of vagueness. I show how it applies equally to the branching-time (BT) supervaluationism (first presented in Thomason 1970), but not to the closely related 'STRL' semantics of Malpass & Wawer (2012).

Keywords: supervaluationism; branching-time; the Thin Red Line; semantics; truth.

RESUMEN: Este artículo se centra en un argumento presentado por Fara (2010) en contra del supervaluacionismo en el contexto de la vaguedad. Muestro cómo dicho argumento es igualmente aplicable al supervaluacionismo de tiempo ramificado (presentado por primera vez por Thomason 1970), pero no a la semántica 'STRL' de Malpass y Wawer (2012), que está estrechamente relacionada.

Palabras clave: supervaluacionismo; tiempo ramificado; la Delgada Línea Roja; semántica; verdad.

1. Introduction

In her (2010) article, "Scope Confusions and Unsatisfiable Disjuncts: Two Problems for Supervaluationism", Delia Graff Fara presents various problems for 'canonical supervaluationism' (i.e. the supervaluationism of Fine 1975), the first of which I shall concentrate on in this paper. The problem consists in a troublesome formula, that I will call 'Fara's Formula' (FF), which she demonstrates is satisfiable in the supervaluationist semantics. I shall show that her problem affects the BT account of supervaluationism that most closely approaches Fine's canonical version (that of Thomason 1970, 2002). However, my main aim will be to demonstrate that there is a closely related theory in the BT context which *avoids* this problem; the Supervaluational Thin Red Line (STRL) semantics of Malpass & Wawer (2012).

2. FF: true disjunctions with impossible disjuncts

Delia Fara presents supervaluationism with a new complaint in her (2010) paper. Fara's complaint is a development of a well known difficulty with the theory. Put simply, the supervaluationist allows for (super-)true disjunctions with (super-)truth-valueless disjuncts. The idea is that it is unacceptable that, for example, it is (super-)true that 'either Juan or Carlos is the shortest person in the room', even though there is no (super-true) answer to 'which one is it?' (see Fara 2010, 376). In general the problem is as follows:

* I would like to give thanks for their comments and contributions to Delia Graff Fara, Peter Øhrstrøm, Jacek Wawer, audiences in both Bristol and Cambridge and two anonymous referees.

(Complaint) Supervaluationism allows some disjunctions to be (super-)true, even though each disjunct fails to be either (super-)true or (super-)false—and this failure is unacceptable.

As Fara articulates, the standard supervaluationist rejoinder to this complaint is that there is no answer to which one it is "because things *could go either way* depending on how you drew precise boundaries" (*ibid*). However, either way it could go, the supervaluationist will continue, it will be one way or the other, which explains why the disjunction itself is super-true.

In general, the response to the objection is as follows:

(Rejoinder) Each disjunct is neither super-true nor super-false because things *could* go either way.

In response to this, Fara offers an objection. She strengthens the original complaint by pointing out that supervaluationism "allows also for true disjunctions where neither disjunct *could* be true" [emphasis mine] (Fara 2010, 376). In general,

(Objection to Rejoinder) Sometimes things could **not** go either way—so *Rejoinder* is **not** sufficient to support the supervaluationist's position on disjunctions.

The aim for Fara therefore is to provide us with a supervaluationally satisfiable formula which features a true disjunction with, not just truth-valueless, but also *unsatisfiable* disjuncts. The argument begins by defining a modal operator, $\Diamond_s$, with the following semantics:

Definition 1: $\Diamond_s$ — Fara's supervaluational satisfiability operator.
$\Diamond_s\varphi$ iff φ is super-true in some supervaluational model. (*ibid*)

Fara envisions '$\Diamond_s$' as the 'supervaluational satisfiability' operator, which as she mentions in a footnote, is not simply a normal modal possibility operator. '$\Diamond_s$' corresponds to a type of meta-level necessity, as it's semantics involve quantifying over *models*, rather than points in a model. The idea is that this makes it suitable to undermine the rejoinder presented above, as the "things *could go either way*" defence (offered above) plausibly rests on this type of modality. The defence is that the model could have been different; that is, the framework for interpreting the language could have had different features. Therefore (the idea is), this type of modality is appropriate.[1]

The content of the objection to the rejoinder is that supervaluationists allow that the following formula ('*Fara's Formula*') is satisfiable:

(FF) $(\varphi \vee \psi) \wedge \neg\Diamond_s\varphi \wedge \neg\Diamond_s\psi$

[1] It might be objected that this notion of satisfiability is not appropriate. A supervaluationist could claim that by 'things could go either way', he means a more usual modal operator, meaning something like: φ is true at some point in this model. This will ultimately not help the supervaluationist though. Fara's operator represents the meta-linguistic notion of 'satisfiability', and even if we grant the supervaluationist a notion of 'possibility' which is immune to the complaint, this will still leave him with the difficulty of explaining how there is a formula which is 'possible' and yet 'not satisfiable'.

The next step is simply to find suitable propositions to substitute in for the variables, φ and ψ. Fara provides us with an example which uses a 'borderline-case' operator, *B*.

Definition 2: *B* — the borderline-case operator.
$B\varphi$ iff φ is neither true nor false (according to supervaluationism) (*ibid*, 377)

The substitution is '$Bp \wedge p$' for φ and '$Bp \wedge \neg p$' for ψ, resulting in:

(FF₀) $((Bp \wedge p) \vee (Bp \wedge \neg p)) \wedge \neg\Diamond_s(Bp \wedge p) \wedge \neg\Diamond_s(Bp \wedge \neg p)$

Thus we have the desired sentence, which the supervaluationist is committed to calling super-true, which consists of a super-true disjunction with impossible (or, 'unsatisfiable') disjuncts. It says:

'[(Either it is borderline that *p*, and *p*) or (it is borderline that *p*, and not-*p*)]

and

[it is not satisfiable that (it is borderline that *p*, and *p*)]

and

[it is not satisfiable that (it is borderline that *p*, and not-*p*)]'

If the formula is satisfiable, then this is a difficult result to bear for the supervaluationist. The idea of true disjunctions with truth-valueless disjuncts is bad enough, but Fara seems to have undermined the traditional first line of defence that supervaluationists fall back on when trying to provide intuitive support for the result. After all, how *can* either disjunct be true when each is unsatisfiable?

