Conditional logics accommodating Stalnaker's thesis

Andrew Bacon*

This is a minimally cleaned up draft from 20/06/2012

Stalnaker's thesis states that the probability of an indicative conditional is the conditional probability of the antecedent on the consequent, whenever the antecedent has non-zero probability. Formally:

$$Pr(A \to B) = Pr(B \mid A) \text{ if } Pr(A) > 0$$

Here $\to$ denotes the indicative conditional.

A conditional logic is incompatible with this thesis if we can derive a contradiction, or some dire restriction on Pr, from the assumption that Pr respects the logic (i.e. assigns probability 1 to every theorem of the logic) and is subject to Stalnaker's thesis.[1] Some conditional logics are incompatible with Stalnaker's thesis – indeed, Stalnaker's own preferred logic of conditionals is incompatible with Stalnaker's thesis (see Stalnaker [**?**], Hajek and Hall [**?**].)

Stalnaker himself rejected the thesis and not the logic; in this paper we investigate the alternative hypothesis. Namely:

> Which conditional logics are compatible with Stalnaker's thesis?

One distinguishing feature of Stalnaker's logic, that plays a significant role in some of the triviality arguments, is that it validates a certain weakening of the transitivity of the conditional.

CSO $(A \leftrightarrow B) \supset ((A \to C) \supset (B \to C))$

Here $\supset$ formalises the material conditional and $A \leftrightarrow B$ is short for $(A \to B) \wedge (B \to A)$. Thus another theme of this paper will be:

> Which conditional logics do not contain CSO?

Whilst no logic containing CSO is compatible with Stalnaker's thesis, the converse connection between CSO is not so clear. We show that out of the conditional logics considered, the logics compatible with Stalnaker's thesis are exactly those which do not prove CSO.

*I wrote these notes about conditional logics accommodating Stalnaker's thesis in 2012, and have had a few queries about it since so I have made it available online. The tenability result in section 5.1 has been subsumed by Bacon 2015.

[1]Usually this dire restriction amounts to showing that Pr is non-zero on at most two disjoint propositions.

The second question is of independent interest too. Over the years several philosophers have proposed counterexamples to CSO or have endorsed theories which do not validate it.[2]

The results of this might be very roughly summarised as follows: the logics among this set that do not collapse into Stalnaker's logic (and thus prove CSO) are compatible with Stalnaker's thesis. At the end we show that all but the strongest of this lattice of logics can satisfy Stalnaker's thesis.

1 Logic and Semantics

We shall work within a modal propositional language, $\mathcal{L}$, consisting of the usual truth functional connectives, $\neg$ and $\supset$ from which the other truth functional connectives are definable, and a special binary modal connective representing the conditional, $\to$. I shall adopt the ordinary definitions of $\wedge, \vee, \bot$ in terms of $\supset$ and $\neg$. I shall also adopt the following shorthands:

$$A \equiv B := (A \supset B) \wedge (B \supset A)$$

$$A \leftrightarrow B := (A \to B) \wedge (B \to A)$$

$$\Box A := (\neg A \to \bot)$$

To increase readability, as is typically done in probability theory, I shall frequently shorten $A \wedge B$ to AB, and $\neg A$ to $\overline{A}$.

A *frame* for $\mathcal{L}$ is a pair $\langle W, f\rangle$ where W is a set of worlds and $f : \mathcal{P}(W)\times W \to \mathcal{P}(W)$ – f is called the 'selection function'.[3] A model is a pair $\langle \mathcal{F}, [\![\cdot]\!]\rangle$ where $\mathcal{F}$ is a frame and $[\![\cdot]\!]$ maps propositional letters to subsets of W. $[\![\cdot]\!]$ extends to a function from the rest of $\mathcal{L}$ to $\mathcal{P}(W)$ as follows:

- $[\![\neg\phi]\!] = W \setminus [\![\phi]\!]$
- $[\![\phi \supset \psi]\!] = (W \setminus [\![\phi]\!]) \cup [\![\psi]\!]$
- $[\![\phi \to \psi]\!] = \{w \mid f([\![\phi]\!], w) \subseteq [\![\psi]\!]\}$

A sentence, ϕ, is true in a model $\langle W, f, [\![\cdot]\!]\rangle$ iff $[\![\phi]\!] = W$, and is valid on a frame iff it's true in every model based on that frame, and valid on a class of frames iff it is valid on every member of that class.

The logic CK denotes the logic consisting of the rules RCN, RCEA and the axiom CK in addition to the usual rules of modus ponens and uniform substitution. CK is the analogue in conditional logic to the smallest normal modal logic K.[4]

RCN if $\vdash \psi$ then $\vdash \phi \to \psi$

[2]See Tichy [?], Martenson [?], Tooley [?], Ahmed [?], for counterexamples, Kvart [?], Schaffer [?], Edgington [?], Schulz [?] for theories which invalidate CSO.

[3]See Chellas [?].

[4]As with the modal logic K we use the same name for the logic and its characteristic axiom.

RCEA if $\vdash \phi \equiv \psi$ then $\vdash (\phi \to \chi) \equiv (\psi \to \chi)$

CK $(\phi \to (\psi \supset \chi)) \supset ((\phi \to \psi) \supset (\phi \to \chi))$

CK is the logic of frames: a sentence is provable in CK iff it is valid on the class of all frames (see Chellas [**?**] p?.) Indeed, we can consider it as a multi-modal logic in which $\phi \to$ is a normal modal operator for each substitution of ϕ. A useful derived rule of CK, which we shall make frequent use of, is the inference $A \to B, A \to C \vdash A \to BC$ (hint: apply RN to the classical theorem $B \supset (C \supset BC)$ and apply CK.) In the proofs that follow I shall not write out derivations that can be done within CK and and propositional logic. When a step is derivable in the logic CK I shall write 'CK' next to it to indicate that the proof has been surpressed.

In addition to these principles it is extremely natural to want to add the following principles

ID $\phi \to \phi$

MP $(\phi \to \psi) \supset (\phi \supset \psi)$

MP effectively states that $\to$ obey modus ponens. Both ID and MP are accepted by almost all theorists working on conditionals.[5] Call the resulting logic L0. L0 is validated by the class of frames which additionally satisfy:

ID $f(A) \subseteq A$

MP $x \in f(A)$ whenever $x \in A$.

So far everything ought to seem pretty uncontroversial. The following hotly contested principle is called 'conditional excluded middle', and is, for example, the primary difference between Lewis's and Stalnaker's approach to subjunctive conditionals.

CEM $(\phi \to \psi) \vee (\phi \to \neg\psi)$

We shall call the result of adding CEM to L0 L1. The primary focus of this paper will be logics extending L1 – or 'L1 logics' for short. CEM is guaranteed if we additionally restrict attention to frames where

CEM $|f(x)| \leq 1$

In the presence of CEM we can modify the semantics in to Stalnaker's original form so that f maps us from a world, w, and a set of worlds, A, to a single possible world (namely x if $f(A, w) = \{x\}$ in the general semantics) or the unique impossible world λ (if $f(A, w) = \emptyset$) in the general semantics) at which every sentence is stipulated to be true. While Chellas' semantics is more general, we shall find it more convenient to use Stalnaker's formulation in section 1.1.

[5]Exceptions include Lowe [**?**] for ID and Lycan [**?**] for MP.

1.1 Frames and Stalnaker's Thesis

This paper is primarily concerned with L1 logics. This allows us to separate what is distinctive about the principle CEM from other features of Stalnaker's logic, most notably CSO.

However, the most important reason for this focus on L1 is that it is the natural setting for studying Stalnaker's thesis. For example, it is easy to show that CEM is true almost-everywhere in every probability frame that obeys Stalnaker's thesis, for antecedents with positive probability. (A proposition is true almost-everywhere iff the set of points where it fails has measure 0.) This follows from a simple probability calculation. Firstly, by finite additivity we have: $Pr((A \to B) \vee (A \to \bar{B})) = Pr(A \to B) + Pr(A \to \bar{B}) - Pr((A \to B) \wedge (A \to \bar{B}))$. Secondly, since in L0, $(A \to B) \wedge (A \to \neg B)$ is logically equivalent to $A \to (B \wedge \neg B)$ we may make that substitution in the last term. Assuming $Pr(A) > 0$, we may also apply Stalnaker's thesis to all the terms, giving us: $Pr(B \mid A) + Pr(\bar{B} \mid A) - Pr(B \wedge \bar{B} \mid A)$, which is just $1 - 0 = 1$.

