Rights, Duties and Commitments between Agents

Leendert van der Torre
Max Planck Institute for Computer Science
Im Stadtwald, D-66123 Saarbrücken
Germany

Yao-Hua Tan
EURIDIS, Erasmus University
P.O. Box 1738, 3000 DR Rotterdam
The Netherlands

Abstract

In this paper we introduce a multi agent deontic update semantics, that builds on a logic of prescriptive obligations (norms) and a logic of descriptive obligations (normative propositions). In this preference-based logic we formalize rights as a new type of strong prescriptive permissions and duties and commitments as prescriptive obligations between agents.

1 Introduction

In groups of agents there is a fundamental difference between an agent *creating* an obligation or permission for another agent and an agent *evaluating* whether such deontic states hold, because the former is an act whereas the latter is an assessment. For example, a purchase contract creates an obligation for the buyer to pay the seller for the goods, when buying a CD via the Internet the buyer can grant permission to the merchant to charge his credit card by sending his card number, and the city counsel can grant a permission to a person to build a house. Usually deontic states are created by performing certain actions like the actual signing of a contract, instructions by a superior, and pressing an 'agree' button on a website.[1] These creations of deontic states are quite different from the evaluation of deontic states where one agent assesses the deontic states of another agent (or himself) to determine which obligations, permissions and rights actually hold. For example, a seller might want to know whether he or the buyer is responsible for paying import taxes for the goods that he shipped to the buyer, and an Internet consumer might want to know whether he has to pay if someone makes fraudulent use of the credit card number he gave to an Internet merchant. Clearly, the most difficult reasoning task concerns the interaction between creation and evaluation of deontic states. Signing a sales contract is easy, but deriving all the legal implications of the contract is a task for legal experts.

In advanced applications of multi-agent systems the agents must be able to reason about the creation as well as the evaluation of deontic states, because in electronic communication the agents should be able to reason whether they commit themselves or create any liability by the messages they exchange. Lack of this reasoning capacity becomes more dangerous as agents become more autonomous. In particular in electronic commerce applications there is an acute lack of adequate formalisms to enable autonomous agents to reason about the deontic states of other agents and to support electronic contracting [Schmid and Selz, 1998].

The most obvious formalism to reason about deontic states is deontic logic, that formalizes logical relations between obligations, permissions, rights etc. Unfortunately, almost all existing deontic logics are based on modal logic and focus exclusively on the evaluation of deontic states, and their ad hoc extensions to create deontic states (as commitments in [Liau, 1998]) only indirectly show logical relations between these creations. In this paper we propose a logic to model the two fundamentally different notions of creating and evaluating deontic states in multi-agent systems by formalizing the *actions* of changing and assessing deontic states. In this way we formalize the logical relations between norms that create deontic states, and between normative propositions that are true or false in a deontic state. In our logic we use the notion *prescriptive obligation* to reason about actions that create a deontic state, and the notion *descriptive obligation* to reason about the obligations that actually hold in a particular deontic state. Prescriptive obligations are motivated in part by a long-standing philosophical discussion on whether it is meaningful to ascribe truth values to norms [Alchourrón and Bulygin, 1981; Makinson, 1998; van der Torre and Tan, 1998b]. The semantics of our logic also cannot be based on truth values, because actions are not true or false. Instead it is based on the so-called update semantics.

Update semantics [Veltman, 1996] is based on the notions of update action and acceptance. An update action changes the information state of a particular person and a formula is accepted by a state if the update with this formula does not change the state. In other words, since this update does not affect the information

[1] Such an active action is typical for contractual obligations, but it is of course not always required. Certain rights, e.g. human rights, always apply and do not have to be created for every person individually.

state, this formula was already known by the person. Similarly, in our deontic logic we do not define a notion of truth for prescriptive and descriptive obligations, but only a notion of acceptance. For example, we check whether a certain deontic creation action was performed by checking whether the corresponding prescriptive obligation is accepted in a particular deontic state, and we check whether a particular obligation holds in a state by checking whether the corresponding descriptive obligation is accepted in this state.

This paper is organized by discussing the various elements of the update system: the language, the deontic states, the updates, the notion of acceptance and finally the validity relation.

