



Available at

[www.ElsevierComputerScience.com](http://www.ElsevierComputerScience.com)

POWERED BY SCIENCE @ DIRECT®

Journal of Applied Logic ●●● (●●●●) ●●●●-●●●

JOURNAL OF  
APPLIED LOGIC[www.elsevier.com/locate/jal](http://www.elsevier.com/locate/jal)

## Agential obligation as non-agential personal obligation plus agency

Paul McNamara

*Department of Philosophy, University of New Hampshire, Durham, NH 03824-3574, USA*

---

### Abstract

I explore various ways of integrating the framework for predeterminism, agency, and ability in [P. McNamara, *Nordic J. Philos. Logic* 5 (2) (2000) 135] with a framework for obligations. However, the agential obligation operator explored here is *defined* in terms of a non-agential *yet personal* obligation operator and a non-deontic (and non-normal) agency operator. This is contrary to the main current trend, which assumes statements of personal obligation always take agential complements. Instead, I take the basic form to be *an agent's being obligated to be such that p*. I sketch some logics for agential obligation based on personal obligation and agency, first in a fairly familiar context that rules out conflicting personal obligations (and derivatively, conflicting agential obligations), and then in contexts that do allow for conflicts (of both sorts).

© 2004 Elsevier B.V. All rights reserved.

---

In [21] (and at DEON'00), I explored a simple framework for agency, predetermination and ability (in the process of developing a framework for agent-evaluative notions like praise and blame).<sup>1</sup> Call this simple framework the “APA framework”. In the current paper, I explore various ways of integrating the APA framework with a framework for personal obligations. To do the latter, I explore a familiar tradition, one that combines a non-agential deontic operator with a non-deontic agential operator to yield a derivative analysis of an agential deontic operator. However, I have in mind an exploration of the strategy suggested at the end of [19], which is critical of the traditional employments of this sort of analysis.

---

*E-mail address:* paulm@cisunix.unh.edu (P. McNamara).

<sup>1</sup> As noted in [21] in greater detail, the core of the agency framework employed here is an expansion of that employed in [18,24,25]. This framework is inspired, in part, by Elgesem [9] (see also [10]). The approach to pre-determination and ability is inspired by Brown [4]. My debt to Chellas [7] is apparent. The debt to Krogh and Herrestad is discussed explicitly in the current text. I also benefited by discussions from an earlier version of this paper with Mark Brown, Jorg Hansen, Andrew Jones, Leon van der Toore, and Peter Vranas.

1570-8683/\$ – see front matter © 2004 Elsevier B.V. All rights reserved.

doi:10.1016/j.jal.2004.01.006

In particular, the deontic operator employed is interpreted as one lying in between an impersonal deontic operator (e.g., “it is obligatory that”, “it ought to be the case that”) and an agential deontic operator (e.g., “it is obligatory for Doe to bring it about that”, “Doe ought to bring it about that”).<sup>2</sup>

In Section 1, I first sketch the predetermination portion of the APA framework from [21], and then integrate this with an Andersonian–Kangerian deontic “reduction” to get an SDL-generating modal framework for non-agential personal obligation. Next, I sketch the agency portion of the APA framework and integrate it with the deontic framework. Because the Andersonian–Kangerian deontic framework is so familiar, it is interesting to first see how in even such a strong framework for obligation, the weak monotonic classical logic for agency constrains the derivative logic for agential obligation. In Section 2, I drop the Andersonian–Kangerian reduction, and turn to weaker and more plausible systems for non-agential personal obligation, especially with an eye to allowing for conflicts of such obligations. I then explore various conflict-allowing systems for non-agential personal obligation with special attention to their impact on the derivative logics for agential obligation. I then briefly reintroduce a deontic constant in our conflict-allowing setting. In this context, the constant will have a restorative rather than reductive role, and the direction of the old reduction is, so to speak, reversed. Finally, I discuss a problem for my approach and sketch a solution.

## 1. A simple deontic-modal-agential framework

### 1.1. The APA framework—predetermination

The main operator in our framework for predetermination is:

PR<sub>p</sub>: It is (as of now) predetermined (for John Doe) that p.<sup>3</sup>

We use standard Kripke structures for modeling “PR”:

$$CO \subseteq W \times W.$$

$CO_{ij}$  iff what happens at  $j$  is consistent with our agent’s current abilities and disabilities at  $i$ .

<sup>2</sup> Horty’s [17] is highly recommended as a complement to [19]. Horty focuses on “it ought to be the case that Doe sees to it that”, and criticizes that as an analysis of what an agent ought to do, as well as making an interesting critical assessment of the literature regarding that strategy. The agency framework employed here is weaker, and “it ought to be the case that”, an impersonal deontic operator, is rather different from the personal but non-agential operator employed here to help in defining full-fledged agential obligations. His objections to that reductive strategy do not readily apply to the Krogh–Herrestad reductive strategy.

<sup>3</sup> Other readings might be: it is (as of now) inevitable/fixed/settled (for John Doe) that p. Note that since we do not consider multiple agents here, we do not bother with a superscript on PR standing for John Doe (e.g., PR<sup>J</sup>) as would be necessary with two or more agents. Nonetheless, as indicated, the intended reading is personal. Similar remarks apply to the subsequent Andersonian–Kangerian deontic constant, obligation operators, and agency operators.

The truth condition for PR is the usual one:

$$M \models_i \text{PR}p \text{ iff } \forall j(CO_{ij} \rightarrow M \models_j p).$$

We introduce the dual, “it is consistent with our agent’s abilities that p”:

$$COp \stackrel{\text{df}}{=} \sim \text{PR} \sim p,$$

$$M \models_i COp \text{ iff } \exists j(CO_{ij} \& \models_j p),$$

and we add a single constraint,

$$CO\text{-RFLX: } CO_{ii}.^4$$

That there is a p-world *consistent with* my abilities does not entail that p *is within* my abilities. Just consider any tautology, or any independent action someone else may or may or not perform.<sup>5</sup>

The normal KT System for PR (PR-KT),

SL: All Tautologies

PR-K:  $\text{PR}(p \rightarrow q) \rightarrow (\text{PR}p \rightarrow \text{PR}q)$

PR-T:  $\text{PR}p \rightarrow p$

MP: If  $\vdash p$  and  $\vdash p \rightarrow q$  then  $\vdash q$

PR-NEC: If  $\vdash p$  then  $\vdash \text{PR}p$ ,

is well-known to be sound and complete in all CO-reflexive models.

## 1.2. An Andersonian–Kangerian deontic-modal framework

We now add an Andersonian–Kangerian constant,  $d$ , to the syntax:

$d$ : The demands on John Doe are all met (or “John Doe’s responsibilities are all met”).

We represent “ $d$ ”’s extension as a set of worlds, DEM,

$$\text{DEM} \subseteq W,$$

and we give “ $d$ ”’s truth-conditions accordingly:

$$M \models_i d \text{ iff } i \in \text{DEM}.$$

We define our *non-agential but personal obligation* operator:

$$OBp \stackrel{\text{df}}{=} \text{PR}(d \rightarrow p) \quad [\text{Df-OB}],$$

<sup>4</sup> Given the intended interpretation, it is plausible to think of CO as an equivalence relation, but we ignore this here.

<sup>5</sup> See [21] for a bit more detail on this module.

and read it as follows:

OB: it is obligatory *for John Doe* to be such that (or “it is obligatory *for John Doe* that it be the case that”).<sup>6</sup>

We add an axiom governing the deontic constant, *d*:

*d*: CO*d* (i.e.,  $\sim$ PR $\sim$ *d*).

Call the resulting system “PR-KT*d*”.

Axiom *d* is validated by the condition that the satisfaction of Doe’s responsibilities is consistent with his abilities:

CO*d*:  $\forall i \exists j (CO_{ij} \ \& \ j \in \text{DEM})$ .

The system PR-KT*d* is characterized by the class of all models satisfying this constraint [20].

Recall that “CO*d*” says that *d*’s truth is *consistent with* John Doe’s abilities, but it does not say it is *within* his abilities, for good reason. John Doe may have delegated the last step in his project to his assistant, and it may now be predetermined for Doe that his project will be completed only if the assistant completes it, which she will. The project’s completion is no longer within Doe’s ability, but it is still consistent with his ability. Now just add that the project’s completion is equivalent to *d*.

It is well known that SDL is contained in PR-KT*d*:

SL: All Tautologies  
 OB-NC:  $OBp \rightarrow \sim OB \sim p$   
 OB-K:  $OB(p \rightarrow q) \rightarrow (OBp \rightarrow OBq)$   
 MP: If  $\vdash p$  and  $\vdash p \rightarrow q$  then  $\vdash q$   
 OB-NEC: If  $\vdash p$  then  $\vdash OBq$ .<sup>7</sup>

Given our intended interpretation of OB, *d* and CO, this means we are engaged in considerable idealization, but we want to begin with this simple system to see how a very familiar system for OB interacts with the less familiar non-normal logic for BA.

### 1.3. Interlude: are all obligations agential?

Our intended reading of “OB”, is suggested at the end of [19]. “OB*p*” is not intended here to express the *impersonal* notion that *it is obligatory that p*. Rather “OB*p*” is intended

<sup>6</sup> There are temporal issues here. For example, “OB*p* & *d*” is satisfiable in KT*d*, and “OB*p* & *d*  $\cdot$   $\rightarrow$  *p*” is a valid theorem. But if all the demands on Doe are met, then how could he (now) have any (outstanding) obligations? Talk of responsibilities is perhaps better here, since we can have standing responsibilities that are all currently satisfied. We must pass over these subtleties. See [5] for a good source on representing diachronic aspects of obligation and fulfillment.

<sup>7</sup> In fact the stronger system that results from adding “OB(OB*p*  $\rightarrow$  *p*)” to SDL corresponds to KT*d*, but we ignore this fact here. See [1,2].

to express a *personal* obligation, one that John Doe, our mock person, is under. Nonetheless, it does not require that John Doe be *the agent of p*. As Krogh and Herrestad in [19] illustrate, suppose,

... the manager of a firm is under an obligation that the companies financial statement is reported to the company board once a month. Let's assume that the manager has a particularly helpful assistant. Without the managers consent this assistant sends the financial status to the board each month, thus seeing to it that the manager's obligation is fulfilled. ... *the manager's obligations are personal, but may be fulfilled by someone else.* (p. 151) [My stress.]

Let's codify the point in a more explicit argument:

If all my obligations are agential, then each of my obligations is an obligation *for me* to bring about some thing. If each of my obligations is an obligation for me to bring about some thing, then none of my obligations can be fulfilled by someone else. But some of my obligations can be fulfilled by someone else. Therefore, not all my obligations are agential.

It certainly appears that the zealous assistant fulfills the manager's obligation but *the manager did not even indirectly (e.g., by delegation) bring that about*. Similarly, it may be obligatory for you that your child does her homework, but it may also be that you are lucky, and she does it on her own, with no prompting from you. If so, your obligation is fulfilled *with no effort on your part*.<sup>8</sup> Similarly, suppose you are obligated to be in Boston tomorrow, and without your involvement, you are kidnapped and taken to Boston. Your obligation to be in Boston is fulfilled without your agency. So "OB" is intended to express a personal obligation, yet one that is not an agential obligation (an obligation to do) since it does not require you to make any effort to fulfill it. Being obligated to do something is not what makes an obligation personal on this account. What makes it personal, presumably, is that if the obligatory thing doesn't come to pass, you are potentially responsible. In contrast, what makes an obligation agential is that only you can fulfill it—what is obligatory is that *you, yourself, do some thing*.

Although I do not pretend that this argument is conclusive, it is interesting because it indicates that

*either obligations are not all agential or fulfillment of obligations is more complex and subtle than has been assumed in deontic logic.*

If obligations are all agential then they are all obligations that John Doe be the agent of some action or state of affairs. But what then is it for an obligation to be *fulfilled by*

<sup>8</sup> We might say that the manager and parent are each obligated to be such that if the thing in question is not otherwise done, then s/he does it, but this just reinforces the next point, that an obligation can be an obligation to be a certain way, not necessarily to act a certain way, and thus an obligation may be not only non-agential in the complement, but may not even derive from such an obligation.

*someone else*? If what is obligatory is that *I myself* do something, how can anyone *else* fulfill *that* obligation? Maintaining that all obligations are agential will make fulfillment more complex than generally assumed. Here we will explore the alternative that keeps fulfillment simple.<sup>9</sup> Let me briefly say a few further things in support of the plausibility of reading “OB” as we do, and of distinguishing personal from agential obligation.

Consider the following:

Personal	Non-Agential	
/	\	/
\	/	\
I'm obligated	to be here.	(1)

This is an obligation *to be* in a location, not *to do* something. The sentential complement is non-agential. (1) can be aptly paraphrased as:

I'm obligated to be such that I am here. (1')

Now consider:

Personal	Agential	
/	\	/
\	/	\
I'm obligated	to bring it about that p.	(2)

This is a paradigm case of what we call an agential obligation. But consider:

I'm obligated to be such that I bring it about that p. (2')

This appears to be equivalent to (2), but this suggests the possibility that personal obligation is the more general form (not to be confused with the more frequently used form), and that agential obligation may be construed as a special case of personal obligation. We will indeed represent an agential obligation as a personal obligation with an agential complement.

True, my obligation *to be* somewhere would typically derive from an obligation *to do* something, one that had being at the location in question as a necessary condition, so that the latter would be a derived obligation. Still, derived obligations are obligations. Secondly, this relationship between obligations to be and obligations to do does not always hold. It seems perfectly alright to assert that a person *is obligated to be* cooperative, just, faithful, honest, punctual, where this means to have the virtues in question, not merely to act in the associated ways. Let's pick one example:

John Doe is obligated to be cooperative (i.e., such that he is a cooperative person).