3. FF, and BT-supervaluationism

In order to evaluate the merits of this argument against STRL logic and semantics, which will be my final goal, it needs to be introduced rigorously into the precise context in which STRL operates. This is what I shall do in this section. Before we get to that goal however, we have to demonstrate that, when brought into the BT context, Fara's argument works against standard BT-supervaluationism (that of Thomason 1970).

3.1 BT and Ockhamism

A BT structure is defined as follows:

Definition 3: S — a BT structure.
$\mathbf{S} =_{\text{def}} \langle M, < \rangle$

The two elements of the structure are as follows:

(1) M is a non-empty set of moments, $\{m, m', m'' ...\}$
(2) $<$ is a binary ordering relation defined on moments, that is *partial* (asymmetric and transitive), and satisfies the following:
   (i) *backwards linear*: $\forall m1, m2, m3{:}[(m2 < m1 \wedge m3 < m1) \Rightarrow (m2 < m3 \vee m3 < m2)])$
   (ii) *connected*: $\forall m1, m2{:}\ [(m1 \not< m2 \wedge m2 \not< m1 \Rightarrow \exists m3{:}\ (m3 < m1 \wedge m3 < m2)]$

The restrictions on the ordering ensure that each moment has a linear past and a branching future. To make a BT structure into a model, we add a valuation function which maps atomic propositions to truth-values:

Definition 4: $\mathbf{M_O}$ — an (Ockhamist) BT model.

$\mathbf{M_O} =_{def} \langle M, <, V \rangle$

V is a valuation function, defined on atomic propositions ($\mathbb{P}$) and moments, $V: M \times \mathbb{P} \rightarrow \{1, 0\}$.

Histories (sometimes called 'branches'), *h*, are maximal linear subsets (ordered with <) of M. Histories therefore represent entire courses of events (or history). The set of all histories in a model is designated *Hist*. The set of histories through some moment *m* is written *H*(*m*).

The language is a classical propositional calculus, extended to include tense operators, F and P (with G = ¬F¬, and H = ¬P¬), and a modal operator, ◊ (with □=¬◊¬).

The semantics known as 'Ockhamist' in the literature provides the base-semantics, which the supervaluationist will use to define his 'post-semantics'[2]. Here are the semantic clauses for Ockhamism:

Definition 5: Ock-truth

$M_O\ m/h \vDash_{Ock} p$	iff	$V(m, p) = 1$
$M_O\ m/h \vDash_{Ock} \neg\varphi$	iff	$M_O\ m/h \nvDash_{Ock} \varphi$
$M_O\ m/h \vDash_{Ock} \varphi \wedge \psi$	iff	$M_O\ m/h \vDash_{Ock} \varphi$ and $M_O\ m/h \vDash_{Ock} \psi$
$M_O\ m/h \vDash_{Ock} P\varphi$	iff	$\exists m'$: $[m' < m$ and $M_O\ m'/h \vDash_{Ock} \varphi]$
$M_O\ m/h \vDash_{Ock} \Diamond\varphi$	iff	$\exists h'$: $[m \in h'$ and $M_O\ m/h' \vDash_{Ock} \varphi]$
$M_O\ m/h \vDash_{Ock} F\varphi$	iff	$\exists m'$: $[m' \in h$ and $m < m'$ and $M_O\ m'/h \vDash_{Ock} \varphi]$

The Ockhamist semantics is distinguished from other accounts of the semantics of such languages by its use of the 'history parameter' on the left of the turnstile. This says that the truth of a formula is dependent not only on what moment is used as the time of evaluation, but also which history (through that moment) is used as the 'history of evaluation'. So, for example, "$\mathbf{M_O}\ m/h \vDash_{Ock} p$" is intended to mean "in model $\mathbf{M_O}$ at the moment/history pair *m*/*h*, it is (Ockhamist-)true that *p*".

Two clauses mention *h* on the right of the biconditional; ◊ and F. However, the details ensure that the only clause which is sensitive to different histories being used as values of the history parameter is that for F. Indeed, in the first clause (that for atomic formulae), we can see that the history parameter is quite redundant, meaning that if an atomic formula is true on one history through a moment, then it is true on all histories through that moment. The following would count as a logical truth, where *p* ranges over atomic formulas: $p \rightarrow \Box p$. While this may seem strange, there is a very clear interpretation of the modality involved according to which it makes perfect sense. If □ is interpreted as meaning 'It is unpreventable that...', then we can see how for all present tensed propositions (as all atomic formulae are), if it is true now, it is now unpre-

[2] The distinction between semantics and 'post-semantics' is just the difference between 'truth' (Ock-truth) and 'super-truth' (Sup-truth or STRL-truth).

ventable. You just made a cup of tea, and although I could earlier have prevented you from doing it, now that you have done it is no longer preventable. Once something *is* (present-tense), it *is necessarily*. This also holds for past-tensed propositions, and for future-tensed ones provided they are non-contingent. This idea is what many writers call 'historical modality' (see Müller 2012 for further discussion).

Another consequence of these Ockhamist definitions is that certain things start off being contingent, but then as time passes they become 'settled' into being necessary (on all histories) or impossible (true on no histories). This seems to represent the existential and philosophical truth that as time passes certain things which used to be possible become impossible. For example, it used to be possible that I would become a child concert pianist. However, now I am a fully grown adult, this possibility has 'gone cold' for me; it is impossible now that I will become a child concert pianist. It is this interaction between time and modality that first led Arthur Prior to develop the Ockhamist semantics in his (1967) book, and is perhaps its most independently interesting philosophical feature.

