DAVID LEWIS MEETS ARTHUR PRIOR AGAIN

KATSUHIKO SANO

Graduate School of Letters
Kyoto University
katsuhiko.sano@kx3.ecs.kyoto-u.ac.jp

1. Introduction and Motivation

This paper proposes a hybridization of David Lewis's counterfactual logic (Lewis 1973). As far as the author knows, in the literature of both hybrid and conditional logic (see, e.g., Areces and ten Cate 2007 and Nute and Cross 2001, respectively), such combination has never been studied. It, however, deserves to be studied since this hybridization enables us to formalize the following inference:

$$\frac{\text{The pig is Mary.} \quad \text{Mary is pregnant.}}{\text{Therefore: The pig is pregnant.}} \tag{1}$$

We can regard this inference as updating the local information, depending on the given situation (e.g., speaker), by using the global information, independent of the situation. To deal with the sentences containing 'the', we make use of David Lewis's egocentric reading of counterfactual connectives. In order to deal with the second sentence, we need the modern hybrid formalism, whose roots trace back to Arthur Prior (see, e.g., Blackburn 2006). In addition, our hybridization has some technical merits: (1) Theorem 1: Completeness and decidability are preserved; (2) Theorem 2: We can characterize *the Limit Assumption* (saying that there are the closest worlds with respect to the possible antecedent of counterfactuals, and David Lewis rejected it metaphysically) by some *proof-rule*, used by hybrid logicians to obtain a general Kripke completeness result for pure formulas (see, e.g., Areces and ten Cate 2007, Theorem 5).

2. David Lewis's Analysis of Contextually Definite Descriptions

It is well known that David Lewis proposed that the counterfactual conditional $\varphi \mathrel{\Box\!\!\rightarrow} \psi$ (read *'If it were the case that φ, then it would be the case that ψ'*) is true at a world w iff $(\varphi \wedge \psi)$-worlds are closer to w than $(\varphi \wedge \neg\psi)$-worlds (Lewis 1973). To define the 'relative closeness' of w rigorously, we need his 'system of spheres' representing a *comparative similarity* between worlds. A pair $\langle W, \$ \rangle$ is *a system of spheres* iff $W \neq \emptyset$ and $\$: W \to \mathcal{P}(\mathcal{P}(W))$ satisfies the following (we write '$\$_w$' instead of '$\$(w)$'): $(S1)$ $\$_w$ *is nested*: $S, T \in \$_w \implies S \subset T$ or $T \subset S$; $(S2)$ $\$_w$ *is closed under unions*: $(S_\lambda)_{\lambda\in\Lambda} \subset \$_w \implies \bigcup_{\lambda\in\Lambda} S_\lambda \in \$_w$; $(S3)$ $\$_w$ *is closed under (nonempty)* *intersections*: $(S_\lambda)_{\lambda\in\Lambda} \subset \$_w$ and $\Lambda \neq \emptyset \implies \bigcap_{\lambda\in\Lambda} S_\lambda \in \$_w$. Given a valuation $V : \mathsf{Prop} \to \mathcal{P}(W)$ (where Prop is the set of all proposition letters), we can formulate the truth condition of the counterfactual conditional as follows:

$$w \in [\![\varphi \mathrel{\Box\!\!\rightarrow} \psi]\!]_{\langle W,\$,V\rangle} \iff \begin{cases} \text{(A)} \bigcup \$_w \cap [\![\varphi]\!]_{\langle W,\$,V\rangle} = \emptyset \text{ or} \\ \text{(B)} \ (\exists S \in \$_w)\,[S \cap [\![\varphi]\!]_{\langle W,\$,V\rangle} \neq \emptyset \text{ and } S \cap [\![\varphi]\!]_{\langle W,\$,V\rangle} \subset [\![\psi]\!]_{\langle W,\$,V\rangle}], \end{cases}$$

where $[\![\varphi]\!]_{\langle W,\$,V\rangle}$ is the denotation of φ defined relatively to $\langle W, \$ \rangle$ and V. We usually drop the subscript from $[\![\varphi]\!]_{\langle W,\$,V\rangle}$ when it is clear from the context.

