Agentive Permissions in Multiagent Systems

Qi Shi
University of Southampton
qi.shi@soton.ac.uk

Abstract

This paper proposes to distinguish four forms of agentive permissions in multiagent settings. The main technical results are the complexity analysis of model checking, the semantic undefinability of modalities that capture these forms of permissions through each other, and a complete logical system capturing the interplay between these modalities.

1 Introduction

Imagine a large factory being built in a city on a river. The factory will dump a pollutant into the river. It is known that a small factory located in another city higher up the river already exists and can dump up to 60g/day of the same pollutant. Also, more than 100g/day of the pollutant dumped into the river combined by the two factories will kill the fish.

Suppose that the large factory will dump 20g/day of the pollutant. Then, the total dumped amount by both factories will not exceed 80g/day no matter how much the other factory dumps and thus the fish in the river will survive *for sure*. In other words, the action that dumping 20g/day ***ensures*** the survival of the fish. On the contrary, if the large factory will dump 60g/day of the pollutant, then the fish will be killed once the other factory dumps more than 40g/day. That is to say, the action that dumping 60g/day does not ensure the survival of the fish. However, it still *leaves the possibility* for the fish to survive, *e.g.* when the other factory dumps no more than 40g/day. In this situation, we say that the action that dumping 60g/day ***admits*** of the survival of the fish.

To ensure and to admit show two different types of ***agency*** of an action. The difference comes from that, in multiagent or nondeterministic settings, the effect of an action of one agent might be affected by the actions of other agents or the nondeterminacy. It is worth noting that, admitting is the *dual* of ensuring. To be specific, if an action of an agent ensures an outcome, then the action does *not* admit of the *opposite* outcome, and vice versa. It is easy to see that their meanings coincide in single-agent deterministic settings.

In this paper, we consider the two types of agency together with permissions. Let us come back to the large factory. Suppose the city government passes a regulation:

> *The factory is permitted to dump any amount of the pollutant as long as the fish is not killed.*

This regulation can be interpreted in two ways corresponding to two types of agency. The first could be

> ***any*** *dumping amount that* ***admits*** *of the survival of the fish is permitted.* (SA)

In this interpretation, the permitted dumping amount of the factory is any amount no more than 100g/day. As long as the factory follows this regulation, there is a chance for the fish to survive if the other factory dumps so little that the total dumped amount does not exceed 100g/day. The second interpretation of the regulation could be

> ***any*** *dumping amount that* ***ensures*** *the survival of the fish is permitted.* (SE)

In this interpretation, the factory should not dump any amount over 40g/day and the fish cannot be killed as long as the factory follows the regulation, no matter how much of the pollutant is dumped by the other factory.

The above two interpretations give the factory two different permissions. It is worth noting that, either of the permissions *enables* the factory to take any of the actions satisfying some criteria. In this paper, we call such permissions "strong". Specifically, we refer to (SA) and (SE) as *strong permission to admit* and *strong permission to ensure*, respectively.

Not all permissions are in the same form as strong permissions. Suppose that, instead of the city's regulation, the factory is under some contractual obligation. To satisfy this obligation, the factory has to dump at least 30g/day of the pollutant. In this case, not every amount that ensures/admits of the survival of the fish (*e.g.* 20g/day) is contractually permitted to be dumped. Nevertheless,

> ***there is*** *a permitted dumping amount that* ***ensures*** *the survival of the fish.* (WE)

For example, 35g/day is a contractually permitted dumping amount that ensures the survival of the fish. Similarly, if the contractual obligation forces the factory to dump at least 50g/day of the pollutant, then no contractually permitted dumping amount ensures the survival of the fish. However,

> ***there is*** *a permitted dumping amount that* ***admits*** *of the survival of the fish.* (WA)

The full version of this paper includes an appendix, which can be found at arXiv: 2404.17053.

In contrast to strong permissions, the above two permissions express the *capability* of the factory to achieve some statements with a permitted action. We call such permissions "weak". Specifically, we refer to (WE) and (WA) as *weak permission to ensure* and *weak permission to admit*, respectively.

As shown in Section 2, the terms "strong permission" and "weak permission" have already existed in the literature for decades [Raz, 1975; Royakkers, 1997; Governatori *et al.*, 2013]. In Section 5, we show the consistency between how these terms are used in our paper and in the literature. However, as far as we see, we are the first to make a clear distinction between "permission to ensure" and "permission to admit", which are ***agentive*** permissions specific to multiagent settings. We are also the first to cross-discuss both strong and weak agentive permissions in multiagent settings.

In this paper, we discuss four forms of permissions in multiagent settings that generalise those expressed in statements (SA), (SE), (WE), and (WA). Our contribution is three-fold. First, we propose a formal semantics for the four corresponding modalities in multiagent transition systems (Section 3). We also consider the model-checking problem (Section 3.1) and the reduction to STIT logic and ATL (Section 3.2). Second, we prove these modalities are semantically undefinable through each other (Section 4). This contrasts to the fact that, when separated from permissions, ensuring and admitting are dual to each other. Third, we give a sound and complete logical system for the four modalities (Section 5 and Section 6). This reveals the interplay between the four forms of permissions and offers an efficient way for permission reasoning.

2 Literature Review and Notion Discussion

Deontic logic [McNamara and van de Putte, 2022] is an appealing approach to solving AI ethics problems by enabling autonomous agents to comprehend and reason about their *obligation*, *permission*, and *prohibition*. It aims at "translating" the deontic statements in natural languages into logical propositions and building up a system for plausible deduction. Von Wright [1951] launched the active development of symbolic deontic logic from the analogies between normative and alethic modalities. Several follow-up works [Anderson, 1956; Prior, 1963] built up the Standard Deontic Logic (SDL), taking obligation as the basic modality and defining permission as the dual of obligation and prohibition as the obligation of the negation. Anderson [1967] and Kanger [1971] reduced this system by defining a propositional constant d for "all (relative) normative demands are met". By this means, obligation (modality O) can be defined as $\mathsf{O}\varphi := \Box(d \to \varphi)$, which is read as "it is *necessary* that φ is true when all normative demands are met". As the dual of obligation, permission (modality P) is defined as

$$\mathsf{P}\varphi := \Diamond(d \wedge \varphi), \tag{1}$$

which is read as "it is *possible* that all normative demands are met and φ is true". In this way, the inference rule

$$\frac{\varphi \to \psi}{\mathsf{P}\varphi \to \mathsf{P}\psi} \tag{2}$$

is valid. The notion of permission that satisfies statement (2) is called **weak permission**. There are two well-known related paradoxes about weak permission:

(i) *Ross's paradox* [Ross, 1944]. The formula

$$\mathsf{P}\varphi \to \mathsf{P}(\varphi \vee \psi) \tag{3}$$

is valid by statement (2). However, in common sense, for a kid "it is permitted to eat an apple" is true but "it is permitted to eat an apple or drink alcohol" should be false, which contradicts statement (3).

