Logical Framework for Normative Systems

Yasuo NAKAYAMA
Osaka University, Graduate School of Human Sciences
1-2 Yamada-oka, Suita, Osaka, Japan
nakayama@hus.osaka-u.ac.jp

In this paper, I propose a new logical framework that can be used to analyze normative phenomena in general. I call this framework a *Logic for Normative Systems* (LNS). I also demonstrate how to solve some paradoxes of Standard Deontic Logic (SDL). A characteristic of LNS is its dynamic behavior. LNS is flexible, hence it can be applied to describe complex normative problems including ethical problems.

1 Definition of Normative Systems

A normative system can be defined as follows, where $\vdash$ means the inference in the first-order logic:

(1a) Let T and OB be sets of sentences. A pair $\langle T, OB\rangle$ consisting of *propositional system* T and *obligation space* OB is called a *normative system* (NS).

(1b) A sentence p belongs to the *propositional context* of normative system $\langle T, OB\rangle$ if and only if (*iff*) $T \vdash p$.

(1c) A sentence p belongs to the *obligation context* of normative system $\langle T, OB\rangle$ (abbreviated as $\mathbf{O}_{\langle T,OB\rangle}p$) *iff* $T \cup OB \vdash p$ & $T \nvdash p$ & $T \cup OB \nvdash \bot$.

(1d) A sentence p belongs to the *prohibition context* of normative system $\langle T, OB\rangle$ (abbreviated as $\mathbf{F}_{\langle T,OB\rangle}p$) *iff* $\mathbf{O}_{\langle T,OB\rangle}\neg p$.

(1e) A sentence p *belongs to the permission context of normative system* $\langle T, OB\rangle$ (abbreviated as $\mathbf{P}_{\langle T,OB\rangle}p$) *iff* $T \cup OB \cup \{p\} \nvdash \bot$ & $T \nvdash p$.

(1f) A group G has (*normative*) *power* of doing act_1 in $\langle T, OB\rangle$ *iff*
$T \vdash \forall x(member(x,G) \to agent(x))$ &
$\mathbf{P}_{\langle T,OB\rangle}\forall x(member(x,G) \to do(x, act_1))$ &
$\mathbf{F}_{\langle T,OB\rangle}\forall x(\neg member(x,G) \to do(x, act_1))$. [1]

These definitions presuppose that we insert *what we believe to be true* into the propositional system and *what we believe ought to be done* into the obligation space.

[1] The notion of (legal) *power* plays an essential role in Hart (1961). This shows that this notion is inevitable for describing legal systems. According to definition (1f), a group G has power of doing act_1 *iff* all members of G and only members of G are allowed to perform act_1.

From these definitions immediately follow the following three fundamental characterizations for LNS.

(2a) If $\mathbf{O}_{\langle T,OB\rangle}p$, then $\mathbf{P}_{\langle T,OB\rangle}p$. [2]

(2b) If $\mathbf{F}_{\langle T,OB\rangle}p$, then not $\mathbf{P}_{\langle T,OB\rangle}p$.

(2c) The propositional context of a NS is independent of its obligation space, while the obligation context depends on the propositional system.

To see how LNS works, let us consider an example.

(3) You should not kill any human beings. Peter is a human being. So you should not kill Peter.

According to our presupposition mentioned before, the first sentence in (3) expresses the content of one component of OB_1 and the second sentence expresses the content of one component in T_1 of normative system $\langle T_1, OB_1\rangle$:

(4a) $\forall x \forall y(agent(x) \wedge human(y) \rightarrow \neg kill(x,y))$ is an element of OB_1.

(4b) $human(Peter)$ is an element of T_1.

From (1a), (1c), and (4a) immediately follow (4c) and (4d), where (4c) corresponds to the conclusion of (3).

(4c) $\mathbf{O}_{\langle T_1,OB_1\rangle}\forall x(agent(x) \rightarrow \neg kill(x, Peter))$.

(4d) For any agent A, $\mathbf{O}_{\langle T_1,OB_1\rangle}\neg kill(A, Peter)$.

This result can be summarized as follows:

(4e) If $\{human(Peter), agent(A)\} \subseteq T_1$ & $\{\forall x \forall y(agent(x) \wedge human(y) \rightarrow \neg kill(x,y))\} \subseteq OB_1$, then
$\mathbf{O}_{\langle T_1,OB_1\rangle}\forall x(agent(x) \rightarrow \neg kill(x, Peter))$ & $\mathbf{O}_{\langle T_1,OB_1\rangle}\neg kill(A, Peter)$ & $\mathbf{F}_{\langle T_1,OB_1\rangle}kill(A, Peter)$.

