A Logical Typology of Normative Systems

Berislav Žarnić
University of Split, Croatia

This paper gives a first order formalization of the proposal put forward by John Broome[1] [2] and develops a typology on that basis. The three-place code function $\mathrm{k} : \mathrm{S} \times \mathrm{A} \times W \longrightarrow \wp\mathcal{L}_{\mathrm{n}}$ delivers the set $\mathrm{k}_{\mathsf{s}}(i, w) \subseteq \mathcal{L}_{\mathrm{n}}$ of propositions in the normative language $\mathcal{L}_{\mathrm{n}}$ that a normative source $\mathsf{s} \in \mathrm{S}$ requires of an agent $i \in \mathrm{A}$ in a world $w \in W$. The value of the code function $\mathrm{k}_{\mathsf{s}}(i, w)$ will be termed the 'set of requirements'. The vocabulary of the normative language $\mathcal{L}_{\mathrm{n}}$ will contain modal operators for belief, B, desire, D, and intention, I. The worlds are construed as subsets of normative language $\mathcal{L}_{\mathrm{n}}$ which are maximal consistent in propositional logic. Possible worlds may violate the laws of modal logics of intentionality according to the philosophical thesis that the essence of the mental is to be subject to norms, not to conform to them (Zangwill [6]).

Definition 1 *The normative language $\mathcal{L}_{\mathrm{n}}$ is built over the base language of propositional logic $\mathcal{L}_{PL}$. Let $i \in A$, $\mathrm{X} = \mathrm{B}, \mathrm{D}, \mathrm{I}$, and $p \in \mathcal{L}_{PL}$*

$$\textit{Sentences of } \mathcal{L}_{\mathrm{n}} ::= p \mid [\mathrm{X}_i]\varphi \mid \neg\varphi \mid (\varphi \wedge \psi)$$

The set of quasi-literals is the set of propositional letters and their negations, and modal formulas and their negations.

The T axiom ($\Box p \rightarrow p$) poses a serious threat to this kind of modeling that keeps modality and world apart. If modalities obeying axiom T were allowed (e.g. epistemic or praxeologic), then possible worlds, being defined as maximal consistent sets in propositional logic, would become intuitively impossible[2]. Since the corresponding T axioms seem to constitute an important part of the meaning of verbs of knowledge and of action, epistemic and praxeologic modalities must be excluded from the language of norms $\mathcal{L}_{\mathrm{n}}$. Von Wright [4] defined 'content of a norm' as "that which ought to or may or must not be or be done". The normative language $\mathcal{L}_{\mathrm{n}}$ departs from von Wright's definition by taking norm-content to be *the psychological state or relation of psychological states that ought to or may or must not be present in the mind of the norm addressee on a particular occasion.* The reduction and the switch may seem

[1] "We must allow for the possibility that the requirements you are under depend on your circumstances. ... There is a set of worlds, at each of which propositions have a truth value. The values of all propositions at a particular world conform to the axioms of propositional calculus. For each source of requirements s, each person i and each world w, there is a set of propositions $\mathrm{k}_{\mathsf{s}}(i, w)$, which is to be interpreted as the set of things that s requires of i at w. Each proposition in the set is a required proposition. The function k_{s} from i and w to $\mathrm{k}_{\mathsf{s}}(i, w)$ I shall call s's *code* of requirements". (Broome [2], p. 14) The symbols in the citation have been changed to match the symbols used in this paper.

[2] For example, although $\{\neg p, [\mathrm{K}]_i p\}$ is pl-consistent set, we do not want to have it included in any world since no false sentence can be known to be true.

drastic but there is a rationale for it. The requirement that agent i knows that p could be replaced by $p \to [\mathrm{B}_i]p$; a required action to see to it that p could be replaced by the required intention, i.e. $[\mathrm{I}_i]p$.