David Lewis (Lewis 1973, sec.5.3) considered Arthur Prior's egocentric reading of sentences and proposed that his counterfactual connective expresses *contextually definite descriptions* (e.g., 'The pig is pregnant'), whose logical form is '*The x such that φ is such that ψ*'. To be more accurate, he used the connective $\varphi \mathrel{\Box\!\!\Rightarrow} \psi$ defined as $\neg(\varphi \mathrel{\Box\!\!\rightarrow} \neg\varphi) \wedge (\varphi \mathrel{\Box\!\!\rightarrow} \psi)$, whose truth condition corresponds exactly to the case (B) above ('$\neg(\varphi \mathrel{\Box\!\!\rightarrow} \neg\varphi)$' means that φ is possible). According to this egocentric reading, the truth of sentence is relativised to a thing or an individual, and so, the truth of sentence φ at x means that the individual x has the property φ. Then, a system of spheres around x represents its *comparative salience*, i.e., x's degree of familiarity between things and individuals. Thus, 'The pig is grunting', formalized as '$\mathsf{Pig} \mathrel{\Box\!\!\Rightarrow} \mathsf{Grunting}$', is true at an individual x iff the grunting pig is more salient to x than the not-grunting pigs.

Furthermore, in Lewis's analysis, we can deal with a *sequence* of egocentric conditionals (Lewis 1973, p.114): Suppose that you are walking past a piggery.

The pig is grunting.
$(\mathsf{Pig} \mathrel{\Box\!\!\Rightarrow} \mathsf{Grunting})$
The pig with floppy ears is not grunting.
$((\mathsf{Pig} \wedge \mathsf{Floppy}) \mathrel{\Box\!\!\Rightarrow} \mathsf{Grunting})$
The spotted pig with floppy ears is grunting.
$((\mathsf{Pig} \wedge \mathsf{Floppy} \wedge \mathsf{Spotted}) \mathrel{\Box\!\!\Rightarrow} \mathsf{Grunting})$

According to the usual analysis of definite description as $\mathsf{Grunting}(\imath x \mathsf{Pig}(x))$, however, we cannot deal with such a sequence, since we never make both $\mathsf{Grunting}(\imath x \mathsf{Pig}(x))$ and $\neg\mathsf{Grunting}(\imath x(\mathsf{Pig}(x) \wedge \mathsf{Floppy}(x)))$ true at the same time.

3. Hybrid Counterfactual Logic: David Lewis Meets Arthur Prior Again

David Lewis's counterfactual logic blends with modern hybrid logic in a surprisingly natural way. This explains the title of the present paper (see Blackburn 2006 for an in-depth explanation of connections between Prior's ideas, description and hybrid logics). Hybrid systems introduce nominals i (names for states) and satisfaction operators $@_i p$ (p is true at the state named by i) and formalize 'Mary is pregnant' as

$@_{\mathsf{MARY}}\mathsf{Pregnant}$. In reformulating Prior's egocentric reading, Lewis also dealt with a similar kind of sentence (Lewis 1973, p.112): '*x* is such that (the Anighito meteorite is an *x* such that *x* is a rock)'. Familiarity with hybrid formalism would allow Lewis to write this sentence in most compact way possible: $@_{\mathsf{ANIGHITO\ METEORITE}}\mathsf{Rock}$. Here we can make David Lewis meet Arthur Prior again.

Thus, we can formalize our motivating inference (1) as follows:

$$[(\mathsf{Pig} \mathrel{\Box\!\!\Rightarrow} \mathsf{MARY}) \wedge @_{\mathsf{MARY}}\mathsf{Pregnant}] \rightarrow (\mathsf{Pig} \mathrel{\Box\!\!\Rightarrow} \mathsf{Pregnant}). \qquad (2)$$

Figures 1 and 2 suggest that this formula is *valid*, i.e., true at any individual $w \in W$ and for any system of spheres $\langle W, \$ \rangle$. Note that the notion of *valuation* is same as before except that $V(i)$ is a singleton for any nominal i and that $w \in [\![@_i\varphi]\!]$ iff $v \in [\![\varphi]\!]$ where v is the denotation of i. In Figure 1, the dotted-lines express that the truth of '$@_{\mathsf{MARY}}\mathsf{Pregnant}$' is independent of the given individual w.