This sounds fine. Furthermore, there is no inconsistency in adding to this claim that John Doe might already be cooperative; it might be a deep and stable part of his character. If so, can his obligation to be cooperative be reduced to an obligation to do something? What exactly? It can't be to become cooperative. He is already there. Surely he is not required to

---

<sup>9</sup> We leave open for future investigation the possibility that obligation is not always agential *and* fulfillment is not always simple and straightforward.

undo his cooperativeness and then reacquire it! Is it to make efforts to remain cooperative? Why? We are imagining that it requires no effort whatsoever on his part to remain cooperative. Surely he does not have to go through the motions, mumbling to himself “I will remain cooperative, by gummy!” in order to fulfill this obligation.

Now consider other evaluatives we need in deontic logic.<sup>10</sup> For example, we can be praiseworthy or blameworthy for having a certain trait. Unless we are being disingenuous, when we praise or blame someone, we do so because we believe they are praiseworthy or blameworthy, and we often praise and blame people sincerely for *traits*. In fact, we do so even for traits that do not appear to be a function of a person’s agency at all. For example, we praise people for being talented, smart, graceful, fast, etc. When we praise someone for being kind or blame someone for being callous, we engage in a form of moral evaluation. So what is the significance of insisting that personal obligation is agential, if other forms of evaluation, even moral forms of evaluation, are not? Do we want logic alone to rule out the substantive moral view that *if I am blameworthy for being F then it is obligatory for me to not be F* because the latter lacks an agential complement? Do we really need this sweeping thesis that obligations are agential, in order to vindicate the importance of agency to moral evaluation?

If these reflections are on track, then it may be the case that connections between agency and obligation have been a bit exaggerated of late.<sup>11</sup> In fact, it may be that the most salient obligations, obligations to do things, reduce to special cases of less salient obligations, obligations to be certain ways. We assume that this perspective is worth exploring, and proceed accordingly. Those who would insist that all obligations are agential need to make their case. I doubt it can be made.

#### 1.4. The APA framework—adding agency

We now introduce an agency operator:

BAp: John Doe Brings it About that p.

We use minimal models for the semantics:

$$BA : W \rightarrow \text{Pow}(\text{Pow}(W)), \quad \text{that is, } BA_i \subseteq \text{Pow}(W),$$

$$M \models_i \text{BAp} \quad \text{iff} \quad \|p\|^M \in BA_i.$$

$BA_i$  denotes the set of propositions (possibly empty) that our agent brings about at  $i$ .

Our basic system for BA, “TECNOCS” (“E” for “RE”), is:

BA-T:  $\vdash \text{BAp} \rightarrow p$

BA-C:  $\vdash (\text{BAp} \ \& \ \text{BAq}) \rightarrow \text{BA}(p \ \& \ q)$

BA-NO:  $\vdash \sim \text{BA} \top$

BA-CS:  $\vdash \text{BA}(p \ \& \ q) \rightarrow (\sim \text{BAp} \rightarrow \text{BAq})$

BA-RE: If  $\vdash p \leftrightarrow q$  then  $\vdash \text{BAp} \leftrightarrow \text{BAq}$ .

<sup>10</sup> For example, see [21].

<sup>11</sup> For example, see “The Restricted Complement Thesis” in [3].

We add these semantic constraints on any world,  $i$ , and proposition,  $X$ , for  $X \subseteq W$ , in a model:

*BA-t*: If  $X \in BA_i$  then  $i \in X$

*BA-c*: If  $X \in BA_i$  &  $Y \in BA_i$  then  $X \cap Y \in BA_i$

*BA-no*:  $W \notin BA_i$

*BA-cs*: If  $X \cap Y \in BA_i$  then  $X \in BA_i$  or  $Y \in BA_i$ .

Call the result of adding our earlier system, *KTd*, to this one, “TECNOCS-*KTd*”.

It is important to keep some things in mind later on. “BA” is meant to have a *strong* agential reading throughout. For example, RM (and thus necessitation) for BA fails: logical truths are not brought about *by* anything our agent now does on the above account. Also, although  $\sim BA\top$  holds, this does not mean that I can’t bring about a conjunction with  $\top$  as a conjunct. On the contrary, if I bring about any  $p$ , I will have thereby brought about  $(p \ \& \ \top)$ . BA-T expresses the success condition for agency, BA-C indicates that if an agent now brings about each of two things, then that agent now brings about both. BA-RE indicates that an agent always brings about anything logically equivalent to what she brings about.

BA-CS deserves separate comment. “Conjunctive Syllogism” for BA says that if I bring about the conjunction of two propositions, but not one of the conjuncts, then I bring about the other conjunct. This sounds right.<sup>12</sup> It would seem that if I do bring about a conjunction, but not (say) its first conjunct, then that is because that conjunct is rendered true independently of my agency.<sup>13</sup> But then the only way the truth of the conjunction could result from my agency is if the truth of the other conjunct results from it. Although it is not validated (or considered) in [9,10,18,24,25], it nonetheless seems quite plausible. So I will include it here.<sup>14</sup>

As noted, the above approach to “BA” is inspired by Elgesem, who rejects BA-K,

BA-K:  $BA(p \rightarrow q) \rightarrow (BAp \rightarrow BAq)$ ,

which would be validated by

BA-k: If  $(X \cup Y) \in BA_i$  &  $\neg X \in BA_i$  then  $Y \in BA_i$ .

In rejecting BA-K [9, p. 83], Elgesem follows an argument from [28]. I have doubts about the argument. However, some accounts of agency operators validate K, others don’t.<sup>15</sup> There is no agreement here. In a few of the initial places where BA-K would have impact, I note this in the footnotes.

<sup>12</sup> It is valid for the *astit*, *cstit* and *dstit* operators. See [3].

<sup>13</sup> This suggests the possibility, passed over here, of deriving BA-CS in a suitably stronger system from mixed principles:  $BA(p \ \& \ q) \ \& \ \sim BA p \ \cdot \rightarrow \ PRp$ ;  $BA(p \ \& \ q) \ \& \ PRp \ \cdot \rightarrow \ BAq$ . However, I ignore possible BA-PR bridge axioms here, in part because principles such as  $BA(p \ \& \ q) \ \& \ \sim BA p \ \cdot \rightarrow \ PRp$  or  $BAp \ \rightarrow \ CO \sim p$ , which have valid *dstit* and *astit* analogs, seem quite contentious to me. This is one reason why I think exploring these thinner agency structures inspired by Elgesem is worthwhile. I hope to discuss problems with the *dstit* and *astit* approach to agency elsewhere. See [13] for an excellent recent critical exposition of approaches to agency.

<sup>14</sup> I ignore various plausible BA-principles that involve embedding occurrences of BA within the scope of BA.

<sup>15</sup> For example, it is valid for the *cstit* and *dstit* operators, but not for the *astit* operator. See [3].



We introduce an operator for ability:

ABp: It is within John Doe's ABility that p,

ABp  $\stackrel{\text{df}}{=} \text{COBAp}$ .

The derivative truth-condition is:

$M \models_i \text{ABp}$  iff  $\exists j(\text{CO}ij \ \& \ M \models_j \text{BAp})$ .

*It is within our agent's ability that p* at *i* iff there is a world consistent with our agent's *i*-based abilities where our agent brings *p* about.

The following are derivable in TECNOCS-KT:<sup>16</sup>

BA-OD:  $\vdash \sim \text{BA}\perp$

BA-NC:  $\vdash \text{BAp} \rightarrow \sim \text{BA} \sim p$

BA-CS':  $\vdash \text{BA}(p \ \& \ q) \rightarrow (\text{BAp} \vee \text{BAq})$

BA-CS'':  $\vdash \text{BAp} \rightarrow \cdot \sim \text{BA}(p \vee q) \rightarrow \text{BA}(p \vee \sim q)$ <sup>17</sup>

CO-T:  $\vdash p \rightarrow \text{CO}p$

N-CO:  $\vdash \text{CO}\top$

BA-CO:  $\vdash \text{BAp} \rightarrow \text{CO}p$

AB-NO:  $\vdash \sim \text{AB}\top$

AB-OD:  $\vdash \sim \text{AB}\perp$

BA-AB:  $\vdash \text{BAp} \rightarrow \text{ABp}$

AB-RE: If  $\vdash p \leftrightarrow q$  then  $\vdash \text{ABp} \leftrightarrow \text{ABq}$ .

Finally, we introduce one more agential notion:

ARp: It is *agentially reflective* on John Doe that p,

ARp  $\stackrel{\text{df}}{=} \text{CO}p \ \& \ \text{PR}(p \rightarrow \text{BAp})$ .

*It is Agentially Reflective on John Doe that p* iff *p* is consistent with Doe's abilities and it is predetermined that: *p* is true only if Doe brings it about that *p*. The following are derivable:

$\vdash \text{ARp} \leftrightarrow (\text{ABp} \ \& \ \text{PR}(p \leftrightarrow \text{BAp}))$

$\vdash \sim \text{AR}\top$

If  $\vdash p \rightarrow q$  then  $\vdash \text{ARq} \rightarrow (\text{BAp} \rightarrow \text{BAq})$  [ARQ-RM].

That's it for the APA framework.<sup>18</sup>

<sup>16</sup> With BA-K, we would also get the following:  $\text{BA}(p \vee q) \ \& \ \text{BA}\neg p \cdot \rightarrow \text{BAq}$ ,  $\text{BAp} \ \& \ \text{BA}(p \rightarrow q) \cdot \rightarrow \text{BA}(p \ \& \ q)$ , and  $\text{CO}(\text{BAp} \ \& \ \text{BA}(p \rightarrow q)) \rightarrow \text{ABq}$ .

<sup>17</sup>  $\text{BA}(p \ \& \ p) \leftrightarrow \text{BA}((p \vee q) \ \& \ (p \vee \sim q))$ . So  $\text{BAp} \rightarrow \cdot \sim \text{BA}(p \vee q) \rightarrow \text{BA}(p \vee \sim q)$ .

<sup>18</sup> See [21] for a bit more detail on this module.

### 1.5. Agential obligation in our reductive framework

Let's explore the implications of coupling our account of personal obligation with our account of agency. We will be especially interested in examining the impact the non-normality of "BA" has on the compounding of "OB" with "BA" to get agential obligation.

We introduce an operator for agential obligation:

AOp: it is agentially obligatory for John Doe that p.

We define it with:

$$\text{AOp} \stackrel{\text{df}}{=} \text{OBBAp} \quad [\text{Df-AO}].$$

The derivative defining equivalence is:

$$| \text{AOp} \leftrightarrow \text{PR}(d \rightarrow \text{BAp}).$$

We offer "AOp" so defined as a tentative analysis of ordinary assertions of the sort:

"it is obligatory for John Doe to bring it about that p".

As Krogh and Herrestad suggest, this strategy holds some promise of avoiding some of the problems that occur as a result of analyzing agential obligation as a combination of an impersonal obligation operator with an agential operator.<sup>19</sup>

So what logical connections do agential obligations have to personal obligations, agency, ability and inevitability according to the current framework? First of all, AOp is not normal. Some of the most salient principles governing "OB" for SDL fail for "AO":

$$\text{AO}(p \rightarrow q) \rightarrow (\text{AO}p \rightarrow \text{AO}q) \quad [\text{AO-K}]$$

$$\text{AO}(p \ \& \ q) \rightarrow (\text{AO}p \ \& \ \text{AO}q) \quad [\text{AO-M}]$$

$$\text{If } \vdash p \rightarrow q \text{ then } \vdash \text{AO}p \rightarrow \text{AO}q \quad [\text{AO-RM}]$$

$$\text{If } \vdash p \text{ then } \vdash \text{AO}p \quad [\text{AO-NEC}].$$

There are no valid wffs or theorems of the form BAp. Given BA-T and BA-RE, such a theorem would entail BA $\top$ , which conflicts with BA-NO—and we are assuming consistency for that system.

Also notice that, although

$$\text{OB}d \quad [\text{OB}d],$$

plainly holds, its AO-analogue,

$$\text{AO}d \quad [\text{AO}d],$$

does not hold. Although it is predetermined that if the demands on John Doe are all met, then they are all met, it is not predetermined that if they are all met, then Doe, himself, brings it about that they are all met.

<sup>19</sup> For example Carmo's leakage problem [19, pp. 138–139] is plugged by this approach, but this is essentially a multi-agent problem, so we do not consider it here.

Although  $OB\top$  is a theorem, its antithesis for AO holds:<sup>20</sup>

$$\vdash \sim AO\top \quad [AO-NO].$$

**Proof.** Given BA-NO,  $\vdash \sim BAT$ , OB-NEC yields  $\vdash OB\sim BAT$ . But from OB-NC,  $\vdash OBBA\top \rightarrow \sim OB\sim BAT$ . So  $\vdash \sim OBBA\top$ .  $\square$

An aggregation principle holds:

$$\vdash (AOp \ \& \ AOq) \rightarrow AO(p \ \& \ q). \quad [AO-C].$$

**Proof.** Assume  $OBBAp \ \& \ OBBAq$ . By OB-C,  $OB(BAp \ \& \ BAq)$ . But BA-M holds:  $\vdash BAp \ \& \ BAq. \rightarrow BA(p \ \& \ q)$ . So by OB-RM,  $OBBA(p \ \& \ q)$ .  $\square$

An RE rule is also derivable:

$$\text{If } \vdash p \leftrightarrow q \text{ then } \vdash AOp \leftrightarrow AOq \quad [RE-AO].$$

As a special case of OB-NC, we obviously get:

$$\vdash AOp \rightarrow \sim OB\sim BAp \quad [NC'-AO].$$

But  $\sim OB\sim BAp$  is not the dual of  $AOp$ . The dual of  $AOp$  is  $\sim AO\sim p$ , i.e.,  $\sim OBBA\sim p$ . So  $NC'-AO$  is not a standard No-Conflicts principle for AO. Rather, this is:

$$\vdash AOp \rightarrow \sim AO\sim p \quad [AO-NC].$$

**Proof.** Assume  $PR(d \rightarrow BAp)$  and  $OBBA\sim p$ . Then  $PR(d \rightarrow BA\sim p)$ . So  $PR(d \rightarrow (BAp \ \& \ BA\sim p))$ . Since  $COd$ ,  $CO(BAp \ \& \ BA\sim p)$ . But by BA-T,  $\vdash (BAp \ \& \ BA\sim p) \rightarrow (p \ \& \ \sim p)$ , and then  $CO(p \ \& \ \sim p)$ , contra  $\vdash \sim CO\perp$ .  $\square$

No agential obligations to do the impossible are allowed:

$$\vdash \sim AO\perp. \quad [AO-D].$$

BA-CS yields:

$$\vdash AO(p \ \& \ q) \rightarrow OB\sim BAp \rightarrow AOq.$$

**Proof.** Assume  $OBBA(p \ \& \ q)$  and  $OB\sim BAp$ . By OB-C,  $OB(BA(p \ \& \ q) \ \& \ \sim BAp)$ . By BA-CS,  $\vdash BA(p \ \& \ q) \ \& \ \sim BAp \rightarrow BAq$ . So by OB-RM,  $AOq$ .  $\square$

Similarly for,

$$\begin{aligned} \vdash AO(p \ \& \ q) &\rightarrow OB(BAp \ \vee \ BAq), \\ \vdash AOp \rightarrow \cdot OB\sim BA(p \ \vee \ q) &\rightarrow AO(p \ \vee \ \sim q). \end{aligned}$$

<sup>20</sup> Some proofs are included to give a sense of how the logic modules interact in generating derivative principles for AO.