2 Deontic logic

The system introduced in this paper combines Deontic Update Semantics (DUS) for prima facie obligations [van der Torre and Tan, 1998c] with Hyper-rational Conditional Logic (HRC) [Weydert, 1994]. HRC is a nested logic for default conditionals that may be seen as an extension of the dyadic deontic logic DSDL3 [Hansson, 1971] with nested obligations. HRC and DSDL3 are traditional monotonic modal logics, in which $O_\alpha p$ is read as 'agent α ought to do p,' $F_\alpha p$ as 'α is forbidden to do p' and $P_\alpha p$ as 'α is permitted to do p.' Moreover, $O^a_{\alpha_1,\alpha_2} p$ can be read as 'according to authority a, α_1 ought to do p towards claimant α_2,' if p is a propositional sentence, then $O_\alpha p$ can be read as 'α ought to see to it that p is the case,' and an ought-to-be obligation Op is read as 'p ought to be the case.' The standard semantics is given in terms of valuations and preference relations over possible worlds: obligations are what hold in the best or 'most preferred' of the accessible worlds. The modal operators evaluate deontic HRC states and are therefore called descriptive operators or normative propositions.

DUS for prima facie obligations formalizes deontic operators that can be overridden by stronger operators.[2] It is based on deontic DUS states, possible worlds structures $\langle W, W*, R, V\rangle$ where W is a set of worlds, $W*$ a subset of W representing the agent's epistemic state, R a ranking function of ordered pairs of worlds (called links) and V a propositional valuation function. Deontic operators can either refer to the so-called context of deliberation $W*$ or the context of justification W. The first contains only states the agent considers to be possible, and implies what should be done *now*. The latter also considers states which were ideal but are no longer reachable, and thus also represents violations. Acts of 'norming' such as commanding, permitting, and derogation change these states and are in this paper written as $\mathsf{oblige}^*_\alpha p$, $\mathsf{permit}^*_\alpha p$ and $\mathsf{forbid}^*_\alpha p$ for the context of deliberation and as $\mathsf{oblige}_\alpha p$, $\mathsf{permit}_\alpha p$ and $\mathsf{forbid}_\alpha p$ for the context of justification. The operators cannot be nested, because norms (e.g. oblige) and what is normed (e.g. p) are completely different.

We combine DUS with HRC by replacing the worlds in the deontic states by HRC models, as illustrated in Figure 1 below. The DUS state contains three worlds and its ranking function assigns 1 to each ordered pair of worlds. The HRC model contains four worlds which are totally ordered. The combined deontic state has as its worlds three not necessarily identical HRC models.

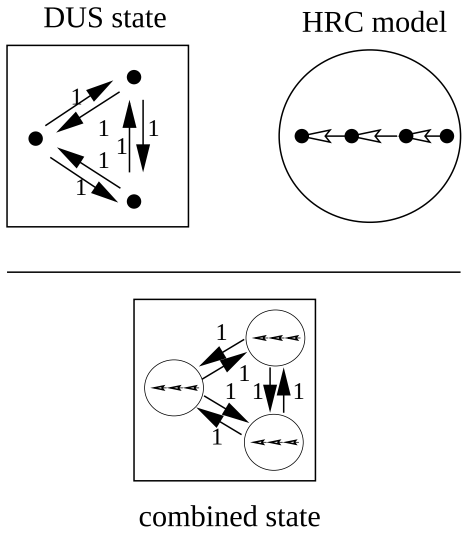

Figure 1: Combining DUS and HRC

The interaction between creation and evaluation of deontic states – as we already mentioned the most difficult reasoning task of the combined logic – is formalized by a reduction of the combined state to a HRC model. This reduction formalizes, among others, how the agent resolves conflicts between the prima facie operators. Descriptive operators that assess the thus constructed HRC model – so-called all-things-considered operators – are called $\mathsf{ideal}^*_\alpha p$, $\mathsf{someideal}^*_\alpha p$ and $\mathsf{noideal}^*_\alpha p$ for the context of deliberation and $\mathsf{ideal}_\alpha p$, $\mathsf{someideal}_\alpha p$ and $\mathsf{noideal}_\alpha p$ for the context of justification.

2.1 Deontic update semantics

We start with the basic definitions of Veltman's update semantics [Veltman, 1996]. To define a deontic update semantics for a deontic language L, one has to specify a set Σ of relevant deontic states (called information states in [Veltman, 1996]), and a function [ ] that assigns to each sentence ϕ an operation $[\phi]$ on Σ. If σ is a state and ϕ a sentence, then we write '$\sigma[\phi]$' to denote the result of updating σ with ϕ. We can write '$\sigma[\psi_1]\ldots[\psi_n]$' for the result of updating σ with the sequence of sentences $\psi_1, \ldots, \psi_n$. Moreover, one of the deontic states has to be labelled as the minimal deontic state, written as $\mathbf{0}$, and another one as the absurd state, written as $\mathbf{1}$.

In this paper we are primarily interested in ought-to-do operators between agents in multi-agent systems. We

[2]If a prima facie principle is overridden, then it can no longer turn into an absolute obligation, but it is still in force as a prima facie obligation. See [van der Torre and Tan, 1998c] for distinctions with other types of defeasible obligations.

therefore introduce sets of authorities and agents in the update system.