(ii) The *free choice permission paradox* [von Wright, 1968; Kamp, 1973]. According to linguistic intuition, if "it is permitted to eat an apple or a banana", then both "eating an apple" and "eating a banana" should be permitted. This shows that disjunctive permission is treated as **free choice permission**, which means the formula

$$\mathsf{P}(\varphi \vee \psi) \to \mathsf{P}\varphi \wedge \mathsf{P}\psi \tag{4}$$

should be valid. However, statement (4) is *not* derivable in SDL. Free choice permission is a form of **strong permission** [Asher and Bonevac, 2005], satisfying the inference rule

$$\frac{\varphi \to \psi}{\mathsf{P}^s\psi \to \mathsf{P}^s\varphi}. \tag{5}$$

Following Anderson and Kanger's way, van Benthem [1979] captured the notion of strong permission as

$$\mathsf{P}^s\varphi = \Box(\varphi \to d), \tag{6}$$

which is read as "it is *necessary* that if φ is true then all normative demands are met". He then gave a complete axiom system for obligation (O) and strong permission (P^s).

Most researchers agree that both weak and strong permission makes sense. As discussed by Lewis [1979], no universal comprehension of permission seems to exist. In general, weak permission is treated as the dual of obligation. Strong permission, as well as free choice permission, is more intractable and arouses more interesting discussions due to its anti-monotonic inference property in statement (5). For instance, Anglberger *et al.* [2015] adopted the notion of strong permission and defined a notion of obligation as the weakest form of (strong) permission. Wang and Wang [2023] axiomatised a logic of strong permission that satisfies some commonly desirable logical properties. Strong permission is also studied in defeasible logic [Asher and Bonevac, 2005; Governatori *et al.*, 2013], which is believed to be able to capture the logical intuition about permission.

The above discussion of permission applies possible-world semantics without specifically considering agents and their agency. However, it is noticed that two kinds of normative statements exist: the *agentless* norms that talk about states (*e.g.* it is permitted to eat an apple) and the *agentive* norms that talk about actions (*e.g.* John is permitted to eat an apple). The possible-world semantics cannot distinguish them. To fill the gap, Chisholm [1964] proposed a transfer from any agentive norm to an agentless norm. For instance, the statement "agent a is permitted to do φ" is transformed into "it is permitted that agent a does φ". Some recent work [Kulicki and Trypuz, 2017; Kulicki *et al.*, 2023] aimed at integrating the agentless and agentive norms in a unified logical frame.

Things become more complicated when agents and their agency are incorporated. In the literature, STIT logic [Chellas, 1969; Belnap and Perloff, 1988; Belnap and Perloff, 1992] is used to express the agency. Horty and Belnap [1995] and Horty [2001, chapter 4] introduced a deontic STIT logic for *ought-to-be* and *ought-to-do* semantics, respectively. The former corresponds to the agentless obligations while the latter corresponds to the agentive obligations in STIT models. Horty [2001, chapter 3] further showed that the transfer proposed by Chisholm does not always work properly. Following Horty, van de Putte [2017] briefly discussed the dual of the ought-to-do obligation, which is the weak permission in deontic STIT logic, and then defined a form of free choice permission following statement (4). Although the agency is considered in deontic STIT logic, the distinction between to ensure and to admit is never discussed there.

In the field of AI, there is a rising interest in applying deontic logic into agents' planning: how to achieve a goal while complying with the deontic constraints [Pandžić *et al.*, 2022; Areces *et al.*, 2023]. There is also some discussion of agents' comprehending and reasoning norms [Arkoudas *et al.*, 2005; Broersen and Ramírez Abarca, 2018]. However, to the best of our knowledge, the agentive weak and strong permissions have never been cross-discussed before.

In this paper, we consider both permission to ensure and permission to admit in both weak and strong forms that follow statements (1) and (6). In a word, we consider four forms of permissions as illustrated in statements (SA), (SE), (WE), and (WA). It is worth mentioning that, our formalisation has a connection with Horty's ought-to-do deontic STIT logic [Horty, 2001]. On the one hand, the notion "ensure" captures the same idea as "see to it that" in STIT logic. On the other hand, our formalisation can be seen as a *reasonable* reduction of Horty's formalisation. Specifically, Horty's approach is to, first, define a preference over the outcomes (*i.e.* "histories" in STIT models) of actions in the *model*, then, apply the *dominance act utiliarianism* to decide which actions are permitted (*i.e.* "optimal" in his work) in *semantics*, and, finally, define the ought-to-do obligation based on the permitted actions in *semantics*. In particular, an action in the STIT frame is the set of outcomes that may follow from the action. An action "sees to it that" φ if and only if φ is true in all the potential outcomes. Then, "do φ" is interpreted as "seeing to it that φ". In this paper, we combine the first two steps of Horty's approach, directly defining the deontic constraints on actions in the *model* and then defining four forms of permissions in *semantics*. Note that, the definition of deontic notions is independent of the process that combines the first two steps in Horty's approach. In other words, our work in this paper can easily be transformed from action-based models into outcome-based models by recovering the step of deciding permitted actions based on preference over outcomes using dominance act utilitarianism. Moreover, we give a reduction of our semantics into STIT logic in Section 3.2.

3 Syntax and Semantics

In this section, we introduce the syntax and semantics of our logical system. Throughout the paper, unless stated otherwise, we assume a fixed set $\mathcal{A}$ of agents and a fixed nonempty set of propositional variables.

Definition 1. *A **transition system** is a tuple* (S, Δ, D, M, π)*:*

1. *S is a (possibly empty) set of **states**;*
2. *$\Delta = \{\Delta_a^s\}_{s\in S, a\in\mathcal{A}}$ is the **action** space, where Δ_a^s is a nonempty set of actions available to agent a in state s;*
3. *$D = \{D_a^s\}_{s\in S, a\in\mathcal{A}}$ is the **deontic constraints**, where D_a^s is a set of permitted actions and $\varnothing \subsetneq D_a^s \subseteq \Delta_a^s$;*
4. *$M = \{M_s\}_{s\in S}$ is the **mechanism**, where a relation $M_s \subseteq \prod_{a\in\mathcal{A}} \Delta_a^s \times S$ satisfies the **continuity** condition: for each **action profile** $\delta \in \prod_{a\in\mathcal{A}} \Delta_a^s$ there is a state $t \in S$ such that $(\delta, t) \in M_s$;*
5. *$\pi(p) \subseteq S$ for each propositional variable p.*

The continuity condition in item 4 above requires the existence of a "next" state t. We say that a transition system is ***deterministic*** if such state t is always unique.