However, as things now stand, the following AO-analog to BA-CS, call it “AO-Conjunctive Syllogism”, is not a theorem of TECNOCS-KTd:

$$\text{AO}(p \ \& \ q) \ \& \ \sim \text{AO}p \ \cdot \rightarrow \ \text{AO}q \quad [\text{AO-CS}].^{21}$$

Neither is it valid given our semantics. For consider the following countermodel, M.

$$\begin{aligned} W &= \{i, j, k, l\}, & CO &= W \times W, & \|p\|^M &= \{i, j, k\}, & \|q\|^M &= \{i, j, l\}, \\ \|d\|^M &= \{i, j\}, & BA_i &= \{\|p\|^M \cap \|q\|^M, \|p\|^M\}, & BA_j &= \{\|p\|^M \cap \|q\|^M, \|q\|^M\}, \\ BA_k &= BA_l = \emptyset. \end{aligned}$$

Here, all the axioms and rules of TECNOCS-KTd are satisfied, as are all our semantic constraints, but not AO-CS. Although in any world where I bring it about that  $p \ \& \ q$ , I do bring about  $p$  or bring about  $q$  (i.e., BA-CS is satisfied), nonetheless, even though in each ideal world I bring about the conjunction of  $p$  and  $q$ , in one I bring about  $p$ , but do not bring about  $q$ , and in the other I bring about  $q$ , but do not bring about  $p$ . So it is not obligatory to bring about either one in particular. Nothing so far rules this out, yet AO-CS is worth considering as a candidate axiom. I have not been able to think of a fully convincing intuitive counterexample. The best I have been able to do is to consider a case where, say, a mindless machine will, by a truly random device, either press button one or press button two, but not both. Let  $p$  = button one is pressed (by someone or thing), and  $q$  = button two is pressed (by someone or thing). Let’s suppose that I am obligated to bring about whichever button-pressing the machine does not cause. Since it is indeterminate as to which button the machine will “choose”, I am neither obligated to bring about  $p$  nor obligated to bring about  $q$ . Questions: Is the obligation for me to bring about whatever the machine doesn’t one I *can* fulfill, and supposing it is, if I comply with this obligation, have I brought about *the conjunction of  $p$  &  $q$*  in doing so?

Call the addition of this BA-CS analog to TECNOCS-KTd, “TECNOCS-Ktd+”. The following constraint validates AO-CS:

$$\text{AO-cs:} \quad \text{If } CO^i \cap \text{DEM} \subseteq BA(X \cap Y) \quad \text{then either } CO^i \cap \text{DEM} \subseteq BAX \\ \text{or } CO^i \cap \text{DEM} \subseteq BAY,$$

where  $CO^i = \{j: COij\}$ , the set of  $i$ -ability-consistent world, and  $BAZ = \{i: Z \in BA_i\}$ , the proposition that our agent brings proposition  $Z$  about.

Analogous to BA-CS’ and BA-CS’’ follow in “TECNOCS-KTd+”:

$$\begin{aligned} \vdash \text{AO}(p \ \& \ q) \ \rightarrow \ (\text{AO}p \ \vee \ \text{AO}q) \quad [\text{AO-CS}'], \\ \vdash \text{AO}p \ \rightarrow \ \cdot \ \sim \text{AO}(p \ \vee \ q) \ \rightarrow \ \text{AO}(p \ \vee \ \sim q) \quad [\text{AO-CS}'']. \end{aligned}$$

Unless otherwise specified, we will focus on just TECNOCS-KTd and its semantics.

It is also the case that what is agentially obligatory is personally obligatory:

$$\vdash \text{AO}p \ \rightarrow \ \text{OB}p \quad [\text{AO-OB}].$$

The converse fails, as we should hope.

<sup>21</sup> Its OB analog,  $\text{OB}(p \ \& \ q) \ \rightarrow \ \cdot \ \sim \text{OB}p \ \rightarrow \ \text{OB}q$ , is of course derivable from OB-M.

Although AO-K fails,<sup>22</sup> a close cousin is derivable:

$$\vdash \text{AO}(p \rightarrow q) \rightarrow (\text{AO}p \rightarrow \text{OB}q) \quad [\text{Weak AO-K}].$$

**Proof.** Assume  $\text{PR}(d \rightarrow \text{BA}(p \rightarrow q))$  and  $\text{PR}(d \rightarrow \text{BA}p)$ . So  $\text{PR}(d \rightarrow (\text{BA}(p \rightarrow q) \& \text{BA}p))$ . By BA-T, it follows that  $\vdash \text{BA}(p \rightarrow q) \& \text{BA}p \cdot \rightarrow q$ , and then by PR-NEC, we get  $\text{PR}(\text{BA}(p \rightarrow q) \& \text{BA}p \cdot \rightarrow q)$ . So  $\text{PR}(d \rightarrow q)$ .  $\square$

Although

$$\vdash \text{PR}(p \rightarrow q) \rightarrow (\text{OB}p \rightarrow \text{OB}q) \quad [\text{OB-RM}']$$

obviously holds, its AO analogue fails:

$$\text{PR}(p \rightarrow q) \rightarrow (\text{AO}p \rightarrow \text{AO}q).$$

However, a weaker cousin holds:

$$\vdash \text{PR}(p \rightarrow q) \rightarrow (\text{AO}p \rightarrow \text{OB}q) \quad [\text{Weak AO-RM}'].$$

Similarly, although AO-RM fails, this cousin holds:

$$\text{If } \vdash p \rightarrow q \text{ then } \vdash \text{AO}p \rightarrow \text{OB}q \quad [\text{Weak AO-RM}].$$

As is well-known, for a KTD system, a version of Kant's Law for OB holds

$$\vdash \text{OB}p \rightarrow \text{CO}p \quad [\text{KL}],$$

and this carries over to AO as well as a special case,

$$\vdash \text{AO}p \rightarrow \text{AB}p \quad [\text{AO-AB}].$$

So, what is agentially obligatory is within the ability of the agent.

Some additional miscellaneous principles follow:

$$\vdash \text{AO}p \rightarrow \text{OBAB}p.$$

$$\vdash \text{AO}p \leftrightarrow \text{AO}(p \& \top)$$

$$\vdash \text{AO}(p \& \top) \rightarrow (\text{AO}p \& \sim \text{AO}\top)$$

$$\text{If } \vdash q \text{ then } \vdash \text{AO}(p \& q) \rightarrow (\text{AO}p \& \text{OB}\sim \text{BA}q).$$

**Proof.** Assume  $\vdash q$ . The BA analogue for this rule holds: if  $\vdash q$ , then  $\vdash \text{BA}(p \& q) \rightarrow (\text{BA}p \& \sim \text{BA}q)$ . So  $\vdash \text{BA}(p \& q) \rightarrow (\text{BA}p \& \sim \text{BA}q)$ . But then by OB-RM,  $\vdash \text{OBBA}(p \& q) \rightarrow \text{OB}(\text{BA}p \& \sim \text{BA}q)$ . But by OB-M,  $\vdash \text{OB}(\text{BA}p \& \sim \text{BA}q) \rightarrow (\text{OBBA}p \& \text{OB}\sim \text{BA}q)$ . So  $\vdash \text{AO}(p \& q) \rightarrow (\text{AO}p \& \text{OB}\sim \text{BA}q)$ .  $\square$

<sup>22</sup> However, given BA-K,  $\text{AO}(p \rightarrow q) \rightarrow (\text{AO}p \rightarrow \text{AO}q)$ , that is AO-K, is derivable with the help of OB-C and OB-RM.

Here are some expected (though contentious) connections between predetermination and obligation:

$$\vdash \text{PRp} \rightarrow \text{OBp} \quad [\text{PR-OB}],$$

$$\vdash \text{PRBAp} \rightarrow \text{AOp}.$$

In passing, what happens if we add an analogue to determinism,  $\vdash p \rightarrow \text{PRp}$ , to our system? Well, in well-known ways, predetermination for John Doe and consistency with John Doe's abilities will both collapse into truth:

$$\text{If } \vdash (p \rightarrow \text{PRp}) \quad \text{then } \vdash (p \leftrightarrow \text{PRp}),$$

$$\text{If } \vdash (p \rightarrow \text{PRp}) \quad \text{then } \vdash (\text{COp} \leftrightarrow p).$$

We then easily get these derivative analogues to hard determinism regarding personal and agential obligations:

$$\text{If } \vdash (p \rightarrow \text{PRp}) \quad \text{then } \vdash (\text{OBp} \leftrightarrow p),$$

$$\text{If } \vdash (p \rightarrow \text{PRp}) \quad \text{then } \vdash (\text{AOp} \leftrightarrow \text{BAp}).$$

Recall  $\text{ARp} \stackrel{\text{df}}{=} \text{COp} \ \& \ \text{PR}(p \rightarrow \text{BAp})$ . Although we have identified a number of invalid principles above, in many cases, agentially reflective qualified versions of them are valid. This tends to confirm the observation in [21] that this agency notion is of interest.<sup>23</sup> However, below, it is primarily the “*only me* if anyone” second clause in the definition of AR that does the work. So let's define an “only me” operator for this component,

$$\text{OMp} \stackrel{\text{df}}{=} \text{PR}(p \rightarrow \text{BAp}),$$

and employ it, noting the corollaries regarding AR.

$$\vdash \text{OMp} \rightarrow (\text{OBp} \rightarrow \text{AOp}) \quad [\text{OM-Qualified OB-AO}].$$

**Corollary.**  $\vdash \text{ARp} \rightarrow (\text{OBp} \rightarrow \text{AOp})$ .

This says that if it is obligatory for me that  $p$  be the case and it is inevitable that  $p$  will occur only if I bring it about, then it is obligatory for me that  $I$  bring it about.

$$\vdash \text{OMq} \rightarrow (\text{PR}(p \rightarrow q) \rightarrow (\text{AOp} \rightarrow \text{AOq})) \quad [\text{OM-Qualified OA-RM}'].$$

**Corollary.**  $\vdash \text{ARq} \rightarrow (\text{PR}(p \rightarrow q) \rightarrow (\text{AOp} \rightarrow \text{AOq}))$ .

$$\text{If } \vdash p \rightarrow q \quad \text{then } \vdash \text{OMq} \rightarrow (\text{AOp} \rightarrow \text{AOq}) \quad [\text{OM-Qualified AO-RM}].$$

**Corollary.**  $\text{If } \vdash p \rightarrow q \quad \text{then } \vdash \text{ARq} \rightarrow (\text{AOp} \rightarrow \text{AOq})$ .

<sup>23</sup> It should not be surprising that in a context that distinguishes personal from agential obligations, the notion of things that can hold only if I bring them about myself would have special significance.

$$\vdash \text{OMp} \ \& \ \text{AO}(p \ \& \ q) \ \rightarrow \ \text{AOp}.$$

$$\vdash (\text{OMp} \ \& \ \text{OMq}) \ \rightarrow \ (\text{AO}(p \ \& \ q) \ \rightarrow \ (\text{AOp} \ \& \ \text{AOq})) \quad [\text{OM-Qualified AO-M}].$$

**Corollary.**  $\vdash (\text{ARp} \ \& \ \text{ARq}) \ \rightarrow \ (\text{AO}(p \ \& \ q) \ \rightarrow \ (\text{AOp} \ \& \ \text{AOq})).$

Note that

$$\text{OM}(p \ \& \ q) \ \rightarrow \ (\text{AO}(p \ \& \ q) \ \rightarrow \ (\text{AOp} \ \& \ \text{AOq}))$$

is not a theorem. This OM-Qualified version of AO-C fails for the reasons that  $\text{OM}(p \ \& \ q) \ \rightarrow \ \text{OMp}$  fails, which, in turn, ultimately hinges on the failure of  $\text{BA}(p \ \& \ q) \ \rightarrow \ \text{BAp}$ . Just let  $p$  be  $\top$ : even if I am obligated to bring about the conjunction, I am not obligated to bring about an independently settled conjunct.

