A note on Brute vs. Institutional Facts: Modal Logic of Equivalence Up To a Signature

Davide Grossi

Institute of Logic, Language and Computation
University of Amsterdam
d.grossi@uva.lu

Abstract. The paper investigates the famous Searlean distinction between "brute" and "institutional" concepts from a logical point of view. We show how the partitioning of the non-logical alphabet—e.g., into "brute" and "institutional" atoms—gives rise to interesting modal properties. A modal logic, called UpTo-logic, is introduced and investigated which formalizes the notion of (propositional) logical equivalence *up to* a given signature.

1 Introduction

In the last decade the logical analysis of constitutive rules, initiated by [9], has focused on a number of aspects: defeasibility [3, 4], contextual and classificatory aspects [7, 8], mental aspects [12]. The prominent view has been to study constitutive rules, or "counts-as statements", as logical conditionals of the form $\varphi_1 \Rightarrow \varphi_2$ where the logic of $\Rightarrow$ was, from case to case, capturing the aforementioned aspects. One aspect, though, that has up to now been neglected concerns the different linguistic nature of the antecedent φ_1 and the consequent φ_2 of such conditionals.

According to Searle [14, 15] a characteristic aspect of constitutive rules is to link brute facts to institutional ones. Antecedent and consequent belong, somehow, to two different sets of concepts into which the language of institutions can be split. Institutional facts are constituted on the top of brute ones, giving to brute ones some sort of 'priority' upon the institutional ones.

The present paper explores, using modal logic, this linguistic aspect of constitutive rules. It develops ideas already introduced and partially investigated in [5,6]. It is structured as follows. Section 2 introduces the notion of equivalence *up to* a given propositional signature. Such notion is then semantically and axiomatically studied in a multi-modal language in Section 3. Section 4 discusses some related work and draws some conclusions.

2 Formal aspects of the brute vs. institutional distinction

In this section Searle's thesis concerning the distinction of brute and institutional facts is related to a specific notion of logical equivalence.

Dagstuhl Seminar Proceedings 09121
Normative Multi-Agent Systems
http://drops.dagstuhl.de/opus/volltexte/2009/1910

2.1 Counts-as conditionals, brute, and institutional facts

Let us start off with one of Searle's paradigmatic examples of a constitutive rule, the one concerning the institution of promising:

> Under certain conditions C anyone who utters the words (sentence) "I hereby promise to pay you, Smith, five dollars" promises to pay Smith five dollars [13, p. 44].

So, in context C the brute fact of uttering "I hereby promise" is a sufficient condition for the institutional fact of promising to occur. Following [7, 8] by interpreting contextual statements as forms of localized propositional validity, this can be semantically rendered as:

$$W_C \models \texttt{utter} \rightarrow \texttt{promise} \tag{1}$$

where W_C is the set of states modeling context C.[1] Now, `utter` belongs to the set BR of "brute" atoms, while `promise` to the set IN of "institutional ones". In the Spirit of Searle, sets BR and INS should obviously be taken to be disjoint, and to cover the set $\mathbf{P}$ of atoms of the language.

So where does the priority of BR in constituting the elements of IN arise? The thesis of this paper—already partially put forth in [6]—is that the priority of BR over IN consists in implications such as $\texttt{utter} \rightarrow \texttt{promise}$ in Formula 1 to cease to be valid once only the "brute" sublanguage, i.e., the atoms in BR, is considered. With respect to Formula 1, this means that counts-as conditionals imply the existence of a state w in context W_C and a state w' such that w and w' are indistinguishable from the point of view of BR (i.e., they satisfy the very same brute facts), and such that $W_C \cup \{w'\} \not\models \texttt{utter} \rightarrow \texttt{promise}$. If such a w' exists, then we can properly say that the truth of `promise` in W_C is constituted by the truth of `utter` since "all brute facts being equal" the implication possibly fails. The paper presents a logic to systematically handle this idea within a modal language.

2.2 Propositional equivalence *up to* a signature

The signature of a propositional language is its non-logial alphabet, that is, its set of propositional atoms. Let $\mathbf{P} = \{p, q, r \ldots\}$ be a countable set of propositional atoms, and let $\mathcal{L}(\mathbf{P})$ be the propositional language built on $\mathbf{P}$ and the usual Boolean connectives. We say that $\mathbf{P}$ is the signature of $\mathcal{L}(\mathbf{P})$.