That L1 is the natural logic for studying Stalnaker's thesis is further evidenced by van Fraassens theorem (see [**?**]) that every theorem of L1 has probability 1 for *any* connective satisfying Stalnaker's thesis and some plausible background premisses (the connective need not necessarily be a connective generated by a selection function.)[6]

Note, however, that there was a complication in the argument for CEM above: we had to assume that the antecedent has non-zero probability. When the conditional probability is undefined, the probability of the conditional is completely unconstrained. It is just extremely natural to distinguish two versions of Stalnaker's Thesis, depending on whether we include a constraint governing the probability of a conditional that has an antecedent with probability zero. Consequently, it also becomes natural to distinguish two forms of probabilistic validity: one in which we consider a sentence valid if it receives probability 1 whenever all the relevant conditional probabilities are defined, and another in which is must receive probability 1 always, but we extend the notion of conditional probability to allow for conditioning on probability zero events.

We begin with the ordinary version of the thesis that leaves the probability of a conditional completely unconstrained when its antecedent has no probability. In what follows we shall focus on probability measures, i.e. countably additive probability functions. (In section [REF] I shall briefly mention ways in which the semantics can be modified to account for failures of countable additivity.) A *probability frame* is a triple $\langle \mathcal{F}, Pr, \Sigma \rangle$, where $\mathcal{F} = \langle W, f \rangle$ is a frame, Σ a set of subsets of W which is closed under complements and countable unions and Pr a probability measure on Σ.[7] Members of Σ are called 'measurable'. A probability model, $\langle \mathcal{F}, Pr, \Sigma, [\![\cdot]\!] \rangle$, based on a probability frame is defined

[6] These background premises are: (i) that if A is consistent and B is inconsistent with C then $A \to B$ is inconsistent with $A \to C$ and (ii) that if two sentences receive the same probability according to any probability function obeying Stalnaker's thesis those two sentences are logically intersubstitutable with one another.

[7] One might also want to stipulate that $\{x \mid f(A, x) \subseteq B\}$ be in Σ whenever A and B are. However, since this condition is entailed by Stalnaker's thesis it is not strictly necessary.

as above. A probability frame satisfies Stalnaker's thesis just in case $Pr(\{w \mid f(A,w) \subseteq B\}) = Pr(B \mid A)$ whenever $A, B \in \Sigma$ and $Pr(A) > 0$.

Note, however, that when $Pr(A) = 0$ the conditional probability $Pr(B \mid A)$ is undefined and so the probability of $Pr(\{w \mid f(A,w) \subseteq B\})$ is completely unconstrained. In order to assess Stalnaker's Thesis within a framework that allows you to talk about arbitrary conditional probabilities we need a more general theory. I shall focus on a natural theory of conditional probabilities due to Popper and Renyi. I shall concentrate on Popper's version (as presented in van Fraassen [**?**]).

Definition 1.0.1. *$\langle W, \Sigma, Pr\rangle$ is a Popper space when Σ is a σ-algebra on W, $Pr : \Sigma \times \Sigma \to [0,1]$ and*

1. *$Pr(\cdot \mid A)$ is either a probability measure or has constant value 1.*
2. *$Pr(AB \mid C) = Pr(A \mid C)Pr(B \mid AC)$*

For convenience we shall shorten $Pr(\cdot \mid \top)$ to $Pr(\cdot)$.

A set $A \in \Sigma$ is said to be normal if $Pr(\cdot \mid A)$ is a probability measure and abnormal if $Pr(\cdot \mid A)$ is constantly 1. A is abnormal if and only if $Pr(A \mid B) = 0$ for every normal B (see [**?**].) A Popper frame, $\langle \mathcal{F}, Pr, \Sigma\rangle$, is as before except that Pr is now a Popper function on W and Σ. Say that 'Strengthened Stalnaker's Thesis' holds in a Popper frame just in case $Pr(\{w \mid f(A,w) \subseteq B\}) = Pr(B \mid A)$ whenever $A, B \in \Sigma$ with no restrictions on A.

1.2 The Limit Assumption

The selection function semantics we have outlined is flexible enough to characterise just about any conditional logic you can think of statable in this language.[8] However there is one controversial principle that the selection function semantics is not neutral on. This is often called the 'limit assumption', and can be stated in an infinitary logic as follows

LIM $\bigwedge_{n\in\omega}(A \to B_n) \supset (A \to \bigwedge_{n\in\omega} B_n)$

LIM is validated in every frame, assuming we extend our definitions to deal with infinite conjunctions in the natural way. In Lewis's semantics it corresponds to the idea that for any condition A there is always a set of maximally close A-worlds to the world of evaluation (see [**?**].) While we can invalidate LIM in Lewis's semantics by dropping this assumption, Lewis's semantics validates CSO and is thus not general enough for our purposes.

It is worth noting that this aspect of the selection function semantics doesn't show up in the finitary logic. However this issue does bear on our discussion of

[8]For example, as well as accommodating the variably strict conditionals of Lewis and Stalnaker, most other theories can be reformulated to conform with this semantics. For example, the material conditional can be obtained by setting $f(A,x) = \{x\}$ if $x \in A$ and $= \emptyset$ otherwise, strict conditionals can be obtained by setting $f(A,x)$ to be the set of A-worlds accessible to x, and so on.

Stalnaker's thesis and on our assumption of countable additivity. If $Pr(\cdot \mid A)$ is not countably additive then there can be cases where $Pr(B_n \mid A) = 1$ for every n but $Pr(\bigcap_n B_n \mid A) = 0$. This shows that the limit rule $\{A \to B_n \mid n \in \mathbb{N}\} \vdash A \to \bigwedge_n B_n$ doesn't preserve probability 1.[9] Conversely, if countable additivity holds, then the limit rule and the limit axiom LIM are valid.

A slight generalisation of the selection function semantics can be introduced that would allow a bit more generality in this direction.

Definition 1.0.2. *An ultrafilter, $U \subset \mathcal{P}(W)$, on a set W is a set such that.*

1. $\emptyset \notin U$
2. *If* $A, B \in U$, $AB \in U$
3. *If* $A \in U$ *and* $A \subseteq B$ *then* $B \in U$
4. *For any* $A \subseteq W$, $A \in U$ *or* $W \setminus A \in U$.

An ultrafilter is principal if and only if $U = \{X \mid x \in X\}$ *for some* $x \in W$.

Let $\mathcal{U}(W)$ be the set of ultrafilters on W. A generalised selection function may then be identified with a function $f : \mathcal{P}(W) \times W \to \mathcal{P}(\mathcal{U}(W))$, and the conditional with antecedent A and consequent B with $\{w \in W \mid B \in f(A, w)\}$. As before, we automatically validate RCEA, RCN and CK. To obtain frame conditions for a generalised selection function from any of the additional principles discussed one must add one of the following constraints.

ID $A \in f(A, x)$

MP If $x \in A$ then $f(A, x) \subseteq \{X \mid x \in X\}$

CEM $|f(A, x)| \leq 1$

LIM If $U \in f(A, x)$ then U is principal.

This semantics deserves further investigation, but it is not the place to do so here. In the rest of the paper we set this issue aside and assume countable additivity and the limit assumption.

2 Strengthening $\mathsf{L1}$

There are, roughly, two ways to strengthen $\mathsf{L1}$:

1. Add principles that govern when a conditional is vacuously true.
2. Add principles governing the logical connections between conditionals with different antecedents.

[9] It is not clear, however, whether failures of countable additivity straightforwardly generate failures of the *axiom* LIM to have probability 1. For example it is possible that whenever $Pr(\bigcap_n B_n \mid A) = 0$ and $Pr(B_n \mid A) = 1$ for each n, $Pr(\bigcap_n (A \to B_n)) = 0$ (a live possibility if countable additivity doesn't hold.)

