Doing the right things – trivalence in deontic action logic

Piotr Kulicki and Robert Trypuz*

John Paul II Catholic University of Lublin

Abstract. Trivalence is quite natural for deontic action logic, where actions are treated as good, neutral or bad. We present the ideas of trivalent deontic logic after J. Kalinowski and its realisation in a 3-valued logic of M. Fisher and two systems designed by the authors of the paper: a 4-valued logic inspired by N. Belnap's logic of truth and information and a 3-valued logic based on nondeterministic matrices. Moreover, we combine Kalinowski's idea of trivalence with deontic action logic based on boolean algebra.

Keywords: deontic action logic, many-valued logic, Kalinowski's deontic logic, Dunn-Belnap's four-valued matrix

Introduction

Deontic logic can be seen as a formal tool for analysing rational agent's behaviour in the context of systems of norms. Two main approaches within it can be pointed out: in one of them deontic notions, such as permission, forbiddance or obligation are attributes of situations (in the language – propositions), in the other they are attributes of actions (in the language – names). We are interested in the latter one – deontic action logic (DAL), introduced in the work of G.H. von Wright [13], which is usually treated as the beginning of modern deontic logic.

The first attempt at defining the formal semantics of DAL made by J. Kalinowski [5] has already taken the form of three-valued tables defining truth values of propositions built with deontic operators (permission, forbiddance and obligation) for different types of actions and negations of actions. M. Fisher [4] used a similar methodology but introduced and analysed more operations on actions, namely conjunction and disjunction. More recently a multi-valued approach to DAL appeared in the paper of A. Kouznetsov [7].

Trivalence is quite natural for deontic action logic. The three values in that context refer to actions that are respectively good, neutral or bad. This general idea needs, of course, more detailed specification to be used as a basis for a formal system. In the present paper we just point out the possible ways of using the techniques of multi-valued logic in the area of norms. A more serious study of their application would require much longer and detailed investigations.

* This research was supported by the National Science Center of Poland (DEC-2011/01/D/HS1/04445).

Although other semantic tools have also been successfully applied to DAL we believe that multi-valued semantics for that logic is worth further research. There are two main advantages of that approach (the same as of the use of many-valued techniques in other branches of logic): it is intuitively clear and produces systems with nice computational properties.

We start the paper from a brief recall of Kalinowski's ideas and Fisher's tri-valued DAL. Then we adapt to DAL the ideas introduced into multi-valued logic in a different context: the bilattice known as $\mathcal{FOUR}$ (bilattice of truth and information) with the respective $\mathcal{FOUR}$-valued matrices introduced by N. Belnap [2] and non-deterministic matrices introduced by A. Avron and I. Lev [1]. Finally, we study the relation between DAL, based on boolean algebra introduced by K. Segerberg in [9], and Kalinowski's ideas.

1 Language of deontic action logic

The language of DAL is defined in Backus-Naur notation in the following way:

$$\varphi \ ::= \ \alpha = \alpha \mid \mathsf{O}(\alpha) \mid \mathsf{P}(\alpha) \mid \mathsf{P}_w(\alpha) \mid \mathsf{F}(\alpha) \mid \neg\varphi \mid \varphi \wedge \varphi \mid \varphi \vee \varphi \mid \varphi \to \varphi \mid \varphi \equiv \varphi \tag{1}$$

$$\alpha \ ::= \ a_i \mid \mathbf{0} \mid \mathbf{1} \mid \overline{\alpha} \mid \alpha \sqcup \alpha \mid \alpha \sqcap \alpha \tag{2}$$

where a_i belongs to a finite set of basic actions Act_0, "$\mathbf{0}$" is the impossible action and "$\mathbf{1}$" is the universal action; "$\alpha = \beta$" means that α is identical with β (the last three elements of the language will be used only in section 4); "$\mathsf{O}(\alpha)$" – α is obligatory; "$\mathsf{P}(\alpha)$" – α is strongly permitted (i.e. its performance is permitted in combination with any action); "$\mathsf{P}_w(\alpha)$" – α is weakly permitted (i.e. its performance is not forbidden);"$\mathsf{F}(\alpha)$" – α is forbidden, "$\alpha \sqcup \beta$" – α or β (a free choice between α and β); "$\alpha \sqcap \beta$" – α and β (parallel execution of α and β); "$\overline{\alpha}$" – not α (complement of α). Further, for fixed Act_0, by *Act* we shall understand the set of formulae defined by (2).