In order to achieve technical clarity we define a first-order metanormative many-sorted language $\mathcal{L}_{\mathrm{meta}}$ with the following extralogical **vocabulary** – individual constants for normative sources, agents and worlds: $\mathrm{s}, \mathrm{s}_1, \ldots$, $\mathrm{a}, \mathrm{a}_1, \ldots$, $\mathrm{v}, \mathrm{v}_1, \ldots$; function symbols for code of requirement, propositional logic consequence, and logic function: k^3, Cn^1, l^1; function symbols for the sentential forms: neg^1, conj^2, and a set of symbols of the type $\mathrm{mod}^1_{\mathrm{X}i}$; monadic predicate symbols expressing properties of being a normative source, an agent, a sentence in $\mathcal{L}_{\mathrm{n}}$, a possible world: $\mathrm{Sr}^1, Ag^1, \mathrm{Sen}^1, W^1$, and dyadic predicates expressing relations of an agent having i-th normative property (corresponding to i-th normative source) in a world, and relation of membership: $\mathrm{K}^2_s, \mathrm{K}^2_{s_1}, \ldots, \in^2$. The **structures** $\mathfrak{M}_{\mathrm{meta}} = \langle \mathrm{D}, \mathcal{I} \rangle$ are built over the domain $\mathrm{D} = \mathrm{S} \cup \mathrm{A} \cup \mathcal{L}_{\mathrm{n}} \cup \wp\mathcal{L}_{\mathrm{n}}$ where S and A are non-empty and disjoint sets, and $\mathcal{L}_{\mathrm{n}}$ is the set already defined (Definition 1). We use variables $w, w_1, \ldots$ to range over worlds; variables $p, p_1, \ldots, q, q_1, \ldots$ to range over sentences in $\mathcal{L}_{\mathrm{n}}$; variables $i, i_1, \ldots$ to range over agents; and variables $x, y, \ldots$ to range over everything. The shorthand notation for sentential form functions uses "Quine quotes", e.g. the shorthand notation for $\mathrm{neg}(x)$ is $\ulcorner \neg x \urcorner$. For the ease of reading, the universal closure of the formula will be notated by formula with free variables. The interpretation of the nonlogical vocabulary is almost straightforward. More complex cases are:

- interpretation of sentential form functions, which we introduce by the way of example – $\mathcal{I}(\mathrm{neg})$ is a function: $\mathcal{L}_{\mathrm{n}} \to \mathcal{L}_{\mathrm{n}}$ such that

$$\mathcal{I}(\mathrm{neg})([\![x]\!]_g^{\mathfrak{M}_{\mathrm{meta}}}) = \begin{cases} \neg \frown [\![x]\!]_g^{\mathfrak{M}_{\mathrm{meta}}} & \text{if } [\![x]\!]_g^{\mathfrak{M}_{\mathrm{meta}}} \in \mathcal{L}_{\mathrm{n}}, \\ \text{undefined}, & \text{otherwise.} \end{cases}$$

  where g is an assignment function and $\frown$ is concatenation operation;

- interpretation of logic function l is function $\mathcal{I}(\mathrm{l}) : \wp\mathcal{L}_{\mathrm{n}} \to \wp\mathcal{L}_{\mathrm{n}}$ such that $\mathcal{I}(\mathrm{l})([\![x]\!]_g^{\mathfrak{M}_{\mathrm{meta}}})$ is the set of all substitutional instances of the formula $[\![y]\!]_g^{\mathfrak{M}_{\mathrm{meta}}} \in \mathcal{L}_{\mathrm{n}}$ for each $y \in x$;

- interpretation of consequence function Cn is a set of consequences in classical propositional logic for a given set, i.e.

$$\mathcal{I}(\mathrm{Cn})([\![x]\!]_g^{\mathfrak{M}_{\mathrm{meta}}}) = \begin{cases} \{y \in \mathcal{L}_{\mathrm{n}} \mid [\![x]\!]_g^{\mathfrak{M}_{\mathrm{meta}}} \vdash_{\mathrm{pl}} y\} & \text{if } [\![x]\!]_g^{\mathfrak{M}_{\mathrm{meta}}} \subseteq \mathcal{L}_{\mathrm{n}}, \\ \text{undefined}, & \text{otherwise.} \end{cases}$$

Definitions 2 *Quantifications over different argument positions in the code function enable a number of interesting type distinctions, some of which will be introduced below using a $\mathcal{L}_{\mathrm{meta}}$ formula in the definiens.*