For example, in some writings Stalnaker states that a conditional is vacuously true exactly when its antecedent is the impossible proposition. This motivates a number of further axioms which are not provable in $\mathsf{L1}$. Similarly, if the selection function is generated by an absolute ordering on worlds then A and B behave exactly the same in the antecedent position if $A \leftrightarrow B$.

To make this precise we shall introduce two definitions

Definition 2.0.1. *Say that A* crashes *at a world x if and only if $f(A, x) = \emptyset$.*

Definition 2.0.2. *Say that A and B are* antecedent equivalents *at a world x if and only if $f(A, x) = f(B, x)$.*

The concept of a sentence, A, crashing can be expressed in the object language by the formula $A \to \bot$ (some people define a similar notion as $A \to \neg A$, although these definitions can come apart if ID is not present.) The concept of two formulae, A and B, being antecedent equivalents cannot be as easily expressed in the object language (at least, not without propositional quantifiers.) Intuitively A and B are antecedent equivalents if and only if, for any choice of C, $A \to C$ is true just in case $B \to C$ is – i.e. A and B conditionally imply exactly the same propositions.

2.1 Antecedent equivalence

The strengthenings of $\mathsf{L1}$ that place further constraints on antecedent equivalence will consist of various combinations of the following principles:

CA $((\phi \to \chi) \wedge (\psi \to \chi)) \supset (\phi \vee \psi \to \chi)$

RCA $(\phi \vee \psi \to \chi) \supset (\phi \to \chi) \vee (\psi \to \chi)$

CV $\neg(\phi \to \neg\psi) \supset ((\phi \to \chi) \supset (\phi \wedge \psi \to \chi))$

RCV $(\phi \to \psi) \supset ((\phi \to \chi) \supset (\phi \wedge \psi \to \chi))$

RT $(\phi \to \psi) \supset ((\phi \wedge \psi \to \chi) \supset (\phi \to \chi))$

CSO $(\phi \leftrightarrow \psi) \supset ((\phi \to \chi) \supset (\psi \to \chi))$

It is worth noting that CV plays a special role in theories that do not contain CEM, such as Lewis's ([**?**]). In the presence of the axiom C2 from the next section CV entails RCV, and for theories such as Lewis's this entailment is strict. However once we have CEM, RCV trivially entails CV (the antecedent of CV entails the antecedent of RCV) so we shall focus on RCV in what follows.

Each of the above axioms can be validated in a frame if additionally place the following constraints on the selection function:

CA $f(A \cup B, x) \subseteq f(A, x) \cup f(B, x)$

RCA $f(A, x) \subseteq f(A \cup B, x)$ or $f(B, x) \subseteq f(A \cup B, x)$

RCV $f(A, x) \subseteq B$ implies $f(A \cap B, x) \subseteq f(A, x)$

RT $f(A,x) \subseteq B$ implies $f(A,x) \subseteq f(A \cap B, x)$

CSO $f(A,x) \subseteq B$ and $f(B,x) \subseteq A$ implies $f(A,x) = f(B,x)$

Since we will be primarily concerned with models satisfying CEM we can simplify the condition for CA to:

CA $f(A \cup B, x) \subseteq f(A,x)$ or $f(A \cup B, x) \subseteq f(B,x)$

This makes a natural dual to RCA.

Many consider the following principles of conditional logic to be undesirable:

AS $(\phi \to \psi) \supset (\phi \wedge \chi \to \psi)$

SDA $(\phi \vee \psi \to \chi) \supset (\phi \to \chi) \wedge (\psi \to \chi)$

TR $(\phi \to \psi) \supset ((\psi \to \chi) \supset (\phi \to \chi))$

CONT $(\phi \to \psi) \supset (\neg\psi \to \neg\phi)$.

MAT $\neg\phi \supset (\phi \to \psi)$

Indeed, given our background logic CK the first two are equivalent, and so have the same frame conditions, AS below. Of course, the most famous theory validating the above principles except for MAT is the strict theory of conditionals which is stated as the frame condition STRICT below: according to the strict theory, each world x is associated with a set of accessible worlds, which may be defined as $f(W,x)$, and $A \to B$ holds at x when every accessible A world is a B world. MAT and MP collapse $\to$ into the material conditional against the background CK.

AS $f(A \cap B, x) \subseteq f(A,x)$

TR If $f(A,x) \subseteq B$ then $f(A,x) \subseteq f(B,x)$.

CONT if $f(A,x) \subseteq B$ then $f(W \setminus B, x) \cap A = \emptyset$

STRICT $f(A,x) = A \cap f(W,x)$

MAT $f(A,x) = \emptyset$ when $x \notin A$

Given other principles we have mentioned above there are several implications between the four consequences of the strict view, AS, SDA, TR and CONT which we note below.

Proposition 2.1. *Against the background logic of* CK *we have*

1. AS *is equivalent to* SDA
2. *Given* CA *and* ID*, AS implies TR.*
3. *Given* ID*,* TR *implies* AS

4. CONT *implies* AS

5. Given CA *and* ID*,* AS *implies* CONT

Since this is not the primary subject of this paper, I briefly sketch the proofs in a footnote.[10]

2.2 Axioms for Crashing

Here are some natural principles governing when an antecedent crashes, in order of strength:

C0. $(A \to \bot) \supset ((B \to C) \equiv (A \vee B \to C))$

C1. $(A \to B) \supset ((B \to \bot) \supset (A \to \bot))$

C2. $(A \to \bot) \supset (AB \to \bot)$

C3. $(A \vee B \to \bot) \supset (A \to \bot) \vee (B \to \bot)$

Once you have CSO C1 and C2 become equivalent (and become redundant with CSO and CEM.) (Note: C0 is really principle governing both crashing and antecedent equivalence: it says that if A crashes then B and $B \vee A$ are antecedent equivalent.)

If we think of crashing as being true in no accessible worlds we naturally get C1 (i.e. $f(A, x) = \emptyset$ iff $R(x) = \emptyset$.) If we also stipulate that $f(A, x) = f(B, x)$ whenever $R(x) \cap A = R(x) \cap B$ we get C0, but it's not clear to me that C0 is straightforwardly valid on my favourite interpretion of the selection function.

I'll also mention the principle, MOD, which is of interest when when CEM is not present:

MOD $(\overline{A} \to \bot) \supset (B \to A)$

MOD entails C1 in L0, but not *vice versa*. However, given CEM C0 and C1 become equivalent.

The above principles correspond to the following frame conditions:

C0. If $f(A, x) = \emptyset$ then $f(A \cup B, x) = f(B, x)$

[10] 1 is derived as follows: $A \to C$ is equivalent given RCEA to $(AB \vee A\overline{B}) \to C$ from which we can infer $AB \to C$ by SDA. Conversely, $(A \vee B) \to C$ implies $(A \vee B) \wedge A \to C$ by AS which implies $A \to C$ by RCEA; a parallel argument establishes $B \to C$. For 2, assume $A \to B$ and $B \to C$. AS and RCEA lets you infer $AB \to C$ from the second assumption. AS applied to the first assumption gives $A\overline{B} \to B$ which given ID and reasoning in CK lets you infer $A\overline{B}$ implies anything: thus $A\overline{B} \to C$. So by CA $AB \vee A\overline{B} \to C$ thus $A \to C$ by RCEA. For 3 we assume $A \to C$. Reasoning in CK with ID we may infer $A \wedge B \to A$ and by TR we get $A \wedge B \to C$. For 4, suppose $A \to C$. By CONT we have $\neg C \to \neg A$. Weakening the consequent we get $\neg C \to \neg(A \wedge B)$, and applying CONT again we get $\neg\neg(A \wedge B) \to \neg\neg C$ which is equivalent to $A \wedge B \to C$ in CK. For 5, assume $A \to B$. By AS, we get $A\overline{B} \to B$ and by ID and CK $A\overline{B}$ crashes and so $A\overline{B} \to \overline{A}$. $\overline{AB} \to \overline{A}$ follows from ID in CK so by CA we get $A\overline{B} \vee \overline{AB} \to \overline{A}$, so by RCEA we get $\overline{B} \to \overline{A}$.