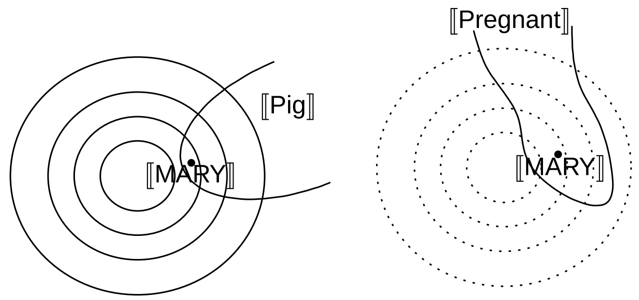

Figure 1 $w \in [\![\mathsf{Pig} \mathrel{\Box\!\!\Rightarrow} \mathsf{MARY}]\!]$ and $w \in [\![@_{\mathsf{MARY}}\mathsf{Pregnant}]\!]$

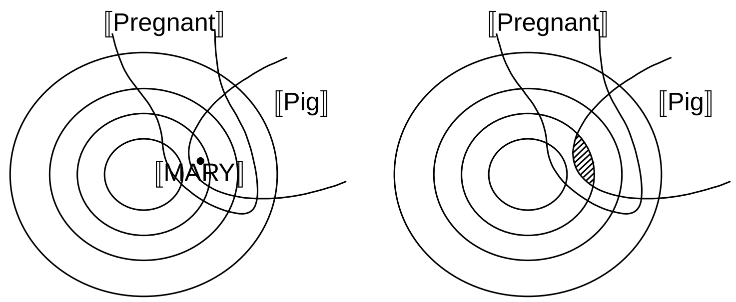

Figure 2 $w \in [\![\mathsf{Pig} \mathrel{\Box\!\!\Rightarrow} \mathsf{Pregnant}]\!]$

Next, we can give an axiomatization of hybrid counterfactual logic $\mathbf{V}_{\mathcal{H}(@)}$ (see the table below) that extends David Lewis's **V** (Lewis 1973, ch.6). Let us derive (2) as

Axioms and rules for $\mathbf{V}_{\mathcal{H}(@)}$

CT	$\vdash \varphi$, for all classical tautologies φ
K@	$\vdash @_i(p \to q) \to (@_i p \to @_i q)$
Self-Dual	$\vdash \neg @_i p \leftrightarrow @_i \neg p$
Ref	$\vdash @_i i$
Intro	$\vdash @_i p \to (i \to p)$
Agree	$\vdash @_i @_j p \to @_j p$
Back	$\vdash @_i p \to (q \Box\!\!\rightarrow @_i p)$
ID	$\vdash p \Box\!\!\rightarrow p$
MOD	$\vdash (\neg p \Box\!\!\rightarrow p) \to (q \Box\!\!\rightarrow p)$
ARR	$\vdash \neg(p \Box\!\!\rightarrow \neg q) \to [((p \wedge q) \Box\!\!\rightarrow r) \leftrightarrow (p \Box\!\!\rightarrow (q \to r))]$
MP	If $\vdash \varphi \to \psi$ and $\vdash \varphi$, then $\vdash \psi$
DwC	If $\vdash (\theta_1 \wedge \cdots \wedge \theta_n) \to \psi$, then $\vdash ((\varphi \Box\!\!\rightarrow \theta_1) \wedge \cdots \wedge (\varphi \Box\!\!\rightarrow \theta_n)) \to (\varphi \Box\!\!\rightarrow \psi)$ $(n \geq 1)$
ILE	If $\vdash \varphi \leftrightarrow \psi$, then $\vdash (\varphi \Box\!\!\rightarrow \theta) \leftrightarrow (\psi \Box\!\!\rightarrow \theta)$.
Nec@	If $\vdash \varphi$, then $\vdash @_i \varphi$
Sub	If $\vdash \varphi$, then $\vdash \varphi\sigma$, where σ denotes a substitution that uniformly replaces proposition letters by formulas and nominals by nominals.

a theorem of $\mathbf{V}_{\mathcal{H}(@)}$. By **Intro**, we have $@_{\mathsf{MARY}}\mathsf{Pregnant} \to (\mathsf{MARY} \to \mathsf{Pregnant})$. Then, we apply **DwC** to this and get:

$$(\mathsf{Pig} \Box\!\!\rightarrow @_{\mathsf{MARY}}\mathsf{Pregnant}) \to [(\mathsf{Pig} \Box\!\!\rightarrow \mathsf{MARY}) \to (\mathsf{Pig} \Box\!\!\rightarrow \mathsf{Pregnant})].$$