- k_{s} *is a* pl-*congruent code iff* $\ulcorner p \leftrightarrow q \urcorner \in \mathrm{Cn}(\varnothing) \to (p \in \mathrm{k}_{\mathsf{s}}(i, w) \leftrightarrow q \in \mathrm{k}_{\mathsf{s}}(i, w))$*;*

- k_{s} *is a* pl-*consistent code iff* $\exists w_2\, \mathrm{k}_{\mathsf{s}}(i, w_1) \subseteq w_2$*;*

- k_{s} *is an achievable code iff* $\exists w\, \mathrm{k}_{\mathsf{s}}(i, w) \subseteq w$*;*

- k_{s} *is a* pl-*deductively closed iff* $\mathrm{k}_{\mathsf{s}}(i, w) = \mathrm{Cn}(\mathrm{k}_{\mathsf{s}}(i, w))$*;*

- $\mathrm{k_s}$ *is a relativistic code iff* $\exists i \exists w_1 \exists w_2\, \mathrm{k_s}(i, w_1) \neq \mathrm{k_s}(i, w_2)$*;*
- *a code is absolute iff it is not relativistic;*
- $\mathrm{k_s}$ *is a socially consistent code iff* $\exists w_2\, \mathrm{k_s}(i_1, w_1) \cup \mathrm{k_s}(i_2, w_1) \subseteq w_2$*;*
- *codes* $\mathrm{k_x}$ *and* $\mathrm{k_y}$ *are realization-equivalent iff* $\mathrm{k_x}(i, w) \subseteq w \leftrightarrow \mathrm{k_y}(i, w) \subseteq w$*;*
- *codes* $\mathrm{k_x}$ *and* $\mathrm{k_y}$ *are compatible iff* $\exists w_2\, \mathrm{k_x}(i, w_1) \cup \mathrm{k_y}(i, w_1) \subseteq w_2$*.*

Consistent and deductively closed codes seem to play an important role in our understanding of the basic normative concepts. For example, deontic KD logic without iterated deontic modalities may be conceived as logic of the specific type of code, namely of consistent pl-deductively closed code.

Definition 3 *Let* $p \in \mathcal{L}_{PL}$ *be a formula of propositional logic:*

$$\textit{Formulas of } \mathcal{L}^{\mathrm{O}}_{\mathrm{KD}} ::= p \mid \mathrm{O}p \mid \mathrm{P}p \mid \neg\varphi \mid (\varphi \wedge \psi)$$

Let us introduce the translation τ^1 from the restricted language $\mathcal{L}^{\mathrm{O}}_{\mathrm{KD}}$ to the metanormative language $\mathcal{L}_{\mathrm{meta}}$, with $\mathrm{O}\,p$ and $\mathrm{P}\,p$ standing for 'i in v has s-obligation (s-permission) to p'.

Definition 4 *Function* τ *maps sentences from the fragment* $\mathcal{L}^{\mathrm{O}}_{\mathrm{KD}} \cap \mathcal{L}_{PL}$ *to the set of sentential variables and sentential function terms of* $\mathcal{L}_{\mathrm{meta}}$*:*

$$\begin{array}{rcll} \tau(l) & \in & \{p, p_1, \ldots, q, q_1, \ldots\} & \textit{for propositional letters } l \in \mathcal{L}_{PL} \\ \tau(\neg\varphi) & = & \neg\tau(\varphi) & \\ \tau(\varphi \wedge \psi) & = & (\tau(\varphi) \wedge \tau(\psi)) & \end{array}$$

Definition 5 *Translation* $\tau^1 : \mathcal{L}^{\mathrm{O}}_{\mathrm{KD}} \to \mathcal{L}_{\mathrm{meta}}$

$$\begin{array}{rcll} \tau^1(p) & = & \ulcorner \tau(p) \urcorner \in \mathrm{v} & \textit{if } p \in \mathcal{L}_{PL} \\ \tau^1(\mathrm{O}\varphi) & = & \ulcorner \tau(\varphi) \urcorner \in \mathrm{k_s}(\mathrm{a}, \mathrm{v}) & \\ \tau^1(\mathrm{P}\varphi) & = & \ulcorner \tau(\neg\varphi) \urcorner \notin \mathrm{k_s}(\mathrm{a}, \mathrm{v}) & \\ \tau^1(\neg\varphi) & = & \neg\tau^1(\varphi) & \\ \tau^1(\varphi \wedge \psi) & = & (\tau^1(\varphi) \wedge \tau^1(\psi)) & \end{array}$$