Consider now the set $2^{\mathbf{P}}$ of all possible sub-signatures of $\mathcal{L}(\mathbf{P})$. Elements of such set will be denoted $P, Q, R, \ldots$ etc. Notice that the set of all sub-signatures of $\mathcal{L}(\mathbf{P})$ naturally yields a set algebra $\left\langle 2^{\mathbf{P}}, \cup, -, \mathbf{P}, \emptyset \right\rangle$. Two propositional models w and w' of $\mathcal{L}(\mathbf{P})$ are propositionally equivalent if they satisfy the same atoms in $\mathbf{P}$. As a consequence, for any formula φ of $\mathcal{L}(\mathbf{P})$: $w \models \varphi$ iff $w' \models \varphi$. If w and w' are equivalent ($w \sim w'$) then there is no set Φ of formulae of $\mathcal{L}(\mathbf{P})$ whose

[1] This is the semantics of what, in [7, 8], is called *classificatory counts-as*.

models contain w but not w', or vice versa. That is to say, the two models are indistinguishable for $\mathcal{L}(\mathbf{P})$.

However, two models which are not equivalent for $\mathbf{P}$ may be equivalent for some sub-signature $P \in 2^{\mathbf{P}}$. In this case, the two models cannot be distinguished by only looking at the atoms in P. The following definition makes such notion formal.

Definition 1. *(Equivalence up to a signature) Two models w and w' for a propositional language $\mathcal{L}$ are equivalent up to signature $P \in 2^{\mathbf{P}}$, or P-equivalent, if and only if for any $p \in P, w \models p$ iff $w' \models p$. If w and w' are P-equivalent we write $w \sim_P w'$.*

Obviously, if $w \sim_P w'$ then for all $\varphi \in \mathcal{L}(P)$: $w \models \varphi$ iff $w' \models \varphi$. The definition makes precise the idea of two propositional models agreeing up to what is expressible on a given signature.

Theorem 1. *(Properties of $\sim_P$) Let W be a set of models for the propositional language $\mathcal{L}(\mathbf{P})$. The following holds:*

(i) For every signature $P \in 2^{\mathbf{P}}$, the relation $\sim_P$ is an equivalence relation on W;
(ii) For all signatures $P, Q \in 2^{\mathbf{P}}$, if $P \subseteq Q$ then $\sim_Q \subseteq \sim_P$;
(iii) For each atom $p \in \mathbf{P}$, the relation $\sim_{\{p\}}$ yields a bipartition of W;
(iv) $\sim_{\mathbf{P}} = \sim$;
(v) $\sim_\emptyset = W^2$.

Proof. (i) The following holds: identity is a subrelation of $\sim_P$ for any sub-signature P; and that $\sim_P \circ \sim_P$ and $\sim_P^{-1}$ are subrelations of $\sim_P$ for any signature P. (ii) If $m \sim_Q m'$ then for all atoms $p \in Q$: $w \models p$ iff $w' \models p$. Therefore, since $P \subseteq Q$, $w \sim_P w'$. (iii) Suppose, per absurdum, that there exist three disjoint equivalence classes: $|w'|_{\sim_{\{p\}}}$, $|w''|_{\sim_{\{p\}}}$ and $|w'''|_{\sim_{\{p\}}}$. For bivalence, we have either $w' \models p$ or $w' \not\models p$. Suppose, without loss of generality, that $w' \models p$. By Definition 1 it follows that $w'' \not\models p$ and $w''' \not\models p$. Hence $|w''|_{\sim_{\{p\}}} = |w'''|_{\sim_{\{p\}}}$, which is impossible. (iv) The set $\mathbf{P}$ is the signature of the propositional language $\mathcal{L}(\mathbf{P})$, hence $\sim_{\mathbf{P}}$ is the propositional equivalence relation for $\mathcal{L}(\mathbf{P})$. (v) Suppose, per absurdum, there exists $w, w' \in W$ such that not $w \sim_\emptyset w'$. For Definition 1, there exists $p \in \emptyset$ such that $w \models p$ and $w' \not\models p$ (or viceversa), which is impossible.

Besides showing that signature-based equivalence is an equivalence relation (i), Theorem 1 shows also that: (ii) the bigger the signature, the more fine-grained is the equivalence relation; (iii) equivalences based on singleton signatured partition the set of states in two classes; (iv) if the propositional language under consideration is $\mathcal{L}(\mathbf{P})$ then relation $\sim_{\mathbf{P}}$ is standard propositional equivalence; (iv) $\sim_\emptyset$ is the universal relation on W. Notice also that from (ii) and (iii) follows that for every signature P it is the case that $\sim \subseteq \sim_P$, that is, propositional equivalence implies signature-based equivalence.