In this way, the informal reasoning in (3) can be formally justified within LNS.

2 Paradoxes of Deontic Logic and Their Solutions

In this section, I propose how to solve some paradoxes of SDL. First, let us consider Ross's paradox (Ross (1941), McNamara (2006) sec. 4.3, Åqvist (2002) sec. 6):

[2] This characterization corresponds to an axiom of SDL, namely $\mathbf{O}p \rightarrow \neg\mathbf{O}\neg p$. However, within LNS the iteration of normative sentences is not possible. This is no failure of LNS, because many normative systems in our ordinary life are expressible without iterative normative expressions. It is interesting that LNS fulfills the *principle of deontic contingency* required by Von Wright (1951). Because of (1c), we can easily prove that no tautology belongs to an obligation context.

(5a) It is obligatory that the letter is mailed.

(5b) It is obligatory that the letter is mailed or the letter is burned.

Within SDL, (5b) seems to follow from (5a), because $\mathbf{O}m \rightarrow \mathbf{O}(m \vee b)$ is a theorem of SDL, where $\mathbf{O}p$ means "It is obligatory that p". However, it seems rather odd to say that an obligation to mail the letter entails an obligation that can be fulfilled by burning the letter. Within LNS, a similar inference is valid:

(6*) If $\mathbf{O}_{\langle T,OB\rangle}p$ & $T \nvdash p \vee q$, then $\mathbf{O}_{\langle T,OB\rangle}(p \vee q)$.

The source of the paradoxical appearance of this example consists in ignoring the incompatibility among two types of actions. In this case, it will be appropriate to assume that mailing a letter is incompatible with burning it. This fact justifies to accept that $\forall x(\mathit{letter}(x) \rightarrow \neg(\mathit{mailed}(x) \wedge \mathit{burned}(x)))$ is a component of the propositional system of the given normative system $\langle T_2, OB_2\rangle$.

(6a) $\mathit{mailed}(l_1)$ is an element of OB_2. (From (5a))

(6b) $T_2 \nvdash \mathit{mailed}(l_1) \vee \mathit{burned}(l_1)$. (Observation)

(6c) $\mathit{letter}(l_1)$ and $\forall x(\mathit{letter}(x) \rightarrow \neg(\mathit{mailed}(x) \wedge \mathit{burned}(x)))$ are elements of T_2. (Observation)

From these conditions follows that $\neg \mathit{burned}(l_1)$ belongs to the obligation context of $\langle T_2, OB_2\rangle$ (see (6d)). Thus, it is forbidden to burn the letter. [3] [4]

(6d) If $\{\mathit{letter}(l_1), \forall x(\mathit{letter}(x) \rightarrow \neg(\mathit{mailed}(x) \wedge \mathit{burned}(x)))\} \subseteq T_2$ & $\{\mathit{mailed}(l_1)\} \subseteq OB_2$ & $T_2 \nvdash (\mathit{mailed}(l_1) \vee \mathit{burned}(l_1))$, then $T_2 \vdash \neg(\mathit{mailed}(l_1) \wedge \mathit{burned}(l_1))$ & $\mathbf{O}_{\langle T_2,OB_2\rangle}(\mathit{mailed}(l_1) \vee \mathit{burned}(l_1))$ & $\mathbf{O}_{\langle T_2,OB_2\rangle}\neg \mathit{burned}(l_1)$ & $\mathbf{F}_{\langle T_2,OB_2\rangle}\mathit{burned}(l_1)$.

The *Good Samaritan Paradox* pointed out by Prior (1958) can be solved in a similar way. Let us consider the following sentences:

(7a) It ought to be the case that John helps Smith who has been robbed.

(7b) John helps Smith who has been robbed *iff* John helps Smith and Smith has been robbed.

$\mathbf{O}(h \wedge r) \rightarrow \mathbf{O}(h)$ is a theorem of SDL. Thus, if we represent sentence (7a) as $\mathbf{O}(h \wedge r)$, (7c) follows from (7a) and (7b). However, (7c) seems hardly right.

(7c) It ought to be the case that Smith has been robbed.

Within LNS, (7a) can be analyzed as the combination of two conditions (8a) and (8c), where (8c) follows from (8a) and (8b).

[3] Belnap et al (2001) also discusses how to solve Ross's paradox within Stit logic (p. 84f). Stit logic can take future developments into consideration and they solve Ross's paradox using this property of Stit logic. They demonstrate that $[\alpha\ \mathit{stit} : p \vee q]$ does not follow from $[\alpha\ \mathit{stit} : p]$, where $[\alpha\ \mathit{stit} : p]$ is an abbreviation of $[\alpha$ sees to it that $p]$.