Definition 1 (MA-DUS) *A multi-agent deontic update system $\langle A_u, A_g, L, \Sigma, [\ ]\rangle$ consists of a set of authorities or normative systems A_u, a set of agents A_g, a logical language L, a set of relevant deontic states Σ and a function $[\ ]$ that assigns to each sentence ϕ of L an operation on Σ. Σ contains the elements* $\mathbf{0}$ *and* $\mathbf{1}$.

2.2 Deontic language

The base language is the language of hyper-rational conditional logic (HRC), relativised for authorities and agents. The conditional or dyadic operators are interpreted as directed obligations $O^a_{\alpha_1,\alpha_2}(p|q)$, and prohibitions and permissions are defined in terms of the obligations. Undirected operators are defined in terms of the directed operators by $O^a_{\alpha,\alpha}$. This base language is extended with dyadic prescriptive operators for obligation, permission, and for new all-things-considered tests ideal and $\mathsf{someideal}$. Prohibitions are again defined in terms of obligations.

We discriminate between propositions that can be influenced by agent action, and ones that cannot. For example, we cannot control whether it will rain, and therefore it does not make sense to say that someone is obliged to see to it that it rains. This is well known from other types of reasoning about agent behavior, for example in reasoning about desires and preferences in decision theory [Boutilier, 1994; Lang, 1996]. The extension of this simple and ad hoc formalization of actions to a full-fledged action logic is beyond the scope of this paper.

Definition 2 (Deontic language) *Let A_u be a set of authorities, A_g a set of agents and A_c and A_n two sets of non-logical symbols (controllable and uncontrollable propositional atoms).*

- *Let L^A_0 be a propositional modal language based on A_c and A_n with the modal operators $O^a_{\alpha_1,\alpha_2}$ for $a \in A_u$ and $\alpha_1, \alpha_2 \in A_g$, such that the first parameter of the dyadic operator does not contain occurrences of A_n. We write $P^a_{\alpha_1,\alpha_2}(p \mid q)$ for $\neg O^a_{\alpha_1,\alpha_2}(\neg p|q)$, $F^a_{\alpha_1,\alpha_2}(p \mid q)$ for $O^a_{\alpha_1,\alpha_2}(\neg p \mid q)$, $O^a_\alpha(p|q)$ for $O^a_{\alpha,\alpha}(p \mid q)$, $O^a_{\alpha_1,\alpha_2}p$ for $O^a_{\alpha_1,\alpha_2}(p \mid \top)$, etc, where $\top$ stands for any tautology.*
- *A string of symbols ϕ is a sentence of L^A_1 if and only if either ϕ is a sentence of L^A_0 or there are $a \in A_u$ and $\alpha_1, \alpha_2 \in A_g$ and two sentences p and q of L^A_0 (where p does not contain any occurrences of A_n) such that $\phi = X^a_{\alpha_1,\alpha_2}(p \mid q)$ for $X =$ oblige, oblige^*, permit, permit^*, ideal, ideal^*, $\mathsf{someideal}$ or $\mathsf{someideal}^*$. We write $\mathsf{forbid}^a_{\alpha_1,\alpha_2}(p|q)$ for $\mathsf{oblige}^a_{\alpha_1,\alpha_2}(\neg p|q)$, $\mathsf{noideal}^a_{\alpha_1,\alpha_2}(p|q)$ for $\mathsf{ideal}^a_{\alpha_1,\alpha_2}(\neg p|q)$, and $\mathsf{oblige}^a_{\alpha_1,\alpha_2}p$ for $\mathsf{oblige}^a_{\alpha_1,\alpha_2}(p|\top)$, etc.*

Some simple examples of the deontic logic literature illustrate the expressive power of the language. The first two formulas below illustrate nested permission (Richard has been permitted to permit Yannis to use the copier) and nested obligation (the minister has been obliged to see to it that the mayor is obliged to declare a state of emergency if there is high water). The next two formulas show how one obligation triggers a new one (the obligation to go to a meeting creates the obligation to write this meeting in your diary) and how one permission triggers a new one (the permission to drive creates the permission to drive a car). Finally, the last formula shows how contrary-to-duty obligations can be formalized (you have to pay a fine when trading drugs).