In our considerations we understand actions as types or descriptions rather then individual events in time and space. Thus e.g. "$\alpha \sqcup \beta$" is a description of all actions that are covered by description "α" or by description "β" (see also [8]).

2 3-valued deontic logics of Kalinowski and Fisher

2.1 Kalinowski's deontic logic

Kalinowski believed that norms like propositions are true or false. In his approach norms describe relations of permission, obligation and prohibition between agents and actions. Thus, if, for instance, permission holds between agent i and action α, then the norm expressing that state of affairs is true. Logical value of norms depends on moral value of actions – *to fulfil norms means to do the right things*. Kalinowski assumed that every action *in genere* is either good (g), or bad (b) or neutral (n), although actions *in concreto* are always good or bad. Good (bad) actions are such by nature and remain good (bad) in all circumstances. On the other hand neutral actions are those which in some circumstances are good and in other circumstances bad. Kalinowski expressed his philosophical intuitions concerning the meaning of deontic concepts of weak permission, obligation and prohibition by the following matrix:

α	$\mathsf{P}_w(\alpha)$	$\mathsf{O}(\alpha)$	$\mathsf{F}(\alpha)$
b	0	0	1
n	1	0	0
g	1	1	0

One can see that obligatory actions are those which are always good, prohibited actions are those which are always bad and weakly permitted actions are (always) good or neutral.

There is only one internal operator in Kalinowski's logic – action complement. Each action α has its complement (or negations) $\overline{\alpha}$. Action negation is defined by the following matrix:

α	$\overline{\alpha}$
b	g
n	n
g	b

A complement of a good action is bad, a complement of a bad action is good and finally a complement of a neutral action is also neutral. Kalinowski states in [6] that his theory can be enriched by other action operators such as parallel execution or indeterministic choice, but at the same time he is very sceptical about the applicability of theories more expressive than his own deontic theory.

2.2 Fisher's trivalent matrices

One of the extensions of Kalinowski's logic is Fisher's deontic logic. Fisher introduced two operators: parallel execution and indeterministic choice (see two matrixes below), which were missing in Kalinowski's theory.

$\sqcap$	b	n	g
b	b	b	b
n	b	n	g
g	b	g	g

$\sqcup$	b	n	g
b	b	b	g
n	b	n	g
g	g	g	g

The two operators are De Morgan duals, i.e.

$$a \sqcup b = \overline{\overline{a} \sqcap \overline{b}} \tag{3}$$

Basic actions are *per se* good, bad or neutral, while the deontic value of other actions, that can be described as combinations of basic actions made with the use of action operators, can be computed using the matrices.

3 Alternative matrix systems

3.1 4-valued deontic semantics

Kalinowski stated in [6] that Fisher's matrixes are "natural". We find them, at least at certain points, problematic. In Fisher's approach a combination (parallel execution) of good and bad actions is bad and free choice between such actions is good. Execution of an action that is composed of two parts of which one is good and the other is bad can be understood as a conflict of norms or values. We want to study formally different possibilities of solving such conflicts.

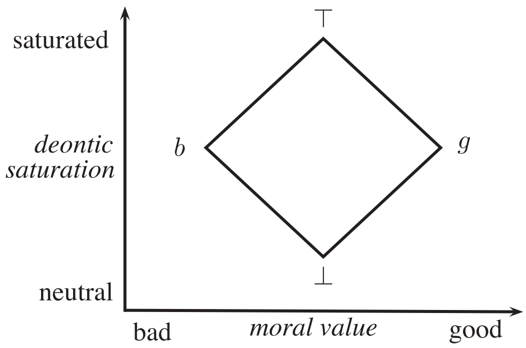
As a first alternative for Fisher's approach let us investigate the solution in which in such a situation a statement about what is good to do is to be made by an agent. In such a situation the action itself can be seen as neutral. That cannot be achieved with trivalent matrices with the preservation of associativity, since we would have:

$$(g \sqcap g) \sqcap b = g \sqcap b = n$$

and

$$g \sqcap (g \sqcap b) = g \sqcap n = g.$$

For that reason we take a structure inspired by Belnap's construction concerning truth information, replacing them respectively by moral value and deontic saturation depicted in the following diagram:



Operator $\sqcap$ is interpreted as supremum in the structure and operator $\sqcup$ as infimum. Negation of $\top$ is $\top$ and negation of $\bot$ is $\bot$.