The language Φ of our logical system is defined by the following grammar:

$$\varphi := p \mid \neg\varphi \mid \varphi \vee \varphi \mid \mathsf{WA}_a\varphi \mid \mathsf{WE}_a\varphi \mid \mathsf{SE}_a\varphi \mid \mathsf{SA}_a\varphi,$$

where p is a propositional variable and $a \in \mathcal{A}$ is an agent. Intuitively, we interpret $\mathsf{WA}_a\varphi$ as "there is a permitted action of agent a that admits of φ", $\mathsf{WE}_a\varphi$ as "there is a permitted action of agent a that ensures φ", $\mathsf{SE}_a\varphi$ as "each action of agent a that ensures φ is permitted", and $\mathsf{SA}_a\varphi$ as "each action of agent a that admits of φ is permitted". We assume that conjunction $\wedge$, implication $\to$, and Boolean constants true $\top$ and false $\bot$ are defined in the usual way. Also, by $\bigwedge_{i\le n}\varphi_i$ and $\bigvee_{i\le n}\varphi_i$ we denote, respectively, the conjunction and the disjunction of the formulae $\varphi_1, \dots \varphi_n$. As usual, we assume that the conjunction and the disjunction of an empty list are $\top$ and $\bot$, respectively.

Definition 2. *For each transition system (S, Δ, D, M, π), each state $s \in S$, and each formula $\varphi \in \Phi$, the **satisfaction relation** $s \Vdash \varphi$ is defined recursively as follows:*

1. *$s \Vdash p$, if $s \in \pi(p)$;*
2. *$s \Vdash \neg\varphi$, if $s \nVdash \varphi$;*
3. *$s \Vdash \varphi \vee \psi$, if $s \Vdash \varphi$ or $s \Vdash \psi$;*
4. *$s \Vdash \mathsf{WA}_a\varphi$, if $(s, i) \not\rightsquigarrow_a \neg\varphi$ for some $i \in D_a^s$;*
5. *$s \Vdash \mathsf{WE}_a\varphi$, if $(s, i) \rightsquigarrow_a \varphi$ for some $i \in D_a^s$;*
6. *$s \Vdash \mathsf{SE}_a\varphi$, if $i \in D_a^s$ for each i such that $(s, i) \rightsquigarrow_a \varphi$;*
7. *$s \Vdash \mathsf{SA}_a\varphi$, if $i \in D_a^s$ for each i such that $(s, i) \not\rightsquigarrow_a \neg\varphi$,*

where the notation $(s, i) \rightsquigarrow_a \varphi$ means that, for each tuple $(\delta, t) \in M_s$, if $\delta_a = i$, then $t \Vdash \varphi$.

Items 4 - 7 above capture the generalised notions of permissions in statements (WA), (WE), (SE), and (SA) in Section 1. Informally, $(s, i) \rightsquigarrow_a \varphi$ means that that action i of agent a in state s *ensures* that φ is true in the next state. Accordingly, $(s, i) \not\rightsquigarrow_a \neg\varphi$ means that action i of agent a in state s *admits* of the situation that φ is true in the next state. Observe that, if a transition system is deterministic and has only one agent a, then $(s, i) \not\rightsquigarrow_a \neg\varphi$ if and only if $(s, i) \rightsquigarrow_a \varphi$. Then, the next lemma follows from items 4 - 7 of Definition 2.

Lemma 1. *If set $\mathcal{A}$ contains only agent a, then for any formula $\varphi \in \Phi$ and state s of a deterministic transition system,*

1. *$s \Vdash \mathsf{WA}_a\varphi$ if and only if $s \Vdash \mathsf{WE}_a\varphi$;*

2. $s \Vdash \mathsf{SA}_a\varphi$ *if and only if* $s \Vdash \mathsf{SE}_a\varphi$.

Note that, in other cases (*i.e.* multiagent or nondeterministic systems), these modalities are not only semantically *inequivalent* but also *undefinable* through each other. We show this in Section 4.

3.1 Model Checking

Following Definition 2, we consider the ***global*** model-checking problem [Müller-Olm *et al.*, 1999] of language Φ. For a finite transition system and a formula $\varphi \in \Phi$, the global model checking determines the *truth set* $[\![\varphi]\!]$ that consists of all states satisfying φ in the transition system. Formally, we define the truth set of a formula as follows.

Definition 3. *For any given transition system and any formula $\varphi \in \Phi$, the **truth set** $[\![\varphi]\!]$ is the set $\{s \mid s \Vdash \varphi\}$.*

The global model checking of formula φ applies a trivial recursive process on its structural complexity. The next theorem shows its time complexity. See Appendix A of the full version [Shi, 2024] for a detailed analysis of the algorithm.

Theorem 1 (time complexity)**.** *For a finite transition system (S, Δ, D, M, π) and a formula $\varphi \in \Phi$, the time complexity of global model checking is $O\big(|\varphi| \cdot (|S| + |M| + |\Delta|)\big)$, where $|\varphi|$ is the size of the formula, $|S|$ is the number of states, $|M| = \sum_{s\in S} |M_s|$ is the size of the mechanism, and $|\Delta| = \sum_{a\in\mathcal{A}} \sum_{s\in S} |\Delta^s_a|$ is the size of the action space.*

3.2 Reduction to Other Logics

Recall statements (1) and (6) in Section 2, which show the way how Anderson and Kanger reduces SDL. Using a similar technique, we can translate our modalities into modalities in STIT logic and ATL [Alur *et al.*, 2002] after properly interpreting the transition system in Definition 1.

Reduction to STIT Instead of being about the states, the statements in STIT logic are about moment-history (m/h) pairs. Due to this fact, there is no exact reduction of our logic to STIT logic. However, we can interpret our modalities in STIT models in the appearance of the necessity and possibility modalities $\Box$ and $\Diamond$[1]. In order to do this, we first incorporate the deontic constraints into the models as atomic propositions. To be specific, $m/h \Vdash d_a$ represents that the action of agent a at moment m that includes history h is permitted. We use the modality XSTIT [Broersen, 2008; Broersen, 2011]. Informally, $m/h \Vdash \mathsf{XSTIT}_a\varphi$ could be interpreted as "the action of agent a at moment m that includes history h sees to it that φ is true at the next moment". Then, our four modalities can be translated as:

$$\mathsf{WA}_a\varphi := \Diamond(d_a \wedge \neg\mathsf{XSTIT}_a\neg\varphi);$$
$$\mathsf{WE}_a\varphi := \Diamond(d_a \wedge \mathsf{XSTIT}_a\varphi);$$
$$\mathsf{SE}_a\varphi := \Box(\mathsf{XSTIT}_a\varphi \to d_a);$$
$$\mathsf{SA}_a\varphi := \Box(\neg\mathsf{XSTIT}_a\neg\varphi \to d_a).$$