An OM-qualified version of K follows as well.<sup>24</sup>

$$\vdash \text{OMq} \ \rightarrow \ (\text{AO}(p \ \rightarrow \ q) \ \rightarrow \ (\text{AOp} \ \rightarrow \ \text{AOq})) \quad [\text{OM-Qualified AO-K}].$$

**Corollary.**  $\vdash \text{AO}(p \ \rightarrow \ q) \ \rightarrow \ (\text{ARq} \ \rightarrow \ (\text{AOp} \ \rightarrow \ \text{AOq})).$

### 1.6. The pure agential obligation fragment of TECNOCS-KTd

What is the logic of the compound operator, AO, in isolation? Consider the “pure” AO principles we derived in TECNOCS-KTd.<sup>25</sup> There were just these:

$$\vdash \text{AOp} \ \rightarrow \ \sim \text{AO} \sim p \quad [\text{AO-NC}]$$

$$\vdash (\text{AOp} \ \& \ \text{AOq}) \ \rightarrow \ \text{AO}(p \ \& \ q). \quad [\text{AO-C}]$$

$$\vdash \sim \text{AO} \top \quad [\text{AO-NO}]$$

$$\vdash \sim \text{AO} \perp \quad [\text{AO-D}]$$

$$\text{If } \vdash p \leftrightarrow q \ \text{ then } \vdash \text{AOp} \leftrightarrow \text{AOq} \quad [\text{AO-RE}].$$

Given AO-C and AO-RE, AO-NC follows from AO-D (but not vice versa, as we will see). So let us zero in on the following system, call it “DEC-NO”:

SL: All Tautologies

AO-D:  $\vdash \sim \text{AO} \perp$

AO-C:  $\vdash (\text{AOp} \ \& \ \text{AOq}) \ \rightarrow \ \text{AO}(p \ \& \ q)$

AO-NO:  $\vdash \sim \text{AO} \top$

MP:  $\vdash \text{If } \vdash p \ \text{and } \vdash p \rightarrow q \ \text{ then } \vdash q$

AO-RE:  $\text{If } \vdash p \leftrightarrow q \ \text{ then } \vdash \text{AOp} \leftrightarrow \text{AOq}.$

<sup>24</sup> From Weak AO-K and OM-Qualified OB-AO.

<sup>25</sup> If we had BA-K, then we would need to add AO-K and make adjustments in what is said here accordingly.

Notice that this classical AO-logic is a proper sub-logic of our BA logic. For the only difference in the axiom systems other than “AO” versus “BA” is that only TEC-NO has BA-T and BA-CS as axioms and only DEC-NO has  $\sim\text{AO}\perp$  as an axiom. But  $\sim\text{BA}\perp$  is plainly derivable from BA-T and SL, but as can be easily shown (e.g., using the semantics below), AO-T is not derivable in DEC-NO.

Comparing DEC-NO for AO to SDL for OB, the only base principle in DEC-NO that has no analogue in SDL is the axiom AO-NO. On the other hand, although OB-NC of SDL has an analogue (as a theorem) in DEC-NO, neither of the AO-analogues of OB-K and OB-NEC are theorems of DEC-NO.

Recall AO-CS from TECNOCS-KTd+:

$$\vdash \text{AO}(p \ \& \ q) \ \& \ \sim\text{AO}p \ \cdot \rightarrow \text{AO}q.$$

This is also a pure-AO principle that we tentatively considered. If we add AO-CS to DEC-NO, we get “DECNOCS”.

Representing agential obligation for the moment with a primitive operator, we can give a direct standard minimal models semantics for these two classic AO-systems in the usual way:

$$\text{AO} : W \rightarrow \text{Pow}(\text{Pow}(W)), \text{ i.e., } \text{AO}_i \subseteq \text{Pow}(W)$$

$$M \models_i \text{AO}p \quad \text{iff} \quad \|p\|^M \in \text{AO}_i.$$

We then introduce the following constraints to validate AO-D, AO-C AO-NO, and AO-CS respectively:

$$\emptyset \notin \text{AO}_i$$

$$\text{If } X \in \text{AO}_i \text{ and } Y \in \text{AO}_i \quad \text{then } X \cap Y \in \text{AO}_i$$

$$W \notin \text{AO}_i.$$

$$\text{If } X \cap Y \in \text{AO}_i \quad \text{then either } X \in \text{AO}_i \text{ or } Y \in \text{AO}_i.$$

We can now also see that given AO-RE and AO-C, although AO-NC is derivable in TECNOCS-KTd from AO-D, the converse does not hold, for just consider a simple model where the only member of  $\text{AO}_i$  is the empty set. RE, AO-NC and AO-C all hold, but not AO-D. So AO-D is stronger than AO-NC even though OB-D and OB-NC are interchangeable for our SDL and KTd systems.

**Conjecture 1.** *DEC-NO (DECNOCS) is determined by the class of all minimal models satisfying the above three (four) constraints.*

**Conjecture 2.** *DEC-NO (DECNOCS) is the pure AO-fragment of TECNOCS-KTd (TECNOCS-KTd+). (This claim can be made more precise following [1,2], and I suspect it can be proved in a similar fashion, which would depend on first proving Conjecture 1.)*



## 2. Agential obligation with weaker deontic bases

### 2.1. Weakening the SDL base for OB

As noted earlier, we engaged in a fair amount of idealization above. We now wish to discharge some of that in order to get a better account of personal obligations, and thereby a better derivative account of agential obligations.

Recall that in system PR-Kd, we defined “OB” as follows:

$$\text{OB}_p = \underset{\text{df}}{\text{PR}}(d \rightarrow p) \quad [\text{Df-OB}].$$

And stipulated this axiom:

$$d: \text{CO}d \quad (\text{i.e., } \sim\text{PR}\sim d).$$

Among other things, this generated SDL for OB:

- SL: All Tautologies
- OB-NC:  $\text{OB}_p \rightarrow \sim\text{OB}\sim p$
- OB-K:  $\text{OB}(p \rightarrow q) \rightarrow (\text{OB}_p \rightarrow \text{OB}_q)$
- MP: If  $\vdash p$  and  $\vdash p \rightarrow q$  then  $\vdash q$
- OB-NEC: If  $\vdash p$  then  $\vdash \text{OB}_q$ .

Now it certainly seems that there are some objectionable things about this account of personal obligation.

Perhaps first and foremost, is the assumption,  $\text{CO}d$ , that it is a logical truth that all of the demands on John Doe can be met. At the level of SDL, this (and Df-OB) generates OB-NC, asserting that it is a logical truth that John Doe never has conflicting personal obligations. Both of these claims seem to be false, and we will be especially interested in the impact of discharging these assumptions. Once we reject the logical necessity of  $\text{CO}d$ , it is not clear that there is any significant role left at all for  $d$  to play. Without  $\text{CO}d$ , if we define  $\text{OB}_p$  as  $\text{PR}(d \rightarrow p)$ , then in all situations where it is predetermined that the demands on John Doe *cannot* all be met (i.e.,  $\text{PR}\sim d$ ), it follows that *absolutely everything* is obligatory. But this is surely not right. So we must reject the Andersonian–Kangerian reduction if we want to allow for non-empty sets of obligations that can’t be jointly realized in situations where this is not because everything is obligatory. We will see toward the end that there may still be a role for  $d$  in a conflict-allowing context, but it will be *restorative*, not *reductive*. For now, let’s set  $d$  aside and focus on SDL independently of its deducibility from  $\text{K}d$ .

We will want to weaken SDL by rejecting OB-NC. However, as is usual in conflict-allowing treatments, we do not want to thereby reject the distinct and plausible claim that there are no personal obligations to do logically impossible things. So we will replace OB-NC with:

$$\sim\text{OB}\perp. \quad [\text{OB-D}].$$

Since we will be considering classical non-normal systems for OB, we will employ minimal models. With that in mind, we introduce a function,  $\text{OB}$ , from worlds to sets of

propositions:

$$OB: W \rightarrow \text{Pow}(\text{Pow}(W)), \text{ i.e., } OB_i \subseteq \text{Pow}(W).$$

For a given world,  $OB_i$  yields a set of propositions (possibly empty) that are obligatory for John Doe at that world. We then model a claim that it is obligatory for John Doe that  $p$  accordingly:

$$M \models_i OBp \text{ iff } \|p\|^M \in OB_i.$$

We can ratify OB-D by constraining models accordingly:

$$OB\text{-d: } \emptyset \notin OB_i.$$

As will be seen, by allowing for conflicts of personal obligation, we will thereby derivatively allow for conflicts of agential obligation. In fact, this is the main goal here. However, a few remarks are in order about the remaining deontic principles for OB, as we will consider a sequence of successively stronger OB-systems and sketch their impact on AO (defined as OBBA).

Let me begin with OB-NEC. This appears to be objectionable on two counts. First, it implies that it is obligatory for John Doe that it is raining or not raining, and similarly for all logical truths. A bit of the sting can perhaps be taken out of this by noting that personal obligation is not agential obligation. For the personal obligation reading of “OB” is “it is obligatory for an individual *that p be the case*”, and not that *that individual make p be the case*. Now any basic obligation that  $p$  is the case can be met only if all  $p$ 's consequences are derivatively met. So it might be suggested that although logical truths are merely logically derivative obligations that are practically not worth stating ordinarily (and thus pragmatically odd to state), their being the case is nonetheless obligatory for an individual. They are the limiting case of things rendered obligatory consequential upon more basic obligations, obligations that are worth stating. But even granting the spirit of this weak defense of OB-NEC for sake of argument, there is still a second objection: namely that a consequence of OB-NEC is that it is a logical truth that some things are obligatory for John Doe. Even if we grant that things *can* be derivatively obligatory, right up to logical truths, that wouldn't yet imply that there must always be such obligatory things, for there might be situations where John Doe has no basic obligations from which to generate such derivative obligatory things. For example, if John Doe is alone on a desert island, with no hope of rescue, he might eventually be under no obligations.

In this respect, OB-RM looks less unattractive than OB-NEC:

$$\text{If } \vdash p \rightarrow q \text{ then } \vdash OBp \rightarrow OBq \text{ [OB-RM].}$$

Relatedly, consider principle M,

$$OB(p \ \& \ q) \rightarrow (OBp \ \& \ OBq) \text{ [OB-M].}$$

This says that if a conjunction is obligatory for John Doe, then so is each conjunct. Although this is plainly problematic for an agential reading, it seems less problematic for the non-agential personal reading we are giving, but once again, given OB-RE, it entails that

OB $\top$  is possible. As is well known,

OB-RE + OB-M and OB-RM are interderivable [7, p. 236].

We will rely on this fact implicitly throughout.

Both OB-M and OB-RM are validated by either of the following equivalent conditions [7, p. 215]:

OB-m: If  $X \cap Y \in OB_i$ , then so is X  
OB-rm: If  $X \subseteq Y$  then if  $X \in OB_i$  then  $Y \in OB_i$ .

RM and M are certainly controversial. However, the main aim here is to highlight the way in which the constraints we have already placed on agency impact the compounding of agency with a personal obligation operator *that allows for conflicts*. So we will include classical systems that contain OB-RM/OB-M. The main OB elements we change here bear on conflicts per se. We postpone to another time the exploration of systems with more plausible and thus restricted versions of RM and M.

What of OB-K?

$OB(p \rightarrow q) \rightarrow (OBp \rightarrow OBq)$  [OB-K]

Although we might argue against AO-K, by arguing against BA-K, at first blush at least, a K principle for personal obligation seems plausible. We will return to the apparent plausibility of OB-K again in a moment.

Consider C:

$(OBp \ \& \ OBq) \rightarrow OB(p \ \& \ q)$  [OB-C].

We must reject K in any systems strong enough to generate RM, for: in the context of RM, C and K are equivalent:

Given RM,  $\vdash K$  iff  $\vdash C$ .

But why do we reject OB-C in the first place? Well the most obvious reason is that it yields a special version of OB-C, which I will call “OB-CD” (since it combines features of C and D):

$OBp \ \& \ OB\sim p \rightarrow OB(p \ \& \ \sim p)$  [OB-CD].

This says that whenever I have directly conflicting obligations, the logical contradiction formed by their conjunction is also obligatory for me. Clearly, we must reject OB-C if we are to both allow for conflicts of obligation but rule out obligations that logically impossible things obtain. Many have thought that a contradiction can never be obligatory for anyone. Even if a contradiction could somehow be obligatory for someone (Romeo solemnly promises Juliet that he will square the circle), such an obligation wouldn’t follow from the mere fact that I had a directly conflicting set of obligations. So it seems OB-CD should be blocked no matter what.

But more can be said against it. Consider a related formula, call it the “Fatal Formula” for a logic of conflicts:

$(OBp \ \& \ OB\sim p) \rightarrow OBq$  [OB-FF].

This says that in the presence of a single direct conflict, every proposition is obligatory, and thus all OB-distinctions between formulas disappear.

Now any SL-inclusive system with OB-RM and OB-CD allows us to conclude FF:

$$\text{OB-RM} + \text{OB-CD} \vdash \text{OB-FF}.$$

**Proof.** Assume  $\text{OB-RM} + \text{OB-CD}$  and  $\text{OB}p \ \& \ \text{OB}\sim p$ . By OB-CD, we have  $\text{OB}(p \ \& \ \sim p)$ . But since  $\vdash (p \ \& \ \sim p) \rightarrow q$ , by OB-RM, we get  $\text{OB}q$ .  $\square$

**Corollary.**  $\text{OB-RE} + \text{M} + \text{OB-CD}, \vdash \text{OB-FF}$ .

So any Classical System with OB-M must avoid OB-CD at all costs. We will revisit OB-CD.

Since we need OB-RM to derive OB-C from OB-K, and RM (and M) are contentious, this suggests the option of keeping OB-K, and dropping OB-RM while retaining the relatively innocuous OB-RE. Some have thought OB-K, even in conflict contexts is not only acceptable, but compelling. Indeed, I also thought so when I began this project, and I was temporarily stymied because I felt it needed to be rejected, yet I couldn't see how to motivate its rejection intuitively. How could it be obligatory for me that if  $p$  then  $q$ , and obligatory for me that  $p$ , yet not be obligatory for me that  $q$ ? As Horty put it, in discussing a classical system with just OB-RE and OB-M,

... in weakening standard deontic logic to allow conflicts, it seems that we have now arrived at a system that is too weak: it fails to validate intuitively desirable inferences. Suppose for example that an agent is subject to the following two norms, you ought either to fight in the army or perform alternative service, you ought not to fight in the army.