The principles of the standard deontic logic[3] hold under the translation τ^1:

- "gaplessness" condition $\mathrm{P}p \vee \mathrm{O}\neg p$ translates to $\ulcorner \neg p \urcorner \notin \mathrm{k_s}(\mathrm{a}, \mathrm{v}) \vee \ulcorner \neg p \urcorner \in \mathrm{k_s}(\mathrm{a}, \mathrm{v})$ and that property obviously holds for any set of requirements;
- K axiom becomes $\ulcorner p \to q \urcorner \in \mathrm{k_s}(\mathrm{a}, \mathrm{v}) \to (p \in \mathrm{k_s}(\mathrm{a}, \mathrm{v}) \to q \in \mathrm{k_s}(\mathrm{a}, \mathrm{v}))$ and that property holds for any pl-deductively closed set;
- D axiom becomes $p \in \mathrm{k_s}(\mathrm{a}, \mathrm{v}) \to \ulcorner \neg p \urcorner \notin \mathrm{k_s}(\mathrm{a}, \mathrm{v})$ and that is just another way of stating pl-consistency;

[3] "...classical deontic logic, on the descriptive interpretation of its formulas, pictures a gapless and contradiction-free system of norms". (Von Wright [5] p. 32)
According to our translation scheme von Wright's claim should be appended: classical deontic logic "pictures a system of norms" that is deductively closed too.

- mutual definability, $\mathrm{P}_1 p \leftrightarrow \neg \mathrm{O} \neg p$ holds if the set of requirements is congruent.

Although iterated deontic operators receive no translation in the scheme proposed above, one may extend the line of thought by giving additional translation rules for language of standard deontic $\mathcal{L}_{\mathrm{KD}}^{\mathrm{OO}}$ restricted to the maximum of two iterations of deontic operators, treating iterated deontic modalities as a sequence of heterogenous operators and introducing the distinction into the syntax:

$$\mathcal{L}_{\mathrm{KD}}^{\mathrm{O_2O}} ::= p \in \mathcal{L}_{\mathrm{KD}}^{\mathrm{O}} \mid \mathrm{O}_2 p \mid \mathrm{P}_2 p \mid \neg\varphi \mid (\varphi \wedge \psi)$$

Definition 6 *Let* $Sub(\varphi)_{[\frac{c_1}{x_1} \cdots \frac{c_n}{x_n}]}$ *denote substitutional instance of* $\varphi \in \mathcal{L}_{\mathrm{meta}}$ *in which constants* $c_1, ..., c_n$ *are replaced by variables* $x_1, ..., x_n$*. Translation* $\tau^2 : \mathcal{L}_{\mathrm{KD}}^{\mathrm{O_2O}} \to \mathcal{L}_{\mathrm{meta}}$

$$\begin{array}{rcll} \tau^2(\mathrm{O}_2 p) & = & \forall i \forall w \; Sub(\tau^1(p))_{[\frac{\mathrm{a}}{i} \frac{\mathrm{v}}{w}]} & \textit{for } p \in \mathcal{L}_{\mathrm{KD}}^{\mathrm{O}} \\ \tau^2(\mathrm{P}_2 p) & = & \exists i \exists w \; Sub(\tau^1(p))_{[\frac{\mathrm{a}}{i} \frac{\mathrm{v}}{w}]} & \textit{for } p \in \mathcal{L}_{\mathrm{KD}}^{\mathrm{O}} \\ \tau^2(\neg\varphi) & = & \neg\tau^2(\varphi) & \\ \tau^2(\varphi \wedge \psi) & = & (\tau^2(\varphi) \wedge \tau^2(\psi)) & \end{array}$$

Such an approach to iterated deontic modalities departs from von Wright's [5] "second order descriptive interpretation" where e.g. O_2 would stand for existence of "normative demands on normative systems" ("norms for the norm givers"). The "first order" translation τ^1 as well as the "second order" translation τ^2 give us statements in metanormative language $\mathcal{L}_{\mathrm{meta}}$ both of which may "picture" some type of "normative system". The difference lies in the fact that τ^1 gives a local picture of a set of requirements (for a particular source, agent and world) while τ^2 gives a more global picture of a code function. In the second case the properties depicted are the properties of a code function for a particular source with respect to any agent and any world.