Reduction to ATL Unlike in STIT logic, in ATL, the statements are about states but there is no way to express the properties of actions. For this reason, we encode deontic constraints into the states. To do this, we expand each state in our original transition system into a set of states in the ATL model. Specifically, each state s in our original transition system corresponds to a set $\{\langle s, \mathcal{D}\rangle \mid \mathcal{D} \subseteq \mathcal{A}\}$ of states in the ATL model. Informally, the tuple $\langle s, \mathcal{D}\rangle$ encodes the information that "state s is reached after the agents in set $\mathcal{D}$ taking permitted actions and the others taking non-permitted actions". Then, $\langle s, \mathcal{D}\rangle \Vdash d_a$ if and only if $a \in \mathcal{D}$. Also, $\langle s, \mathcal{D}\rangle \Vdash p$ if and only if $s \Vdash p$ in our original transition system. Correspondingly, the transition $(\langle s, *\rangle, \langle t, \mathcal{D}\rangle)$ exists in the ATL model if there is a tuple $(\delta, t) \in M_s$ in our original transition system such that $\mathcal{D} = \{a \in \mathcal{A} \mid \delta_a \in D^s_a\}$ and $*$ is a wildcard. Note that, ATL requires the transitions to be deterministic. Thus, if needed, we incorporate a dummy agent $Nature$ into the agent set $\mathcal{A}$ to achieve determinacy. Then, we can translate our modalities into standard ATL syntax as:

$$\mathsf{WA}_a\varphi := \langle\!\langle \mathcal{A}\rangle\!\rangle \mathsf{X}(d_a \wedge \varphi);$$
$$\mathsf{WE}_a\varphi := \langle\!\langle a\rangle\!\rangle \mathsf{X}(d_a \wedge \varphi);$$
$$\mathsf{SE}_a\varphi := \neg\langle\!\langle a\rangle\!\rangle \mathsf{X}\neg(\varphi \to d_a);$$
$$\mathsf{SA}_a\varphi := \neg\langle\!\langle \mathcal{A}\rangle\!\rangle \mathsf{X}\neg(\varphi \to d_a),$$

where $\langle\!\langle \mathcal{C}\rangle\!\rangle \mathsf{X}\varphi$ is informally interpreted as "the agents in set $\mathcal{C}$ can cooperate to enforce φ in the next state" and $\langle\!\langle a\rangle\!\rangle$ is the abbreviation for $\langle\!\langle \{a\}\rangle\!\rangle$.

4 Mutual Undefinability

As we define four modalities in the language, we would like to figure out if all of them are necessary to express the corresponding notions of permission. Specifically, if some of these modalities are semantically definable through the others, then the definable ones are not necessary for the language. As an example, a well-known result in Boolean logic is De Morgan's laws, which say conjunction and disjunction are interdefinable in the presence of negation. Therefore, to consider a "minimal" system for propositional logic, it is not necessary to include both conjunction and disjunction.

In this section, we consider the definability of modalities in the same way as De Morgan's laws (*i.e.* semantical equivalence). For example, modality WA is definable through modalities WE, SE, and SA if every formula in language Φ is semantically equivalent to a formula using only modalities WE, SE, and SA. Formally, in the transition systems, we define semantical equivalence as follows.

Definition 4. *Formulae φ and ψ are **semantically equivalent** if $[\![\varphi]\!] = [\![\psi]\!]$ for each transition system.*

We prove that none of the modalities WA, WE, SE, and SA is definable through the other three. To do this, it suffices to show that, for each modality $\odot$ of the four modalities, there exists a formula $\odot\varphi \in \Phi$ and a transition system where $[\![\odot\varphi]\!] \neq [\![\psi]\!]$ for each formula ψ not using modality $\odot$. In particular, we use the *truth set algebra* technique [Knight *et al.*, 2022]. This technique uses one model (*i.e.* transition system) and shows that, in this model, the truth sets of all formulae ψ not using modality $\odot$ form a *proper subset* of the family of all truth sets in language Φ, while the truth set of the formula $\odot\varphi$ does not belong to this subset. We formally state the undefinability results in the next theorem. A detailed explanation of the technique and proof can be found in Appendix B of the full version.

[1] $m/h \Vdash \Box\varphi$ iff $m/h' \Vdash \varphi$ for each history h' such that $m \in h'$.

Theorem 2 (undefinability of WA)**.** *The formula* $\mathsf{WA}_a p$ *is not semantically equivalent to any formula in language* Φ *that does not use modality* WA*.*

The formal statements of the undefinability results for modalities WE, SE, and SA are the same as Theorem 2 except for using the corresponding modalities instead of WA, see Theorem 5, Theorem 6, and Theorem 7 in Appendix B of the full version.

Note that, all four undefinability results, as presented in Appendix B of the full version, require that our language contains at least two agents. In single-agent settings, if a transition system is nondeterministic, these undefinability results still hold. This can be observed by modifying the two-agent transition systems in the proofs into single-agent nondeterministic transition systems by treating one of the agents as the nondeterministic factor. If a single-agent transition system is deterministic, then, as observed in Lemma 1, modalities WA and WE are semantically equivalent, so as modalities SE and SA. However, *modalities* WA *(*WE*) and* SA *(*SE*) are not definable through each other*, see Appendix B.5 of the full version.

5 Axioms

In addition to the tautologies in language Φ, our logical system contains the following schemes of axioms for all agents $a, b \in \mathcal{A}$ and all formulae $\varphi, \psi \in \Phi$:

A1. $\neg\mathsf{WA}_a\bot$;

A2. $\mathsf{WE}_a\top$;

A3. $\mathsf{SA}_a\bot$;

A4. $\mathsf{SE}_a\top \to \mathsf{SA}_a\top$;

A5. $\mathsf{WA}_a(\varphi\vee\psi) \to \mathsf{WA}_a\varphi \vee \mathsf{WA}_a\psi$;

A6. $\mathsf{SA}_a\varphi \wedge \mathsf{SA}_a\psi \to \mathsf{SA}_a(\varphi\vee\psi)$;

A7. $\mathsf{WE}_a\varphi \wedge \neg\mathsf{WE}_a\psi \to \mathsf{WA}_a(\varphi\wedge\neg\psi)$;

A8. $\neg\mathsf{SE}_a\varphi \wedge \mathsf{SE}_a\psi \to \neg\mathsf{SA}_a(\varphi\wedge\neg\psi)$;

A9. $\neg\mathsf{WA}_a\varphi \wedge \mathsf{SA}_a\psi \to \neg\mathsf{WA}_b(\varphi\wedge\psi) \wedge \mathsf{SA}_b(\varphi\wedge\psi)$.

Axiom A1 says agent a does not have a permitted action that has no next state. This is true because of the continuity property of the mechanism (item 4 of Definition 1). Axiom A2 says agent a always has a permitted action that ensures a next state. This is true because of the continuity property and the nonempty set of permitted actions (item 3 of Definition 1). Axiom A3 says every action that may have no next state is permitted. This is true because no such actions exist again due to the continuity property. Axiom A4 is true because both $\mathsf{SE}_a\top$ and $\mathsf{SA}_a\top$ mean that every action of agent a is permitted.