3 A modal logic of propositional equivalence *up to* a signature

The present section presents a modal logic—which we call UpTo—characterizing the notion of propositional equivalence *up to* a given signature.

3.1 Syntax of UpTo.

Let $\mathbf{P} = \{p, q, r \ldots\}$ be a countable set of propositional atoms. The language $\mathcal{L}_{\mathsf{UpTo}}(\mathbf{P})$[2] of logic UpTo on $\mathbf{P}$ is defined by the following BNF:

$$\mathcal{L}_{\mathsf{UpTo}} : \varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid [P]\varphi$$

where p ranges over $\mathbf{P}$ and P over $2^{\mathbf{P}}$. The Boolean connectives $\top, \vee, \rightarrow, \leftrightarrow$ and the dual operators $\langle P \rangle$ are defined as usual.

3.2 Semantics of UpTo.

Let us first define frames and models built on the notion of equivalente *up to* a given signature, in short, UpTo-frames and UpTo-models.

Definition 2. *(UpTo-frames) An UpTo-frame $\mathcal{F} = \langle W, \{\sim_P\}_{P \in 2^{\mathbf{P}}} \rangle$ for the propositional language $\mathcal{L}(\mathbf{P})$ is a tuple such that:*

- *W is a non-empty set of states;*
- *Each $\sim_P$ is an equivalence relation based on signature $P \in 2^{\mathbf{P}}$.*

Intuitively, an UpTo-frame fixes a particular arrangement of the equivalence classes available given a propositional language $\mathcal{L}(\mathbf{P})$. To make a simple example, suppose $W = \{w', w''\}$, $\mathbf{P} = \{p\}$ and $\sim_{\{p\}} = \{(w', w'), (w'', w'')\}$. Such frame for $\mathcal{L}(\{p\})$ states that w' and w'' are equivalent up to signature $\{p\}$ only to themselves. The valuation function will then say whether it is w' that satisfies p while w'' does not, or vice versa. This brings us to the notion of UpTo-model.

Definition 3. *(UpTo-models) An UpTo-model $\mathcal{M} = \langle \mathcal{F}, \mathcal{I} \rangle$ for the modal language $\mathcal{L}_{\mathsf{UpTo}}(\mathbf{P})$ is a tuple such that:*

- *$\mathcal{F}$ is an UpTo-frame for the propositional language $\mathcal{L}(\mathbf{P})$;*
- *$\mathcal{I} : \mathbf{P} \longrightarrow 2^W$ is an interpretation function.*

It may be instructive to notice that for each UpTo-frame there are exactly $2^{\mathbf{P}}$ different UpTo-models since Definition 1 requires that, for any atom p in $\mathbf{P}$, each element in the bipartition yielded by p coincides either with the truth-set of p or with the truth-set of $\neg p$.

The satisfaction relation is defined as follows.

Definition 4. *(Satisfaction for UpTo-models) Let $\mathcal{M}$ be an UpTo-model for $\mathcal{L}_{\mathsf{UpTo}}(\mathbf{P})$, $w \in W$ and $\varphi, \psi \in \mathcal{L}_{\mathsf{UpTo}}(\mathbf{P})$.*

$$\begin{aligned}
\mathcal{M}, w \models p \quad &\textit{iff} \quad w \in \mathcal{I}(p); \\
\mathcal{M}, w \models \neg\varphi \quad &\textit{iff} \quad \mathcal{M}, w \not\models \varphi; \\
\mathcal{M}, w \models \varphi \wedge \psi \quad &\textit{iff} \quad \mathcal{M}, w \models \varphi \;\&\; \mathcal{M}, w \models \psi; \\
\mathcal{M}, w \models [P]\varphi \quad &\textit{iff} \quad \forall w' \in W, w \sim_P w' : \mathcal{M}, w' \models \varphi
\end{aligned}$$

[2] In what follws we will often drop the reference to $\mathbf{P}$ and denote the language of UpTo simply by $\mathcal{L}_{\mathsf{UpTo}}$.