The same definitions of the operators can be expressed by the following matrices:

$\sqcap$	b	$\bot$	$\top$	g
b	b	b	$\top$	$\top$
$\bot$	b	$\bot$	$\top$	g
$\top$	$\top$	$\top$	$\top$	$\top$
g	$\top$	g	$\top$	g

$\sqcup$	b	$\bot$	$\top$	g
b	b	$\bot$	b	$\bot$
$\bot$	$\bot$	$\bot$	$\bot$	$\bot$
$\top$	b	$\bot$	$\top$	g
g	$\bot$	$\bot$	g	g

a	$\overline{a}$
b	g
$\bot$	$\bot$
$\top$	$\top$
g	b

The interpretation of deontic atoms is defined by the following matrix:

a	$\mathsf{P}_w(a)$	$\mathsf{O}(a)$	$\mathsf{F}(a)$
b	0	0	1
$\bot$	1	0	0
$\top$	1	0	0
g	1	1	0

The matrix shows that intuitively values $\bot$ and $\top$ are treated both as *neutral*. Thus, in a sense, the system remains trivalent, though formally there are $\mathcal{FOUR}$ values that can be attached to actions.

3.2 A system based on non-deterministic matrices

Another possible approach solves the problem of the conflict of norms by stating that what should be done depends on situation and cannot be predicted deterministicly. We formalise it by using nondeterministic matrices. The idea of nondeterministic attachment of deontic value to actions was present in [7] but we use different content of matrix and a more elaborated formalism introduced in [1]. Definition of indeterministic matrices after Avron and Lev is as follows.

- N-matrix for language L is a triple $M = \langle V, D, O\rangle$, where:
  - V is a non-empty set (of truth values);
  - D is a non-empty proper subset of V (designated values);
  - O includes an n-ary function $\underline{x} : V^n \longrightarrow 2^V - \emptyset$ for every n-ary operator x of L
- Let M be N-matrix for L. An M-valuation v is a function $v : F_L \longrightarrow V$ such that for every n-ary operator x of L and every $\alpha_1, \ldots, \alpha_n$,

$$v(x(\alpha_1, \ldots, \alpha_n)) \in \underline{x}(v(\alpha_1), \ldots, v(\alpha_n)).$$

- A model in M is defined in a usual way.

The following nondeterministic matrices define the deontic value of combined actions:

a	$\overline{a}$
b	g
n	n
g	b

$\sqcap$	b	n	g
b	$\{b\}$	$\{b\}$	$\{b,n,g\}$
n	$\{b\}$	$\{n\}$	$\{g\}$
g	$\{b,n,g\}$	$\{g\}$	$\{g\}$

$\sqcup$	b	n	g
b	$\{b\}$	$\{b,n\}$	$\{b,g\}$
n	$\{b,n\}$	$\{n\}$	$\{n,g\}$
g	$\{b,g\}$	$\{n,g\}$	$\{g\}$

The interpretation of deontic atoms is as in Kalinowski's formalisation.

The most interesting cases are those in which good and bad actions are combined with the operations of parallel execution and free choice. The result in the former case is treated as fully indeterministic. Depending on the estimation of how good is a good component of an action and how bad is its bad component one can treat such a compound action as good, neutral and bad. For the latter case we assume, as a general rule for creating the matrix, that a collection of values of two actions is a value of free choice between them. Thus a free choice between a good and a bad action cannot be neutral, as it was defined in the 4-valued system from the previous section. However, we may also use an alternative non-deterministic matrix in which value n is also possible in this case.

