Modal Algebraic Models of Counterfactuals

Stefano Bonzio[1], Tommaso Flaminio[2], and Giuliano Rosella[3]

[1] Department of Mathematics and Computer Science, University of Cagliari, Italy
stefano.bonzio@unica.it
[2] Artificial Intelligence Research Institute (IIIA), CSIC, Bellaterra, Spain
tommaso@iiia.csic.es
[3] Department of Philosophy and Education Sciences, University of Turin, Italy
giuliano.rosella@unito.it

The aim of the present work is to put forward an algebraic approach to counterfactual conditionals (or simply *counterfactuals* from now on) based on Boolean Algebras of Conditionals as defined in [3]. A Boolean algebra of Conditionals (BAC), $\mathcal{C}(\mathbf{A})$, is a Boolean algebra obtained starting from any Boolean algebra $\mathbf{A} = \langle A, \wedge, \vee, \neg, \top, \bot \rangle$ and taking a certain quotient of the free Boolean algebra generated by the pairs $(a, b) \in A \times A$ with $b \neq \bot$. Each basic element in a BAC is identified with $(a \mid b)$ which is intended to represent the conditional event "a given b", where b is the *antecedent* and a the *consequent*. The framework of BACs offers an innovative and privileged perspective on conditionals events: as it is shown in [3], BACs are a valuable tool to analyze the algebraic properties of conditionals events, their logic and their relation with probability measures.

The framework of BACs, although promising, is not yet fully developed in all its potentialities. Our goal is to extend BACs in order to account for *counterfactual* conditional events. More precisely, we consider a normal modal operator $\Box$ on a BAC so defining *modal* Boolean Algebras of Conditionals $\langle \mathcal{C}(\mathbf{A}), \Box \rangle$ that we name *Lewis algebras*. We investigate the properties of these new structures and the resulting logic of counterfactuals. Our idea is motivated by the fact that, although counterfactuals are not captured by BACs, a normal modal operator $\Box$, when combined with the algebraic properties of conditional events, logically behaves very similarly to the counterfactual conditional operator $\Box\!\!\rightarrow$ in David Lewis's semantics for counterfactuals (see [5] and [6]) so as to interpret a Lewis' counterfactual $b \Box\!\!\rightarrow a$ as $\Box(a \mid b)$ in a modal BAC. By doing so, already basic properties of counterfactuals can be proved to hold in our modal framework. For instance, $\Box(a \mid b) \wedge \Box(c \mid b) = \Box(a \wedge c \mid b)$ holds in every modal BAC and analogously $((b \Box\!\!\rightarrow a) \wedge (b \Box\!\!\rightarrow c)) \leftrightarrow (b \Box\!\!\rightarrow (a \wedge c))$ is valid in Lewis' semantics for counterfactuals.

Starting from this construction, we analyze the properties of the dual Kripke frame of Lewis algebras, in the sense of Jónsson-Tarski (see [7]). In particular, for a Lewis algebra $\langle \mathcal{C}(\mathbf{A}), \Box \rangle$, we discuss what conditions should be imposed on $\Box$, in order to characterize Lewis' different logics for counterfactuals, and what properties these conditions imply on the dual frame. In particular, we show that if a Lewis algebra $\langle \mathcal{C}(\mathbf{A}), \Box \rangle$, satisfies the following identities:

(1) $\Box(a \mid \top) = (a \mid \top)$

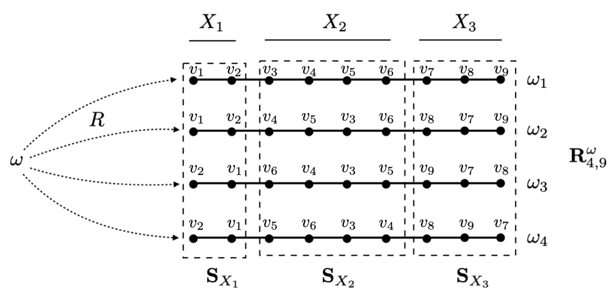
(2) $\Box(a \mid a \vee b) \vee \Box(b \mid a \vee b) \vee (\Box(c \mid a \vee b) \to \Box((c \mid a) \wedge (c \mid b))) = 1$

then, the resulting modal logic of conditionals corresponds to a slightly stronger logic than the system C1, that Lewis himself claims to be the "correct logic of counterfactual conditionals" (see [5, p. 80]).

Slightly more formally, for every Boolean algebra $\mathbf{A}$, we call *Lewis algebra* any modal BAC $\mathcal{L}(\mathbf{A}) = \langle \mathcal{C}(\mathbf{A}), \Box \rangle$ satisfying (1) and (2). The dual frame $\langle \mathrm{at}(\mathcal{C}(\mathbf{A})), R \rangle$ of $\mathcal{L}(\mathbf{A})$ will be called a *Lewis Frame* and we denote it by $\mathrm{F}_{\mathcal{L}(\mathbf{A})}$. We then show how the above (1) and (2) characterize specific properties of Lewis frames. In particular, if $\mathbf{A}$ is a finite Boolean algebra with atoms

$v_1, \ldots, v_n$, a Lewis frame takes the form $\mathcal{F}_{\mathcal{L}(\mathbf{A})} = \langle \mathrm{at}(\mathcal{C}(\mathbf{A})), \mathrm{R} \rangle$ where $\mathrm{at}(\mathcal{C}(\mathbf{A}))$ denotes the finite set of atoms of the BAC $\mathcal{C}(\mathbf{A})$. In such a case, $\mathrm{F}_{\mathcal{L}(\mathbf{A})}$ validates (1) iff R is serial and each $\omega \in \mathrm{at}(\mathcal{C}(\mathbf{A}))$ only accesses to worlds that have the same initial element as ω. As for the latter, recall that the atoms of a BAC, $\mathcal{C}(\mathbf{A})$, can be identified with strings of maximal length of atoms in $\mathbf{A}$, i.e. $\omega = \langle v_1, v_2, \ldots, v_n \rangle$.