We can represent these norms through the formulas  $\text{O}(F \vee S)$  and  $\text{O}\sim F$ . Now it seems intuitively that the agent should conclude from these premises that he ought to perform alternative service. However, the inference from  $\text{O}(F \vee S)$  and  $\text{O}\sim F$  to  $\text{O}S$  is not valid in the logic EM. [16, p. 21].

Plainly, the principle behind the inference in question is OB-K in disguise (i.e.,  $\text{OB}(\sim F \rightarrow S) \rightarrow (\text{OB}\sim F \rightarrow \text{O}S)$ ). Unfortunately, despite OB-K's initially plausible ring, it is quite unacceptable on reflection, since it entails the validity of OB-CD as a special case in *all* classical systems, and thus *even without OB-M/RM*:

Given OB-RE, if  $\vdash \text{OB-K}$  then  $\vdash \text{OB-CD}$ .

**Proof.** By SL,  $\vdash \sim p \leftrightarrow (p \rightarrow (p \ \& \ \sim p))$ . So by RE,  $\text{OB}\sim p \leftrightarrow \text{OB}(p \rightarrow (p \ \& \ \sim p))$ . By K,  $\text{OB}(p \rightarrow (p \ \& \ \sim p)) \rightarrow \cdot \text{OB}p \rightarrow \text{OB}(p \ \& \ \sim p)$ . So  $\text{OB}\sim p \rightarrow \cdot \text{OB}p \rightarrow \text{OB}(p \ \& \ \sim p)$ , i.e.,  $(\text{OB}p \ \& \ \text{OB}\sim p) \rightarrow \text{OB}(p \ \& \ \sim p)$ .  $\square$

Seen semantically, OB-K amounts to this:

$$\text{If } \neg X \cup Y \in \text{OB}_i \ \& \ X \in \text{OB}_i \ \text{ then } Y \in \text{OB}_i.$$

But this implies the following special case:

$$\text{If } \neg X \cup \emptyset \in OB_i \text{ \& } X \in OB_i \text{ then } \emptyset \in OB_i.$$

Since  $\neg X \cup \emptyset = \neg X$ , and  $X \cap \neg X = \emptyset$ , this is equivalent to:

$$\text{If } \neg X \in OB_i \text{ \& } X \in OB_i \text{ then } X \cap \neg X \in OB_i,$$

which is just a natural expression of OB-CD's truth-condition.

So although  $K + RE$  does not entail  $C$  or  $FF$  ( $M$  is needed for these), it only takes maximally plausible  $RE$ , to derive  $CD$  from  $K$ . So  $OB-K$  is not acceptable at all. When we see it as plausible, it is because we naturally think of the wffs, " $p \rightarrow q$ " and " $p$ ", governed by the first two "OB"s in  $OB-K$  as *mutually consistent*, as in Horty's example where we assume not fighting and alternative service are jointly satisfiable. But this need not be so, as the instance in the proof above makes clear. There we saw that  $K$  automatically generates a *contradictory obligation* from a pair of feebly disguised *conflicting obligations*, using only  $OB-RE$ . So  $OB-K$  is a maximally implausible principle for conflict-allowing contexts, and far more problematic than  $RM$ .

Now note that we have modal operators that allow for the expression of conditions sufficient for logical consistency. With that in mind, consider two weakened versions of  $C$  and  $K$ :

$$CO(p \ \& \ q) \rightarrow \cdot (OBp \ \& \ OBq) \rightarrow OB(p \ \& \ q) \quad [OB-C'],$$

$$CO(p \ \& \ q) \rightarrow \cdot OB(p \rightarrow q) \rightarrow (OBp \rightarrow OBq) \quad [OB-K']^{26}$$

Since we saw above that  $OB-C$  and  $OB-K$  are equivalent in logics with  $OB-RE$  &  $OB-M$ , we get:

**Corollary.** Given  $OB-RE$  and  $OB-M$ ,  $\vdash C' \leftrightarrow K'$ .

$OB-C'$  is plausible. The only reason we seem to have for blocking full-fledged  $OB-C$  is to prevent cases where a pair of things individually obligatory for me cannot be jointly met—they conflict. So if a pair of propositions are individually obligatory for me, and it is consistent with my abilities that the pair jointly occur, then there is no conflict involved in their joint occurrence, and so nothing to prevent its being obligatory for me that they jointly occur, or so it would seem. We will explore this qualified version of  $OB-C$ , and ask later if it is qualification enough.

We can validate  $OB-C'$  and  $OB-K'$  by either of the following equivalent conditions:

$$OB-c': \text{ If } X \cap Y \in CO_i \text{ \& } X, Y \text{ are in } OB_i \text{ then } X \cap Y \in OB_i,$$

$$OB-k': \text{ If } X \cap Y \in CO_i, \sim X \cup Y \in OB_i, \text{ and } X \in OB_i, \text{ then } Y \in OB_i,$$

where  $CO_i = \{X: X \cap \{j : CO_{ij}\} \neq \emptyset\}$ , the propositions consistent with our agent's abilities at  $i$ .

<sup>26</sup> Since  $\vdash [(p \rightarrow q) \ \& \ p] \leftrightarrow (p \ \& \ q)$ ,  $OB-K'$  is equivalent to  $CO[(p \rightarrow q) \ \& \ p] \rightarrow \cdot OB(p \rightarrow q) \rightarrow (OBp \rightarrow OBq)$ .

So I suggest that we first briefly consider the impact on AO of splicing the following basic classical systems for personal obligation to our agential-predetermination systems.

The basic system is OB-E:

- SL: All Tautologies  
 MP: If  $\vdash p$  and  $\vdash p \rightarrow q$  then  $\vdash q$   
 OB-RE: If  $\vdash p \leftrightarrow q$  then  $\vdash OBp \leftrightarrow OBq$ .

We can then turn to the impact on AO of the result of adding OB-D, to get system ED, and then OB-M to get the still stronger system, MED.<sup>27</sup> Let's see how these weaker conflict-allowing logics interact first with our logic for BA, especially with an eye on AO. Then we can go on to weave principles involving the CO and PR operators back in, tweaking the logics a bit when we do so in order to get more integrated PR-inclusive systems. Then we will briefly consider the possible role of a deontic constant like "d". Finally, we will return to OB-C'.

## 2.2. Agential obligation again

Imagine that we conjoin the E system for OB with our BA system, TECNOCS (enriching the language in obvious ways). Call the result E-TECNOCS:

- SL: All Tautologies  
 BA-T:  $BAp \rightarrow p$   
 BA-C:  $(BAp \ \& \ BAq) \rightarrow BA(p \ \& \ q)$   
 BA-CS:  $BA(p \ \& \ q) \rightarrow (\sim BAp \rightarrow BAq)$   
 MP: If  $\vdash p$  and  $\vdash p \rightarrow q$  then  $\vdash q$   
 BA-RE: If  $\vdash p \leftrightarrow q$  then  $\vdash BAp \leftrightarrow BAq$   
 OB-RE: If  $\vdash p \leftrightarrow q$  then  $\vdash OBp \leftrightarrow OBq$ .

Similarly, by adding OB-D to the above system, we get ED-TECNOCS, and by then adding OB-M to that system, we get MED-TECNOCS.

We define "AO" as before, but consider it tied to our new account of "OB":

$$AOp \underset{df}{=} OBBAp \quad [Df-OB].$$

Obviously the first sub-logic E, for OB is much weaker than SDL (and hence *Kd*). Using our prior work on AO in the *Kd*/SDL setting as a guide, what if anything is left over by way of theorems? Not much. OB-RE alone generates AO-RE, given Df-AO:

$$\text{If } \vdash p \leftrightarrow q \text{ then } \vdash AOp \leftrightarrow AOq \quad [AO-RE].$$

Now consider some results of adding OB-D to get ED-TECHNO:

$$\vdash \sim AO \perp. \quad [AO-D].$$

<sup>27</sup> Cf. MED to the weakest system for "O" proposed in [7, Chapter 6].

**Proof.**  $\vdash \sim BA \perp$ . So,  $\vdash BA \perp \leftrightarrow \perp$ . So by OB-RE,  $\vdash OBBA \perp \leftrightarrow OB \perp$ . But by OB-D,  $\vdash \sim OB \perp$ . So  $\vdash \sim OBBA \perp$ , i.e.,  $\vdash \sim AO \perp$ .  $\square$

$\vdash \sim AO \top$  [AO-NO].

$\vdash AO(p \ \& \ \top) \rightarrow (AOp \ \& \ \sim AO \top)$ .

Note that although OB-NO is not a theorem, nor is it valid (on the associated semantics), its agential analogue, AO-NO, is a theorem and is valid. Similarly for the last theorem.

Although adding M to ED-TECNOCS boosts its strength significantly, the result still falls far short of adding SDL to TECHNOCS. Much of what was derivable for AO in the fairly strong Andersonian–Kangerian framework is no longer derivable, nor valid. For example, these all fail:

$AO(p \rightarrow q) \rightarrow (AOp \rightarrow OBq)$  [Weak AO-K],

$AOp \rightarrow \sim OB \sim BAp$  [NC'-AO],

$AOp \rightarrow \sim AO \sim p$  [i.e.,  $OBBAp \rightarrow \sim OBBA \sim p$ ] [AO-NC].

Similarly, the following principles are not derivable and are invalid even though their OB-analogues are derivable and valid:

$AO(p \ \& \ q) \rightarrow (AOp \ \& \ AOq)$  [AO-M]

$AOp \rightarrow AO(p \vee q)$  [Ross Paradox]

If  $\vdash p \rightarrow q$  then  $\vdash AOp \rightarrow AOq$  [AO-RM].

Consider the latter. Suppose  $OBBAp$  holds at  $i$ . For any  $p$ ,  $\vdash p \rightarrow \top$ . Now  $BA \top$  fails at all worlds in all models, so  $\|BA \top\|^M = \emptyset$  in all models. But then by OB-d and the truth clause for “OB”,  $OBBA \top$  fails at all worlds in all models.

In fact, the only AO-pure principles we noted are those we saw in Part I other than AO-C and AO-CS, namely:

If  $\vdash p \leftrightarrow q$  then  $\vdash AOp \leftrightarrow AOq$  [AO-RE]

$\vdash \sim AO \perp$ . [AO-D]

$\vdash \sim AO \top$  [AO-NO]

$\vdash AO(p \ \& \ \top) \rightarrow (AOp \ \& \ \sim AO \top)$ .

The latter is easily derivable from the first two, so it looks like the pure AO-system we get might be DE-NO, a sublogic of DEC-NO discussed in our earlier section, “The Pure Agential Obligation Fragment of TECHNOCS-KTd”. The only missing pure AO-principle from that system is AO-C. A minimal model semantics and conjectures like those in the earlier section are easily adapted. Were we to add AO-CS (recall TECHNOCS-KTd+),

$AO(p \ \& \ q) \rightarrow \cdot \sim AOp \rightarrow AOq$ ,

given the new semantics for OB, we would need to validate it with this constraint:

If  $BA(X \cap Y) \in OB_i$  then either  $BAX \in OB_i$  or  $BAY \in OB_i$ .

As before, the following are still easily derived:

$$\begin{aligned} &\vdash \text{AO}(p \ \& \ q) \rightarrow (\text{AO}p \vee \text{AO}q) \quad [\text{AO-CS}'], \\ &\vdash \text{AO}p \rightarrow \cdot \sim \text{AO}(p \vee q) \rightarrow \text{AO}(p \vee \sim q) \quad [\text{AO-CS}'']. \end{aligned}$$

Set AO-CS aside henceforth.

What positive links are there between OB and AO in the setting of MED-TECNOCS? Well, that something is obligatory for me does not imply that *I* am obligated to bring it about, but the converse, that what is obligatory for *me* to bring about is obligatory for me is provable:

$$\vdash \text{AO}p \rightarrow \text{OB}p \quad [\text{AO-OB}].$$

Although AO-M is invalid, OB-M does generate a weak AO-M analogue:

$$\vdash \text{OB}(\text{BA}p \ \& \ \text{BA}q) \rightarrow (\text{AO}p \ \& \ \text{AO}q) \quad [\text{Weak AO-M}].$$

Let's note one important principle:

$$\vdash \text{OB}(\text{BA}p \ \& \ \text{BA}q) \rightarrow \text{AO}(p \ \& \ q) \quad [\text{OB-AO C}].$$

Here are a few more derivable principles:

$$\begin{aligned} &\text{If } \vdash p \rightarrow q \text{ then } \vdash \text{AO}p \rightarrow \text{OB}q \quad [\text{Weak AO-RM}] \\ &\vdash \text{AO}p \rightarrow \text{OB}T \quad (\text{From AO-OB and OB-RM}) \\ &\text{If } \vdash q \text{ then } \vdash \text{AO}(p \ \& \ q) \rightarrow (\text{AO}p \ \& \ \text{OB}\sim\text{BA}q) \\ &\vdash \text{AO}p \rightarrow \text{OB}\sim\text{BA}\sim p \quad (\text{From BA-NC and OB-RM}). \end{aligned}$$

We also get some cousins of K:

$$\begin{aligned} &\vdash \text{OB}(\text{BA}(p \rightarrow q) \ \& \ \text{BA}p) \rightarrow \text{OB}q. \\ &\vdash \text{OB}(\text{BA}(p \rightarrow q) \ \& \ \text{BA}p) \rightarrow \cdot \text{AO}(p \rightarrow q) \ \& \ \text{AO}p \ \& \ \text{OB}q. \end{aligned}$$

Also, we have:

$$\vdash \text{AO}(p \rightarrow q) \rightarrow \text{OB}(\text{BA}p \rightarrow q).$$

Now we need to explore possible links between PR, our notion of predetermination, and OB and AO. If we just graft on our KT system for PR to MED-TECNOCS, we get MED-TECNOCS-KT:

$$\begin{aligned} \text{SL:} & \quad \text{All Tautologies} \\ \text{BA-T:} & \quad \text{BA}p \rightarrow p \\ \text{BA-C:} & \quad (\text{BA}p \ \& \ \text{BA}q) \rightarrow \text{BA}(p \ \& \ q) \\ \text{BA-CS:} & \quad \text{BA}(p \ \& \ q) \rightarrow (\sim \text{BA}p \rightarrow \text{BA}q) \\ \text{PR-T:} & \quad \text{PR}p \rightarrow p \\ \text{PR-K:} & \quad \text{PR}(p \rightarrow q) \rightarrow (\text{PR}p \rightarrow \text{PR}q) \\ \text{OB-D:} & \quad \sim \text{OB}\perp \end{aligned}$$



- OB-M:  $OB(p \ \& \ q) \rightarrow (OBp \ \& \ OBq)$   
 MP: If  $\vdash p$  and  $\vdash p \rightarrow q$  then  $\vdash q$   
 BA-RE: If  $\vdash p \leftrightarrow q$  then  $\vdash BAp \leftrightarrow BAq$   
 PR-NEC: If  $\vdash p$  then  $\vdash PRq$   
 OB-RE: If  $\vdash p \leftrightarrow q$  then  $\vdash OBp \leftrightarrow OBq$ .