Let us consider KD45 deontic logic! The τ^2 translations of reinterpreted axioms 4, $\mathrm{O}_1 p \to \mathrm{O}_2 \mathrm{O}_1 p$ and 5, $\mathrm{P}_1 p \to \mathrm{O}_2 \mathrm{P}_1 p$ amount to stating that any s-obligation and any s-permission holds universally. So, the reinterpreted axioms will hold only if s-code is absolute.

Definition 7 *An agent* i *at world* w *has an "all-or-nothing" normative property* K_{s} *that corresponds to the source* s *iff the set of requirements* $\mathrm{k}_{\mathsf{s}}(i, w)$ *is satisfied in* w*, i.e.* $\mathrm{K}_{\mathsf{s}}(i, w) \leftrightarrow \mathrm{k}_{\mathsf{s}}(i, w) \subseteq w$*.*

If the only way to satisfy some relativistic code and some absolute code is to satisfy them simultaneously, then these codes define the same normative property. The question arises as to whether (non)absoluteness of a code function introduces a difference with respect to normative properties. The next theorem provides a negative answer.

Theorem 8 *For any code there is a realization equivalent absolute code.*

The proof requires extension of the normative language $\mathcal{L}_{\mathrm{n}}$ to the language $\mathcal{L}_{\mathrm{n}(\omega_1)}$ of a variant of infinitary logic which has the same vocabulary as $\mathcal{L}_{\mathrm{n}}$, but in $\mathcal{L}_{\mathrm{n}(\omega_1)}$ the conjunction symbol $\bigwedge$ may be applied to subsets of the set of *quasi-literals*. A function $\mathrm{k}_{\mathsf{s}}^{cond}$ is a conditionalized variant of a code k_{s} iff

$$\forall p \forall w_1 (p \in \mathrm{k}_{\mathsf{s}}^{cond}(i, w_1) \leftrightarrow \exists q \exists w_2 (p = \ulcorner \bigwedge \mathrm{lit}(w_2) \to q \urcorner \wedge q \in \mathrm{k}_{\mathsf{s}}(i, w_2)))$$

where $\mathrm{lit}(w_2)$ is the set of all *quasi-literals* belonging to w_2. The existence of conditionalized variant for any code proves the theorem. In the light of theorem 8, world and agent generalizing translation of axioms 4 and 5 do not introduce distinctions into logical typology of normative properties.

There are several plausible principles of intentionality and normativity: intentionality is normative, i.e. subjected to norms of different sources (e.g. [6]); rationality is one of the normative sources; some norms of rationality are based on logic of psychological modalities. If we accept these principles, then the codes that deliver some "logical" set of sentences deserve our attention. A number of authors take the closure under equivalence to be either unproblematic (e.g. [2]) or at least plausible minimal logical property of a code. In other words, the codes inherit some of the easily noticeable logical properties of the language in which norm-contents are stated. But then a question arises as to which properties are to be preserved in any code. E.g. if the truth-functional equivalence should be inherited, should not the modal congruence[4] be inherited as well, especially in the light of the widely accepted principle that propositions, and not sentences, are the objects of intentionality?