C1. If $f(A,x) \subseteq B$ and $f(B,x) = \emptyset$, $f(A,x) = \emptyset$

C2. If $A \subseteq B$ and $f(B,x) = \emptyset$, $f(A,x) = \emptyset$

C3. If $f(A \cup B, x) = \emptyset$ then $f(A,x) = \emptyset$ or $f(B,x) = \emptyset$.

MOD. If $f(\overline{A},x) = \emptyset$, $f(B,x) \subset A$ for every B

We shall show that when you add C1 or stronger (i.e. C0 or MOD) then the operator $\Box A$, defined as $\overline{A} \to \bot$, behaves like a normal modal operator. This raises the question of how it iterates; two natural principles to consider, in this regard, are:[11]

4 $(A \to \bot) \supset (B \to (A \to \bot))$

B $A \supset (A \to \bot) \to \bot$

Three more principles to consider (instances of RCV, RT and CA respectively. Note that by this naming convention C1 = CT, a special instance of T transitivity, C2 = CAS, a special instance of antecedent strengthing, (AS), C3 = CRCA, a special instance of RCA.)

CRCV $(A \to B) \supset ((A \to \bot) \supset (AB \to \bot))$

CRT $(A \to B) \supset ((AB \to \bot) \supset (A \to \bot))$

CCA $((A \to \bot) \wedge (B \to \bot)) \supset (A \vee B \to \bot)$

3 Strengthening L1

Here we consider supplementing L1 with principles governing crashing, and principles governing antecedent equivalence.

3.1 Crashing

Our first order of business is to show that the principles C0-C3 really are ordered by strength, with C0 being the strongest and C3 the weakest.

Proposition 3.1. *Assuming* L0*, C0 entails C1, C1 entails C2 and C2 entails C3.*

The entailments from C1 to C2 to C3 are completely straightforward, so it remains to show that C0 entails C1. First we show that C0 entails C2.

1. $(A \to \bot) \supset ((AB \to \bot) \equiv ((A \vee AB) \to \bot))$ by C0.
2. $(A \vee AB) \to \bot) \equiv (A \to \bot)$ by RCEA
3. $(A \to \bot) \supset ((AB \to \bot) \equiv (A \to \bot))$ from 1 and 2 substituting equivalents.

[11] The particular formulations of these principles are due to Cian Dorr.

4. $(A \to \bot) \supset (AB \to \bot)$ by propositional logic.

Then:

1. $A \to B$ assumption.
2. $B \to \bot$ assumption.
3. $AB \to \bot$ from 2 by C2.
4. $AB \to \bot \supset ((A\bar{B} \to \bar{B}) \supset ((AB \vee A\bar{B}) \to \bar{B}))$ by C0.
5. $(AB \vee A\bar{B}) \to \bar{B}$ from 3 and 4.
6. $A \to \bar{B}$ from 5
7. $A \to \bot$ by 1 and 6.

Proposition 3.2. *The entailments between C0, C1, C2 and C3 are all strict.*

We present here, for example, a model of C1 without C0: $W = \{a, b, c\}$, $A = \{a\}, B = \{b, c\}$.To refute C0R let $C = \{b\}$, for C0L let $C = \{c\}$

- $f(X, x) = \{x\}$ whenever $x \in X$
- $f(\{a\}, d) = \emptyset$
- $f(\{a, b, c\}, d) = \{c\} \neq \{b\} = f(\{b, c\}, d)$.
- $f(\{a, b\}, d) = f(\{b\}, d) = \{b\}$.
- $f(\{a, c\}, d) = f(\{c\}, d) = \{c\}$.
- Fill out $f(\cdot, a)$, $f(\cdot, b)$ and $f(\cdot, c)$ in any consistent way (e.g. $f(X, x) = \{x\}$ if $x \in X$ and $= \emptyset$ otherwise.)

3.1.1 Normality

Recall the operator $\Box A$ which we defined as $\bar{A} \to \bot$. Say that $\Box$ is *normal* in a logic just in case the rule RN is admissible and the axiom K is derivable:

RN If $\vdash \phi$ then $\vdash \Box\phi$

K $\Box(\phi \supset \psi) \supset (\Box\phi \supset \Box\psi)$

It is also factive in a logic if you can prove all instances of

T $\Box\phi \to \phi$

So long as you have MP, $\Box$ will be factive. In L1 the operator $\Box A$ is not normal. While it necessitates (i.e. RN is an admissible rule) you cannot prove that it obeys the closure axiom K. When $\rightarrow$ represents an indicative conditional, $\Box$ naturally represents a kind of epistemic necessity (perhaps knowledge, or being entailed by the agents evidence) suggesting it should be both factive and normal.

Since $\Box$ is not normal in L1 it is perhaps better to define a notion of epistemic necessity using propositional quantifiers: $\Box A := \forall P(P \rightarrow A)$. While this operator will always behave normally it involves augmenting the language. Thus, in this section we explore the conditions under which $\Box$, as defined initially, is normal.

However once you have a reasonably strong logic of crashing you can show that $\Box A$ defined the former way is normal.

Proposition 3.3. *$\Box$ is normal in L0+C2+CCA.*

1. $\bar{A} \rightarrow \bot$ assumption
2. $A\bar{B} \rightarrow \bot$ assumption
3. $\bar{A}\bar{B} \rightarrow \bot$ from 1 by C2
4. $(\bar{A}\bar{B} \vee A\bar{B}) \rightarrow \bot$ from 2 and 3 by CCA (an instance of CA)
5. $\bar{B} \rightarrow \bot$ by RCEA

Proposition 3.4. *In L1 + C1 $\Box$ is normal.*

1. $\bar{A} \rightarrow \bot$ assumption
2. $A\bar{B} \rightarrow \bot$ assumption
3. $(\bar{B} \rightarrow A) \vee (\bar{B} \rightarrow \bar{A})$ CEM
4. $(\bar{B} \rightarrow A) \supset (\bar{B} \rightarrow A\bar{B})$ ID and CK reasoning.
5. $(\bar{B} \rightarrow A\bar{B}) \supset (\bar{B} \rightarrow \bot)$ from 2 by C1
6. So $(\bar{B} \rightarrow A) \supset (\bar{B} \rightarrow \bot)$ from 4 and 5
7. $(\bar{B} \rightarrow \bar{A}) \supset (\bar{B} \rightarrow \bot)$ from 1 and C1
8. $\bar{B} \rightarrow \bot$ from 3, 6, 7 and reasoning by cases.

In the logics you get by adding C2 or C3 to L1, a number of natural principles remain unprovable, including the closure principle for $\Box$, but also the desirable principles CCA, CRT, CRCV. This fact changes once we reach C1 (and thus, C0); closure and all of the aforementioned principles become derivable.

Proposition 3.5. C1 *entails* CCA, CRT, CRCV.

The trickiest case is CCA:

1. $A \to \bot$
2. $B \to \bot$
3. $(A \vee B \to A) \vee (A \vee B \to \neg A)$ CEM
4. $(A \vee B \to A) \supset (A \vee B \to \bot)$ by 1 and C1
5. $(A \vee B \to \neg A)$ (Supposition)
6. $(A \vee B \to B) \vee (A \vee B \to \neg B)$ CEM
7. $(A \vee B \to B) \supset (A \vee B \to \bot)$ by 2 and C1.
8. $(A \vee B \to \neg B) \supset (A \vee B \to \neg A \neg B)$ by 5 and CK.
9. $(A \vee B \to \neg B) \supset (A \vee B \to \bot)$ from 8
10. $A \vee B \to \bot$ from 6, 7, 9 and reasoning by cases.
11. $(A \vee B \to \neg A) \supset (A \vee B \to \bot)$ from 5-10 conditional proof.
12. $A \vee B \to \bot$ from 3, 4, 11 reasoning by cases

Recall the fact mentioned in section 1.1 that Stalnaker's thesis guarantees, in a very natural sense, the logic $\mathsf{L1}$. Note, however, that this does not extend to any of the principles discussed in this section. Roughly speaking, this is because Stalnaker's thesis says nothing about conditionals with antecedents that have probability 0.