Axiom A5 says, if agent a has a permitted action that admits of $\varphi\vee\psi$, then agent a either has a permitted action that admits of φ or has a permitted action that admits of ψ. This is true because the permitted action that admits of $\varphi\vee\psi$ indeed either admits of φ or admits of ψ (item 3 of Definition 2). Axiom A6 says, if every action of agent a that admits of φ is permitted and every action of agent a that admits of ψ is permitted, then every action of agent a that admits of $\varphi\vee\psi$ is permitted. This is true because any action that admits of $\varphi\vee\psi$ either admits of φ or admits of ψ (item 3 of Definition 2).

Axiom A7 says, if agent a has a permitted action that ensures φ and has no permitted action that ensures ψ, then agent a has a permitted action that admits of $\varphi\wedge\neg\psi$. This is true because the permitted action i that ensures φ does not ensure ψ. Hence, action i admits of $\neg\psi$ while φ is ensured to happen. Axiom A8 says, if agent a has a non-permitted action that ensures φ and every action that ensures ψ is permitted, then agent a has a non-permitted action that admits of $\varphi\wedge\neg\psi$. This is true because the non-permitted action j that ensures φ does not ensure ψ. Hence, action j admits of $\neg\psi$ while φ is ensured to happen.

Axiom A9 says, if agent a has no permitted action that admits of φ and every action that admits of ψ is permitted, then agent b has no permitted action that admits of $\varphi\wedge\psi$ and every action of agent b that admits of $\varphi\wedge\psi$ is permitted. This is true because the antecedent means agent a's permitted actions ensure $\neg\varphi$ and non-permitted actions (if existing) ensure $\neg\psi$. Thus, every action of agent a ensures $\neg\varphi\vee\neg\psi$ (item 3 of Definition 2). Hence, $\neg(\varphi\wedge\psi)$ is *unavoidable* in the next state. This implies that agent b has no permitted action that admits of $\varphi\wedge\psi$, and any action of agent b that admits of $\varphi\wedge\psi$ is permitted because no such action of agent b exists.

We write $\vdash\varphi$ and say that formula φ is a ***theorem*** of our logical system if it can be derived from the axioms using the following four inference rules:

IR1. $\dfrac{\varphi,\quad \varphi\to\psi}{\psi}$ (Modus Ponens);

IR2. $\dfrac{\varphi\to\psi}{\mathsf{WA}_a\varphi\to\mathsf{WA}_a\psi}$;

IR3. $\dfrac{\varphi\to\psi}{\mathsf{SA}_a\psi\to\mathsf{SA}_a\varphi}$;

IR4. $\dfrac{\varphi_1\wedge\dots\wedge\varphi_m\to\neg\psi_1\vee\dots\vee\neg\psi_n}{\mathsf{WE}_{a_1}\varphi_1\wedge\dots\wedge\mathsf{WE}_{a_m}\varphi_m\to\mathsf{SE}_{b_1}\psi_1\vee\dots\vee\mathsf{SE}_{b_n}\psi_n}$, where agents $a_1,\dots,a_m,b_1,\dots,b_n$ are distinct.

Rule IR2 is the **monotonicity** rule for modality WA. It is valid because, in each state of each transition system, the permitted action of agent a that admits of φ also admits of ψ, as $\varphi\to\psi$ is universally true. By this rule, modality WA represents a form of ***weak permission*** following statement (2). Rule IR3 is the **anti-monotonicity** rule for modality SA. It is valid because, in each state of each transition system, the set of actions that admits of ψ is a *superset* of the set of actions that admits of φ, as $\varphi\to\psi$ is universally true. Hence, as long as the actions in the former set are all permitted, those in the latter set are also permitted. This rule shows that modality SA represents a form of ***strong permission*** following statement (5). It can be derived that modality WE represents a form of weak permission and modality SE represents a form of strong permission, see Appendix C.1 of the full version.

Rule IR4 is a conflict-preventing rule following the notion of "ensure" in semantics. The premise says, if every one of $\varphi_1,\dots,\varphi_m$ is true, then at least one of $\psi_1,\dots,\psi_n$ is false. The conclusion says, for a set of distinct agents $\{a_1,\dots,a_m,b_1,\dots,b_n\}\subseteq\mathcal{A}$, if every agent a_i has a permitted action to ensure φ_i, then for at least one agent b_j, every action that ensures ψ_j is permitted. This is valid because there would be a conflict otherwise. To be specific, if there is a state s where the conclusion of the rule is false, then, each agent a_i has a permitted action k_i that ensures φ_i and each agent b_j has a non-permitted action ℓ_j that ensures ψ_j. Consider an action profile δ such that $\delta_{a_i} = k_i$ for each $i \leq m$ and $\delta_{b_j} = \ell_j$ for each $j \leq n$. By the continuity condition in item 4 of Definition 1, there is a state t such that $(\delta, t) \in M_s$. In such state t, each of the formulae $\varphi_1, \dots, \varphi_m, \psi_1, \dots, \psi_n$ is true. However, this conflicts with the premise of the rule.

We write $X \vdash \varphi$ if a formula φ can be derived from the *theorems* of our logical system and an additional set of assumptions X using *only* the Modus Ponens rule. Note that statements $\varnothing \vdash \varphi$ and $\vdash \varphi$ are equivalent. We say that a set of formulae X is *consistent* if $X \nvdash \bot$.

Lemma 2 (deduction). *If $X, \varphi \vdash \psi$, then $X \vdash \varphi \to \psi$.*

See the proof of Lemma 2 in Appendix C.2 of the full version.

Lemma 3 (Lindenbaum). *Any consistent set of formulae can be extended to a maximal consistent set of formulae.*

The standard proof of this lemma can be found in [Mendelson, 2009, Proposition 2.14].

Lemma 4. $\vdash \mathsf{WE}_a\varphi \wedge \neg\mathsf{WA}_a\psi \to \mathsf{WE}_a(\varphi \wedge \neg\psi)$.

Lemma 5. $\vdash \neg\mathsf{SE}_a\varphi \wedge \mathsf{SA}_a\psi \to \neg\mathsf{SE}_a(\varphi \wedge \neg\psi)$.

See the proofs of the above two lemmas in Appendix C.3 of the full version. The next theorem follows from the above discussion of the axioms and the inference rules.

Theorem 3 (soundness). *If $\vdash \varphi$, then $s \Vdash \varphi$ for each state s of each transition system.*

6 Completeness

In this section, we prove the strong completeness of our logical system. As usual, at the core of the completeness theorem is the canonical model construction. In our case, it is a canonical transition system.

6.1 Canonical Transition System

In this subsection, we define the canonical transition system (S, Δ, D, M, π).

Definition 5. *Set S of states is the family of all maximal consistent subsets of our language Φ.*

For each formula $\varphi \in \Phi$, we introduce two actions: a permitted action φ^+ and a non-permitted action φ^-. Formally, by φ^+ and φ^- we mean the pairs $(\varphi, +)$ and $(\varphi, -)$, respectively. By item 7 of Definition 2, the formula $\mathsf{SA}_a\top$ expresses that agent a is permitted to use *every* action. In other words, if $\mathsf{SA}_a\top$ is true, then there are no non-permitted actions available to agent a in the current state. This explains the intuition behind the following definition.