Formula φ is valid in $\mathcal{M}$, noted $\mathcal{M} \models \varphi$, if and only if for all w in W, $\mathcal{M}, w \models \varphi$. Formula φ is valid in $\mathcal{F}$, noted $\mathcal{F} \models \varphi$, if and only if it is valid in all models built on $\mathcal{F}$. Finally, φ is UpTo*-valid, noted $\models_{\mathsf{UpTo}} \varphi$, iff it is valid in all* UpTo*-frames. The logical consequence of formula φ from a set of formulae, noted $\Phi \models_{\mathsf{UpTo}} \varphi$, can be defined as usual.*

Intuitively, the *up to* operator $[P]$ means that φ holds in all states that are equivalent to the state of evaluation up to signature P.

3.3 Axiomatics of UpTo.

Logic UpTo is axiomatized by the following schemata.

$$
\begin{array}{rl}
(\mathtt{P}) & \text{all tautologies of propositional calculus} \\
(\mathtt{K}) & [P](\varphi \to \psi) \to ([P]\varphi \to [P]\psi) \\
(\mathtt{T}) & [P]\varphi \to \varphi \\
(4) & [P]\varphi \to [P][P]\varphi \\
(5) & \langle P \rangle \varphi \to [P]\langle P \rangle \varphi \\
(\mathtt{PO}) & [P]\varphi \to [Q]\varphi \text{ IF } P \subseteq Q \\
(\mathtt{Bipart}) & [\{p\}]p \vee [\{p\}]\neg p \\
(\mathtt{Dual}) & \langle P \rangle \varphi \leftrightarrow \neg [P] \neg \varphi \\
(\mathtt{MP}) & \text{IF } \vdash \varphi_1 \text{ AND } \vdash \varphi_1 \to \varphi_2 \text{ THEN } \vdash \varphi_2 \\
(\mathtt{N}) & \text{IF } \vdash \varphi \text{ THEN } \vdash [P]\varphi
\end{array}
$$

where P, Q range over $2^{\mathbf{P}}$, φ, ψ over $\mathcal{L}_{\mathsf{UpTo}}(\mathbf{P})$ and p over $\mathbf{P}$. The *up to* operators are **S5** operators with the addition of axioms `PO` (partial order) and `Bipart` (bipartition). Axiom `PO` orders the strength of the operators according to the relation of set-inclusion on the set of signatures. Notice that it consists of a transposition, in modal logic, of property (ii) in Theorem 1. Axiom `Bipart` states that if the signature considered consists of only atom p then it is either necessarily the case that p, or it is necessarily the case that $\neg p$. In other words, the equivalence up to p determines a bipartition of the set of states where the one cluster coincides with the set of p-states and the other with the set of $\neg p$ states. This axiom rephrases property (iii) of Theorem 1. Notice that from `PO`, `Bipart` and `P` follows that $[P]p \vee [P]\neg p$ if $p \in P$. [3]

Provability of a formula φ, noted $\vdash_{\mathsf{UpTo}} \varphi$, and derivability of a formula φ from a set of formulae Φ, noted $\Phi \vdash_{\mathsf{UpTo}} \varphi$ can be defined as usual. Appendix A offers a proof of the soundness and strong completeness of the proposed axiomatics with respect to the class of models built on UpTo-frames.

[3] A slightly different version of such schema has been used as an axiom in [5], where it is called `NoCross`. Notice that it forces the accessibility relation not to cross the bipartitions of the domain W yielded by each atom p, when p does belong to signature in the modal operator.

3.4 Embedding UpTo into S5

Take the standard modal language $\mathcal{L}_\Box(\mathbf{P})$ with one modal operator $\Box$ defined on the set of atoms $\mathbf{P}$. If we allow only *up to* operators $[P]$ where P is finite, it is possible to define an EXPtime truth-preserving reduction $f : \mathcal{L}_{\mathsf{UpTo}}(\mathbf{P}) \longrightarrow \mathcal{L}_\Box(\mathbf{P})$ as follows:

$$\begin{aligned} f(p) &= p \\ f(\neg\varphi) &= \neg f(\varphi) \\ f(\varphi \wedge \psi) &= f(\varphi) \wedge f(\psi) \\ f([\emptyset]\varphi) &= \Box f(\varphi) \\ f([P]\varphi) &= \bigwedge_{\pi_i \in 2^P} \Big(\big(\bigwedge \pi_i^+ \wedge \bigwedge \pi_i^-\big) \to \Box\big(\big(\bigwedge \pi_i^+ \wedge \bigwedge \pi_i^-\big) \to f(\varphi)\big)\Big) \end{aligned}$$

where $\pi_i^+ = \pi_i$ and $\pi_i^- = \{\neg p \mid p \in P \;\&\; p \notin \pi_i\}$. Intuitively, the *up to* operators are translated by taking care of all the possible truth-value combinations of the atoms in the signature P. If a given combination, e.g., $\bigwedge \pi_i^+ \wedge \bigwedge \pi_i^-$, is true at the given state, then in all accessible states, if that combination is true, than φ is also true. In addition, this should be the case for any combination drawn from a non-empty P, which explains $\bigwedge_{\pi_i \in 2^P - \emptyset}$. If P is empty, than $[P]$ is taken to be $\Box$. As a consequence, $\Box$ has to be interpreted as a universal modality (Theorem 1).