These prior principles of the strong  $Kd$  system no longer hold:

- $PRp \rightarrow OBp$   
 $OBp \rightarrow (PRq \rightarrow OBq)$ .  
 $PR(p \rightarrow q) \rightarrow (AOp \rightarrow OBq)$   
 $PRBAp \rightarrow AOp$   
 $AO(p \rightarrow q) \rightarrow \cdot ARq \rightarrow (AOp \rightarrow AOq)$   
 If  $\vdash p \rightarrow PRp$  then  $\vdash BAp \rightarrow AOp$ .

Here are two other prior theorems:

- $OBp \rightarrow COp$  [KL (for “Kant’s Law”)],  
 $AOp \rightarrow ABp$ .

The latter is just a special case of the former. I think these are contentious principles, since an at least plausible argument that they are false can be made. Raising doubts about the latter principle will suffice. Suppose I have promised you to be at a meeting at noon, and thereby acquire an agential obligation to attend. Now add that my car breaks down on the way, rendering me unable to make the meeting. In that case, it seems that my obligation to be at the meeting at noon is unfulfillable. For example, can’t I say truly that “I *have* to be at the meeting at noon, but there is just no way I can make it.”? Consider debts. Many people have debts they are unable to pay. Surely some of these cases are cases where the person has an obligation to pay some money to someone that they can’t pay. So to accept  $AOp \rightarrow ABp$  (or  $OBp \rightarrow COp$ ) appears to amount to denying that there can be *unfulfillable obligations*. But the adjective appears to go with the noun seamlessly enough. No contradiction is apparent. How can this be so if the above principles are true? These principles are at least contentious, and they are neither valid nor derivable. We will return to them later. To validate them we would need to add the following semantic constraints, respectively (though the first entails the second),

- If  $X \in OB_i$  then  $X \in CO_i$ ,  
 If  $BAX \in OB_i$  then  $BAX \in CO_i$ .

Are there any positive links? The following do hold:

- $\vdash AOp \rightarrow OBCOp$ ,  
 $\vdash AOp \rightarrow OBABp$ .

Conjoined with the two prior non-validities, the position is that although being obligated to bring about  $p$  does not entail that  $p$  is consistent with, much less within my abilities, it does entail that it is obligatory that  $p$  is both consistent with and within my abilities.

At this point, let's add in OB-C',

$$\vdash \text{CO}(p \ \& \ q) \rightarrow \cdot (\text{OB}p \ \& \ \text{OB}q) \rightarrow \text{OB}(p \ \& \ q) \quad [\text{OB-C}'],$$

to get MEDC'-TECHNO-KT.

We have already noted that with OB-RE & OB-M, we can derive OB-K':

$$\vdash \text{CO}(p \ \& \ q) \rightarrow \cdot \text{OB}(p \rightarrow q) \rightarrow (\text{OB}p \rightarrow \text{OB}q) \quad [\text{OB-K}'].$$

These yield the following AO-analogues:

$$\vdash \text{CO}(\text{BA}p \ \& \ \text{BA}q) \rightarrow \cdot (\text{AO}p \ \& \ \text{AO}q) \rightarrow \text{OB}(\text{BA}p \ \& \ \text{BA}q) \quad [\text{Weak AO-C}'],$$

$$\vdash \text{CO}(\text{BA}p \ \& \ \text{BA}q) \rightarrow \cdot \text{OB}(\text{BA}p \rightarrow \text{BA}q) \rightarrow (\text{AO}p \rightarrow \text{AO}q) \quad [\text{Weak OB-K}'].$$

Notice the latter and recall Horty's "fight or do alternative service" example. First, we assume tacitly that it is consistent with the agent's abilities that he both brings it about that he does not fight and brings it about that he does alternative service. Then, relying on this, we infer that he must perform alternative service. For (1) it is obligatory for him to be such that if he brings about his not fighting, then he performs alternative service, and (2) it is obligatory for him to be such that he brings it about that he doesn't fight. Notice also that it is less plausible to say that the agent is *obligated to bring it about* that (as opposed to being *obligated to be such that*) either he will fight or he will do alternative service, for that may already be a deeply settled part of his character. Again, it looks like we are obligated to not only do things, but to be certain ways, and the latter do not reduce to the former.

We also get the following from Weak AO-C' and Weak AO-M:

$$\vdash \text{CO}(\text{BA}p \ \& \ \text{BA}q) \rightarrow \cdot \text{OB}(\text{BA}p \ \& \ \text{BA}q) \leftrightarrow (\text{AO}p \ \& \ \text{AO}q) \quad [\text{Weak AO-R}'].$$

Similarly, we get an important AO-analogue to OB-C':

$$\vdash \text{CO}(\text{BA}p \ \& \ \text{BA}q) \rightarrow \cdot (\text{AO}p \ \& \ \text{AO}q) \rightarrow \text{AO}(p \ \& \ q) \quad [\text{AO-C}'].$$

**Proof.** Assume  $\text{CO}(\text{BA}p \ \& \ \text{BA}q)$ . By Weak AO-C', we get  $(\text{AO}p \ \& \ \text{AO}q) \rightarrow \text{OB}(\text{BA}p \ \& \ \text{BA}q)$ . But by OB-AO C,  $\vdash \text{OB}(\text{BA}p \ \& \ \text{BA}q) \rightarrow \text{AO}(p \ \& \ q)$ . So we get  $(\text{AO}p \ \& \ \text{AO}q) \rightarrow \text{AO}(p \ \& \ q)$  from our assumption.  $\square$

However this is neither derivable nor valid:

$$\text{CO}(\text{BA}p \ \& \ \text{BA}q) \rightarrow \cdot \text{AO}(p \ \& \ q) \rightarrow (\text{AO}p \ \& \ \text{AO}q).$$

Suppose it is consistent with my abilities that I both bring it about that p and bring it about that q. Now add that if I don't bring about q, you will. It may then be obligatory for me to bring about p *after* you have settled q instead of me. I will then be obligated to bring it about that the conjunction of p and q holds by bringing about the remaining conjunct, p, but it need not be obligatory for me to bring it about that q.

Other than OB-C', we have no real bridging principles linking PR and OB. So let's briefly consider adding that inevitably equivalent propositions are jointly obligatory if at all:

$$\text{PR}(p \leftrightarrow q) \rightarrow (\text{OB}p \leftrightarrow \text{OB}q) \quad [\text{OB-E}'].$$

With OB- $E'$ , we no longer need OB-RE as a basic principle. So let's consider this system, ME'DC'-TECNOCS-KT, briefly.

SL:	All Tautologies
BA-T:	$BAp \rightarrow p$
BA-C:	$(BAp \ \& \ BAq) \rightarrow BA(p \ \& \ q)$
BA-CS:	$BA(p \ \& \ q) \rightarrow (\sim BAp \rightarrow BAq)$
PR-T:	$PRp \rightarrow p$
PR-K:	$PR(p \rightarrow q) \rightarrow (PRp \rightarrow PRq)$
OB-D:	$\sim OB \perp$
OB-M:	$OB(p \ \& \ q) \rightarrow (OBp \ \& \ OBq)$
OB-C':	$CO(p \ \& \ q) \rightarrow \cdot (OBp \ \& \ OBq) \rightarrow OB(p \ \& \ q)$
OB-E':	$PR(p \leftrightarrow q) \rightarrow (OBp \leftrightarrow OBq)$
MP:	If $\vdash p$ and $\vdash p \rightarrow q$ then $\vdash q$
BA-RE:	If $\vdash p \leftrightarrow q$ then $\vdash BAp \leftrightarrow BAq$
PR-NEC:	If $\vdash p$ then $\vdash PRq$ .

The increased strength of  $E'$  over OB-RE is impactful when we consider the analogues to the earlier OM/AR qualified principles. Recall that, by BA-T and PR-RE,  $OMP$  (i.e.,  $PR(p \rightarrow BAp)$ ) entails  $PR(p \leftrightarrow BAp)$ . Now consider the first such important analogue:

$$\vdash OMp \rightarrow (OBp \leftrightarrow AOp) \quad [OM\text{-Qualified OB-AO}].$$

**Proof.** Assume  $OMP$ . So by BA-T and PR-NEC,  $PR(p \leftrightarrow BAp)$ . So by  $E'$ ,  $OBp \leftrightarrow OBBAp$ .  $\square$

**Corollary.**  $\vdash ARp \rightarrow (OBp \leftrightarrow AOp)$ .

In other words, under circumstances where it is predetermined that something will be the case only if I bring it about, then that thing is personally obligatory for me iff it is agentially obligatory for me. Consider some analogues to the other earlier  $Kd$ -based theorems involving OM/AR:

$$\vdash OMQ \rightarrow (PR(p \rightarrow q) \rightarrow (AOp \rightarrow AOq)) \quad [OM\text{-Qualified RM}'].$$

$$\text{If } \vdash p \rightarrow q \text{ then } \vdash OMQ \rightarrow (AOp \rightarrow AOq) \quad [OM\text{-Qualified AO-RM}].$$

**Corollary.** If  $\vdash p \rightarrow q$  then  $\vdash ARq \rightarrow (AOp \rightarrow AOq)$ .

Although AO-M is invalid, an OM-qualified version is valid:

$$\vdash (OMp \ \& \ OMQ) \rightarrow \cdot AO(p \ \& \ q) \rightarrow (AOp \ \& \ AOq) \quad [OM\text{-Qualified AO-M}]$$

**Corollary.**  $\vdash (ARp \ \& \ ARq) \rightarrow (AO(p \ \& \ q) \rightarrow (AOp \ \& \ AOq))$ .

Let us consider one further, still stronger, bridge principle, that the inevitable consequences of obligatory things are obligatory,

$$\text{PR}(p \rightarrow q) \rightarrow (\text{OB}p \rightarrow \text{OB}q) \quad [\text{OB-RM}'].$$

Notice that the following is then easily derivable:

$$\text{OB}p \ \& \ \sim\text{CO}p \ \cdot \rightarrow \text{OB}q.^{28}$$

So if we add OB-RM', we need to add our previously discussed

$$\text{OB}p \rightarrow \text{CO}p \quad [\text{KL}]$$

in tandem. We assume that the principle what is necessitated for me by what is obligatory is obligatory only makes sense in contexts where what is obligatory is consistent with my abilities.

Call the resulting system KLC'RM'–TECNOCS-KT:

SL:	All Tautologies
BA-T:	BAp $\rightarrow$ p
BA-C:	(BAp & BAq) $\rightarrow$ BA(p & q)
BA-CS:	BA(p & q) $\rightarrow$ ( $\sim$ BAp $\rightarrow$ BAq)
PR-T:	PRp $\rightarrow$ p
PR-K:	PR(p $\rightarrow$ q) $\rightarrow$ (PRp $\rightarrow$ PRq)
KL:	OBp $\rightarrow$ COp
OB-C':	CO(p & q) $\rightarrow$ $\cdot$ (OBp & OBq) $\rightarrow$ OB(p & q)
OB-RM':	PR(p $\rightarrow$ q) $\rightarrow$ (OBp $\rightarrow$ OBq)
MP:	If $\vdash$ p and $\vdash$ p $\rightarrow$ q then $\vdash$ q
BA-RE:	If $\vdash$ p $\leftrightarrow$ q then $\vdash$ BAp $\leftrightarrow$ BAq
PR-NEC:	If $\vdash$ p then $\vdash$ PRq.

First note that the following are now all derivable:

$\vdash$ $\sim$ OB $\perp$	[OB-D]
$\vdash$ AOp $\rightarrow$ ABp	
$\vdash$ PR(p $\leftrightarrow$ q) $\rightarrow$ (OBp $\leftrightarrow$ OBq)	[E']
If $\vdash$ p $\leftrightarrow$ q then OBp $\leftrightarrow$ OBq	[OB-RE]
If $\vdash$ p $\rightarrow$ q then $\vdash$ OBp $\rightarrow$ OBq	[OB-RM]
$\vdash$ OB(p & q) $\rightarrow$ (OBp & OBq)	[OB-M].

We also get these:

$\vdash$ PR(p $\rightarrow$ q) $\rightarrow$ (AOp $\rightarrow$ OBq),
$\vdash$ PR(BAp $\rightarrow$ BAq) $\rightarrow$ (AOp $\rightarrow$ AOq).

<sup>28</sup> This no doubt reflects one of the difficulties in facing unfulfillable obligations.

These have at least some attraction. Consider the last principle. Suppose I am obligated to get myself to the meeting, and I can only do so if I also get myself to my car. Then I am obligated to get myself to my car.