Definition 9 *Let $x \subseteq \mathcal{L}_{\mathrm{n}}$. The set of sentences $\mathrm{l}(x) \subseteq \mathcal{L}_{\mathrm{n}}$ is an axiomatic basis for a set of modal operators occurring in sentences in x ($\mathrm{l}(x)$ is the set of all the substitutional instances of sentences in the set x).*

Let us suppose that $\mathrm{l}(x)$ is also an axiomatic basis for the set of modal operators occurring in sentences in the sets of requirements delivered by $\mathrm{k_s}$. Then we may distinguish several interesting types of codes that do not violate a logic of the modal part of its language:

- code is consistent with respect to $\mathrm{l}(x)$ iff $\exists w_2\ \mathrm{Cn}(\mathrm{l}(x) \cup \mathrm{k_s}(i, w_1)) \subseteq w_2$;
- code is a logic iff $\exists x\ \mathrm{k_s}(i, w) = \mathrm{Cn}(\mathrm{l}(x))$;
- code is "more than a logic" iff $\exists x \exists y(\neg y \subseteq \mathrm{Cn}(\mathrm{l}(x)) \wedge \mathrm{k_s}(i, w) = \mathrm{Cn}(\mathrm{l}(x) \cup y))$;
- code is "less than a logic" iff $\exists x \exists y(\neg y \subseteq \mathrm{Cn}(\mathrm{l}(x)) \wedge \mathrm{k_s}(i, w) = \mathrm{Cn}(\mathrm{l}(x) \cup y) - \mathrm{Cn}(\mathrm{l}(x)))$.

The second type of the logical code could be termed 'formal code', the third and the fourth — 'material codes'. All the four types exhibit some kind of "internal logicality".

One may distinguish two types of logical properties that a code may have. On the one hand, there are external properties of sets of requirements and code functions, like those given in the definitions 2. On the other hand, there is also an internal logicality of a code, pertaining to the modal logic of the code contents.

This approach relaxes the burden of unrealistic logical models of intentionality by their relocation to the normative side; e.g. it is nonsensical to attribute logical omniscience to real agents with finite resources available for reasoning, but one might argue that it is not nonsensical to consider logical omniscience as a normative requirement. The interaction between the normative and the real takes place on the level of agent properties. The widely accepted "ought

[4] If p and q are truth-functionally equivalent, then $[\mathrm{X}_i]p \in \mathrm{k_s}(i, w)$ iff $[\mathrm{X}_i]q \in \mathrm{k_s}(i, w)$.

implies can" principle holds if a code is achievable, i.e. if it is possible for an agent to have the normative property that corresponds to to the source of the code. A straightforward definition of the "all-or-nothing" normative property has been proposed by Broome (see definition 7 above). It is commonly held that rationality as a normative property is not all-or-nothing property but a matter of degree (e.g. Davidson [3]). Therefore, the set of requirements satisfied by an agent having the property of rationality need not include all the requirements delivered by rationality as a normative source. Consequently, the definition of achievability of the code should be modified for "extensive properties".

Further work. The typology of normative systems seems to need a supplementary typology of normative properties, most notably of those that are defined in terms of partial satisfaction. The motivation for the theory of belief revision came from legal context. AGM theory *inter alia* described the logical ways in which consistency of a theory should be maintained. The logical properties that define the state of equilibrium for "homeostatic dynamics" of normative codes should be determined. *Prima facie*, a number of other properties besides "external consistency" like social consistency, achievability, "internal consistency" should be included[5].

References

[1] Carlos E. Alchourrón and Eugenio Bulygin. The expressive conception of norms. In R. Hilpinen (ed.)*New Studies in Deontic Logic*, pp. 95–125, D. Reidel Publishing Company, Dordrecht, 1981.

[2] John Broome. Requirements. In T. Rønnow-Rasmussen, B. Petersson, J. Josefsson, and D. Egonsson, editors,*Homage a Wlodek: Philosophical Papers Dedicated to Wlodek Rabinowicz*, pages 1–41. Lunds universitet, Lund, 2007. `http://www.fil.lu.se/hommageawlodek`.

[3] Donald Davidson. *Problems of Rationality.* Clarendon Press, Oxford, 2004.

[4] Georg Henrik von Wright. *Norm and Action : A Logical Enquiry.* Routledge and Kegan Paul, London, 1963.

[5] Georg Henrik von Wright. Deontic logic: a personal view. *Ratio Juris*, **12**: 26–38, 1999.

[6] Nick Zangwill. The normativity of the mental. *Philosophical Explorations*, **8**: 1–19, 2005.

[5]The author would like to thank anonymous referees for helpful comments. The investigation presented in the paper is supported by MZOŠ project "Logical structures and intentionality".