Definition 6. *For each state $s \in S$ and each agent $a \in \mathcal{A}$,*

$$\Delta_a^s = \begin{cases} \{\varphi^+ \mid \varphi \in \Phi\}, & \text{if } \mathsf{SA}_a\top \in s, \\ \{\varphi^+, \varphi^- \mid \varphi \in \Phi\}, & \text{otherwise;} \end{cases}$$

$$D_a^s = \{\varphi^+ \mid \varphi \in \Phi\}.$$

The next definition is the key part of the canonical transition system construction. It specifies the mechanism of the transition system. Recall that $\neg\mathsf{WA}_a\varphi$ means that agent a is not permitted to use any action that admits of φ. Hence, each permitted action of a must ensure $\neg\varphi$. We capture this rule in item 1 of the definition below. Recall that $\mathsf{WE}_a\varphi$ means that agent a has at least one permitted action that ensures φ. In the canonical transition system, this action is defined to be φ^+. The rule captured in item 2 below guarantees φ in the next state whenever agent a uses action φ^+. Next, $\neg\mathsf{SE}_a\varphi$ means that agent a is not permitted to use at least one action that ensures φ. We denote such action by φ^-. The rule captured by item 3 stipulates that action φ^- ensures φ. Finally, $\mathsf{SA}_a\varphi$ means that agent a is permitted to use all actions that admit of φ. In other words, $\mathsf{SA}_a\varphi$ means that all non-permitted actions ensure $\neg\varphi$. This is captured by item 4 below.

Definition 7. *$(\delta, t) \in M_s$ when for each agent $a \in \mathcal{A}$ and each formula $\varphi \in \Phi$,*

1. *if $\delta_a \in D_a^s$ and $\mathsf{WA}_a\varphi \notin s$, then $\neg\varphi \in t$;*
2. *if $\delta_a = \varphi^+$ and $\mathsf{WE}_a\varphi \in s$, then $\varphi \in t$;*
3. *if $\delta_a = \varphi^-$ and $\mathsf{SE}_a\varphi \notin s$, then $\varphi \in t$;*
4. *if $\delta_a \in \Delta_a^s \setminus D_a^s$ and $\mathsf{SA}_a\varphi \in s$, then $\neg\varphi \in t$.*

Note that, each state s is a maximal consistent set by Defintion 5. Hence, for item 1 above, the statement $\mathsf{WA}_a\varphi \notin s$ is equivalent to $\neg\mathsf{WA}_a\varphi \in s$. The same goes for item 3.

Definition 8. $\pi(p) = \{s \in S \mid p \in s\}$ *for each p.*

This concludes the definition of the canonical transition system (S, Δ, D, M, π). Next, we show that it satisfies the *continuity* condition in item 4 of Definition 1.

Lemma 6 (continuity). *For each state $s \in S$ and each action profile $\delta \in \prod_{a \in \mathcal{A}} \Delta_a^s$, there is a state t such that $(\delta, t) \in M_s$.*

Proof. Consider a partition $\{A, B\}$ of the set $\mathcal{A}$ of agents:

$$A = \{a \in \mathcal{A} \mid \delta_a \in D_a^s\}; \tag{7}$$

$$B = \{b \in \mathcal{A} \mid \delta_b \in \Delta_b^s \setminus D_b^s\}. \tag{8}$$

Then, $\mathsf{SA}_b\top \notin s$ for each agent $b \in B$ by Definitions 6. Hence, $\neg\mathsf{SA}_b\top \in s$ for each agent $b \in B$ because s is a maximal consistent set. Then, by the contrapositive of axiom A4,

$$\neg\mathsf{SE}_b\top \in s. \tag{9}$$

Consider the set of formulae

$$\begin{aligned} X = &\{\neg\psi \mid \exists a \in A\, (\neg\mathsf{WA}_a\psi \in s)\} \\ &\cup \{\sigma \mid \exists a \in A\, (\delta_a = \sigma^+, \mathsf{WE}_a\sigma \in s)\} \\ &\cup \{\neg\chi \mid \exists b \in B\, (\mathsf{SA}_b\chi \in s)\} \\ &\cup \{\tau \mid \exists b \in B\, (\delta_b = \tau^-, \neg\mathsf{SE}_b\tau \in s)\}. \end{aligned} \tag{10}$$

Claim 1. *Set X is consistent.*

Proof of Claim. Suppose the opposite. Then, by axiom A2, statements (9) and (10), there are formulae

$$\begin{aligned} &\neg\mathsf{WA}_{a_1}\psi_{11}, \dots, \neg\mathsf{WA}_{a_1}\psi_{1k_1}, \mathsf{WE}_{a_1}\widehat{\sigma}_1 \in s, \\ &\dots \\ &\neg\mathsf{WA}_{a_m}\psi_{m1}, \dots, \neg\mathsf{WA}_{a_m}\psi_{mk_m}, \mathsf{WE}_{a_m}\widehat{\sigma}_m \in s, \end{aligned} \tag{11}$$

and

$$\begin{aligned} &\mathsf{SA}_{b_1}\chi_{11}, \dots, \mathsf{SA}_{b_1}\chi_{1\ell_1}, \neg\mathsf{SE}_{b_1}\widehat{\tau}_1 \in s, \\ &\dots \\ &\mathsf{SA}_{b_n}\chi_{n1}, \dots, \mathsf{SA}_{b_n}\chi_{n\ell_n}, \neg\mathsf{SE}_{b_n}\widehat{\tau}_n \in s, \end{aligned} \tag{12}$$

where

$$a_1, \dots, a_m, b_1, \dots, b_n \text{ are distinct agents,} \tag{13}$$

$$\widehat{\sigma}_i = \begin{cases} \sigma_i, & \text{if } \delta_{a_i} = \sigma_i^+ \text{ and } \mathsf{WE}_{a_i}\sigma_i \in s, \\ \top, & \text{otherwise,} \end{cases}$$

for each $i \leq m$, and

$$\widehat{\tau}_i = \begin{cases} \tau_i, & \text{if } \delta_{b_i} = \tau_i^- \text{ and } \neg\mathsf{SE}_{b_i}\tau_i \in s, \\ \top, & \text{otherwise,} \end{cases}$$

for each $i \leq n$, such that

$$\bigwedge_{i \leq m} \Big(\widehat{\sigma}_i \wedge \bigwedge_{j \leq k_i} \neg\psi_{ij}\Big) \wedge \bigwedge_{i \leq n} \Big(\widehat{\tau}_i \wedge \bigwedge_{j \leq \ell_i} \neg\chi_{ij}\Big) \vdash \bot. \tag{14}$$

By multiple application of the countrapositive of axiom A5 and propositional reasoning, statement (11) implies that

$$s \vdash \neg\mathsf{WA}_{a_i}\Big(\bigvee_{j \leq k_i} \psi_{ij}\Big) \text{ for each } i \leq m. \tag{15}$$