Formal results of the definitions are also pleasing. Intuitive validities include: $\varphi \rightarrow \mathrm{PF}\varphi$, $\mathrm{F}\varphi \vee \neg\mathrm{F}\varphi$, $\mathrm{FF}\varphi \rightarrow \mathrm{F}\varphi$ and $\mathrm{F}\Diamond\varphi \rightarrow \Diamond\mathrm{F}\varphi$, etc. Discussions of Ockhamism can be found in Prior (1967), Thomason (1970), Burgess (1979) Øhrstrøm & Hasle (1995) and Belnap et al. (2001). A full axiomatisation of the Ockhamist logic has been given in Reynolds (2003).

Despite these intuitive results, the notion of *truth* (Ockhamist-truth) is not acceptable as it stands. The problem is precisely the dependency on the history of evaluation, and for this reason I call Ock-truth 'history-dependent' truth. I have criticised this notion elsewhere (see Malpass & Wawer 2012, sec. 3.1), but the basic problem is that one needs to specify a value to the history parameter to give a truth-value to a prediction, and yet no way of privileging one history over the others is acceptable to Ockhamists, who characteristically reject the notion of an 'actual future'. So it needs to be done, but no way of doing it is acceptable. This means that we should look for a notion of truth that is not history-dependent.

3.2 BT-Supervaluationism

The Ockhamist semantics was proposed by Prior in his classic text on tense logic, *Past, Present and Future* (1967), and within 3 years criticisms of the history-dependent truth were to be found in the literature (i.e. Thomason 1970). One way to get a notion of truth that is history-independent is to use supervaluationism. The supervaluational (post)semantics is defined by quantifying over the histories, so that a formula is super-true (Sup-true) iff it is Ock-true on values of *h*:

Definition 6: Super-truth (and falsity)

$$\mathbf{M_O}\ m \vDash_{Sup} \varphi \quad =_{def} \quad \forall h\colon [h \in H(m) \quad \Rightarrow \quad \mathbf{M_O}\ m/h \vDash_{Ock} \varphi]$$
$$\mathbf{M_O}\ m \Dashv_{Sup} \varphi \quad =_{def} \quad \forall h\colon [h \in H(m) \quad \Rightarrow \quad \mathbf{M_O}\ m/h \vDash_{Ock} \neg\varphi]$$

This supervaluational semantics, which corresponds to that found in Thomason (1970, 2002), says that all formulae are Sup-true or Sup-false apart from those of the form F*p*, where *p* is a future contingent.

The merits of this theory are worth reciting here:

(1) Firstly, it is a notion of truth that is 'history-*in*dependent', in the sense that there is no history parameter on the left of the biconditional in Definition 6.

(2) Secondly, it retains all the validities of the Ockhamist semantics; i.e. if a formula is Ock-valid, then it is Sup-valid. This means that the supervaluationist can interpret $\Box$ as 'it is now unpreventable that...', and thus he can represent historical modality very well.

(3) The third point is related to the second: it respects the twin ideas that future contingents are i) neither-true-nor-false at the time they are uttered, and that ii) when looking back on predictions of future contingents they have determinate truth values. Thomason's supervaluational theory accommodates both, in the sense that future contingents have no truth-value at their time of utterance, and yet $\varphi \rightarrow \mathrm{PF}\varphi$ is valid (i.e. Sup-true at every moment in every model).

(4) Lastly, Thomason's theory incorporates the idea that future contingents have no truth value, while at the same time as respecting historical modality, i.e. by endorsing results like $p \rightarrow \Box p$. This seems to do remarkably well at formalising Aristotle's comments in *On Interpretation* #9, where he seems to argue for both the thought that future contingents have no truth value at their time of utterance (the sea battle argument), and the $p \rightarrow \Box p$ result ("if it is true to say that a thing is white, it must necessarily be white", *On Interpretation* #9 paragraph 3). Thus, Thomason's theory seems to be the modern incarnation of Aristotle's position on future contingents.[3]

3.3 BT-Supervaluationism and Fara's Formula

Now that we have introduced BT, Ockhamist semantics and Thomason's supervaluationism, let us return to Fara's Formula. Before we can state it properly, we need to introduce the B operator and the $\Diamond_s$ operator.

In the BT setting, the B operator corresponds to a 'future contingent' operator (in this setting the only formulae that are 'neither true nor false according to supervaluationism' are predictions of future contingents). The semantics for the operator should be introduced at the Ock-level:

Definition 7: B in the BT setting.

$\mathbf{M_O}\ m/h \vDash_{Ock} B\,\varphi$ iff $\exists h'$: $[\mathbf{M_O}\ m/h' \vDash_{Ock} \varphi]$ and $\exists h''$: $[\mathbf{M_O}\ m/h'' \vDash_{Ock} \neg\varphi]$
iff $\mathbf{M_O}\ m/h \vDash_{Ock} \Diamond\,\varphi \wedge \Diamond\neg\varphi$

The only proposition we will be able to substitute in for φ in the above condition is Fp, where on one history Fp is true, and on another $\neg$Fp is true, as they are the only formulae that fulfil the condition for formulae on the right.

[3] Aristotle's position on future contingents is one of the most discussed topics throughout the history of logic and philosophy, and there are those who would argue that Thomason's position is not actually very faithful to Aristotle. Nevertheless, his position is very similar to one that Aristotle could be argued to take.

An immediate consequence of this definition of B is that if $B\varphi$ is Ock-true at some moment history pair m/h, then it is Sup-true at that moment; i.e. if $\mathbf{M_O}\ m/h \vDash_{Ock} B\varphi$, then $\mathbf{M_O}\ m \vDash_{Sup} B\varphi$. If it is borderline on one history through a moment, then it is borderline on all histories through that moment.

The final element which we need to add to the picture before we can restate Fara's Formula is her 'satisfiablility operator', $\Diamond_s$. This operator worked by quantifying existentially over "supervaluational models". In this BT context, this means to quantify existentially over different BT-models. This makes the analysis slightly less than standard, as quantifying over models is to quantify into a somewhat unusual domain of reference; the domain is a domain whose elements are domains. This is a consequence of the fact that the satisfiability is usually a meta-linguistic notion, employed in the proof theory for a logic, in which it makes more sense to quantify over models. Therefore, the introduction of satisfiability as an object-language operator is itself somewhat non-standard. The procedure is intuitive enough however, and poses no outright logical difficulties of which I am aware. So, having noted the oddity, I shall proceed to define the operator thusly:

Definition 8: $\Diamond_s$ in the BT setting.