One might thus expect that Strengthened Stalnaker's Thesis should offer a stronger logic in this regard. In fact, it entails every principle mentioned in this section except for C0 and the iteration principles, B and $\mathsf{4}$. Given propositions [REF] and [REF] we know that any $\mathsf{L1}$ logic containing C1 contains all the other crashing principles except for C0, B and $\mathsf{4}$. Thus it suffices to show that C1 is guaranteed by Strengthened Stalnaker's Thesis.

Theorem 3.6. *In any Popper frame satisfying Strengthened Stalnaker's Thesis C1 is true $Pr(\cdot \mid \top)$-almost everywhere.*

Proof. For ease of evaluation we rewrite C1 as: $(A \to \bot) \supset ((B \to A) \supset (B \to \bot))$. We shall show that $Pr(C1 \mid \top) = 1$ (recall that $Pr(\cdot)$ is shorthand for $Pr(\cdot \mid \top)$.) We split the argument into two cases depending on whether A is normal.

A is normal. So $Pr(\bot \mid A) = Pr(A \to \bot) = 0$ and thus $Pr((A \to \bot) \supset D) = 1$ for any D, in particular, for $D = (B \to A \supset B \to \bot)$.

A is abnormal. So $Pr(\bot \mid A) = Pr(A \to \bot) = 1$. So $Pr(C1) = 1$ if and only if $Pr((B \to A) \supset (B \to \bot)) = 1$. Now we split into subcases depending on whether B is normal.

If B is normal then $Pr(A \mid B) = Pr(B \to A) = 0$, so $Pr((B \to A) \supset (B \to \bot)) = 1$.

If B is abnormal $Pr(B \to \bot) = Pr(\bot \mid B) = 1$ so $Pr((B \to A) \supset (B \to \bot)) = 1$.

□

3.2 Antecedent equivalence

Theorem 3.7. CSO *entails* CA, RCA, RT *and* RCV *(in* L1.*)*

Proposition 3.8. $A \to (B \vee C)) \supset (A \to B) \vee (A \to C)$ *is a theorem of* L1*.*

1. $A \to (B \vee C)$ (ASS)
2. $(A \to \neg B) \supset (A \to (\neg B \wedge (B \vee C))$ by 1, ID and agglomeration (CK)
3. $(A \to (\neg B \wedge (B \vee C))) \supset (A \to C)$ by CK
4. $(A \to \neg B) \supset (A \to C)$ by 2, 3 and transitivity of $\supset$
5. $(A \to B) \vee (A \to \neg B)$ by CEM
6. $(A \to B) \vee (A \to C)$ from 4, 5 and reasoning by cases.

Proposition 3.9. CSO *entails* CA *(in* L1.*)*

1. $A \to C$ ASS
2. $B \to C$ ASS
3. $A \vee B \to A \vee B$
4. $(A \vee B \to A) \vee (A \vee B \to B)$
5. $A \to A \vee B$
6. $B \to A \vee B$
7. So either $A \leftrightarrow (A \vee B)$ or $B \leftrightarrow (A \vee B)$ by 4, 5, 6.
8. In the former case $A \vee B \to C$ by 1 and CSO.
9. In the latter case $A \vee B \to C$ by 2 and CSO.

Proposition 3.10. CSO *entails* RCA *(in* L1.*)*

1. $(A \vee B) \to C$ (ASS)
2. $A \to A \vee B$ by L0
3. $B \to A \vee B$ by L0
4. $A \vee B \to A \vee B$ by ID
5. $(A \vee B \to A) \vee (A \vee B \to B)$ from 4 and CEM by above lemma.
6. $A \vee B \to A, A \to A \vee B, A \vee B \to C \vdash A \to C$ instance of CSO
7. $(A \vee B \to A) \supset (A \to C)$ from 1, 2 and 6.
8. $A \vee B \to B, B \to A \vee B, A \vee B \to C \vdash B \to C$ instance of CSO

9. $(A \vee B \to B) \supset (B \to C)$ from 1, 3 and 8.
10. $(A \to C) \vee (B \to C)$ reasoning by cases on 5, 7 and 9

Proposition 3.11. CSO *entails* RT *and* RCV *(in* $\mathsf{L1}$*.)*

It suffices to show that $(A \to B) \supset ((A \to C) \equiv (AB \to C))$

1. $A \to B$ assumption
2. $A \to AB$ by ID and CK.
3. $AB \to A$ by CK
4. $(AB \to C) \supset (A \to C)$ from 2,3 and CSO
5. $(A \to C) \supset (AB \to C)$ from 2,3 and CSO

Theorem 3.12. *The entailments from* CSO *to* CA, RCA, RT *and* RCV *are all strict.*

Here we construct a frame validating $\mathsf{L1}$+RT+RCA which does not entail CSO. Let $W = \mathbb{N}$. We shall pick a member, $j \in W$, which we shall call the joker (for concreteness sake, $j = 3$.) We now define f:

If $j \in A$ then

- $f(A, x) = \emptyset$ if $x \notin A$
- $f(A, x) = \{x\}$ if $x \in A$

If $j \notin A$ then

- $f(A, x) = \{y\}$ where y is the smallest number $\geq x$ with $y \in A$ if there is such a number
- $f(A, x) = \emptyset$ otherwise.

Provided that A does not contain the joker f behaves exactly like a Stalnaker selection function, based on an ordering of the worlds. However, if A does contain the joker A crashes whenever isn't true. We have constructed the model to ensure that (i) either $f(A, x) \subseteq f(A \cup B, x)$ or $f(B, x) \subseteq f(A \cup B, x)$ and (ii) if $f(A, x) \subseteq B$ then $f(A, x) \subseteq f(AB, x)$. However note that $f(\{2, 3\}, 1) = \emptyset$, since 3 is the joker, and $f(\{2\}, 1) = \{2\} \neq \emptyset$ even though $f(\{2, 3\}, 1) \subseteq \{2\}$ and $f(\{2\}, 1) \subseteq \{2, 3\}$. Thus the frame condition for CSO fails in this model.

Here is a frame validating $\mathsf{L1}$+RCV+CA but which does not validate CSO. Again we let $W = \mathbb{N}$ and pick a member, $j \in W$, arbitrarily, which we shall call the joker. We now define f:

If $j \notin A$ then

- $f(A, x) = \emptyset$ if $x \notin A$

- $f(A, x) = \{x\}$ if $x \in A$

If $j \in A$ then

- $f(A, x) = \{y\}$ where y is the smallest number $\geq x$ with $y \in A$ if there is such a number
- $f(A, x) = \emptyset$ otherwise.

f is exactly as before except that A crashes whenever it *doesn't* contain the joker and isn't true.

This time the construction is such that both (i) either $f(A \cup B, x) \subseteq f(A, x)$ or $f(A \cup B, x) \subseteq f(B, x)$ and (ii) if $f(A, x) \subseteq B$ then $f(AB, x) \subseteq f(A, x)$. Yet once again CSO fails. Suppose that $j = 3$ and let $A = \{2\}$ and $B = \{2, 3\}$. So $f(A, 1) = \emptyset$ and $f(B, 1) = \{2\}$. So $f(A, 1) \neq f(B, 1)$ even though $f(A, 1) \subseteq B$ and $f(B, 1) \subseteq A$. Thus CSO fails in this model.

4 The structure of L1 logics

Someone sympathetic to Stalnaker's thesis should be particularly interested in the logics between L1 and Stalnaker's own logic (denoted S below.) As we have argued already, L1 is a lower bound given Stalnaker's thesis (and L1+C1 a lower bound given Strengthened Stalnaker's Thesis.) On the other hand, due to the result of Stalnaker [?], S is a strict upper bound: S is incompatible with Stalnaker's thesis. Any logic compatible with Stalnaker's thesis will be found somewhere inbetween.

We have considered two ways to strengthen L1: adding principles governing crashing and adding principles governing antecedent equivalence. In section [REF] we shall prove a tenability result that shows we can add all of the principles governing crashing mentioned in section 2.2 to L1 in a way that is compatible with Stalnaker's thesis.