Theorem 2. *(f preserves satisfiability) Let $\mathcal{M} = \langle W, \{\sim_P\}_{P \in 2^{\mathbf{P}}}, \mathcal{I}\rangle$ be an* UpTo*-model for language $\mathcal{L}_{\mathsf{UpTo}}(\mathbf{P})$ and $\mathcal{M}' = \langle W', R', \mathcal{I}'\rangle$ be an* **S5** *model for $\mathcal{L}_\Box(\mathbf{P})$ such that:*

- $W' = W$;
- $R' = \sim_\emptyset$;
- $\mathcal{I}' = \mathcal{I}$.

For any $w \in W$ and $\varphi \in \mathcal{L}_{\mathsf{UpTo}}(\mathbf{P})$, $\mathcal{M}, w \models \varphi$ iff $\mathcal{M}', w \models f(\varphi)$.

Proof. The Boolean clauses and the clause for $[\emptyset]$ are obvious. As to the the last clause, by induction hypothesis (IH): $\mathcal{M}, w \models \varphi$ iff $\mathcal{M}', w \models f(\varphi)$. By IH, the semantics of $[P]$ and $\Box$, and Definition 1, the following expressions are all equivalent to $\mathcal{M}, w \models [P]\varphi$:

$$\begin{aligned} &\forall w' \in W, w \sim_P w' : \mathcal{M}, w' \models \varphi \\ &\forall w' \in W, w \sim_P w' : \mathcal{M}', w' \models f(\varphi) \\ &\forall w' \in W, \forall \pi_i \in 2^P \text{ IF } \mathcal{M}', w \models \bigwedge \pi_i^+ \wedge \bigwedge \pi_i^- \text{ THEN } \mathcal{M}', w' \models \big(\bigwedge \pi_i^+ \wedge \bigwedge \pi_i^-\big) \to f(\varphi) \\ &\forall \pi_i \in 2^P \text{ IF } \mathcal{M}', w \models \bigwedge \pi_i^+ \wedge \bigwedge \pi_i^- \text{ THEN } \mathcal{M}', w' \models \Box\big(\big(\bigwedge \pi_i^+ \wedge \bigwedge \pi_i^-\big) \to f(\varphi)\big) \\ &\mathcal{M}', w' \models \bigwedge_{\pi_i \in 2^P - \emptyset} \Big(\big(\bigwedge \pi_i^+ \wedge \bigwedge \pi_i^-\big) \to \Box\big(\big(\bigwedge \pi_i^+ \wedge \bigwedge \pi_i^-\big) \to f(\varphi)\big)\Big) \end{aligned}$$

This completes the proof.

As a consequence, we also obtain the following result.

Corollary 1. *(Decidability) The satisfiability problem for* UpTo *is decidable.*

Proof. The satisfiability problem for **S5** is decidable [2]. The result follows from Theorem 2.

Translation f makes explicit how the *up to* operators enable a compact representation of rather rich logical information. What can be expressed by UpTo can as well be expressed in **S5**, but not as easily.

4 Related work and conclusions

In these last two sections we relate the results presented in this paper to existing work in modal logic, and we finally draw some conclusions pointing at future research directions.

4.1 Related work: up to, release and ceteris paribus logics

The logic presented in Section 3 is a strict relative of the so-called release logics, first introduced and studied in [10, 11] in order to provide a modal logic characterization of a general notion of irrelevancy. Modal operators in release logics are **S5** operators indexed by an abstract set denoting the issues that are taken to be irrelevant while evaluating the formula in the scope of the operator. In [5] a special release logic is studied where the potentially irrelevant issues are precisely the propositional atoms of the language. This allows for the characterization of a notion of equivalence *modulo* a given signature. Instead of studying formulae $[P]\varphi$, whose intuitive meaning is "φ is the case" *up to* signature P, that logic studies formulae $[P]\varphi$ whose intuitive meaning is "φ is the case" *modulo* signature P, that is, if we abstract from the atoms in P. Therefore, in order to obtain a truth-preserving translation f of this logic to UpTo we just need to require: $f([P]\varphi) = [-P]f(\varphi)$, where $-$ is the set-theoretic complement. The UpTo logic can therefore be considered to belong to the family of release logics.[4]