$\mathbf{M_O}\ m/h \vDash_{Ock} \Diamond_s\ \varphi$ iff $\exists \mathbf{M_O}', \exists m' \in \mathbf{M_O}$: $[\mathbf{M_O}'\ m' \vDash_{Sup} \varphi]$

This reads: in the model $\mathbf{M_O}$ at the moment/history pair m/h the formula '$\Diamond_s\ \varphi$' is Ock-true iff there is a model $\mathbf{M_O}$' containing a moment m' such that 'φ' is Sup-true there.

The question now is whether the argument still holds in the BT context, i.e. whether the following formula is supervaluationally true at some moment in some BT model (i.e. if Fara's formula is satisfiable in this context):

$$\textit{(FF}_1\textit{)}\ \mathbf{M_O}\ m \vDash_{Sup} ((B\mathrm{F}p \wedge \mathrm{F}p) \vee (B\mathrm{F}p \wedge \neg \mathrm{F}p)) \wedge \neg\Diamond_s(B\mathrm{F}p \wedge \mathrm{F}p) \wedge \neg\Diamond_s(B\mathrm{F}p \wedge \neg \mathrm{F}p)$$

If the formula is satisfiable, then there will be a moment in a BT model at which it is Sup-true. So we have to find such a moment. In order to make this task more manageable, let's break up each of the tree conjuncts and go through them one by one to see if they can each be true at the same moment.

Firstly, the disjunction: can $(B\mathrm{F}p \wedge \mathrm{F}p) \vee (B\mathrm{F}p \wedge \neg \mathrm{F}p)$ Sup-true at any moment m? And, if so, what sort of moment is it? Well, any moment which makes this disjunction Ock-true on all histories (i.e. Sup-true) must have $B\mathrm{F}p$ true at it, because it features in each disjunct. If $B\mathrm{F}p$ is Ock-true (on all histories), then $\Diamond \mathrm{F}p \wedge \Diamond\neg\ \mathrm{F}p$ will be also be Ock-true (Definition 7). So the relevant moment must be one on which has p true at some moment on one history h, and another history h' which has no moment where p is true; i.e. p must be a future contingent at m. At such a moment, on each history either $\mathrm{F}p$ will be Ock-true (if it is a history like h), or $\neg\mathrm{F}p$ will be Ock-true (if it is a history like h'). Therefore, on each history $B\mathrm{F}p$ will be Ock-true and either $\mathrm{F}p$ or $\neg\mathrm{F}p$ will be Ock-true; therefore (by distribution) 'either $B\mathrm{F}p$ and $\mathrm{F}p$, or $B\mathrm{F}p$ and $\neg\mathrm{F}p$' will be Ock-true will be true at m, on all histories; and thus also Sup-true at m. This establishes that the first conjunct can be true at this type of m. So far, so good for FF in BT.

Secondly, $\neg\Diamond_s(BFp \wedge Fp)$. Can this formula also be super-true at the same moment? Given Definition 8, the effect of negating the $\Diamond_s$ operator is to negate the existential quantification to which the operator corresponds. Thus, the negated formula would be true iff there were *no* moment in any alternative model we could use that would result in the formula on its own (i.e. $BFp \wedge Fp$) being classed as Sup-true. It is fairly easy to see that no moment in any model satisfies this: because B is effectively a contingency operator, for BFp to be Sup-true at some moment, Fp *cannot be* Ock-true on all histories through that moment—but for Fp to be Sup-true, Fp *has to be* true on all histories.

Exactly the same reasoning provides the result that the third disjunct is holds also (after all, they only differ over the negation of Fp).

Therefore, Fara's complaint imports itself perfectly into the BT supervaluational setting. All this reasoning gives can be condensed into the following result:

Fact 1:
FF_1 is BT-Sup-satisfiable.

By showing that Fara's formula is satisfiable in the supervaluational BT setting, I have also shown that this semantics is committed to the truth of disjunctions with unsatisfiable disjuncts. Therefore, Fara's complaint is a serious complaint against supervaluationism in general - and this is no less so for its BT-supervaluational cousin.

4. FF and TRL

The Thin Red Line (TRL) theory (originally proposed by Peter Øhrstrøm in his PhD thesis, which was published in (1981), and was developed in his (1984)) has models that are supplemented to have a designated history which is thought of as the 'actual course of events'. This provides another history-independent notion of truth. The motivation for the TRL was precisely to have a notion of 'moment-truth', and do away with history-dependent truth.

A TRL model is defined as follows:

Definition 9: $\mathbf{M_{TRL}}$ — TRL model
$\mathbf{M_{TRL}} = \langle M, <, V, \mathit{TRL}\rangle$

$\langle M, <, V\rangle$ is a usual BT model, and *TRL* picks out a history as the 'actual history'. A model, $\mathbf{M_{TRL}} = \langle M, <, V, \mathit{TRL}\rangle$, is therefore 'based on' a BT structure $\mathbf{S} = \langle M, <\rangle$ by adding both a V-function and *TRL*.

Øhrstrøm's original paper (1981) gives the following semantic definition of the future tense:

Definition 10: TRL future tense.
$\mathbf{M_{TRL}}\ m \vDash_{TRL} F\varphi$ iff $\exists m'$: $[m' \in \mathit{TRL}$ and $m < m'$ and $\mathbf{M_{TRL}}\ m' \vDash_{TRL} \varphi]$

According to this definition, at m, it will be that φ is TRL-true iff there is a later moment m' in the TRL at which φ is TRL-true. TRL theory differs primarily from supervaluationism by assigning truth-values to predictions of future contingents. A consequence of this is that a distinction is drawn between actual and non-actual predictions:

(i) At all moments on the TRL, some future contingents are true and the rest are false (i.e. if, on the TRL, you say that there will be a sea battle and I say that there won't be, then one of us has spoken a truth and the other a falsity).