3.3 Some formulas

The following formulas are tautologies in all of the systems defined in the previous sections.

$$\mathsf{F}(\alpha) \equiv \neg \mathsf{P}_w(\alpha) \tag{4}$$

$$\mathsf{O}(\alpha) \to \mathsf{P}_w(\alpha) \tag{5}$$

$$\mathsf{O}(\alpha) \equiv \neg \mathsf{P}_w(\overline{\alpha}) \tag{6}$$

$$\mathsf{P}_w(\alpha) \wedge \mathsf{P}_w(\beta) \to \mathsf{P}_w(\alpha \sqcap \beta) \tag{7}$$

$$\mathsf{P}_w(\alpha) \wedge \mathsf{P}_w(\beta) \to \mathsf{P}_w(\alpha \sqcup \beta) \tag{8}$$

$$\mathsf{P}_w(\alpha \sqcup \beta) \to \mathsf{P}_w(\alpha) \vee \mathsf{P}_w(\beta) \tag{9}$$

In contrast the following formulas are valid respectively only in Fisher's logic and only in the 4-valued logic:

$$\mathsf{F}(\alpha) \to \mathsf{F}(\alpha \sqcap \beta) \tag{10}$$

$$\mathsf{O}(\alpha) \to \mathsf{P}_w(\alpha \sqcup \beta) \tag{11}$$

4 Trivalence and deontic action logics based on boolean algebra

4.1 Fundamentals of DAL based on boolean algebra

Deontic action logic based on boolean algebra was introduced by K. Segerberg in [9] and recently studied in [3, 10–12]. Fundamental axioms of $\mathcal{DAL}$ systems are those of Segerberg for $\mathcal{B.O.D.}$, i.e.:

$$\mathsf{P}(\alpha \sqcup \beta) \equiv \mathsf{P}(\alpha) \wedge \mathsf{P}(\beta) \tag{12}$$

$$\mathsf{F}(\alpha \sqcup \beta) \equiv \mathsf{F}(\alpha) \wedge \mathsf{F}(\beta) \tag{13}$$

$$\alpha = \mathbf{0} \equiv \mathsf{F}(\alpha) \wedge \mathsf{P}(\alpha) \tag{14}$$

It is important to note that permission (prohibition) axiomatised above characterise permitted (prohibited) actions as always permitted (prohibited), i.e. permitted (prohibited) in combination with any action:

$$\mathsf{P}(\alpha) \to \mathsf{P}(\alpha \sqcap \beta) \tag{15}$$

$$\mathsf{F}(\alpha) \to \mathsf{F}(\alpha \sqcap \beta) \tag{16}$$

Deontic action model for $\mathcal{DAL}$ is a structure $\mathcal{M} = \langle \mathcal{DAF}, \mathcal{I} \rangle$, where $\mathcal{DAF} = \langle E, Leg, Ill \rangle$ is a *deontic action frame* in which $E = \{e_1, e_2, \ldots, e_n\}$ is a *nonempty* set of possible outcomes (events), *Leg* and *Ill* are subsets of E and should be understood as sets of legal and illegal outcomes respectively. The basic assumption is that there is no outcome which is legal and illegal:

$$Ill \cap Leg = \emptyset \tag{17}$$

$\mathcal{I} : Act \longrightarrow 2^E$ is an interpretation function for $\mathcal{DAF}$ defined as follows:

$$\mathcal{I}(a_i) \subseteq E, \text{ for } a_i \in Act_0 \tag{18}$$

$$\mathcal{I}(\mathbf{0}) = \emptyset \tag{19}$$

$$\mathcal{I}(\mathbf{1}) = E \tag{20}$$

$$\mathcal{I}(\alpha \sqcup \beta) = \mathcal{I}(\alpha) \cup I(\beta) \tag{21}$$

$$\mathcal{I}(\alpha \sqcap \beta) = \mathcal{I}(\alpha) \cap I(\beta) \tag{22}$$

$$\mathcal{I}(\overline{\alpha}) = E \setminus \mathcal{I}(\alpha) \tag{23}$$

Additionally we accept that the interpretation of every atom is a singleton:

$$\overline{\overline{\mathcal{I}(\delta)}} = 1 \tag{24}$$

where δ is an atom of *Act*. A basic intuition is such that an atomic action corresponding to (a set with) one event/outcome is indeterministic. It is important to note two things in this place. The first one is that $\mathcal{I}(\delta)$ is a subset of either *Leg*, or *Ill* or $-Leg \cap -Ill$ and the second one is that in every situation an agent's action has only one outcome, which means in practice that what agents really do is to carry out atomic actions.