$\mathsf{permit}^a_R(P^a_Y copy)$
$\mathsf{oblige}^a_M(O^a_{m,M}(em-state|high-water)|\top)$
$\mathsf{oblige}^a_M(O^a_{m,M}(em-state|\top)|high-water)$
$\mathsf{oblige}^a_{\alpha_1,\alpha_2}(diary|O^a_{\alpha_1,\alpha_2}meeting)$
$\mathsf{permit}^a_\alpha(drive-car|P^a_\alpha drive)$
$\mathsf{oblige}^a_{\alpha_1,\alpha_2}(pay-fine|drugs \wedge F^a_{\alpha_1,\alpha_2}drugs)$

2.3 Deontic states

Most conditional or dyadic logics are extensions of Hansson's DSDL3 [Hansson, 1971]. It is based on models $\langle W, \leq, V\rangle$ where W is a set of worlds, $\leq$ a binary reflexive, transitive and totally connected relation on W, and V a propositional valuation function for each world. The model satisfies $O(p \mid q)$ if the minimal (or preferred) q worlds satisfy p. HRC is an extension of Hansson's logic with a second accessibility relation R to give meaning to nested conditionals. Moreover, HRC models $\langle w*, W, \leq, R, \models\rangle$ also contain an explicit actual world $w*$, the valuation function is replaced by a propositional satisfaction function $\models$ for the worlds, and an additional local stopperedness condition is imposed. The model satisfies $O(p \mid q)$ if the $\leq$-minimal (or preferred) R-accessible q worlds satisfy p. The nested operators are axiomatized by the axioms $O(p \mid q \wedge O(p \mid q))$ and $P(\top|P(p|q)) \rightarrow P(p|q \wedge P(p|q))$.[3]

A deontic state $\langle W, R, \models_{\mathrm{HRC}}\rangle$ contains a set of worlds W (HRC models), a ranking function R on ordered pairs of worlds of W, and a satisfaction function $\models_{\mathrm{HRC}}$ for HRC formulas at the worlds (at $w*$ of the HRC model). The ranking R is a mapping of $W \times W$ to the set of positive integers $I\!N$ plus infinity, $I\!N \cup \{\infty\}$, with infinity larger than any element of $I\!N$, i.e. $\forall x \in I\!N(x \neq \infty \rightarrow x < \infty)$. The rank of a pair of worlds (w_1, w_2) represents the strength of the prima facie obligation that prefers world w_2 to w_1. If there is no such obligation then its rank is 0, and if there are several of such obligations, then its rank is the strength of the strongest of the obligations. We call an ordered pair of worlds a link. In particular, we call an ordered pair (w_1, w_2) a (p_1, p_2) link if $w_1 \models p_1$ and $w_2 \models p_2$.

[3] The logic has the usual properties of rational conditionals such as the conjunction rule for the consequent and the disjunction rule for the antecedent, but it does not trivialize nested operators by for example $O(O(p|q)|r) \leftrightarrow O(p|q \wedge r)$.

The deontic state is extended with an epistemic state, which is a sub-state of the former. This sub-state contains all HRC models the agent holds possible. The deontic state is used for the context of justification and the epistemic state is used for the context of deliberation, see [van der Torre and Tan, 1998b]. Whereas in Kripke semantics a unique world is singled out, called the actual world, in DUS a set of worlds is singled out, called the context of deliberation.

For the initial state we assume that everything is permitted.

Definition 3 (Deontic state) *Let A_u be a set of authorities, A_g a set of agents and L_1^A a deontic language. Assume a set of worlds W (HRC models) and a satisfaction function $\models_{HRC}$ for L_0^A such that for every interpretation of L_0^A there is at least one corresponding $w \in W$. A deontic state is a tuple $\Sigma = \langle W, W^*, R, \models_{HRC}\rangle$ consisting of the set of worlds W, a possibly empty subset $W^* \subseteq W$, the satisfaction function $\models_{HRC}$ and for each combination of $a \in A_u$ and $\alpha_1, \alpha_2 \in Ag$ an integer (or ∞) valued ranking function $R^a_{\alpha_1,\alpha_2}$ on $W \times W$.*

0, *the* minimal state, *is* $\langle W, W, W \times W \to 0, V\rangle$, *and*
1, *the* absurd state, *is* $\langle W, \emptyset, W \times W \to \infty, V\rangle$.

2.4 Deontic updates

The deontic updates are operations on the deontic states that either zoom in on the deontic state (for HRC sentences), or increase the ranks of links (for prescriptive operators). The prescriptive operators have the dynamic component of creating a new deontic state. The general principle is that in case of conflict later operators are stronger than earlier ones. The operator $\mathsf{permit}_\alpha p$ is not defined in terms of absence of $\mathsf{oblige}_\alpha \neg p$ and it is therefore a so-called strong permission operator.