Note that, in the specific case where $k_i = 0$, statement (15) follows directly from axiom A1. By Lemma 4, statement (15) and the part $\mathsf{WE}_{a_i}\widehat{\sigma}_i \in s$ of statement (11) imply

$$s \vdash \mathsf{WE}_{a_i}\Big(\widehat{\sigma}_i \wedge \neg \bigvee_{j \leq k_i} \psi_{ij}\Big) \text{ for each } i \leq m. \tag{16}$$

Meanwhile, by multiple application of Lemma 2 and propositional reasoning, statement (14) can be reformulated to

$$\vdash \bigwedge_{i \leq m} \Big(\widehat{\sigma}_i \wedge \neg \bigvee_{j \leq k_i} \psi_{ij}\Big) \to \bigvee_{i \leq n} \neg\Big(\widehat{\tau}_i \wedge \bigwedge_{j \leq \ell_i} \neg\chi_{ij}\Big). \tag{17}$$

By statement (13) and rule IR4, statement (17) implies

$$\vdash \bigwedge_{i \leq m} \mathsf{WE}_{a_i}\Big(\widehat{\sigma}_i \wedge \neg \bigvee_{j \leq k_i} \psi_{ij}\Big) \to \bigvee_{i \leq n} \mathsf{SE}_{b_i}\Big(\widehat{\tau}_i \wedge \bigwedge_{j \leq \ell_i} \neg\chi_{ij}\Big).$$

Then, by statement (16) and the Modus Ponens rule,

$$s \vdash \bigvee_{i \leq n} \mathsf{SE}_{b_i}\Big(\widehat{\tau}_i \wedge \neg \bigvee_{j \leq \ell_i} \chi_{ij}\Big). \tag{18}$$

At the same time, by multiple application of axiom A6 and propositional reasoning, statement (12) implies

$$s \vdash \mathsf{SA}_{b_i}\Big(\bigvee_{j \leq \ell_i} \chi_{ij}\Big) \text{ for each } i \leq n. \tag{19}$$

Note that, in the specific case where $\ell_i = 0$, statement (19) follows directly from axiom A3. By Lemma 5, statement (19) and the part $\neg\mathsf{SE}_{b_i}\widehat{\tau}_i \in s$ of statement (12) imply

$$s \vdash \neg\mathsf{SE}_{b_i}\Big(\widehat{\tau}_i \wedge \neg \bigvee_{j \leq \ell_i} \chi_{ij}\Big) \text{ for each } i \leq n,$$

which contradicts statement (18). ⊠

Let t be any maximal consistent extension of set X. By Lemma 3, such t must exist. Hence, $t \in S$ by Definition 5.

Claim 2. $(\delta, t) \in M_s$.

Proof of Claim. It suffices to verify that conditions 1 – 4 of Definition 7 are satisfied for the tuple (δ, t) for each agent $x \in \mathcal{A}$. Recall that sets A and B form a partition of the agent set $\mathcal{A}$. Hence, it suffices to consider the following two cases.

Case 1: $x \in A$. Condition 1 of Definition 7 follows from line 1 of statement (10) because $X \subseteq t$. Condition 2 follows from line 2 of statement (10). Conditions 3 and 4 trivially follow from $\delta_x \in D_x^s$ by statement (7).

Case 2: $x \in B$. Conditions 1 and 2 of Definition 7 trivially follow from $\delta_x \notin D_x^s$ by statement (8). Condition 3 follows from line 4 of statement (10) because $X \subseteq t$. Condition 4 follows from line 3 of statement (10). ⊠

The statement of this lemma follows from Claim 2. □

6.2 Strong Completeness Theorem

As usual, at the core of the proof of completeness is a truth lemma proven by induction on the structural complexity of a formula. In our case, it is the Lemma 7. The completeness result, as shown in Theorem 4, is proved with Lemma 7 in the standard way. We put the formal proofs in Appendix D of the full version.

Lemma 7. *$s \Vdash \varphi$ if and only if $\varphi \in s$ for each state s of the canonical transition system and each formula $\varphi \in \Phi$.*

Theorem 4 (strong completeness)**.** *For each set of formulae $X \subseteq \Phi$ and each formula $\varphi \in \Phi$ such that $X \nvdash \varphi$, there is a state s of a transition system such that $s \Vdash \chi$ for each $\chi \in X$ and $s \nVdash \varphi$.*

7 Conclusion and Future Research

We are the first to classify the agentive permissions in multiagent settings into permissions to ensure and permissions to admit and cross-discuss them in both weak and strong forms. To do this, we propose and formalise four forms of agentive permissions in multiagent transition systems, analyse the time complexity of the model checking algorithm, prove their semantical undefinability through each other, and give a sound and complete logical system that reveals their interplay.

Future research could be in two directions. One is to extend the deontic constraints from one-step actions to multi-step actions. Indeed, multi-step deontic constraints are commonly seen in application scenarios. For example, if a child is permitted to eat only one ice cream per day, then whether to eat an ice cream in the morning affects whether she is permitted to eat one in the afternoon. This is closely related to conditional norms (*e.g.* conditional obligations) discussed in the literature [van Fraassen, 1973; Chellas, 1974; DeCew, 1981; Rulli, 2020] but is interpreted in multiagent transition systems instead of possible-world semantics. The other direction could be the interaction between permission and responsibility in multiagent settings. It might have been noticed that our introductory example about factories and fish is a variant of [Halpern, 2015, Example 3.11] and [Halpern, 2016, Example 6.2.5], which talk about causality and responsibility in multiagent settings. Indeed, the connection between obligation and responsibility in linguistic intuition has already been noticed by philosophers [van de Poel, 2011]. However, a formal investigation of the interaction among norm, causation, and responsibility in multiagent settings is still lacking.

Acknowledgements

I would like to express my sincere gratitude to Dr. Pavel Naumov, my PhD supervisor, for his selfless help in completing this paper. The research is funded by the China Scholarship Council (CSC No.202206070014).

References

[Alur *et al.*, 2002] Rajeev Alur, Thomas A Henzinger, and Orna Kupferman. Alternating-time temporal logic. *Journal of the ACM*, 49(5):672–713, 2002.

[Anderson, 1956] Alan Ross Anderson. The formal analysis of normative systems. Technical report, Office of Naval Research, 1956.

[Anderson, 1967] Alan Ross Anderson. Some nasty problems in the formal logic of ethics. *Noûs*, pages 345–360, 1967.

[Anglberger *et al.*, 2015] Albert JJ Anglberger, Nobert Gratzl, and Olivier Roy. Obligation, free choice, and the logic of weakest permissions. *The Review of Symbolic Logic*, 8(4):807–827, 2015.

[Areces *et al.*, 2023] Carlos Areces, Valentin Cassano, Pablo F Castro, Raul Fervari, and Andrés R Saravia. A deontic logic of knowingly complying. In *Proceedings of the 22nd International Conference on Autonomous Agents and MultiAgent Systems*, volume 23, 2023.