### 2.3. Andersonian–Kangerian constants again and some glimpses ahead

Although we had to reject the Andersonian–Kangerian reduction in order to model conflicting obligations, here we briefly explore reintroducing  $d$  into the language with its usual informal reading:

$d$ : The demands on John Doe are all met (or “John Doe’s responsibilities are all met”).

Having this in our language facilitates being able to talk about situations where all our obligations are met or can be met, and so situations where there are no conflicts of obligation at all. But since we have introduced an independent conflict-allowing notion of obligation,  $d$  can now be thought of as a proposition asserting that all our obligations, *as already independently construed*, have been met. Indeed, semantically, the classic dependence we see in the Andersonian–Kangerian reduction will be reversed:  $d$ ’s truth-condition will be defined via our now independent notion of Doe’s obligations:

$$M \models_i d \text{ iff } i \in \bigcap OB_i, \text{ i.e., } \forall X(X \in OB_i \rightarrow i \in X) \quad [d].$$

This says  $d$  is true at a world iff all of our agent’s obligations at that world are jointly met there. Plainly,  $d$  can’t be true where our obligations logically conflict, since that would require a world where an inconsistent set of propositions holds. Consider the following fundamental valid formula:

$$d \rightarrow (OBp \rightarrow p) \quad [dOB \text{ Truth}].$$

This says that if all of the demands on Doe are satisfied, then any particular obligation he is under is satisfied.

Couched this way, our Andersonian–Kangerian constant can now be used in a *restorative* rather than reductive way. We presumably want something at least as strong as:

$$\text{If } \|d\|^M \in OB_i \text{ then } \bigcap OB_i \neq \emptyset.$$

This says that if the proposition that *all the demands on John Doe are met* is itself one of the obligatory things for John Doe at  $i$  then the propositions obligatory for John Doe at  $i$  are jointly logically consistent. This yields the following valid rule:

$$\text{If } \vdash p \rightarrow \sim q \text{ then } \vdash OBd \rightarrow (OBp \rightarrow \sim OBq).$$

This in turn entails as a special case:

$$\vdash OBd \rightarrow (OBp \rightarrow \sim OB\sim p) \quad [dOB\text{-NC}],$$

thus restoring our earlier no-conflicts principle in a qualified form.

Relatedly, the following stronger formula sounds plausible:

$$COd \rightarrow OBd.$$

It asserts that if it is merely consistent with my abilities that all the demands on me are met, then one of the obligatory things for me is that all the demands on me are met. This formula is validated by the following constraint:

$$\text{If } \|d\|^M \in CO_i \text{ then } \|d\|^M \in OB_i.^{29}$$

This says that if at some world consistent with my abilities here, all the demands made on me are met there, then meeting all the demands made on me is obligatory here.

This leads naturally to thinking about  $d$  in the context of intuitions about the transfer or “traveling” of our obligations across the situations consistent with our abilities.<sup>30</sup>

Consider the claim that what is obligatory is settled obligatory:

$$OB_p \rightarrow PROBP.$$

This is validated by

$$\forall j(CO_{ij} \rightarrow OB_i \subseteq OB_j) \quad [CO-OB \text{ Export}].$$

*CO-OB* Export says that my obligations at this world “travel” (perhaps with additions) to all worlds consistent with my abilities here. Intuitively it amounts to saying that it is not consistent with my abilities that my obligations (now) contract. This would validate

$$COd \rightarrow \cdot (OB_p \ \& \ OB_q) \rightarrow CO(p \ \& \ q)$$

as well. For suppose  $COd$ ,  $OB_p$  and  $OB_q$  hold at  $i$ . Then for some  $j$ , such that  $CO_{ij}$ ,  $d$  holds at  $j$ . But then by *OB-CO* Export,  $OB_p$  and  $OB_q$  must hold at that  $j$  as well. But then by  $d$ 's truth-conditions,  $p$  and  $q$  must each hold at  $j$ . So their conjunction holds there, and hence  $CO(p \ \& \ q)$  must hold back at  $i$ .

Also, with the above formula, from  $OB - C'$ ,  $CO(p \ \& \ q) \rightarrow \cdot (OB_p \ \& \ OB_q) \rightarrow OB(p \ \& \ q)$ , we get a new qualified version on *OB-C*:

$$COd \rightarrow \cdot (OB_p \ \& \ OB_q) \rightarrow OB(p \ \& \ q),$$

along with a semantic analog to *OB-c'*, if  $X \cap Y \in CO_i$  &  $X, Y$  are in  $OB_i$  then  $X \cap Y \in OB_i$ , namely,

$$\text{If } \|d\|^M \in CO_i \ \& \ X, Y \in OB_i \text{ then } X \cap Y \in OB_i.^{31}$$

Before closing with a brief discussion of a problem, let me note some avenues for further exploration. We could easily introduce another deontic constant,  $d'$ , to model “all Doe’s *agential* obligations are met” (by him, of course). Then, for example, we would want to validate  $d \rightarrow d'$ , but not vice versa. Similarly, we could impose an *importance* ordering on the propositions that constitute Doe’s responsibilities,<sup>32</sup> and define the notion of a strict obligation, an overridden obligation, etc., along with additional deontic constants. We could then

<sup>29</sup> Recall that  $CO_i = \{X: X \cap \{j: CO_{ij}\} \neq \emptyset\}$ , the ability-consistent propositions.

<sup>30</sup> See the related discussion of the *CO* relation and the definition of ability in [21] for the need to address related issues.

<sup>31</sup> This is a consequence of *OB-m*, and the more recent constraint, if  $\|d\|^M \in CO_i$  then  $\|d\|^M \in OB_i$ .

<sup>32</sup> The resources in [21] might adapt well for this.

express claims like “all the strict (most important) demands on Doe can be met, but not all demands on him can be met” and “if all of Doe’s demands can be satisfied, then none of his obligations are overridden”. Clearly, the generalization to *multi-agent contexts* would also be of interest, so that we could model Carmo’s slippage problem explicitly, and show how the current strategy escapes it. Finally, with the introduction of predicates and names, we could represent obligations of the form: S is obligated to be F (e.g., to be in Boston or to be cooperative). But we must also face a problem requiring further qualification on principles of aggregation, using a more fine-grained approach to obligations.

#### 2.4. Refining aggregation in conflict-allowing contexts

Recall our qualification of OB-C (simple aggregation):<sup>33</sup>

$$\text{OB-C}' : \text{CO}(p \ \& \ q) \rightarrow \cdot (\text{OB}p \ \& \ \text{OB}q) \rightarrow \text{OB}(p \ \& \ q)$$

We indicated that we would ask later if OB-C' is qualification enough of full-fledged aggregation. It is not.

*Hansen’s objection.* Suppose  $\text{PR}(p \rightarrow p')$ ,  $\text{PR}(q \rightarrow q')$ , and  $\sim\text{CO}(p \ \& \ q)$ , but  $\text{CO}(p' \ \& \ q')$ . Then by OB-RM', from  $\text{OB}p \ \& \ \text{OB}q$ , we have  $\text{OB}p' \ \& \ \text{OB}q'$ . But then by OB-C', we get  $\text{OB}(p' \ \& \ q')$ .<sup>34</sup>

For example, let

- p: I keep an appointment this morning in Montreal.
- p': I travel to Montreal this morning.
- q: I keep an appointment in London this afternoon.
- q': I travel to London this morning.

Imagine that although keeping the appointments in Montreal and in London respectively necessitate traveling to Montreal this morning and traveling to London this morning, and although keeping both appointments is not open to me, nonetheless traveling to both places this morning is open to me (e.g., I could drive directly to a Montreal airport and fly to London late this morning). However, since the traveling obligations derive respectively and exclusively from two obligations that conflict, despite the joint realizability of these two derivative traveling obligations, no singular conjunctive obligation to travel to both places actually follows from them. Notice that the problem can’t be easily solved by restricting OB-RM', since in the example above, p' and q' express contingent truths associated with actions in my power that are practical prerequisites to p and to q respectively. Even the most reasonably restricted version of OB-RM' must allow such inferences to go through. So it appears that OB-C' is at fault. Though a significant improvement over OB-C, it nonetheless allows us to indirectly derive pointless conjunctive obligations from obligations that conflict. A faithful representation of conflicting obligations must disallow such pointless derivative obligations.

<sup>33</sup> OB-C is  $(\text{OB}p \ \& \ \text{OB}q) \rightarrow \text{OB}(p \ \& \ q)$ .

<sup>34</sup> Jorg Hansen brought this objection to my attention at DEON'02. The basic idea for the solution proposed here occurred to me there, but I note that it is similar to that arowed by Hansen himself in Section 5 of [12].

Here I give a rapid impressionistic sketch of a solution to be developed elsewhere. I ignore agency.

*Intuition* In conflict-allowing contexts, we need to restrict the application of deontic aggregation to mutually consistent basic obligations and deontic consequences of *unproblematic pedigree*. To do so, it will help to distinguish basic from non-basic (derivative) obligations, and to distinguish unary-sourced deontic consequences from problematic/unproblematic multi-sourced deontic consequences.

Assume we have the earlier system KT for PR and CO. Suppose also that we have propositional quantifiers in our language. Now add one undefined operator, BO, where BOp means p is a *basic obligation* (as opposed to a derived one). We assume that BO satisfies just these principles:

$$\text{BO-CO: } \text{BOp} \rightarrow \text{COp}^{35}$$

$$\text{BO-RE: } \text{If } \vdash p \leftrightarrow q \text{ then } \vdash \text{BOp} \leftrightarrow \text{BOq}^{36}$$

We use minimal models again

$$\text{BO: } W \rightarrow \text{Pow}(\text{Pow}(W)) \text{ i.e. } \text{BO}_i \subseteq \text{Pow}(W),$$

$$M \models_i \text{BOp} \text{ iff } \|p\|^M \in \text{BO}_i.$$

For BO-CO, we stipulate:

$$\text{If } X \in \text{BO}_i \text{ then } \exists j(\text{CO}_{ij} \ \& \ j \in X).$$

Although not necessary, we assume that the number of *basic* obligations is finite:

$$\forall i \in W, \exists X_1 \dots X_n \text{ such that } \text{BO}_i = \{X_1, \dots, X_n\}, \text{ where } n \geq 0.$$

These humble beginnings are deceptively powerful, as we will try to quickly illustrate.

Let p be *obligatory* iff there is an ability-consistent (finite) set of basic obligations that necessitate p:

$$\text{OBp} \stackrel{\text{df}}{=} \exists q_1 \dots q_n [(\text{BO}_{q_1} \ \& \ \dots \ \& \ \text{BO}_{q_n}) \ \& \ \text{CO}(q_1 \ \& \ \dots \ \& \ q_n) \\ \& \ \text{PR}((q_1 \ \& \ \dots \ \& \ q_n) \rightarrow p)].$$

The clause, CO( $q_1 \ \& \ \dots \ \& \ q_n$ ), is necessary, else any conflicting obligations would render everything obligatory. Note some theorems:

$$\text{BOp} \rightarrow \text{OBp}$$

$$\text{OBp} \rightarrow \text{COp}$$

$$\text{PR}(p \rightarrow q) \rightarrow (\text{BOp} \rightarrow \text{OBq})$$

$$\text{PR}(p \rightarrow q) \rightarrow (\text{OBp} \rightarrow \text{OBq}).$$

<sup>35</sup> This axiom need not be basic, but it facilitates a fast sketch.

<sup>36</sup> Not including BO-M is defensible, but that defense must wait.



Where  $(BO_{q_1} \& \dots \& BO_{q_n}) \& CO(q_1 \& \dots \& q_n) \& PR((q_1 \& \dots \& q_n) \rightarrow p)$  we will say that “ $(q_1 \& \dots \& q_n)$  is a *basis for OBp*”.

We note in passing that the increased expressive resources are significant. To illustrate:

$$\begin{aligned} SOP &\stackrel{\text{df}}{=} \exists q(BO_q \& PR(q \rightarrow p)) && \text{[Singly Obligatory]} \\ JOp &\stackrel{\text{df}}{=} OBp \& \sim SOP && \text{[Jointly Obligatory]} \\ IOp &\stackrel{\text{df}}{=} OBp \& \sim BOp && \text{[Indirectly Obligatory].} \end{aligned}$$

Here are some derivable relationships:

$$\begin{aligned} BOp &\rightarrow SOP && JOp &\rightarrow IOp \\ SOP &\rightarrow \sim JOp && IOp &\rightarrow OBp \\ BOp &\rightarrow \sim JOp && BOp &\rightarrow \sim IOp \\ OBP &\leftrightarrow (SOp \vee JOp) && OBp &\leftrightarrow (BOp \vee IOp). \end{aligned}$$

Let us now return to Hansen’s problem. Suppose I am under exactly two basic obligations,  $BOp$  &  $BOq$ .<sup>37</sup> Each is ability-consistent with itself, and the first necessitates  $p'$ , and the second necessitates  $q'$ . So  $OBp'$  and  $OBq'$ . However it does not follow that  $OB(p' \& q')$ . For there is no finite set of ability-consistent *basic* obligations that necessitates  $(p' \& q')$ .

Note that the following aggregation principle for *basic* obligations is directly derivable from our new definition of OB (and our KT logic for PR):

$$BO-C': \quad CO(p_1 \& \dots \& p_n) \cdot \rightarrow (BO_{p_1} \& \dots \& BO_{p_n}) \rightarrow OB(p_1 \& \dots \& p_n).^{38}$$

But a crucial question is how does aggregation work in other cases, and how might we define the notion of a (non-basic) obligation of aggregation-unproblematic pedigree?

First define permissibility:

$$PEp \stackrel{\text{df}}{=} \sim OB \sim p,$$

which entails:

$$PEp \leftrightarrow \forall q_1 \dots q_n [BO_{q_1} \& \dots \& BO_{q_n} \& CO(q_1 \& \dots \& q_n) \cdot \rightarrow CO(p \& q_1 \& \dots \& q_n)].$$

Clearly, not all things that are obligatory are permissible. This is a theorem:

$$OBp \& OBq \& \sim CO(p \& q) \rightarrow (\sim PEp \& \sim PEq).$$

Now what of *permissible obligations*,

$$POp \stackrel{\text{df}}{=} OBp \& PEp,$$

<sup>37</sup> This is sloppy, since BO-RE gives us infinitely many equivalent formulas, but it is convenient, since these represent just two obligations.