On the other hand, one might wonder if we can add antecedent equivalence principles to L1 in a way that retains compatibility with Stalnaker's thesis. In each case the answer is negative. One way to go about showing this is to directly produce a triviality result for each such principle.[12] The approach we shall adopt here, however, is to show if one adds any of CA, RCA, RCV, RT or CSO with an appropriate C principle (i.e. C0, C1, C2 or C3) to L1 the logic collapses into S.

This is of independent interest as it leads to five very simple axiomatisations of Stalnaker's logic. Of particular importance, I think, is an axiomatisation of Stalnaker's logic consists of adding to L1 only the two principles

CA. $((A \to C) \wedge (B \to C)) \supset (A \vee B \to C)$

C2. $(A \to \bot) \supset (AB \to \bot)$

[12] For example, one can modify the argument in Hajek and Hall [?] to directly show that RCV leads to triviality. Edgington [?]hints at an argument using RCA.

The principles CV, RCV, RT, CSO are quite complex and hard to motivate. While they can all be seen as restrictions of transitivity and antecedent strengthening, motivating them on the basis of these connections would be unwise since both transitivity and antecedent strengthening are subject to counterexamples. RCA is perhaps easier to directly reason about, however our intuitions in favour of RCA seem to generalise to the much stronger principle SDA which is a form of antecedent strengthening.

Unlike these principles, CA has direct intuitive appeal and is quite easy to evaluate. Inferences, such as the following, seem to be hard to deny and call out for the kind of general explanation that CA provides

> If John goes to the party we will have a great time,
>
> If Mary goes to the party we will have a great time,
>
> Therefore if John or Mary goes to the party we will have a great time.

Thus a reduction of the less transparent theses of Stalnaker's logic to CA might be seen as casting a favourable light on these more controversial theses.

However given the modest assumptions of L1 and C2 we know that with CA comes Stalnaker's logic, and with Stalnaker's logic goes Stalnaker's thesis. It is therefore worth saying some more about CA. We might begin by reminding ourselves that CSO, RCV and CV have been the subject of a number of counterexamples (see [REF].) Thus the fact that these are consequences of CA could equally be seen as casting doubt on CA.

I think the most salient point to be made in this regard is that we should be cautious when assessing the intuitions in favour of principles involving disjunctive antecedents such as CA, RCA and SDA. Of these SDA is the most obviously problematic since it states a form of antecedent strengthening allowing one to infer $A \to C$ from a conditional with a weaker antecedent, namely $A \vee B \to C$ (see our earlier discussion) and indeed, is equivalent to the problematic conjunctive form of antecedent strengthening given RCEA: $(A \to C) \supset (AB \to C)$. Since we have extremely good reasons to reject antecedent strengthening, we have good reason to reject the validity of SDA. We need some other explanation of why the move from $A \vee B \to C$ to $A \to C/B \to C$ is always acceptable, one that does not appeal to its validity.

Anyone subscribing to Stalnaker's thesis has a satisfying explanation immediately available:

> Anyone who knows that $A \vee B \to C$ is in a position to know that $A \to C$ (and that $B \to C$.) Thus one should not assert that $A \vee B \to C$ unless one is also in a position to assert that $A \to C$ and that $B \to C$.

The principle appealed to is not strictly speaking guaranteed by Stalnaker's thesis, but is an obvious extension to knowledge of an analogous principle for evidential probabilities: anyone who has an evidential probability of 1 in $A \vee B \to C$ must have an evidential probability of 1 in $A \to C$ and in $B \to C$.[13]

However this justification immediately generalises to CA and RCA. Focussing on CA (the case of RCA trivially follows from our discussion of SDA) we have

> Anyone who knows that $A \to C$ and that $B \to C$ is in a position to know that $A \vee B \to C$. Thus one should not assert both that $A \to C$ and that $B \to C$ unless one is also in a position to assert that $A \vee B \to C$.

Once again, the principle that anyone who knows the antecedent of CA knows the consequent is extremely natural if you are sympathetic to Stalnaker's thesis since whenever $A \to C$ and $B \to C$ have evidential probability 1 so does $A \vee B \to C$.

4.1 Stalnaker's logic

We shall work with the following axiomatisation, S, of Stalnaker's logic modified from Nute [**?**]RCEA, RCN, ID, MP, CEM, MOD, CSO, CV. Thus S is the result of adding

MOD $(\bar{A} \to \bot) \supset (B \to A)$

CV $\neg(\phi \to \neg\psi) \supset ((\phi \to \chi) \supset (\phi \wedge \psi \to \chi))$

CSO $(\phi \leftrightarrow \psi) \supset ((\phi \to \chi) \supset (\psi \to \chi))$

to L1.

The main theorem of this section delivers five distinct axiomatisations of S.

Theorem 4.1. S *can be axiomatised by adding to* L1 *any of the following*

1. CSO
2. CA, C2
3. RCV, C1
4. RT, C2
5. RCA, C0

To show 1 it is sufficient to show that both CV and MOD are redundant, meaning that we can axiomatise S simply as L1 + CSO.

Proposition 4.2. S = L1 + CSO

[13] If $Pr(C \mid A \vee B) = 1$, then $Pr(C \mid A) = 1$ and $Pr(C \mid B) = 1$. Note, however, that this cannot be the whole story of why SDA seems plausible, since it generalises to antecedent strengthening. It's quite natural to think that a full explanation of why SDA seems valid while antecedent strengthening doesn't would isolate that phenomenon that is specific to conditionals with disjunctive antecedents (this phenomenon seems to be closely tied to the paradoxes of free choice permission; see Fox [REF].)

As we noted already, given CEM, CV is straightforwardly derivable from RCV, and by proposition **??** CSO entails RCV. Thus all that remains to show is MOD.

1. $\bar{A} \to \bot$ assumption.
2. $(B \to A) \vee (B \to \bar{A})$ CEM
3. $B \to \bar{A}$ (if the left disjunct held we'd be done.)
4. $\bar{A} \to B$ from 1 by CK.
5. $B \to \bot$ from 3, 4, 1 by CSO
6. $B \to A$ from 5 by CK

It suffices, then to show that any of the other combinations of axioms in theorem **??** prove CSO. We next show that you can you can prove CSO from RCV + C1 and from RT + C2, establishing parts 3 and 4 of theorem **??**.

Proposition 4.3. S = RCV + C1 + $\mathsf{L1}$

1. $A \to B$ assumption.
2. $B \to A$ assumption.
3. $A \to C$ assumption.
4. $A \wedge B \to C$ RCV.
5. $\neg(B \to C)$ assumption
6. $(B \to \neg C)$ CEM
7. $(A \wedge B \to \neg C)$ from 2, 6 and RCV.
8. $(A \wedge B \to \bot)$ 4, 7 and CK.
9. $B \to A \wedge B$ by 2 and ID.
10. $B \to \bot$ by 8, 9 and C1
11. $B \to C$ by CK. Contradicting 5.
12. $B \to C$ by reductio.

Proposition 4.4. S = RT + C2 + $\mathsf{L1}$

1. $A \to B$ assumption.
2. $B \to A$ assumption.
3. $A \to C$ assumption.

4. $A \wedge B \to C$ RCV.
5. $\neg(B \to C)$ assumption
6. $(B \to \neg C)$ CEM
7. $(A \wedge B \to \neg C)$ from 2, 6 and RCV.
8. $(A \wedge B \to \bot)$ 4, 7 and CK.
9. $B \to A \wedge B$ by 2 and ID.
10. $B \to \bot$ by 8, 9 and C1
11. $B \to C$ by CK. Contradicting 5.
12. $B \to C$ by reductio.

The most complicated part of theorem **??** is showing that S is the $\mathsf{L1}$ logic: C2 + CA.