But, from **Back**, we have:

$$@_{\mathsf{MARY}}\mathsf{Pregnant} \to (\mathsf{Pig} \Box\!\!\rightarrow @_{\mathsf{MARY}}\mathsf{Pregnant}).$$

Thus, from two formulas above, we can derive:

$$@_{\mathsf{MARY}}\mathsf{Pregnant} \to [(\mathsf{Pig} \Box\!\!\rightarrow \mathsf{MARY}) \to (\mathsf{Pig} \Box\!\!\rightarrow \mathsf{Pregnant})].$$

By recalling the definition of $\varphi \Box\!\!\Rightarrow \psi := \neg(\varphi \Box\!\!\rightarrow \neg\varphi) \wedge (\varphi \Box\!\!\rightarrow \psi)$ and using some inference of propositional logic, we can derive (2).

4. Technical Merits

In the previous section, we have revealed that (2) is semantically valid and that (2) is a theorem of $\mathbf{V}_{\mathcal{H}(@)}$. In this section, we will connect the notion of *validity* with the notion of *theorem*. That is, we will establish completeness (and decidability at the same time) of our logic. First of all, we can easily prove the soundness of our logic by induction on $\vdash \varphi$.

Proposition 1 (Soundness)**.** $\mathbf{V}_{\mathcal{H}(@)}$ is sound with respect to the class of sphere models. That is, for any φ, $\vdash_{\mathbf{V}_{\mathcal{H}(@)}} \varphi \implies [\llbracket\varphi\rrbracket = W$ for any sphere model $\langle W, \$, V \rangle]$.

We can also prove the following completeness result:

Theorem 1 (Completeness and Decidability)**.** $\mathbf{V}_{\mathcal{H}(@)}$ is complete with respect to the class of finite sphere models. That is, for any φ, $[\llbracket\varphi\rrbracket = W$ for any finite sphere model $\langle W, \$, V \rangle] \implies \vdash_{\mathbf{V}_{\mathcal{H}(@)}} \varphi$. Therefore, $\mathbf{V}_{\mathcal{H}(@)}$ is decidable.

Here '*a finite sphere model*' means a sphere model whose domain is a finite set.

Sketch of Proof. In sum, our completeness proof is a combination of Lewis's completeness proof for counterfactual logic via the *selection functions* (Lewis 1973, pp.132-4) (roughly, multimodal Kripke frame having a binary relation R_φ for every formula φ) and ten Cate et al. 2005's technique for completeness proof for hybrid logic. First, we prove that $\mathbf{V}_{\mathcal{H}(@)}$ is complete with respect to the class of models based on selection function by ten Cate et al. 2005's technique. Counterfactual vocabulary fits this argument. Second, we construct a sphere model from a countermodel based on a selection function in a truth-preserving way (for this construction, see (Lewis 1973, sec.2.7)). Hybrid vocabulary does not affect this technique at all. Finally, we filter our sphere model down to a finite sphere model by filtration technique (Lewis 1973, sec.6.2). Hybrid vocabulary also fits this technique. □

Another merit of our hybridization is related to *the Limit Assumption* saying that there are the closest worlds with respect to the possible antecedent of counterfactuals. To be more precise, a system of sphere $\langle W, \$ \rangle$ satisfies *the Limit Assumption* (LA) iff, for any $w \in W$ and any $X \subset W$,

$$\bigcup \$_w \cap X \neq \emptyset \implies \bigcap \{ S \in \$_w \mid S \cap X \neq \emptyset \} \cap X \neq \emptyset.$$