- For the update with the obligation $\mathsf{oblige}^a_{\alpha_1,\alpha_2}(p|q)$ there is a conflict if all the $(p \wedge q, \neg p \wedge q)$ links are non-zero. If there is no conflict then the rank of the $(\neg p \wedge q, p \wedge q)$ links is at least 1. Otherwise, their rank is higher than the minimum of the reverse $(p \wedge q, \neg p \wedge q)$ links.
- Analogously, for the permission $\mathsf{permit}^a_{\alpha_1,\alpha_2}(p \mid q)$ there is a conflict if all the $(p \wedge q, \neg p \wedge q)$ links are non-zero. If there is no conflict then the rank of the $(\neg p \wedge q, p \wedge q)$ links does not change. Otherwise, their rank is the minimum of the $(p \wedge q, \neg p \wedge q)$ links.

To evaluate all-things-considered ideal and $\mathsf{someideal}$ operators we reduce the deontic state to a set of HRC models. It is here that the prescriptive and descriptive logics interact. In this paper we only consider a simple reduction, in which the constructed models only contain the actual worlds of the HRC models.

Definition 4 (σ reduction) *The reduction of σ to HRC models, written as HRC(σ), is defined as follows.*

- *Each element of HRC(σ) contains all actual worlds of the HRC models of σ.*
- *The actual world of an element of HRC(σ) is one of the actual worlds of the HRC models of σ.*
- *The accessibility relation R of an element of HRC(σ) is the universal relation.*
- *The starting point to construct the preference ordering is that w_1 is preferred to w_2, i.e. $w_1 < w_2$, if there is a prima facie preference for w_1, i.e. $R(w_2, w_1) > 0$. Then the following two procedures are carried out.*

  Cycle elimination. *For all cycles $w_1 < w_2 < \ldots < w_n < w_1$ simultaneously remove the weakest links of the cycles. If all links of a cycle are equally strong then they are all removed. For example, if there are only two-step cycles then $w_1 < w_2$ if the prima facie preference for w_1 is stronger than the preference for w_2, i.e. $R(w_2, w_1) > R(w_1, w_2)$.*

  Extension. *Afterwards, take the transitive closure and construct a totally connected order.*

A cycle represents a deontic conflict: there are arguments for and against an issue. Eliminating cycles is therefor the formal counterpart of conflict resolution (and taking the transitive closure of the deontic state is the formal counterpart of dealing with incomplete information). Conflicts are resolved by weighing the arguments, because only the weakest links are removed. This is the obvious and most simple construction for conflict resolution. There are of course more sophisticated mechanisms, that for example eliminate cycles in some order or that do cycle elimination and taking the transitive closure simultaneously. Due to lack of space we cannot discuss alternatives here, but obviously they can immediately be used in our logical framework.

Finally, von Wright's contingency principle, i.e. the obligation 'p ought to be (done) if q is (done)' implies the consistency of $p \wedge q$ and $\neg p \wedge q$, is formalized by a test on the existence of $p \wedge q$ and $\neg p \wedge q$ worlds. The operators oblige^*, permit^*, ideal^* and $\mathsf{someideal}^*$ refer to the epistemic state or the context of deliberation W^* instead of the context of justification W.

Definition 5 (Deontic updates) *Let $\sigma = \langle W, W^*, R, \models_{HRC}\rangle$ be a deontic state, and let $\min^a_{\alpha_1,\alpha_2}(p|q)$ be the minimum of*

$$\{R^a_{\alpha_1,\alpha_2}(w_1, w_2) \mid \sigma, w_1 \models p \wedge q \text{ and } \sigma, w_2 \models \neg p \wedge q\}$$

if this set is non-empty, undefined otherwise. The update function $\sigma[\phi]$ is defined as follows.

- *if ϕ is a HRC sentence of L_0^A, then*
  - *if $W' = \{w \in W^* \mid \sigma, w \models \phi\} \neq \emptyset$, then $\sigma[\phi] = \langle W, W', R, V\rangle$;*
  - *otherwise, $\sigma[\phi] = \mathbf{1}$.*
- *if $\phi = \mathsf{oblige}^a_{\alpha_1,\alpha_2}(p|q)$, then*
  - *if $W' = \{w \in W^* \mid \sigma, w \models O^a_{\alpha_1,\alpha_2}(p|q)\} \neq \emptyset$ and there are $w_1, w_2 \in W$ such that $\sigma, w_1 \models \neg p \wedge q$ and $\sigma, w_2 \models p \wedge q$, then $\sigma[\phi] = \langle W, W', R', V\rangle$ with for all $w_1, w_2 \in W$*

* if $\sigma, w_1 \models \neg p \wedge q$ and $\sigma, w_2 \models p \wedge q$ then $R'^{a}_{\alpha_1,\alpha_2}(w_1, w_2) = \max(R^a_{\alpha_1,\alpha_2}(w_1, w_2), \min^a_{\alpha_1,\alpha_2}(p|q) + 1)$;
    * otherwise $R'^{a}_{\alpha_1,\alpha_2}(w_1, w_2) = R^a_{\alpha_1,\alpha_2}(w_1, w_2)$;
  - otherwise, $\sigma[\phi] = \mathbf{1}$.