We begin by showing that RCA is provable from CA and the weakest C axiom, $\mathsf{C3}$: if $A \vee B$ crashes then either A crashes or B crashes. C3 is, of course, just a special case of RCA. If you have C3 (and therefore if you have C2) you can prove RCA from $\mathsf{CA}' = \neg(A \vee B \to \neg C) \vdash (A \to C) \vee (B \to C)$ (which is obtained from CA by CEM):

Proposition 4.5. CA + C3 *entails* RCA *(in* $\mathsf{L1}$.*)*

1. $(A \vee B) \to C$ (assumption)
2. $(A \vee B) \to \neg C) \supset (A \vee B) \to \bot)$ from 1 and CK.
3. $((A \vee B) \to \bot) \supset (A \to \bot) \vee (B \to \bot)$ by C3
4. $(A \to \bot \vee B \to \bot) \supset (A \to C) \vee (B \to C)$ by CK and reasoning by cases.
5. $(A \vee B \to \neg C) \supset (A \to C) \vee (B \to C)$ 2, 3 and 4.
6. $\neg(A \vee B \to \neg C) \supset (A \to C) \vee (B \to C)$ by CA' (or CA+CEM.)
7. $(A \to C) \vee (B \to C)$ by 5, 6 and reasoning by cases

Theorem 4.6. $\mathsf{S} = \mathsf{CA} + \mathrm{C2} + \mathsf{L1}$

Observation: C2 gives us C3, and CA+C3 gives RCA by proposition [REF]. We begin by showing RCV is a theorem of $\mathsf{L1}$ + CA + C2.

1. $A \to B$ Ass
2. $A \to C$ Ass
3. $A \to CB$
4. $(AB \to CB) \vee (A\bar{B} \to CB)$ from 3 by RCA

5. $(A\bar{B} \to CB)$ (if the first disjunct is true, we are done.)
6. $(A\bar{B} \to \neg C)$ (since $A\bar{B} \to \bot$ by 5.)
7. $(AB \to C) \vee (AB \to \neg C)$ CEM
8. $AB \to \neg C$ (if the first disjunct is true we are done.)
9. $A \to \neg C$ from 6 and 8 and CA.
10. $A \to \bot$ from 9 and 2
11. $AB \to \bot$ by C2
12. $AB \to C$

Now we are in a position to prove CSO.[14]

1. $A \to B$ assumption
2. $B \to A$ assumption
3. $A \to C$ assumption
4. $AB \to C$ by RCV
5. $AB \to (A \supset C)$ by CK
6. $\bar{A}B \to \bar{A}B$ by ID
7. $\bar{A}B \to (A \supset C)$ by CK.
8. $B \to (A \supset C)$ from 5, 7 and CA
9. $B \to C$ from 2,8, and CK

The final aspect of theorem **??** is that $\mathsf{S} = \mathsf{RCA} + \mathrm{C0}$. It is sufficient to to show that RCA + C0 entails CA + C2. Proposition [REF] guarantees that C0 entails C2. So

1. $A \to C$ assumption
2. $B \to C$ assumption
3. $(A \vee B \to C) \vee (A \vee B \to \bar{C})$ by CEM
4. $A \vee B \to \bar{C}$ (left disjunct of 3 gives result.)
5. $A \to \bar{C} \vee B \to \bar{C}$ by RCA.
6. $A \to \bar{C}$ assumption
7. $A \to \bot$ from 1 and 6.

[14]The following argument, from CA and RCV to CSO is due to Burgess [REF].

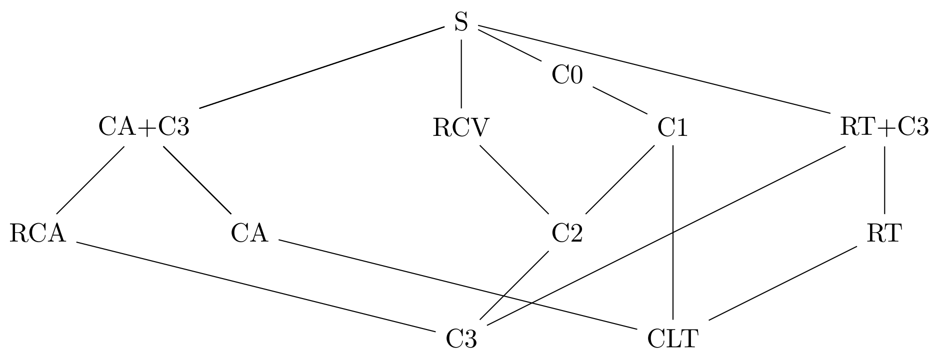
8. $A \vee B \to C$ from 2, 7 and C0.

9. $(A \to \bar{C}) \supset (A \vee B \to C)$ from 6 to 8.

10. $(B \to \bar{C}) \supset (A \vee B \to C)$ repeating 6 to 8 with B.

11. $A \vee B \to C$ from 5, 9, 10 and reasoning by cases.

4.2 The lattice of L1 logics



Note that RT does not entails C3, so RT is strictly weaker than RT+C3. Model of RT without C3: $f(\{a, b\}) = \emptyset$, $f(\{a\}) = \{a\}$, $f(\{b\}) = \{b\}$.

- CA = CA + CLT

5 Tenability

Here we define a class of probability frames that validate L1, C0 and S5 for crashing, in addition to satisfying Stalnaker's thesis. Here it will be useful to use Stalnaker's semantics in which f represents a selection function into W.

Given a probability space $\langle W, \Sigma, \mu\rangle$ we define a subspace of W to be those spaces of the form $\langle X, \Sigma \cap \mathcal{P}(X), \mu(\cdot \mid X)\rangle$ with $X \in \Sigma$. I shall write μ_X for $\mu(\cdot \mid X)$.

We need to employ a notion from measure theory – that of a measure-preserving map:

Definition 5.0.1. *Let X and Y be subspaces of W. A map, $t : X \to Y$, is measure preserving on the spaces $\langle X, \mu_X\rangle$, $\langle Y, \mu_Y\rangle$ iff (i) $t^{-1}(A)$ is measurable in X when A is in Y and (ii) $\mu_X(t^{-1}(A)) = \mu_Y(A)$ for each A in Y's sigma-algebra.*

As usual, the preimage of a set A under the function f, written $f^{-1}(A)$, is defined as $\{x \mid f(x) \in A\}$.

Definition 5.0.2. *A stretchy-rubber-model is a tuple $\langle W, \Sigma, \mu, t_A\rangle$ where*

1. *Σ is a σ-algebra over W,*

2. *μ a probability measure over Σ and for each non-empty $A \subseteq W$,*

3. *$t_A : \overline{A} \to A$, for each $A \subseteq W$,*

4. *t_A is measure preserving whenever $\mu(A) \in (0,1)$*

In a stretchy-rubber-model any set A which has measure in $(0,1)$ is stretchy: it can be stretched out (possibly cutting it up into pieces or gluing bits together) onto its complement in a way that preserves the measure of its measurable subsets. (In the models we consider any pair of sets, X and Y, with measures in $(0,1]$, can be stretched on to the other.)

We get a selection function from such a model by identifying $f(A,\cdot)$ with $id_A \cup t_A$ where id_A is the identity function on A, i.e. by setting:

$f(A,x) = x$ if $x \in A$ and $f(A,x) = t_A(x)$ if $x \in \bar{A}$ provided A is non-empty

$f(\emptyset, x) = \lambda$.

By construction $f(A,x) \in A$ and $f(A,x) = x$ whenever $x \in A$ so we automatically get L1. But notice further that A crashes ($f(A,x) = \lambda$) only if $A = \emptyset$, so the principles C0, B and 4 for crashing are validated in this kind of model as well.[15] So the logic of stretchy-rubber-models is at least L1+C0+S5; whether the logic of stretchy-rubber-models is exactly this logic bears further investigation.

It should be clear that Stalnaker's thesis holds in any stretchy-rubber-model. If $\mu(A) = 0$ then Stalnaker's thesis vacuously holds. If $\mu(A) = 1$ then (i) $\mu(B \mid A) = \mu(B)$ and (ii) $t_A^{-1}(B) \subseteq \bar{A}$ has measure 0 so $\mu(f^{-1}(A,B)) = \mu(AB) + \mu(t_A^{-1}(B)) = \mu(AB) = \mu(B)$. Suppose that $\mu(A) \in (0,1)$. Note that $f^{-1}(A,B) = id_A^{-1}(B) \cup t_A^{-1}(B) = AB \cup t_A^{-1}(B)$. Note that $\mu(t^{-1}(B) \mid \overline{A}) = \mu(B \mid A)$, since t_A is measure preserving, so $\mu(t^{-1}(B)) = \mu(B \mid A)\mu(\overline{A})$. Thus $f^{-1}(A,B)$ has a measure of $\mu(AB) + \mu(B \mid A)\mu(\overline{A}) = \mu(B \mid A)$.