(ii) At all moments off the TRL, all future contingents get assigned falsity by default (i.e. if, off the TRL, one of us predicts that there will be a sea battle, and the other denies this, then we will have both spoken a falsity).

The TRL theory assigns a truth-value to all formulae (including future contingents). Technically, this means that the theory is not vulnerable to Fara's formula, i.e. it is not possible to satisfy FF in TRL[4]. This should come as no surprise when we consider that Fara's argument is only supposed to be applied to versions of supervaluationism, with truth-value gaps.

The idea that future contingents at non-TRL moments are all false strikes many as a rather large flaw in the theory (see Belnap et al. 2001, 162-163). However, this point could be debated; for instance, recently Alan Hájek (ms.) has independently argued for the thesis that all contingent 'would'-counterfactuals are false, which is very close to the thought that all non-actual future contingents are false[5]. We will discuss this more later on.

Nevertheless, TRL theory is in more trouble; for all moments off the TRL, the theory does not just classify all future contingents as false, but classifies *all future tensed formulae* as false. So even future tensed tautologies, like 'either there will be a sea battle or there will not be a sea battle' come out as false (see Malpass & Wawer 2012, sec. 6). This result is too much to bear, and indeed goes beyond even Hájek's radical position, and so for this reason (among others) the TRL theory will not do as it stands.

Just because we have abandoned the initial TRL-definition of the future tense (above) does not mean that we should abandon the TRL altogether. I will (in a moment) propose a theory in which the TRL plays an active role. In order to define this properly, it will help to note a fact mentioned in a paper by Øhrstrøm about the TRL theory (adapted very slightly to fit my purpose). In his (2009) paper, *In Defence of the Thin Red Line*, Øhrstrøm describes one potential (though extremely basic) TRL theory, in which the TRL "does not play any active role in the semantics" (Øhrstrøm 2009, 13). The idea is simply to use the Ockhamist semantic clauses (which do not mention a TRL), but using TRL-models as interpretations of the language. The clauses look as follows (where "$\mathbf{M}_{\mathbf{O}}$" has been replaced by "$\mathbf{M}_{\mathbf{TRL}}$"):

---

[4] FF is obviously not satisfiable at any moment *off* the TRL, as it contains a future contingent in the first conjunct, and these are all false at such moments on the TRL theory. It is not satisfiable at any moments on the TRL either. This is because at these moments, one of the disjuncts in the first conjunct will be true, which in turn falsifies either the second or the third conjuncts (the counter-model is the model of evaluation itself). So TRL is safe from FF.

[5] His manuscript is entitled '*Most Counterfactuals are False*', and is currently unpublished.

Definition 11: Ock-truth in TRL models

$\mathbf{M}_{\mathbf{TRL}}\ m/h \vDash_{Ock} p$	iff	$V(m, p) = 1$
$\mathbf{M}_{\mathbf{TRL}}\ m/h \vDash_{Ock} \neg\varphi$	iff	$\mathbf{M}_{\mathbf{TRL}}\ m/h \nvDash_{Ock} \varphi$
$\mathbf{M}_{\mathbf{TRL}}\ m/h \vDash_{Ock} \varphi \wedge \psi$	iff	$\mathbf{M}_{\mathbf{TRL}}\ m/h \vDash_{Ock} \varphi$ and $\mathbf{M}_{\mathbf{TRL}}\ m/h \vDash_{Ock} \psi$
$\mathbf{M}_{\mathbf{TRL}}\ m/h \vDash_{Ock} \mathrm{P}\varphi$	iff	$\exists m'$: $[m' < m$ and $\mathbf{M}_{\mathbf{TRL}}\ m'/h \vDash_{Ock} \varphi]$
$\mathbf{M}_{\mathbf{TRL}}\ m/h \vDash_{Ock} \Diamond\varphi$	iff	$\exists h'$: $[m \in h'$ and $\mathbf{M}_{\mathbf{TRL}}\ m/h' \vDash_{Ock} \varphi]$
$\mathbf{M}_{\mathbf{TRL}}\ m/h \vDash_{Ock} \mathrm{F}\varphi$	iff	$\exists m'$: $[m' \in h$ and $m < m'$ and $\mathbf{M}_{\mathbf{TRL}}\ m'/h \vDash_{Ock} \varphi]$

As Øhrstrøm says, "the solution is just a simple addition of the TRL to the Priorean Ockhamist model, and it does not lead to any proper change in the semantics, since exactly the same formulae will be valid as in Priorean Ockhamism" (*ibid*). This can be summarised as the following theorem:

Theorem:
If $\mathbf{M}_{\mathbf{TRL}}$ and $\mathbf{M}_{\mathbf{O}}$ are both models based on some structure S,
$\mathbf{M}_{\mathbf{TRL}}\ m/h \vDash_{Ock} \varphi$ iff $\mathbf{M}_{\mathbf{O}}\ m/h \vDash_{Ock} \varphi$

This theorem is correct because TRL-models are 'based on' BT structures, just like BT models are (except that they add a *TRL*, which does not "play any active role in the semantics"). Using this, we are able to rewrite the BT definitions of Fara's *B* operator (i.e. Definition 7), replacing "$\mathbf{M}_{\mathbf{O}}$" with "$\mathbf{M}_{\mathbf{TRL}}$":

Definition 12: *B* in TRL-models

$\mathbf{M}_{\mathbf{TRL}}\ m/h \vDash_{Ock} B\varphi$ iff $\exists h'$: $[\mathbf{M}_{\mathbf{TRL}}\ m/h' \vDash_{Ock} \varphi]$ and $\exists h''$:$[\mathbf{M}_{\mathbf{TRL}}\ m/h'' \vDash_{Ock} \neg\varphi]$

At this point, as Øhrstrøm notes, the semantics pays no attention to the TRL. Therefore, it is time to introduce the (post)semantic pseudo-supervaluationist definition which *does* pay attention to it. The reason for introducing Øhrstrøm's simple theory is that this new (post)semantic definition requires us to use both the Ockhamist clauses and TRL-models. What follows is the crucial definition, which is precisely the novel idea presented in Malpass & Wawer (2012). It is a modification of the supervaluational approach, but with a *disjunctive* clause which mentions the TRL:

Definition 13: STRL-truth

$\mathbf{M}_{\mathbf{TRL}}\ m \vDash_{STRL} \varphi =_{def} \forall h$:$[h \in H(m) \Rightarrow \mathbf{M}_{\mathbf{TRL}}\ m/h \vDash_{Ock} \varphi]$ or $\mathbf{M}_{\mathbf{TRL}}\ m/TRL \vDash_{Ock} \varphi$
$\mathbf{M}_{\mathbf{TRL}}\ m \dashv_{STRL} \varphi =_{def} \forall h$:$[h \in H(m) \Rightarrow \mathbf{M}_{\mathbf{TRL}}\ m/h \vDash_{Ock} \neg\varphi]$ or $\mathbf{M}_{\mathbf{TRL}}\ m/TRL \vDash_{Ock} \neg\varphi$

This definition says that φ is STRL-true iff either φ is Ock-true on all histories, or if it is Ock-true using the TRL as the value of the history parameter. If the semantics can use the TRL as the history of evaluation then it does, otherwise it quantifies universally over histories.

An immediate consequence is that actual predictions of future contingents have STRL-truth-values, but non-actual ones do not. This result is intended as a reflection of the thought that the passing of time resolves contingents, and there is no such passage at non-actual moments (this view is elaborated somewhat in Malpass & Wawer 2012, sec. 7.1.2).

The STRL verdict on non-actual predictions may seem strange, but they may well be connected to issues in the logic of counterfactuals (which is obviously a closely related area). In order to motivate the result, observe that the STRL verdict on non-actual predictions is similar to independent positions in the counterfactuals literature, notably that of Hájek amongst others. A consequence of the STRL definition is that *only* non-actual future contingents come out as neither-true-nor-false; all other non-actual formulae (including predictions of tautologies etc) get given truth-values. In the case of predictions of tautologies, these come out true on STRL (unlike on TRL), and therefore STRL can be seen as an improvement on TRL. This also means that the STRL assignment of truth to non-actual formulae is closer to Hájek's proposal than TRL was, as he also considers 'necessary would-counterfactuals' to be true. The difference between Hájek's position and STRL is that the latter calls non-actual future contingents neither-true-nor-false, and the former calls them false. The STRL verdict is a position Hájek considers (on page 24-26) not too unsympathetically as a 'fall-back' for his own view, as they both agree that contingent would-counterfactuals are 'not true'.

This observation is only meant to highlight that the particularities of STRL could be considered independently plausible. However, the details of precisely how to relate STRL semantics to a theory of branching-time counterfactuals will have to wait for now though (I plan to write a follow-up paper in which this issue is addressed head-on).

4.1 STRL and FF

Let us now turn our attention back to Fara's Formula, and see how STRL copes. It is important to note that according to STRL, there can be situations where Fp is STRL-true, and yet p is a future contingent; i.e. if p happens in the TRL but not every other history, for instance. The following diagram illustrates this situation:

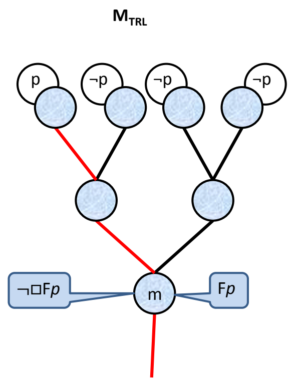

Figure 1

Given the STRL definition, we can add the final piece to the picture, $\Diamond_s$ in TRL-models:

Definition 14: $\Diamond_s$ in TRL-models

$\mathbf{M_{TRL}}\ m/h \vDash_{Ock} \Diamond_s\varphi$ iff $\exists \mathbf{M_{TRL}}', \exists m' \in \mathbf{M_{TRL}}': [\mathbf{M_{TRL}}'\ m' \vDash_{STRL} \varphi]$

What we need to do now is evaluate Fara's Formula at a point of evaluation in a TRL model, and see if it is STRL-true. If it is, then Fara's argument is equally crippling to the STRL theory (Definition 13) as it is to the previously defined supervaluationism (Definition 6). If it cannot be satisfied at *any* point, then the STRL theory avoids the problem she outlines. Here is the formula:

(FF₂) $\mathbf{M_{TRL}}\ m \vDash_{STRL}((B\mathrm{F}p \wedge \mathrm{F}p) \vee (B\mathrm{F}p \wedge \neg \mathrm{F}p)) \wedge \neg\Diamond_s(B\mathrm{F}p \wedge \mathrm{F}p) \wedge \neg\Diamond_s(B\mathrm{F}p \wedge \neg \mathrm{F}p)$

Let us go through the three conjuncts of FF_2 and see if they are all STRL-true.

Again, we must look at moments where there is a chance of finding the formula true. The Fara formula is not *valid*, nor is that Fara's claim. So there are plenty of points where it is false. For instance, at any point m/h at which $\Box \mathrm{F}p$ is Ock-true, $B\mathrm{F}p$ will be false. So, if we want the first conjunct $((B\mathrm{F}p \wedge \mathrm{F}p) \vee (B\mathrm{F}p \wedge \neg \mathrm{F}p))$ to be (at least) Ock-true, then we need only look to use points of evaluation where $\mathrm{F}p$ is not necessary (i.e. where it is contingent).

The original motivation behind Fara's complaint was the dissatisfaction with the result that the disjunction $(\varphi \vee \psi)$ could be true even though each disjunct lacked a truth-value. If we use a moment like m in Figure 1 (above), then the disjunction $(B\mathrm{F}p \wedge \mathrm{F}p) \vee (B\mathrm{F}p \wedge \neg \mathrm{F}p)$ will have one STRL-true disjunct and one STRL-false disjunct; in the example above $B\mathrm{F}p \wedge \mathrm{F}p$ is true because p is a future contingent (hence $B\mathrm{F}p$ is STRL-true), and p happens in the TRL (hence $\mathrm{F}p$ is STRL-true). It is only off the TRL that the STRL-semantics generates results that mimic the supervaluationist treatment of 'borderline cases' (i.e. disjuncts having no super-truth-value). Therefore, if we want to evaluate Fara's formula at a moment that is as similar as possible to the original setting (with super-truth-valueless disjuncts etc), then we should try moments *off* the TRL.