- if $\phi = \mathsf{permit}^a_{\alpha_1,\alpha_2}(p|q)$, then
  - if $W' = \{w \in W^* \mid \sigma, w \models P^a_{\alpha_1,\alpha_2}(p|q)\} \neq \emptyset$ and there are $w_1, w_2 \in W$ such that $\sigma, w_1 \models \neg p \wedge q$ and $\sigma, w_2 \models p \wedge q$, then $\sigma[\phi] = \langle W, W', R', V\rangle$ with for all $w_1, w_2 \in W$
    * if $\sigma, w_1 \models \neg p \wedge q$ and $\sigma, w_2 \models p \wedge q$ then $R'^{a}_{\alpha_1,\alpha_2}(w_1, w_2) = \max(R^a_{\alpha_1,\alpha_2}(w_1, w_2), \min^a_{\alpha_1,\alpha_2}(p|q))$;
    * otherwise $R'^{a}_{\alpha_1,\alpha_2}(w_1, w_2) = R^a_{\alpha_1,\alpha_2}(w_1, w_2)$;
  - otherwise, $\sigma[\phi] = \mathbf{1}$.
- $\mathsf{oblige}^{*a}_{\alpha_1,\alpha_2}(p \mid q)$ and $\mathsf{permit}^{*a}_{\alpha_1,\alpha_2}(p \mid q)$ are defined analogously to $\mathsf{oblige}^a_{\alpha_1,\alpha_2}(p|q)$ and $\mathsf{permit}^a_{\alpha_1,\alpha_2}(p|q)$, replacing W by W^*.
- if $\phi = \mathsf{ideal}^a_{\alpha_1,\alpha_2}(p|q)$, then
  - if all HRC(σ) models satisfy $O^a_{\alpha_1,\alpha_2}(p|q)$, then $\sigma[\phi] = \sigma$;
  - otherwise, $\sigma[\phi] = \mathbf{1}$.
- if $\phi = \mathsf{someideal}^a_{\alpha_1,\alpha_2}(p|q)$, then
  - if all HRC(σ) models satisfy $P^a_{\alpha_1,\alpha_2}(p|q)$, then $\sigma[\phi] = \sigma$;
  - otherwise, $\sigma[\phi] = \mathbf{1}$.
- $\mathsf{ideal}^{*a}_{\alpha_1,\alpha_2}(p \mid q)$ and $\mathsf{someideal}^{*a}_{\alpha_1,\alpha_2}(p \mid q)$ are defined analogously to $\mathsf{ideal}^a_{\alpha_1,\alpha_2}(p \mid q)$ and $\mathsf{someideal}^a_{\alpha_1,\alpha_2}(p|q)$, replacing W by W^*.

2.5 Acceptance

A crucial notion of update systems is acceptance. The formula ϕ is accepted in a deontic state σ, written as $\sigma \Vdash \phi$, if the update by ϕ results in the same state. In that case, the information conveyed by ϕ is already subsumed by σ. Acceptance is the counterpart of satisfaction in standard semantics.

Definition 6 (Acceptance) *Let σ be an deontic state and ϕ a formula of the logical language L. $\sigma \Vdash \phi$ if and only if $\sigma[\phi] = \sigma$.*

If an update is accepted, then the deontic state usually has a specific content. For example, the following proposition is easily checked.

Proposition 1 *A fact α is accepted if all the worlds of $W^* \neq \emptyset$ satisfy α, or $\sigma = \mathbf{1}$. Moreover, an obligation $\mathsf{oblige}(p \mid q)$ is accepted if the rank of all $(\neg p \wedge q, p \wedge q)$ links is higher than the smallest rank of the reverse links, and a permission $\mathsf{permit}(p \mid q)$ is accepted if the rank of all $(\neg p \wedge q, p \wedge q)$ links is at least as high as the smallest rank of the reverse links.*

These two definitions can be reinterpreted in standard possible worlds semantics $\langle W, \leq, V\rangle$ if we consider operators that cannot be overridden, because in that case the rank can only take two values. The obligation is a weakened version of the descriptive operator of Prohairetic Deontic Logic (PDL) [van der Torre and Tan, 1998a], and the permission operator is new.[4]

Definition 7 *A possible worlds model $\langle W, \leq, V\rangle$ satisfies $O^s(p \mid q)$ if there are no $p \wedge q$ world w_1 and $\neg p \wedge q$ world w_2 such that $w_2 \leq w_1$, and it satisfies $P^s(p \mid q)$ if there are no $p \wedge q$ world w_1 and $\neg p \wedge q$ world w_2 such that $w_2 < w_1$.*

Obligation implies permission, $O^s(p \mid q) \rightarrow P^s(p \mid q)$. However, permission is not implied by the absence of obligation, because $\neg(O^s(p \mid q) \vee P^s(p \mid q))$ is consistent. Thus P^s is a strong permission operator.