5.1 Existence of a model

It suffices to prove that there are models, $\langle W, \Sigma, \mu, t_A \rangle$, of the form described. We shall work in a simple model in which W is just a some measurable set of reals with positive but finite measure, Σ as the Lebesgue measurable subsets of W, and μ the Lebesgue measure, λ, renormalised so that the measure of W is 1. For concreteness sake we shall set $W = [0,1]$ so that no renormalisation is needed.

For existence it suffices to prove the following[16]

Theorem 5.1. *Given any two measurable sets of reals, X and Y, of positive and finite measure there is a measure preserving function, f,from X to Y.*

[15] In my view neither C0 nor S5 are valid; however for the purposes of showing that a reasonable logic is consistent with Stalnaker's thesis this does not matter as every sublogic is also shown to be consistent.

[16] I am indebted to Gareth Davies here for some helpful suggestions regarding this proof.

This suffices to show that $\langle [0,1], \Sigma, \lambda \rangle$ can be extended with functions t_A to form a stretchy-rubber-model. If $\bar{A}$ and A have positive measure we can use this theorem to choose a measure preserving map t_A from $\bar{A}$ to A.

The basic idea for the proof of this theorem is to construct a pair of measure preserving maps, $f : X \to [0,1]$ and $h : [0,1] \to Y$, which can be composed to form a measure preserving map from X to Y. Things are more transparent if we define h in terms of a another measure preserving map, $g : Y \to [0,1]$. Here is how we define them:

- $f : X \to [0,1]$
- $g : Y \to [0,1]$
- $h : [0,1] \to Y$
- $f(x) = \mu_X((-\infty, x] \cap X)$
- $g(y) = \mu_Y((-\infty, y] \cap Y)$
- $h(\alpha) = \begin{cases} y & \text{if there is exactly one } y \text{ such that } g(y) = \alpha \\ w & \text{otherwise} \end{cases}$

here w is a simply a randomly selected member of Y. We will also make use of the following property of the Lebesgue measure:

Nifty fact: the Lebesgue measure, λ, is regular. This means that:

1. $\lambda(S) = \inf\{\lambda(O) \mid S \subseteq O, O \text{ is open}\}$
2. $\lambda(S) = \sup\{\lambda(C) \mid C \subseteq S, C \text{ is closed}\}$

Lemma 5.2. *f and g are measure preserving on open (and therefore closed) sets.*

Proof. Since g is defined exactly analogously to f it suffices to show that f is measure preserving on open sets.

Firstly note that by construction $\mu_X(f^{-1}((a,b))) = b - a$.

Let O be an open set. Since O is open, it may be written as a countable union of disjoint intervals, $\bigcup_i (a_i, b_i)$. So $\mu_X(f^{-1}(O)) = \mu_X(f^{-1}(\bigcup_i (a_i, b_i))) = \mu_X(\bigcup_i f^{-1}((a_i, b_i))) = \Sigma_i \mu_X(f^{-1}((a_i, b_i))) = \Sigma_i (b_i - a_i) = \lambda(O)$ as required.

Now let C be a closed set, so $C = [0,1] \setminus O$ for some open set O. So $\lambda(C) = 1 - \lambda(O) = 1 - \mu_X(f^{-1}(O)) = 1 - \mu_X(f^{-1}([0,1] \setminus C)) = 1 - (1 - \mu_X(f^{-1}(C))) = \mu_X(f^{-1}(C))$. So f is measure preserving on closed sets too. $\square$

Theorem 5.3. *f and g are measure preserving.*

Proof. Let $S \subset [0,1]$ be a measurable set. Then by regularity (form 1) and the fact that f is measure preserving on opens sets we have:

> $\lambda(S) = \inf\{\lambda(O) \mid S \subseteq O, O \text{ is open}\} = \inf\{\mu_X(f^{-1}(O)) \mid S \subseteq O, O \text{ is open}\} \geq \mu_X(f^{-1}(S))$

Then by regularity (form 2) and the fact that f is measure preserving on closed sets we have:

> $\lambda(S) = \sup\{\lambda(C) \mid C \subseteq S, C \text{ is closed}\} = \sup\{\mu_X(f^{-1}(C)) \mid C \subseteq S, C \text{ is closed}\} \leq \mu_X(f^{-1}(S))$

So $\lambda(S) = \mu_X(f^{-1}(S))$ as required. The argument that g is measure-preserving is exactly analogous. $\square$

Now to finish the argument we have

Theorem 5.4. *h is measure preserving.*

Proof. Suppose that $Z \subseteq Y$.

Our strategy will be to show that $\mu_Y(Z) = \mu_Y(g^{-1}(h^{-1}(Z)))$. This suffices since $\mu_Y(g^{-1}(h^{-1}(Z))) = \lambda(h^{-1}(Z))$ by the fact that g is measure preserving. Here goes.

$g^{-1}(h^{-1}(Z)) = \{y \mid g(y) \in h^{-1}(Z)\} = \{y \mid \exists! z : g(z) = g(y) \text{ and } z \in Z\} = Z \setminus \{y \mid g(y) = g(z) \text{ for some } z \neq y\} = Z \setminus g^{-1}(\{\alpha \mid |g^{-1}(\{\alpha\})| > 1\})$.

Now note that the set $S := \{\alpha \mid |g^{-1}(\{\alpha\})| > 1\}$ is countable. We can map S injectively into $\mathbb{Q}$ as follows: if $\alpha \in S$, then since $|g^{-1}(\{\alpha\})| > 1$ there is a rational number, q, strictly inside the convex hull of $g^{-1}(\{\alpha\})$. So we can map α to q. This mapping is injective because g is increasing: if $\alpha < \beta$ then the convex hull of $g^{-1}(\{\alpha\})$ and of $g^{-1}(\{\beta\})$ overlap at most at a boundary point (since, if $\alpha < \beta$, $g(x) = \alpha$ and $g(y) = \beta$ then $x \leq y$) and we have chosen q not to be a boundary point.

Now, of course, $\{\alpha\}$ has Lebesgue measure 0, so $\mu_Y(g^{-1}(\{\alpha\})) = 0$ since g is measure preserving. So $g^{-1}(\{\alpha \mid |g^{-1}(\{\alpha\})| > 1\})$ is a countable union of null sets, and is thus a null set. So putting this all together we have $\mu_Y(g^{-1}(h^{-1}(Z))) = \mu_Y(Z \setminus g^{-1}(\{\alpha \mid |g^{-1}(\{\alpha\})| > 1\})) = \mu_Y(Z) - 0 = \mu_Y(Z)$.

So $\mu_Y(Z) = \mu_Y(g^{-1}(h^{-1}(Z)))$. $\square$

This completes the proof. To obtain a measure preserving map, t, from X to Y we simply let $t = h \circ f$.

References

[1] B.F. Chellas. Basic conditional logic. *Journal of philosophical logic*, 4(2):133–153, 1975.

[2] A. Hájek and N. Hall. The hypothesis of the conditional construitl of conditional probability. *Probability and conditionals: belief revision and rational decision*, page 75, 1994.

[3] D. Lewis. Counterfactuals oxford, 1973.

[4] M. Schulz. Counterfactuals and arbitrariness.

[5] R. Stalnaker. Letter to van fraassen. *WL Harper and CA Hooker (1976)*, pages 302–306, 1976.

[6] R.C. Stalnaker. Probability and conditionals. *Philosophy of Science*, pages 64–80, 1970.

[7] B.C. Van Fraassen. Probabilities of conditionals. *Foundations of probability theory, statistical inference, and statistical theories of science*, 1:261–308, 1976.