So, assume that p is Ock-true at some later moment m' in some, but not all, histories through m (i.e. that p is a future contingent at m), and that m is *off* the TRL. Now, each history will make $B\mathrm{F}p$ Ock-true, and therefore it will also be STRL-true. Just as we reasoned before (on page 4), each history will either make $\mathrm{F}p$ Ock-true or $\neg \mathrm{F}p$ Ock-true, which results in the disjunction of $(B\mathrm{F}p \wedge \mathrm{F}p) \vee (B\mathrm{F}p \wedge \neg \mathrm{F}p)$ being Ock-true on every history. Therefore, the disjunction will also be STRL-true, even though each disjunct if evaluated separately would have no STRL-truth-value. Here is a diagram of the model, $\mathbf{M_{TRL}}$, which satisfies $((B\mathrm{F}p \wedge \mathrm{F}p) \vee (B\mathrm{F}p \wedge \neg \mathrm{F}p))$ at some moment m off the TRL (note that the TRL has branched off earlier than m, and so is dashed to indicate its irrelevance for the evaluation at hand):

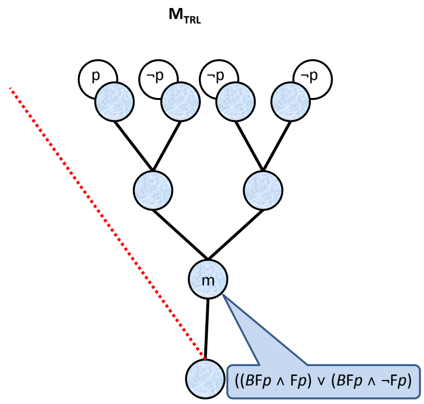

Figure 2

Next, let us look at the second disjunct. This essentially requires that there is no moment *m'* in any TRL-model, $\mathbf{M_{TRL}}$**'**, at which ($BFp \wedge Fp$) is STRL-true. If we can find such a moment in a model, then the second conjunct is false and we can avoid the trap.

The whole point of STRL definition of truth is to allow that there can be true (actual) future contingents. If there is a sea battle in the actual future, but in no other possible future, it is contingent, but still simply true that there will be a sea battle. Given Definition 14, we can always find such a model using the $\lozenge_s$ operator; essentially all we have to do is consider a model with a different TRL. Here is what the new model, $\mathbf{M_{TRL}}$**'**, would look like:

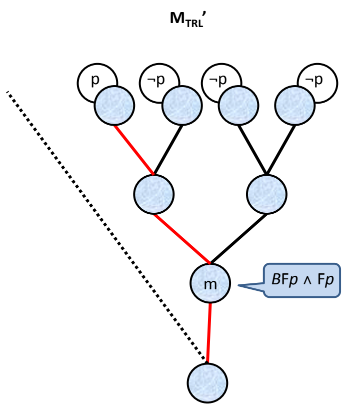

Figure 3

Thus, p is true only at some later moment in the TRL and in no other history[6], and m is in the TRL, so we get that BFp is STRL-true (because there is a history through m where Fp is Ock-false and a history where is it Ock-true) but we also get that Fp is STRL-true (because m/TRL makes Fp Ock-true). Therefore, there **is** a model, such as $\mathbf{M_{TRL}}$**'**, in which (BFp ∧ Fp) is STRL-true at m, and so $\neg\Diamond_s$(BFp ∧ Fp) is not STRL-true at $\mathbf{M_{TRL}}$ m. The fact that one conjunct is true and another false is all I need to show, in order to demonstrate that her formula is not satisfied in this model. However, it also follows that the final conjunct is also STRL-false (for entirely similar reasons).

There are no other cases to consider. The only formulae that ever lack an STRL-truth-value are non-actual predictions of future contingents. But in these situations, while we can satisfy the first conjunct, we can always find a model like in Figure 3, meaning that we never get the second or third conjuncts.

This means that I have completed my task of showing that the STRL semantics is not vulnerable to the same problem as 'canonical' supervaluationism or Thomason's BT supervaluationism. That leads to the final Fact of the paper:

Fact 2:
FF_2 is not STRL-satisfiable.

4. Conclusion

The conclusion is that Fara's ingenious and troublesome formula cannot be satisfied in the STRL logic, and so creates no problem for the theory. This is not to say that her original target is off the hook. The original supervaluational theory, and its application to BT, have both been successfully hit by the attack.

This invites the question of whether STRL counts as a version of supervaluationism or not. If the former, then the conclusion should be:

(i) Supervaluationism about future contingents can survive only in a modified form (i.e. STRL).

If the latter, then the conclusion should be:

(ii) Supervaluationism must be abandoned, because only theories that are not supervaluational survive (like STRL).

Here are some reasons to think that STRL is a version of supervaluationism, and that (i) should be the advice. To begin with, STRL can simply be considered to be Thomason's theory with two modifications; the relatively minor addition of the TRL to the models (which to begin with makes "no proper change to the [Ockhamist] semantics"), and the addition of the TRL-sensitive disjunct in the clause for super-truth. So, STRL seems close to supervaluationism, as it still retains the distinctive ideas of having a distinction between (sub-)truth and super-truth, and involving universal quantification over histories in the latter.

On the other hand, STRL does not retain every feature of supervaluationism, as it obviously sacrifices the idea that predictions of future contingents should receive no

[6] Although, this is the strongest case—as noted above, all that is required is that p is not true in every history.

truth-value (at least as far as actual predictions are concerned). It also sacrifices the idea that super-truth is truth on all histories; in STRL super-truth is not *simply* truth on all histories. Also, STRL obviously has a TRL in it, which seems to indicate that it is really a version of TRL-theory, rather than a version of supervaluationism per se. It might be suggested that this departure from the conventional supervaluationism is enough to preclude it from being considered as a type of supervaluationism.