<sup>38</sup> This is just a generalized BO-analog of our earlier OB-C', but the restriction to BO is crucial.

obligations that are consistent with any jointly consistent basic obligations? Are those fully unproblematic for aggregation? Let's consider this proposal for a moment:

$p$  is an aggregation unproblematic obligation iff POP.

This doesn't work since this fails:

$$\text{POp} \ \& \ \text{POq} \ \cdot \rightarrow \ \text{PO}(p \ \& \ q).$$

For example, let my basic obligations be just these (and their equivalents): BO $p$ , BO $q$ , BO( $\sim p \vee \sim q$ ). So OB $p$ , OB $q$ . Assume that CO $p$ , CO $q$ , CO( $\sim p \vee \sim q$ ), CO( $p \ \& \ q$ ), CO( $p \ \& \ \sim q$ ), and CO( $q \ \& \ \sim p$ ). So PE $p$  & PE $q$ , since by exhaustion of cases,  $p$  and  $q$  are each compatible with any jointly consistent basic obligations. Hence, POP & PO $q$ . What of PO( $p \ \& \ q$ )? Consider OB( $p \ \& \ q$ ) first. Since BO $p$  & BO $q$  & CO( $p \ \& \ q$ ) and PR( $(p \ \& \ q) \rightarrow (p \ \& \ q)$ ), OB( $p \ \& \ q$ ). But consider PE( $p \ \& \ q$ ):  $\forall q_1 \dots q_n [(BO_{q_1} \ \& \ \dots \ \& \ BO_{q_n}) \ \& \ CO(q_1 \ \& \ \dots \ \& \ q_n) \rightarrow CO(p \ \& \ q \ \& \ q_1 \ \& \ \dots \ \& \ q_n)]$ . Does this hold? No. For BO( $\sim p \vee \sim q$ ) & CO( $\sim p \vee \sim q$ ), yet clearly  $\sim CO(p \ \& \ q \ \& \ \sim p \vee \sim q)$ . So  $\sim PO(p \ \& \ q)$ .

Instead, let us define an *unproblematic obligation* not as a permissible one, but as one with a permissible basis:

$$\begin{aligned} \text{UOp} \equiv \exists p_1 \dots p_n [ & (\text{BOp}_1 \ \& \ \dots \ \& \ \text{BOp}_n) \ \& \ \text{CO}(p_1 \ \& \ \dots \ \& \ p_n) \ \& \\ \text{PR}((p_1 \ \& \ \dots \ \& \ p_n) \rightarrow p) \ \& \ \text{PE}(p_1 \ \& \ \dots \ \& \ p_n)]. \end{aligned}$$

Fully ticketed, the definiens is:

$$\begin{aligned} \exists p_1 \dots p_n [ & (\text{BOp}_1 \ \& \ \dots \ \& \ \text{BOp}_n) \ \& \ \text{CO}(p_1 \ \& \ \dots \ \& \ p_n) \ \& \\ & \text{PR}((p_1 \ \& \ \dots \ \& \ p_n) \rightarrow p)] \ \& \\ \forall q_1 \dots q_n [ & (\text{BO}_{q_1} \ \& \ \dots \ \& \ \text{BO}_{q_n}) \ \& \ \text{CO}(q_1 \ \& \ \dots \ \& \ q_n) \\ & \cdot \rightarrow \ \text{CO}(p_1 \ \& \ \dots \ \& \ p_n \ \& \ q_1 \ \& \ \dots \ \& \ q_n)]. \end{aligned}$$

These are easily derived:

$$\text{UOp} \rightarrow \text{OBp}$$

$$\text{UOp} \rightarrow \text{PEp}.$$

Our new definition is very close to the one just rejected, but the crucial difference is that with OB $p$  & PE $p$ , no link is made between *the basis for* OB $p$  and permissibility. This difference is crucial, as the following is now provable

$$\text{UO-C':} \quad \text{UOp} \ \& \ \text{UOq} \rightarrow \ \text{UO}(p \ \& \ q).$$

**Proof.** We assume (1) UOp and (2) UOq. We need to show UO( $p \ \& \ q$ ), i.e.,  $\exists r_1 \dots r_n [(\text{BOR}_1 \ \& \ \dots \ \& \ \text{BOR}_n) \ \& \ \text{CO}(r_1 \ \& \ \dots \ \& \ r_n) \ \& \ \text{PR}((r_1 \ \& \ \dots \ \& \ r_n) \rightarrow (p \ \& \ q))] \ \& \ [\text{PE}(r_1 \ \& \ \dots \ \& \ r_n)]$ .

Since UOp and UOq, there is a permissible basis for each:

- (1)  $\exists p_1 \dots p_n [(\text{BOp}_1 \ \& \ \dots \ \& \ \text{BOp}_n) \ \& \ \text{CO}(p_1 \ \& \ \dots \ \& \ p_n) \ \& \ \text{PR}((p_1 \ \& \ \dots \ \& \ p_n) \rightarrow p) \ \& \ \text{PE}(p_1 \ \& \ \dots \ \& \ p_n)]$  and

$$(2) \exists q_1 \dots q_n [(BOq_1 \& \dots \& BOq_n) \& CO(q_1 \& \dots \& q_n) \& PR((q_1 \& \dots \& q_n) \rightarrow q) \& PE(q_1 \& \dots \& q_n)].$$

Fixing  $p_1 \dots p_n$ , and  $q_1 \dots q_n$ , we get:

$$(1a) (BOP_1 \& \dots \& BOP_n) \& CO(p_1 \& \dots \& p_n) \& PR((p_1 \& \dots \& p_n) \rightarrow p) \&$$

$$(1b) PE(p_1 \& \dots \& p_n) \text{ and}$$

$$(2a) (BOq_1 \& \dots \& BOq_n) \& CO(q_1 \& \dots \& q_n) \& PR((q_1 \& \dots \& q_n) \rightarrow q) \&$$

$$(2b) PE(q_1 \& \dots \& q_n).$$

We will show  $OU(p \& q)$  by showing an instance of its existential definiens, namely an instance of the form:

$$(i) (BOR_1 \& \dots \& BOR_n) \& CO(r_1 \& \dots \& r_n) \& PR((r_1 \& \dots \& r_n) \rightarrow (p \& q))] \&$$

$$(ii) PE(r_1 \& \dots \& r_n)].$$

(i) First, from (2a),  $(BOq_1 \& \dots \& BOq_n) \& CO(q_1 \& \dots \& q_n)$ , and from (1b),  $PE(p_1 \& \dots \& p_n)$ . From these, it follows that  $CO(p_1 \& \dots \& p_n \& q_1 \& \dots \& q_n)$ . But then given  $BOP_1 \& \dots \& BOP_n$  from (1a), we get  $(BOP_1 \& \dots \& BOP_n \& BOq_1 \& \dots \& BOq_n) \& CO(p_1 \& \dots \& p_n \& q_1 \& \dots \& q_n)$ . But since by (1a) and (2a),  $PR((p_1 \& \dots \& p_n) \rightarrow p)$  and  $PR((q_1 \& \dots \& q_n) \rightarrow q)$ , it follows that

$$PR(p_1 \& \dots \& p_n \& q_1 \& \dots \& q_n) \rightarrow (p \& q).$$

So we have  $(BOP_1 \& \dots \& BOP_n \& BOq_1 \& \dots \& BOq_n) \& CO(p_1 \& \dots \& p_n \& q_1 \& \dots \& q_n) \& PR((p_1 \& \dots \& p_n \& q_1 \& \dots \& q_n) \rightarrow (p \& q))$ .<sup>39</sup>

(ii) It remains to be shown that  $PE(p_1 \& \dots \& p_n \& q_1 \& \dots \& q_n)$ , i.e.,

$$\forall s_1 \dots s_n [(BOS_1 \& \dots \& BOS_n) \& CO(s_1 \& \dots \& s_n) \rightarrow CO(p_1 \& \dots \& p_n \& q_1 \& \dots \& q_n \& s_1 \& \dots \& s_n)].$$

Assume  $(BOS_1 \& \dots \& BOS_n) \& CO(s_1 \& \dots \& s_n)$ . We need to show that  $CO(p_1 \& \dots \& p_n \& q_1 \& \dots \& q_n \& s_1 \& \dots \& s_n)$ . Since by (2b),  $PE(q_1 \& \dots \& q_n)$ , it follows from our assumption that  $CO(q_1 \& \dots \& q_n \& s_1 \& \dots \& s_n)$ . But then from our assumption and (2a), it follows that  $(BOq_1 \& \dots \& BOq_n \& BOS_1 \& \dots \& BOS_n)$ . So we have  $(BOq_1 \& \dots \& BOq_n \& BOS_1 \& \dots \& BOS_n) \& CO(q_1 \& \dots \& q_n \& s_1 \& \dots \& s_n)$ . But then this, along with (1b), namely  $PE(p_1 \& \dots \& p_n)$ , implies  $CO(p_1 \& \dots \& p_n \& q_1 \& \dots \& q_n \& s_1 \& \dots \& s_n)$ .  $\square$

Using  $UO-C'$ , an induction shows that this generalization holds:

$$UOp_1 \& \dots \& UOp_n \rightarrow UO(p_1 \& \dots \& p_n)$$

And since we have  $UOp \rightarrow OBp$ , the latter implies:

$$UOp_1 \& \dots \& UOp_n \rightarrow OB(p_1 \& \dots \& p_n).$$

<sup>39</sup> This already suffices for  $OB(p \& q)$ , but the stronger  $UO(p \& q)$  is needed to generate the last generalization.

We hope this sketch suggests that the approach holds considerable promise. The approach allows for the definition of a wide variety of notions, including other notions connected with an obligation's pedigree and aggregation. Finally, we believe that it readily solves versions of van Fraassen's largely neglected puzzle.<sup>40</sup>

## References

- [1] L. Aqvist, Deontic logic, in [11], pp. 605–713.
- [2] L. Aqvist, Introduction to Deontic Logic and the Theory of Normative Systems, Bibliopolis, Napoli, 1987.
- [3] N. Belnap, M. Perloff, M. Xu, Facing the Future: Agents and Choices in Our Indeterminist World, Oxford University Press, Oxford, 2001.
- [4] M. Brown, Normal bimodal logics of ability and action, *Studia Logica* 51 (1992) 519–532.
- [5] M. Brown, Agents with changing and conflicting commitments: a preliminary study, in [22].
- [6] M. Brown, J. Carmo, Deontic Logic, Agency and Normative Systems: DEON'96, Springer-Verlag, Berlin, 1996.
- [7] B. Chellas, Modal Logic: An Introduction, Cambridge University Press, Cambridge, 1980.
- [8] E. Ejerhed, S. Lindstrom (Eds.), Logic, Action and Cognition—Essays in Philosophical Logic, Kluwer Academic, Dordrecht, 1997.
- [9] D. Elgesem, Action theory and modal logic, Dissertation, Department of Philosophy, University of Oslo, 1993.
- [10] D. Elgesem, The modal logic of agency, *Nordic J. Philos. Logic* 2 (2) (1997) 1–46.
- [11] D. Gabbay, F. Guenther, Handbook of Philosophical Logic, vol. II, D. Reidel, Dordrecht, 1984.
- [12] J. Hansen, Problems and results for logics about imperatives, *J. Appl. Logic*, in press; doi: 10.1016/j.jal.2004.01.003.
- [13] R. Hilpinen, On Action and agency, in [8], pp. 3–27.
- [14] G. Holmstrom-Hintikka, R. Tuomela (Eds.), Contemporary Action Theory: vol. II, Kluwer Academic, Dordrecht, 1997.
- [15] J. Horty, Moral dilemmas and nonmonotonic logic, *J. Philos. Logic* 23 (1) (1994) 35–65.
- [16] J. Horty, Non-monotonic foundations for deontic logic, in [23], pp. 17–44.
- [17] J. Horty, Agency and Deontic Logic, Oxford University Press, Oxford, 2001.
- [18] A. Jones, M. Sergot, A formal characterization of institutionalized power, *Logic J. IGPL* 4 (3) (1996) 429–445.
- [19] C. Krogh, H. Herrestad, Getting personal: Some notes of the relationship between personal and impersonal obligation, in [6], pp. 134–153.
- [20] P. McNamara, Doing well enough in an Andersonian–Kangerian framework, in [22], pp. 181–198.
- [21] P. McNamara, Toward a framework for agency, inevitability, praise and blame, *Nordic J. Philos. Logic* 5 (2) (2000) 135–160.
- [22] P. McNamara, H. Prakken (Eds.), Norms, Logics and Information Systems: New Studies on Deontic Logic and Computer Science, IOS Press, Amsterdam, 1999.
- [23] D. Nute (Ed.), Defeasible Deontic Logic, Kluwer Academic, Dordrecht, 1997.
- [24] F. Santos, J. Carmo, Indirect action, influence and responsibility, in [6], pp. 194–215.
- [25] F. Santos, A. Jones, J. Carmo, Responsibility for action in organizations: a formal model, in [14], pp. 333–350.
- [26] L. van der Toore, Y.-H. Tan, Two-phase deontic logic, *Logique et Analyse* 171–172 (2000) 411–456.
- [27] B.C. van Fraassen, Values and the heart's command, *J. Philos.* 70 (1973) 5–19.
- [28] D. Walton, Modal logic and agency, *Logique et Analyse* 69–70 (1975) 103–111.

<sup>40</sup> For the puzzle, see the final section of [27]; van Fraassen credits Stalnaker with raising a problem that led to the puzzle. Some excellent exceptions to the puzzle's neglect are [12], [26] (where a nice variant of the puzzle is presented), and [15,16].