David Lewis rejected it metaphysically (Lewis 1973, sec.1.4), but stated that there exists no *characteristic axiom* associated with it (Lewis 1973, sec.6.1, p.121). The same situation also occurs in our hybrid counterfactual logic. We say that formula φ *corresponds to* a property Q of systems of sphere if, for any $\langle W, \$ \rangle$, $\langle W, \$ \rangle$ satisfies $Q \iff \llbracket\varphi\rrbracket_{\langle W, \$, V \rangle} = W$ for any valuation V on $\langle W, \$ \rangle$. Note that any finite system of spheres trivially satisfies (LA) by $(S1)$: $\$_w$ is nested. Then, we can prove that there exists *no* formula φ of hybrid counterfactual logic such that φ corresponds to (LA): Suppose for contradiction that there exists such a formula φ. Consider $\langle \mathbb{R}, \$ \rangle$ where $\$_r := \{ (r-\varepsilon, r+\varepsilon), [r-\varepsilon, r+\varepsilon] \mid \varepsilon > 0 \} \cup \{ \{ r \}, \emptyset, \mathbb{R} \}$. $\langle \mathbb{R}, \$ \rangle$ is a system of sphere but fails to satisfy (LA). By definition, for some valuation V on $\langle \mathbb{R}, \$ \rangle$, $\llbracket\varphi\rrbracket_{\langle \mathbb{R}, \$, V \rangle} \neq \mathbb{R}$. From Proposition 1, $\nvdash_{\mathbf{V}_{\mathcal{H}(@)}} \varphi$. Then, by Theorem 1, $\llbracket\varphi\rrbracket_{\langle W', \$', V' \rangle} \neq W'$ for some finite sphere model $\langle W', \$', V' \rangle$. However, since $\langle W', \$', V' \rangle$ satisfies (LA) trivially, $\llbracket\varphi\rrbracket_{\langle W', \$', V' \rangle} = W'$. Contradiction.

We can, however, characterize (LA) by the following *proof-rule*:

CBG If $\vdash @_i\neg(\psi \Box\!\!\rightarrow \neg j) \rightarrow @_j\varphi$, then $\vdash @_i(\psi \Box\!\!\rightarrow \varphi)$, where $i \neq j$ and j does not appear in φ and ψ.

We say that $\langle W, \$ \rangle$ *admits* **CBG** if every valuation falsifying the consequent $@_i(\psi \Box\!\!\rightarrow \varphi)$ can be extended to a valuation falsifying the antecedent $@_i\neg(\psi \Box\!\!\rightarrow \neg j) \rightarrow @_j\varphi$.

Theorem 2. $\langle W, \$ \rangle$ satisfies (LA) $\Longleftrightarrow$ $\langle W, \$ \rangle$ admits **CBG**.

This characterization is inspired by ten Cate and Litak 2007's characterization of the topological equivalent of the relational **S4**-frames (i.e., Alexandrov spaces) by the proof-rule called **BG**. We can prove this theorem as in (ten Cate and Litak 2007, Theorem 3.4). By this result, we claim that Lewis's rejection of (LA) would result in his non-acceptance of **CBG**.

5. Conclusion

We have argued that nominals fit naturally into Lewis formalism and their introduction is a desirable step. If our main thesis is true and we reject (LA) as Lewis, it means that work on topological and neighborhood semantics for hybrid logic opens new perspective for Lewis's counterfactual semantics.

Acknowledgements

I would like to thank Mamoru Kaneko, Graham Priest, Takeshi Sakon, Tomoyuki Kaneko, Kohei Kishida, Nobu-Yuki Suzuki, Tadeusz Litak for fruitful discussions and Patrick Blackburn for his encouragement.

Bibliography

Areces, C. and ten Cate, B.: 2007, Hybrid logics, in P. Blackburn, J. van Benthem, and F. Wolter (eds.), *Handbook of Modal Logic*, pp 821–868, Elsevier

Blackburn, P.: 2006, Arthur Prior and hybrid logic, *Synthese* 150(3), 329–372

Lewis, D.: 1973, *Counterfactuals*, Blackwell Publishing, 2 edition, 2nd edition was pubulished in 1986.

Nute, D. and Cross, C. B.: 2001, Conditional logic, in D. Gabbay and F. Guenthner (eds.), *Handbook of Philosophical Logic*, Vol. 4, pp 1–98, Kluwer Academic Publishers, 2 edition

ten Cate, B. and Litak, T.: 2007, Topological perspective on the hybrid proof rules, *Electronic Notes in Theoretical Computer Science* 174(6), 79–94, Proceeding of International workshop on hybrid logic 2006

ten Cate, B., Marx, M., and Viana, P.: 2005, Hybrid logics with Sahlqvist axioms, *Logic Journal of the IGPL* 13(3), 293–300