[4] One of the reviewers pointed out my misinterpretation of Ross's paradox in the first draft of this paper. According to him, Ross's paradox is $\mathbf{O}m \rightarrow \mathbf{O}(m \vee b)$ itself. In this case, I will state that (6*) is not so bad, because realizing p remains still as an obligation after realizing q.

(8a) $agent(John)$ and $robbed(Smith)$ are elements of T_3.

(8b) $\forall x \forall y(agent(x) \wedge robbed(y) \to help(x, y))$ is an element of OB_3.

(8c) $\mathbf{O}_{\langle T_3, OB_3 \rangle}(robbed(Smith) \to help(John, Smith))$.

(8d) If $\{agent(John), robbed(Smith)\} \subseteq T_3$ & $\{\forall x \forall y(agent(x) \wedge robbed(y) \to help(x, y))\} \subseteq OB_3$, then $\mathbf{O}_{\langle T_3, OB_3 \rangle}(robbed(Smith) \to help(John, Smith))$ & $\mathbf{O}_{\langle T_3, OB_3 \rangle} help(John, Smith)$.

From (8a) and (8c), follows that "John helps Smith" belongs to the obligation context of $\langle T_3, OB_3 \rangle$ ((8d)). This result means that *It ought to be the case that John helps Smith*, which is the result we sought.

Note that a normative system can express explicitly who the bearers of an obligation are, where we assume that they always try to fulfill their obligations, if they accept them:

(9a) Given a normative system $\langle T, OB \rangle$, "Teachers should prepare for their lectures" can be expressed as follows: $\mathbf{O}_{\langle T, OB \rangle}\ \forall x \forall y(agent(x) \wedge teacher(x) \wedge lecture\text{-}of(y, x) \to prepare(x, y))$.

(9b) Given a normative system $\langle T, OB \rangle$, "Students should study hard" can be expressed as follows: $\mathbf{O}_{\langle T, OB \rangle}\ \forall x(agent(x) \wedge student(x) \to study\text{-}hard(x))$.

3 Conflicts in Normative Systems

In normative system $\langle T, OB \rangle$, two kinds of contradictions are distinguished: the contradiction in propositional system T and that in $\langle T, OB \rangle$. T is *inconsistent as the propositional system of* $\langle T, OB \rangle$ *iff* T is inconsistent. However, $\langle T, OB \rangle$ is *inconsistent iff* $T \cup OB$ is inconsistent.

A characteristic of NSs is the property that any violation of the presupposed obligations produces inconsistency in NSs. Let us consider the following example:

(10a) It ought to be that Jones does go (to the assistance of his neighbors).

(10b) Jones doesn't go.

This situation can be represented as follows:

(10c) $\{agent(Jones), \neg go(Jones)\} \subseteq T_4$ & $\{go(Jones)\} \subseteq OB_4$.

This kind of violation of obligations could threaten the significance of a NS. If people perform their actions without respecting a given NS, that system can lose its significance. However, a small violation of a NS can sometimes be managed through taking the dynamic aspect of the reality into consideration (see section 4).

Next, let us consider a case of conflicts in an obligation space. Suppose that Tom is required to do act_1 as a member of group A. Furthermore, suppose that he is also required not to do act_1 as a member of group B. This situation produces inconsistency in the presupposed normative system $\langle T_5, OB_5 \rangle$:

(11) If $\{agent(Tom), member(Tom, A), member(Tom, B)\} \subseteq T_5$ &
$\{\forall x(agent(x) \wedge member(x, A) \rightarrow do(x, act_1)),$
$\forall x(agent(x) \wedge member(x, B) \rightarrow \neg do(x, act_1))\} \subseteq OB_5$, then
$\mathbf{O}_{\langle T_5, OB_5 \rangle} do(Tom, act_1)$ & $\mathbf{O}_{\langle T_5, OB_5 \rangle} \neg do(Tom, act_1)$.

One possible solution for Tom is to drop out from one of these groups. In that case, this decision reproduces a consistent NS. For example, if Tom drops out from group B, he need no more consider the prohibition of doing act_1.

4 Dynamic Aspects and Future Orientation of LNS

A normative system $\langle T, OB \rangle$ can function like a *conversational score* proposed by David Lewis (1979). A normative system can be updated to describe a development of a situation. Let us reconsider the case of a violating action described in (10a) and (10b). To describe the shift of time, we introduce here the *past-tense operator* P. Then, we can consider the shift of time and update normative system $\langle T_4, OB_4 \rangle$ and create $\langle T_{4up}, OB_4 \rangle$:

(12a) $\{agent(Jones), \neg go(Jones)\} \subseteq T_4$ & $\{go(Jones)\} \subseteq OB_4$.