We need to be clear about what drives the supervaluationist to his definition. I think that it would be wrong to suggest that the supervaluationist is really *that* concerned with ensuring that super-truth is nothing other than truth on all histories as such; rather this is a means to his end. Thomason developed his early (1970) paper in his entry in the *Handbook of Philosophical Logic*, which was updated in his (2002), and in this article he spends time explaining what motivated his creation of his supervaluationism for future contingents. He notes that what is good about the Ockhamist semantics is that

> [...] indeterminist frames can be accommodated without sacrificing any orthodox validities. This is good for those who (like me) are not determinists, but feel that these validities are intuitively plausible. (Thomason 2002, 215)

However, Thomason feels dissatisfied with Ockhamism mainly because of the "entirely prima facie" manner in which the history parameter has to be used (see Thomason 1970, 270-271). What is desired is a way of keeping the "intuitive validities" while doing away with the "prima facie" history parameter. This is exactly what STRL delivers, as it too contains all the Ockhamist validities (see Malpass & Wawer 2012, sec. 7.3.1), and has a notion of truth that is history-independent. To the extent then that the supervaluationist about future contingents is motivated by these concerns, STRL should be an attractive option for them, regardless of whether it counts as a version of supervaluationism or not.

Perhaps, if STRL is not a version of supervaluationism, it should be counted as a version of TRL theory. If so, this would support the conclusion being (ii) rather than (i). However, STRL does not retain every feature of TRL theory, as it obviously sacrifices the idea that the truth-value of future-tensed propositions is determined solely by the actual future. One might like to interpret predictions in Øhrstrøm's theory as 'it *actually* will be that...', but STRL cannot be interpreted simply in the same manner. In situations where at a non-TRL moment a coin is flipped and the sentence 'it will land heads or tails' is evaluated, STRL would classify the sentence as super-true, even though (on the assumption that the coin is not flipped on the TRL) the coin will not *actually* land either way.

This means that if we are uncomfortable with calling STRL a version of supervaluationism, we might also feel uncomfortable calling it a version of TRL theory. STRL, it seems to me, is clearly a combination, or a cross-breed, of both supervaluationism and TRL theory. It belongs only to some extent in either rival camp, as it does not provide every feature of either. Then again, it represents a fairly notable improvement on each of its predecessors. It also seems to have independent philosophical motivation, which may well be connected with issues in the logic of counterfactuals.

The final advice then is that supervaluationists about future contingents should abandon Thomason's theory and adopt STRL. The reasons for this are that STRL is a theory which avoids the problems that beset their previous theory, and is delivers results that are acceptable to the original motivations for creating supervaluationism about future contingents in the first place. It is therefore philosophically close to supervaluationism, regardless of whether it is a version of supervaluationism, or TRL theory, or neither.

REFERENCES

Aristotle, 1975. *Categories and De Interpretatione.* J. L. Ackrill, ed. and trans. Oxford: Clarendon Press.

Belnap, Nuel, Perloff, Michael and Xu, Ming. 2001. *Facing the Future: Agents and Choices in Our Indeterminist World.* USA: Oxford University Press.

Burgess, John. 1979. Logic and Time. *Journal of Symbolic Logic* 44/4: 566-582

Graff Fara, Delia. 2010. Confusions and Unsatisfiable Disjuncts: Two Problems for Supervaluationism. In *Cuts and Clouds: Vaguenesss, its Nature, and its Logic* (pp. 373-382), Oxford: Oxford University Press

Fine, Kit. 1975. Vagueness, Truth and Logic. *Synthese* 30/3: 265-300

Hájek, Alan. (unpublished ms.) Most Counterfactuals are False. Available at: http://philrsss.anu.edu.au/people-defaults/alanh/papers/MCF.pdf

Malpass, Alex and Wawer, Jacek. 2012. A Future for the Thin Red Line. *Synthese Special Issue: Branching Time and Indeterminacy*. Eds. Alex Malpass and Chris Gifford

Müller, Thomas. 2012. Branching in the Landscape of Possibilities. *Synthese Special Issue: Branching Time and Indeterminacy*

Øhrstrøm, Peter. 1981. Problems Regarding the Future Operator in an Indeterministic Tense Logic. *Danish Yearbook of Philosophy*, 18: 91-95.

—. 1984. Anselm, Ockham and Leibniz on Divine Foreknowledge and Human Freedom. *Erkenntnis* 21: 209-222.

—. 2009. In Defence of the Thin Red Line: A Case for Ockhamism. *Humana.mente* 8: 17-32.

—. and Hasle, Per. 1995. *Temporal Logic: From Ancient Ideas to Artificial Intelligence*. Studies in Linguistics and Philosophy 57. Netherlands: Kluwer Academic Publishers.

Prior, Arthur. 1967. Past, Present and Future. Oxford: Oxford University Press.

Reynolds, Mark. 2003. An Axiomatization of Prior's Ockhamist Logic of Historical Necessity. *Advances in Modal Logic 4* (pp. 335-370), eds. Balbiani, Suzuki, Wolter and Zakharyaschev, Kings College Publications.

Thomason, Richmond. 1970. Indeterminist Time and Truth-Value Gaps. *Theoria* 36/3: 264-281.

—. 2002. Combinations of Tense and Modality. *Handbook of Philosophical Logic*, Vol 7 (205-234). eds. Dov Gabbay and Franz Guenthner. Netherlands: Kluwer Academic Publishers.

ALEX MALPASS is a lecturer in philosophy at the University of Glasgow. He has also taught at the University of Bristol, where he was awarded his PhD on the logic of time and modality. He has edited a special issue of *Synthese* and the book *The History of Formal and Philosophical Logic: From Aristotle to Tarski*.

ADDRESS: 69 Oakfield Avenue, Department of Philosophy, University of Glasgow, Glasgow G12 8QQ, UK. E-mail: a.p.malpass@gmail.com