[Arkoudas *et al.*, 2005] Konstantine Arkoudas, Selmer Bringsjord, and Paul Bello. Toward ethical robots via mechanized deontic logic. In *AAAI Fall Symposium on Machine Ethics*, pages 17–23. The AAAI Press Menlo Park, CA, USA, 2005.

[Asher and Bonevac, 2005] Nicholas Asher and Daniel Bonevac. Free choice permission is strong permission. *Synthese*, 145:303–323, 2005.

[Belnap and Perloff, 1988] Nuel Belnap and Michael Perloff. Seeing to it that: a canonical form for agentives. *Theoria*, 54(3):175–199, 1988.

[Belnap and Perloff, 1992] Nuel Belnap and Michael Perloff. The way of the agent. *Studia Logica*, pages 463–484, 1992.

[Broersen and Ramírez Abarca, 2018] Jan Broersen and Aldo Iván Ramírez Abarca. Formalising oughts and practical knowledge without resorting to action types. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*, pages 1877–1879, 2018.

[Broersen, 2008] Jan Broersen. A complete STIT logic for knowledge and action, and some of its applications. In *Proceedings of the 6th International Workshop on Declarative Agent Languages and Technologies*, pages 47–59. Springer, 2008.

[Broersen, 2011] Jan Broersen. Deontic epistemic STIT logic distinguishing modes of mens rea. *Journal of Applied Logic*, 9(2):137–152, 2011.

[Chellas, 1969] Brian Farrell Chellas. *The logical form of imperatives*. Stanford University, 1969.

[Chellas, 1974] Brian F Chellas. Conditional obligation. In *Logical Theory and Semantic Analysis: Essays Dedicated to Stig Kanger on His Fiftieth Birthday*, pages 23–33. Springer, 1974.

[Chisholm, 1964] Roderick M Chisholm. The ethics of requirement. *American Philosophical Quarterly*, 1(2):147–153, 1964.

[DeCew, 1981] Judith Wagner DeCew. Conditional obligation and counterfactuals. *Journal of Philosophical Logic*, pages 55–72, 1981.

[Governatori *et al.*, 2013] Guido Governatori, Francesco Olivieri, Antonino Rotolo, and Simone Scannapieco. Computing strong and weak permissions in defeasible logic. *Journal of Philosophical Logic*, 42:799–829, 2013.

[Halpern, 2015] Joseph Y Halpern. A modification of the halpern-pearl definition of causality. In *Proceedings of the 24th International Conference on Artificial Intelligence*, pages 3022–3033, 2015.

[Halpern, 2016] Joseph Y Halpern. *Actual causality*. MIT Press, 2016.

[Horty and Belnap, 1995] John F Horty and Nuel Belnap. The deliberative STIT: a study of action, omission, ability, and obligation. *Journal of Philosophical Logic*, 24(6):583–644, 1995.

[Horty, 2001] John F Horty. *Agency and Deontic Logic*. Oxford University Press, 2001.

[Kamp, 1973] Hans Kamp. Free choice permission. In *Proceedings of the Aristotelian Society*, volume 74, pages 57–74. JSTOR, 1973.

[Kanger, 1971] Stig Kanger. *New Foundations for Ethical Theory*. Springer, 1971.

[Knight *et al.*, 2022] Sophia Knight, Pavel Naumov, Qi Shi, and Vigasan Suntharraj. Truth set algebra: a new way to prove undefinability. *arXiv:2208.04422*, 2022.

[Kulicki and Trypuz, 2017] Piotr Kulicki and Robert Trypuz. Connecting actions and states in deontic logic. *Studia Logica*, 105:915–942, 2017.

[Kulicki *et al.*, 2023] Piotr Kulicki, Robert Trypuz, Robert Craven, and Marek J Sergot. A unified logical framework for reasoning about deontic properties of actions and states. *Logic and Logical Philosophy*, pages 1–35, 2023.

[Lewis, 1979] David Lewis. A problem about permission. In *Essays in Honour of Jaakko Hintikka*, pages 163–175. Springer, 1979.

[McNamara and van de Putte, 2022] Paul McNamara and Frederik van de Putte. Deontic logic. In Edward N. Zalta and Uri Nodelman, editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, 2022.

[Mendelson, 2009] Elliott Mendelson. *Introduction to Mathematical Logic*. CRC press, 2009.

[Müller-Olm *et al.*, 1999] Markus Müller-Olm, David Schmidt, and Bernhard Steffen. Model-checking: a tutorial introduction. In *Static Analysis: Proceedings of the 6th International Symposium*, pages 330–354. Springer, 1999.

[Pandžić *et al.*, 2022] Stipe Pandžić, Jan Broersen, and Henk Aarts. Boid*: autonomous goal deliberation through abduction. In *Proceedings of the 21st International Conference on Autonomous Agents and MultiAgent Systems*, pages 1019–1027, 2022.

[Prior, 1963] Arthur N Prior. The logic of modality. In *Formal Logic*, pages 186–229. Oxford University Press, 1963.

[Raz, 1975] Joseph Raz. Permissions and supererogation. *American Philosophical Quarterly*, 12(2):161–168, 1975.

[Ross, 1944] Alf Ross. Imperatives and logic. *Philosophy of Science*, 11(1):30–46, 1944.

[Royakkers, 1997] Lamber MM Royakkers. Giving permission implies giving choice. In *Proceedings of the 8th International Conference on Database and Expert Systems Applications*, pages 198–203. IEEE, 1997.

[Rulli, 2020] Tina Rulli. Conditional obligations. *Social Theory and Practice*, pages 365–390, 2020.

[Shi, 2024] Qi Shi. Agentive permissions in multiagent systems. *arXiv:2404.17053*, 2024.

[van Benthem, 1979] Johan van Benthem. Minimal deontic logics. *Bulletin of the Section of Logic*, 8(1):36–42, 1979.

[van de Poel, 2011] Ibo van de Poel. The relation between forward-looking and backward-looking responsibility. In *Moral Responsibility: Beyond Free Will and Determinism*, pages 37–52. Springer, 2011.

[van de Putte, 2017] Frederik van de Putte. Free choice permission in STIT. In *Logica yearbook 2016*, pages 289–303. College Publications, 2017.

[van Fraassen, 1973] Bas C van Fraassen. The logic of conditional obligation. In *Exact Philosophy*, pages 151–172. Springer, 1973.

[von Wright, 1951] Georg Henrik von Wright. Deontic logic. *Mind*, 60(237):1–15, 1951.

[von Wright, 1968] Georg Henrik von Wright. *An Essay in Deontic Logic and the General Theory of Action: With a Bibliography of Deontic and Imperative Logic*. North-Holland Pub. Co, 1968.

[Wang and Wang, 2023] Zilu Wang and Yanjing Wang. Strong permission bundled: first steps. In *Proceedings of the 16th International Conference on Deontic Logic and Normative Systems*, pages 217–234. Springer, 2023.