Another work coming very close to the spirit of the present paper is [1]. In that paper a logic is presented for *ceteris paribus* preferences, that is to say, for preferences under the"all other things being equal" condition. Leaving the preferential component of such logic aside, its ceteris paribus fragment concerns sentences of the form $\langle \Gamma \rangle \varphi$ whose intuitive meaning is "there exists a state which is equivalent to the evaluation state with respect to all the formulae in the (finite) set Γ and which satisfies φ", where the formulae in Γ are drawn from the full language. At this point it is easy to see that logic UpTo is, in fact, the fragment of the ceteris paribus logic where Γ is allowed to consist only of a set of atoms. It is, we could say, the logic of "everything else being equal which you can express on this signature". From the semantic point of view, this means that UpTo-models contain considerably less equivalence classes than ceteris paribus models.

[4] See [5] for more details.

4.2 Conclusions

The paper has introduced and studied modal logic UpTo characterizing the notion of equivalence up to a given propositional signature. Soundness and completeness of the axiomatics, as well as the decidability of the satisfaction problem has been proven.

To conclude, let us go back to the beginning of Section 2 and show how Formula 1 can be appropriately extended in order to capture the "brute vs. institutional" distinction:

$$(2) \qquad W_C \models \texttt{utter} \rightarrow \texttt{promise} \text{ AND } \texttt{W}_C \not\models [\texttt{BR}](\texttt{utter} \rightarrow \texttt{promise})$$

Using the syntax of the modal context logic Cxt developed in [7, 8], Formula 2 could be expressed in the object-language as follows:

$$(3) \qquad [C](\texttt{utter} \rightarrow \texttt{promise}) \wedge \neg[\texttt{C}][\texttt{BR}](\texttt{utter} \rightarrow \texttt{promise})$$

where $[C]$ denotes the context operator. A systematic study of the interaction of logics Cxt and UpTo is left for future work.

References

1. J. van Benthem, P. Girard, and O. Roy. Everything else being equal: A modal logic for ceteris paribus preferences. *Journal of Philosophical Logic*, 38:83–125, 2009.
2. P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic*. Cambridge University Press, Cambridge, 2001.
3. G. Boella and L. Van der Torre. Regulative and constitutive norms in normative multiagent systems. In D. Dubois, C. A. Christopher A. Welty, and M. Williams, editors, *Proceedings of KR2004, Whistler, Canada*, pages 255–266, 2004.
4. G. Governatori, J. Gelati, A. Rotolo, and G. Sartor. Actions, institutions, powers. preliminary notes. In *International Workshop on Regulated Agent-Based Social Systems: Theories and Applications (RASTA'02)*, pages 131–147, 2002.
5. D. Grossi. Linguistic relevance in modal logic. In A. Nijholt, M. Pantic, M. Poel, and G. Hondorp, editors, *Proceedings of the 20th Belgian-Netherlands Conference on Artificial Intelligence (BNAIC'08)*. University of Twente, 2008.
6. D. Grossi. Pushing Anderson's envelope: The modal logic of ascription. In R. van der Meyden and L. van der Torre, editors, *Proceedings of the 9th International Conference on Deontic Logic in Computer Science (DEON 2008), Luxembourg, Luxembourg, July 15-18, 2008th International Workshop on Deontic Logic in Computer Science (DEON 2008), Luxembourg, Luxembourg, July 15-18, 2008*, number 5076/2008 in LNAI, pages 263–277. Springer, 2008.
7. D. Grossi, J.-J.Ch. Meyer, and F. Dignum. Classificatory aspects of counts-as: An analysis in modal logic. *Journal of Logic and Computation*, 16(5):613–643, 2006. Oxford University Press.
8. D. Grossi., J.-J.Ch. Meyer, and F. Dignum. The many faces of counts-as: A formal analysis of constitutive-rules. *Journal of Applied Logic*, 6(2):192–217, 2008.
9. A. J. I. Jones and M. Sergot. A formal characterization of institutionalised power. *Journal of the IGPL*, 3:427–443, 1996.