(12b) $T_{4up} = (T_4 - \{\neg go(Jones)\}) \cup \{P(\neg go(Jones))\}$. (Information updated)

This example shows that a violated NS can sometimes automatically recover its consistency when its propositional system is updated.

Next, let us consider a case where an obligation becomes applicable through a change of the given situation.

(13a) We should help suffering neighbors.

(13b) Mary, who is a neighbor of John, was not suffering, but is now suffering.

(13c) So John should help Mary now.

Within LNS, the informal inference "(13c) follows from (13a) and (13b)" can be explicitly described as the inference of (13f) from (13d) and (13e):

(13d) $\{agent(John, neighbor(Mary, John)\} \subseteq T_6$ &
$\{\forall x \forall y(agent(x) \wedge neighbor(y, x) \wedge suffering(y) \rightarrow help(x, y))\} \subseteq OB_6$.

(13e) $T_{6up} = T_6 \cup \{suffering(Mary)\}$. (Information updated)

(13f) $\mathbf{O}_{\langle T_{6up}, OB_6 \rangle} help(John, Mary)$.

Normative sentences are normally future oriented. Some problems can be solved by considering this property. Consider the following sentences: [5]

(14a) It should be the case that from now on any two countries have peaceful relations ($\{\forall t \forall x \forall y(t_{now} \leq t \wedge country(x) \wedge country(y) \wedge x \neq y \rightarrow peaceful(x, y, t))\} \subseteq OB_7$).

[5] This problem is pointed out by one of the reviewers.

(14b) A and B are countries ($\{country(A) \wedge country(B)\} \subseteq T_7$).

(14c1) A and B have now peaceful relations ($\{peaceful(A, B, t_{now})\} \subseteq T_7$).

(14c2) A and B have always peaceful relations ($\{\forall t\ peaceful(A, B, t)\} \subseteq T_7$)

(14d) It should be the case that from now on A and B have peaceful relations ($\mathbf{O}_{\langle T_7, OB_7 \rangle} \forall t(t_{now} \leq t \rightarrow peaceful(A, B, t))$).

Within LNS, (14d) follows from (14a), (14b), and (14c1), while (14d) does not follow from (14a), (14b), and (14c2). However, the second invalidity can be justified, because (14c2) expresses that the goal of (14d) has been already satisfied. [6]

5 Conclusions

It is well known that SDL has many theoretical difficulties (Åqvist (2002), McNamara (2006)). Recently, Stit logic was proposed and researchers have shown many interesting results (Belnap et al (2001), Horty (2001)). LNS is an alternative framework that can explicitly express both propositional and normative constraints. This property of explicitness makes LNS applicable to numerous classes of normative problems. For example, a legal system could be described as a normative system in the sense of LNS. The method developed in this paper can be modified to explain inferences among speech acts. However, this remains a future task. [7]

References

[1] Åqvist, L. (2002) "Deontic Logic", D. Gabbay and F. Guenthner (eds.) *Handbook of Philosophical Logic*, Vol. 8, Kluwer Academic Pub., pp. 147-264.

[2] Belnap, N., Perloff, M. and Xu, M. (2001) *Facing the Future: Agents and Choices in Our Indeterminist World*, Oxford University Press.

[3] Hart, H. L. A. (1961) *The Concept of Law*, Clarendon Press.

[4] Horty, J. F. (2001) *Agency and Deontic Logic*, Oxford University Press.

[5] Lewis, D. (1979) "Scorekeeping in a Language Game", *Journal of Philosophical Logic* 8, pp. 339-359.

[6] McNamara, P. (2006) "Deontic Logic", *Stanford Encyclopedia of Philosophy.*

[7] Prior, A. N. (1958) "Escapism: The Logical Basis of Ethics", In A.I. Melden (1958), *Essays in Moral Philosophy*, University of Washington Press, pp. 135-146.

[8] Ross, A. (1941) "Imperatives and Logic", *Theoria* 7, pp. 53-71.

[9] Von Wright, G. H. (1951) "Deontic Logic", *Mind* 60, pp. 1-15.

[6] As this section shows, LNS can deal with interactions between deontic states and temporal developments. Thus, LNS can be seen as a framework that satisfies the following requirements mentioned in Åqvist (2002): "for any serious perposes of application, the expressive resources of deontic languages must be enriched so as to include temporal and quantificational ones" (p. 150).

[7] I would like to thank two reviewers for many insightful comments.