Thus, (1) characterizes the following properties of $\mathrm{F}_{\mathcal{L}(\mathbf{A})}$: (*i*) for all $\omega = \langle v_1, v_2, \ldots, v_n \rangle \in \mathrm{at}(\mathcal{C}(\mathbf{A}))$ there is a $\omega' = \langle v'_1, v'_2, \ldots, v'_n \rangle$ such that $\omega R \omega'$; and (*ii*) if $\omega = \langle v_1, v_2, \ldots, v_n \rangle$ and $\omega' = \langle v'_1, v'_2, \ldots, v'_n \rangle$ are such that $\omega R \omega'$, then $v_1 = v'_1$. Condition (2) characterizes a property on Lewis frames that we call *sphericity*. This property defines, for each $\omega \in \mathrm{at}(\mathcal{C}(\mathbf{A}))$, a certain composition of the set $R[\omega] = \{\omega' \in \mathrm{at}(\mathcal{C}(\mathbf{A})) \mid \omega \mathrm{R} \omega'\}$ of accessible worlds from ω. More precisely, one can easily display the elements of $R[\omega]$ in a finite matrix (see the figure below for an example). This will be called the *matrix generated by* $R[\omega]$, and it will be denoted by $\mathbf{R}^{\omega}_{k,n}$. If $\mathrm{F}_{\mathcal{L}(\mathbf{A})}$ satisfies (2) then, for all $\omega \in \mathrm{at}(\mathcal{C}(\mathbf{A}))$, $\mathbf{R}^{\omega}_{k,n}$ can be partitioned into submatrices such that they do not share any element with each other and each of them contains the same elements in all its rows and its columns. Although sphericity has an intricate formulation, it is easier to grasp with a graphical example:



The matrix $\mathbf{R}^{\omega}_{k,n}$ is induced by the sphericity condition as it can be partioned into disjoint cells $\mathbf{S}_{X_1}, \mathbf{S}_{X_2}$ and $\mathbf{S}_{X_3}$ and each of them contains the same elements in its columns and its rows. Hence, we get that a Lewis frames validates (2) iff it satisfies sphericity.

The semantic conditions (with respect to a Lewis frames) for a counterfactual of the form $\Box(\varphi \mid \psi)$ correspond to the usual modal Kripke-semantic conditions: $\Box(\varphi \mid \psi)$ is true at ω iff $(\varphi \mid \psi)$ is true at all the ω' such that $\omega R \omega'$[1], and the semantic conditions for Boolean combinations of formulas are the usual classical ones. Given the characterization of the class of Lewis frames, we show how to go back and forth from Lewis frames to sphere models for counterfactuals[2]. In particular, we show that each Lewis frame $\mathcal{F}_{\mathcal{L}(\mathbf{A})}$ corresponds to a sphere model $\mathcal{M}$ satisfying exactly the same counterfactuals formulas as $\mathcal{F}_{\mathcal{L}(\mathbf{A})}$, and viceversa. This correspondence between the two semantic frameworks allows us to prove soundness and completeness of the logic $\mathsf{C1}^+ = \mathsf{C1} + \Box(\varphi \mid \psi) \to \Diamond(\varphi \mid \psi)$ (for all ψ such that $\psi \not\leftrightarrow \bot$) with respect to Lewis frames and Lewis algebra.

The above results represent a step towards an algebraic approach to counterfactual conditionals. Although the research on the semantics of counterfactuals has been prolific (see for instance [1], [2] and [4]), an algebraic framework to analyze counterfactual conditionals is, to the best of the authors' knowledge, still missing. In the present work, we have tried to start filling this gap. Finally, we will see how Lewis algebras can contribute to understanding the uncertain quantifications of counterfactuals by analyzing how a belief function P behaves on a Lewis algebra, so as to represent the uncertainty of a counterfactual as $P(\Box(a \mid b))$.

[1] For an analysis of the truth conditions of a conditional $(\varphi \mid \psi)$ with respect to a $\omega = \langle v_1, \ldots, v_n \rangle$, see [3].

[2] See [6] for more details the sphere-based semantics for counterfactuals.

References

[1] Fausto Barbero and Gabriel Sandu. Team semantics for interventionist counterfactuals: Observations vs. interventions. *Journal of Philosophical Logic*, 50(3):471–521, dec 2020.

[2] Rachael Briggs. Interventionist counterfactuals. *Philosophical Studies*, 160(1):139–166, may 2012.

[3] Tommaso Flaminio, Lluis Godo, and Hykel Hosni. Boolean algebras of conditionals, probability and logic. *Artificial Intelligence*, 286:103347, sep 2020.

[4] Joseph Y. Halpern. Axiomatizing causal reasoning. *Journal of Artificial Intelligence Research*, 12:317–337, 2000.

[5] David Lewis. Completeness and decidability of three logics of counterfactual conditionals. *Theoria*, 37(1):74–85, 1971.

[6] David Lewis. *Counterfactuals*. Cambridge, MA, USA: Blackwell, 1973.

[7] Hiroakira Ono. *Proof Theory and Algebra in Logic*. Singapore: Springer Singapore, 2019.