10. J. Krabbendam and J.-J. C. Meyer. Contextual deontic logics. In P. McNamara and H. Prakken, editors, *Norms, Logics and Information Systems*, pages 347–362, Amsterdam, 2003. IOS Press.
11. J. Krabbendam and J.-J.Ch. Meyer. Release logics for temporalizing dynamic logic, orthogonalising modal logics. In M. Barringer, M. Fisher, D. Gabbay, and G. Gough, editors, *Advances in Temporal Logic*, pages 21–45. Kluwer Academic Publisher, 2000.
12. E. Lorini and D. Longin. A logical account of institutions: from acceptances to norms via legislators. In G. Brewka and J. Lang, editors, *International Conference on Principles of Knowledge Representation and Reasoning (KR), Sidney, Australia, 16/09/08-19/09/08*, pages 38–48, http://www.aaai.org/Press/press.php, 2008. AAAI Press.
13. J. Searle. How to derive "ought" from "is". *The Philosophical Review*, 73(1):43–58, 1964.
14. J. Searle. *Speech Acts. An Essay in the Philosophy of Language*. Cambridge University Press, Cambridge, 1969.
15. J. Searle. *The Construction of Social Reality*. Free Press, 1995.

A Soundness and completeness of UpTo

Soundness is easily proven.

Theorem 3. *(Soundness of UpTo) For any $\varphi \in \mathcal{L}_{\mathsf{UpTo}}$, if $\vdash_{\mathsf{UpTo}} \varphi$ then $\models_{\mathsf{UpTo}} \varphi$.*

Proof. It is well-known that inference rules `MP` and `N` preserve validity on any class of frames, and that axioms `T`, `4` and `5` are valid on models built on equivalence relations[5]. The validity of `PO` and of `Bipart` follows from Theorem 1.

As to completeness, we make use of the standard canonical model technique.

Lemma 1. *Logic UpTo is strongly complete w.r.t. the class of UpTo-frames iff every UpTo-consistent set Φ of formulae is satisfiable on some model built on an UpTo-frame.*

Proof. From right to left we argue by contraposition. If UpTo is not strongly complete w.r.t. the class then there exists a set of formulae $\Phi \cup \{\varphi\}$ s.t. $\Phi \models_{\mathsf{UpTo}} \varphi$ and $\Phi \nvdash_{\mathsf{UpTo}} \varphi$. It follows that $\Phi \cup \{\neg\varphi\}$ is UpTo-consistent but not satisfiable on any UpTo-model. From left to right we argue per absurdum. Let us assume that $\Phi \cup \{\neg\varphi\}$ is UpTo-consistent but not satisfiable in any sublanguage equivalent model built on a frame in class UpTo. It follows that $\Phi \models_{\mathsf{UpTo}} \varphi$ and hence $\Phi \cup \{\neg\varphi\}$ is not UpTo-consistent, which is impossible.

Now let $\mathcal{M}^{\mathsf{UpTo}}$ be the canonical model of logic UpTo in language $\mathcal{L}_{\mathsf{UpTo}}(\mathbf{P})$. Model $\mathcal{M}^{\mathsf{UpTo}}$ is the structure $\left\langle W^{\mathsf{UpTo}}, \{R_P^{\mathsf{UpTo}}\}_{P\in\mathbf{P}}, \mathcal{I}^{\mathsf{UpTo}} \right\rangle$ where:

1. The set W^{UpTo} is the set of all maximal UpTo-consistent sets.
2. The canonical relations $\{R_P^{\mathsf{UpTo}}\}_{P\in\mathbf{P}}$ are defined as follows: for all $w, w' \in W^{\mathsf{UpTo}}$, if for all formulae φ, $\varphi \in w'$ implies $\langle P\rangle\varphi \in w$, then $wR_P^{\mathsf{UpTo}}w'$.
3. The canonical interpretation $\mathcal{I}^{\mathsf{UpTo}}$ is defined by $\mathcal{I}^{\mathsf{UpTo}}(p) = \{w \in W^{\mathsf{UpTo}} \mid p \in w\}$.

[5] See [2].

We have now to prove the Existence and Truth Lemmata for logic UpTo.