2.6 Validity relations

Different notions of validity can be based on the notion of acceptance (see [Veltman, 1996] for an overview). We use the following two. An argument is $\Vdash_1$ valid if updating the minimal state $\mathbf{0}$ with the premises $\psi_1, \ldots, \psi_n$, in that order, yields a deontic state in which the conclusion is accepted. An argument is $|\!\sim$ valid if all deontic states constructed by updating the minimal state $\mathbf{0}$ with the premises $\psi_1, \ldots, \psi_n$ *in some order such that the premises are accepted*, also accept the conclusion.

Definition 8 (Validity) *Let $\psi_1, \ldots, \psi_n$ and ϕ be formulas of the deontic language L_1^A. The argument of ϕ from the premises $\psi_1, \ldots, \psi_n$ is valid, written as $\psi_1, \ldots, \psi_n \Vdash_1 \phi$, if and only if $\mathbf{0}[\psi_1] \ldots [\psi_n] \Vdash \phi$. $\psi_1, \ldots, \psi_n \,|\!\sim \phi$ if and only if for all permutations π of $1 \ldots n$ such that $\psi_{\pi(1)}, \ldots, \psi_{\pi(n)} \Vdash_1 \psi_i$ for $1 \leq i \leq n$ we have $\psi_{\pi(1)}, \ldots, \psi_{\pi(n)} \Vdash_1 \phi$.*

We end with a few properties of the operators and the relations between them. The prescriptive and descriptive operators have some properties in common. For example, they are both closed under the conjunction rule. However, they are in another sense complementary. For example, the operator oblige has strengthening of the antecedent but not weakening of the consequent, and O vice versa. This expresses that the prescriptive obligation is applicable in all states that imply its antecedent (unless it is overridden) whereas the descriptive obligation only evaluates the state of its antecedent.

$$\mathsf{oblige}^a_{\alpha_1,\alpha_2}(p|q) \,|\!\sim \mathsf{oblige}^a_{\alpha_1,\alpha_2}(p|q \wedge r)$$
$$\mathsf{oblige}^a_{\alpha_1,\alpha_2}(p|q) \,|\!\not\sim \mathsf{oblige}^a_{\alpha_1,\alpha_2}(p \vee r|q)$$
$$O^a_{\alpha_1,\alpha_2}(p|q) \,|\!\not\sim O^a_{\alpha_1,\alpha_2}(p|q \wedge r)$$
$$O^a_{\alpha_1,\alpha_2}(p|q) \,|\!\sim O^a_{\alpha_1,\alpha_2}(p \vee r|q)$$

[4] The $\leq$ relation can be reflexive and transitive. When it is also totally connected, as for example DSDL3 models, then counterintuitive conclusions follow, see [van der Torre and Tan, 1998a].

Prescriptive operators are in a sense stronger than descriptive operators, because we have the following due to the check in the definition of the prescriptive updates.

$$\mathsf{oblige}^{a}_{\alpha_1,\alpha_2}(p|q) \mid\!\mid\!\sim O^{a}_{\alpha_1,\alpha_2}(p|q)$$
$$\mathsf{permit}^{a}_{\alpha_1,\alpha_2}(p|q) \mid\!\mid\!\sim P^{a}_{\alpha_1,\alpha_2}(p|q)$$

Prescriptive permissions and obligations can interact, for example in the following conflict between them.

$$\mathsf{oblige}^{a}_{\alpha_1,\alpha_2} p, \mathsf{permit}^{a}_{\alpha_1,\alpha_2} \neg p \mid\!\mid\!\sim \perp$$

It is shown in [van der Torre and Tan, 1998c] that more specific and conflicting obligations are only accepted if they are later than more general ones. Hence, more specific and conflicting obligations are stronger than more general ones and override them. Moreover, the new prescriptive permissions introduced in this paper can override obligations analogously.

$$\mathsf{oblige}^{a}_{\alpha_1,\alpha_2}(p|\top) \mid\!\mid\!\sim \mathsf{oblige}^{a}_{\alpha_1,\alpha_2}(p|q)$$
$$\mathsf{oblige}^{a}_{\alpha_1,\alpha_2}(p|\top), \mathsf{oblige}^{a}_{\alpha_1,\alpha_2}(\neg p|q) \mid\!\mid\!\not\sim \mathsf{oblige}^{a}_{\alpha_1,\alpha_2}(p|q)$$
$$\mathsf{oblige}^{a}_{\alpha_1,\alpha_2}(p|\top), \mathsf{permit}^{a}_{\alpha_1,\alpha_2}(\neg p|q) \mid\!\mid\!\not\sim \mathsf{oblige}^{a}_{\alpha_1,\alpha_2}(p|q)$$

3 Conclusions

Rights, duties and commitments are important for multi agent systems. For example, in agent oriented programming [Shoham, 1993] commitments play an important role, though their semantics is not given. This paper contributes to their formal foundations by extending deontic logic with the following elements:

Combination. Rights are formalized as prescriptive permissions between agents and duties and commitments are formalized as prescriptive obligations between agents in an update semantics. The prescriptive operators could easily be combined with standard descriptive operators and mixed all-things-considered operators, because they are all defined in a preference-based framework.

Strong permissions. The framework contains strong permissions needed in multi agent environments. Makinson [1998] shows that one of the merits of prescriptive deontic logics is that they also enable natural distinctions between weak and strong permissions and various ways of relating them to obligations, but thus far no strong permission for traditional possible worlds semantics had been given.

Controllability and contexts. The framework distinguishes controllable and uncontrollable propositions to formalize actions and circumstances, and deontic and epistemic states to formalize the context of justification and the context of deliberation [Thomason, 1981].

Acknowledgments

Thanks to Emil Weydert for commenting on a previous version of this paper.

References

[Alchourrón and Bulygin, 1981] C.E. Alchourrón and E. Bulygin. The expressive conception of norms. In R. Hilpinen, editor, *New Studies in Deontic Logic: Norms, Actions and the Foundations of Ethics*, pages 95–124. D. Reidel, 1981.

[Boutilier, 1994] C. Boutilier. Toward a logic for qualitative decision theory. In *Proceedings of the KR'94*, pages 75–86, 1994.

[Hansson, 1971] B. Hansson. An analysis of some deontic logics. In R. Hilpinen, editor, *Deontic Logic: Introductory and Systematic Readings*, pages 121–147. D. Reidel, Dordrecht, Holland, 1971.

[Lang, 1996] J. Lang. Conditional desires and utilities - an alternative approach to qualitative decision theory. In *Proceedings of the ECAI'96*, pages 318–322, 1996.

[Liau, 1998] C.-J. Liau. A logic for reasoning about action, preference and commitment. In H. Prade, editor, *Proceedings of the ECAI'98*, pages 552–556, 1998.

[Makinson, 1998] D. Makinson. On a fundamental problem of deontic logic. In P. McNamara and H. Prakken, editors, *Norms, Logics and Information Systems. New Studies on Deontic Logic and Computer Science*, pages 29–53. IOS Press, 1998.

[Schmid and Selz, 1998] Electronic contracting (focus theme). In B. Schmid and D. Selz, editors, *International Journal of Electronic Markets*, volume 8:3, 1998. http://www.electronicmarkets.org.

[Shoham, 1993] Y. Shoham. Agent-oriented programming. *Artificial Intelligence*, 60:51–92, 1993.

[Thomason, 1981] R. Thomason. Deontic logic as founded on tense logic. In R. Hilpinen, editor, *New Studies in Deontic Logic: Norms, Actions and the Foundations of Ethics*, pages 165–176. D. Reidel, 1981.

[van der Torre and Tan, 1998a] L. van der Torre and Y. Tan. Prohairetic Deontic Logic (PDL). In *Logics in Artificial Intelligence*, LNAI 1489, pages 77–91. Springer, 1998.

[van der Torre and Tan, 1998b] L. van der Torre and Y. Tan. An update semantics for deontic reasoning. In P. McNamara and H. Prakken, editors, *Norms, Logics and Information Systems. New Studies on Deontic Logic and Computer Science*, pages 73–90. IOS Press, 1998.

[van der Torre and Tan, 1998c] L. van der Torre and Y. Tan. An update semantics for prima facie obligations. In *Proceedings of the ECAI'98*, pages 38–42, 1998.

[Veltman, 1996] F. Veltman. Defaults in update semantics. *Journal of Philosophical Logic*, 25:221–261, 1996.

[Weydert, 1994] E. Weydert. Hyperrational conditionals: monotonic reasoning about nested default conditionals. In *Foundations of Knowledge Representation and Reasoning*, LNAI 810, pages 310–332. Springer, 1994.