Lemma 2. *(Existence lemma) For all states in W^{UpTo}, if $\langle P\rangle\varphi \in w$ then there exists a state $w' \in W^{\mathsf{UpTo}}$ s.t. $R_P^{\mathsf{UpTo}}(w, w')$ and $\varphi \in w'$.*

Proof. The claim is proven by construction. Assume $\langle P\rangle\varphi \in w$ and let $w'_0 = \{\varphi\} \cup \{\psi \mid [P]\psi \in w\}$. The set w'_0 must be UpTo-consistent since otherwise there would exist $\psi_1, \ldots, \psi_m \in w'_0$ such that $\vdash_{\mathsf{UpTo}} (\psi_1 \wedge \ldots \wedge \psi_m) \rightarrow \neg\varphi$, from which we obtain $\vdash_{\mathsf{UpTo}} ([P]\psi_1 \wedge \ldots \wedge [P]\psi_m) \rightarrow [P]\neg\varphi$. Since $[P]\psi_1, \ldots, [P]\psi_m \in w$ we have that $\neg\langle P\rangle\varphi \in w$, which contradicts our assumption. Therefore, w'_0 is UpTo-consistent and can be extended to a maximal UpTo-consistent set (for Lindenbaum's Lemma[6]). By construction, w' contains φ and is such that for all ψ, if $[P]\psi \in w$ then w' contains ψ. From this it follows $R_P^{\mathsf{UpTo}}(w, w')$ since, if this was not the case, then there would exist a formula ψ' s.t. $\psi' \in w'$ and $\langle P\rangle\psi' \notin w$. Since w is maximal UpTo-consistent, $[P]\neg\psi' \in w$ and hence $\neg\psi' \in w'$, which contradicts the UpTo-consistency of w'.

Lemma 3. *(Truth lemma) For any formula $\varphi \in \mathcal{L}_{\mathsf{UpTo}}(\mathbf{P})$ and $w \in W^{\mathsf{UpTo}}$: $\mathcal{M}^{\mathsf{UpTo}}, w \models \varphi$ iff $\varphi \in w$.*

Proof. The claim is proven by induction on the complexity of φ. The Boolean case follows by the properties of maximal UpTo-consistent sets. As to the modal case, it follows from the definition of the canonical relations R_P^{UpTo} and Lemma 2.

Everything is now put into place to prove the strong completeness of UpTo.

Theorem 4. *(Strong completeness of UpTo) For any formula $\varphi \in \mathcal{L}_{\mathsf{UpTo}}(\mathbf{P})$ and set of formulae Φ, if $\Phi \vdash_{\mathsf{UpTo}} \varphi$ then $\Phi \models_{\mathsf{UpTo}} \varphi$.*

Proof. By Proposition 1, given an UpTo-consistent set Φ of formulae, it suffices to find a model state pair $(\mathcal{M}, w)$ such that: (a) $\mathcal{M}, w \models \Phi$, (b) $\mathcal{M}$ is an UpTo-model. Let $\mathcal{M}^{\mathsf{UpTo}} = \left\langle W^{\mathsf{UpTo}}, \{R_P^{\mathsf{UpTo}}\}_{P \in 2^{\mathbf{P}}}, \mathcal{I}^{\mathsf{UpTo}} \right\rangle$ be the canonical model of UpTo, and let Φ^+ be any maximal UpTo-consistent set in W^{UpTo} extending Φ. By Lemma 3 it follows that $\mathcal{M}^{\mathsf{UpTo}}, \Phi^+ \models \Phi$, which proves (a). To prove (b), we show that $\mathcal{M}^{\mathsf{UpTo}}$ is s.t.: (b.1) the frame on which $\mathcal{M}$ is based is an UpTo-frame; and (b.2) for all $p \in \mathbf{P}$, $R_{\{p\}}^{\mathsf{UpTo}}(w, w')$ iff it is the case that $p \in w$ iff $p \in w'$. As to (b.1), it is well-known that axioms `T`, `4` and `5` force the relations R_P^{UpTo} to be equivalence relations. It remains to be shown that if $P \subseteq Q$ then $R_Q^{\mathsf{UpTo}} \subseteq R_P^{\mathsf{UpTo}}$. Assume $R_Q^{\mathsf{UpTo}}(w, w')$. It follows that for all φ, if $\varphi \in w'$ then $\langle Q\rangle\varphi \in w$ and hence, by the contrapositive of axiom `P0`, $\langle P\rangle\varphi \in w$. Therefore, $R_P^{\mathsf{UpTo}}(w, w')$. As to (b.2), form left to right. Assume $R_{\{p\}}^{\mathsf{UpTo}}(w, w')$. For axioms `T` and `Bipart`, $p \in w$ iff $p \in w'$. From right to left, we assume $p \in w$ iff $p \in w'$. If $p \in w'$, by axioms `T` and `Bipart`, $\langle\{p\}\rangle p \in w$ and therefore $R_{\{p\}}^{\mathsf{UpTo}}(w, w')$. This completes the proof.

[6] See [2].