The definition of an interpretation function makes it clear that every action generator is interpreted as a set of (its) possible outcomes, the impossible action has no outcomes, the universal action brings about all possible outcomes, operations "$\sqcup$", "$\sqcap$" between actions and "-" on a single action are interpreted as set-theoretical operations on interpretations of actions. A class of models defined as above will be represented by **C**. Satisfaction conditions for the primitive formulas of $\mathcal{DAL}$ in any model $\mathcal{M} \in \mathbf{C}$ are defined as follows:

$$\begin{array}{lll} \mathcal{M} \models \mathsf{P}(\alpha) & \iff & \mathcal{I}(\alpha) \subseteq Leg \\ \mathcal{M} \models \mathsf{F}(\alpha) & \iff & \mathcal{I}(\alpha) \subseteq Ill \\ \mathcal{M} \models \alpha = \beta & \iff & \mathcal{I}(\alpha) = \mathcal{I}(\beta) \\ \mathcal{M} \models \neg\varphi & \iff & \mathcal{M} \not\models \varphi \\ \mathcal{M} \models \varphi \wedge \psi & \iff & \mathcal{M} \models \varphi \text{ and } \mathcal{M} \models \psi \end{array}$$

Action α is strongly permitted iff all of its possible outcomes are legal. It means in practice that if α is permitted, then it is permitted in combination with any action (cf. thesis 15). The same is true for forbiddance.

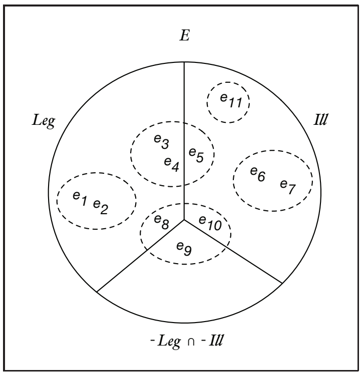


Fig. 1. Five dashed line ovals illustrate some interpretations of $\mathcal{DAL}$ actions. This model includes events which are neither legal nor illegal.

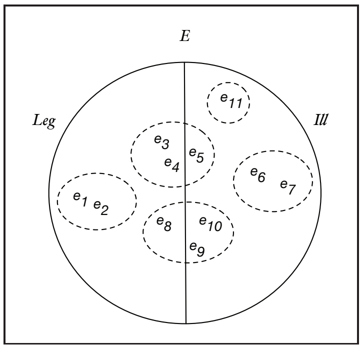


Fig. 2. This model of $\mathcal{DAL}$ is closed in the sense that every event is either legal or illegal.

4.2 Embedding Kalinowski's deontic logic into $\mathcal{DAL}$

To embed Kalinowski's deontic logic into $\mathcal{DAL}$ we need to make a few additional assumptions. First we assume that *Leg* and *Ill* are sets of good and bad events respectively. Then we need to exclude from the models the events which are neither good nor bad, since in Kalinowski's approach each action event is either good or bad. As a result we'd like to obtain models similar to the illustrated below in figure 2.

To obtain the intended result we add a new axiom to $\mathcal{DAL}$ which states that an action is either good or bad or has only good or bad components:

$$\mathsf{P}(\delta) \vee \mathsf{F}(\delta), \text{ for } \delta \text{ being an atom} \tag{25}$$

Axiom 25 explicitly says that action atoms are good or bad. Additionally we assume that there are actions which are neither good nor bad, to make room for neutral ones:

$$\neg\mathsf{F}(\mathbf{1}) \wedge \neg\mathsf{P}(\mathbf{1}) \tag{26}$$

Finally, we express Kalinowski's assumption that the complement of a good action is bad:

$$\mathsf{P}(\alpha) \equiv \mathsf{F}(\overline{\alpha}) \tag{27}$$

The last axiom restricts models of $\mathcal{DAL}$ to the ones illustrated in figure 3. It also shows that "P" and "F" refer to Kalinowski's obligation and forbiddance respectively. They also satisfy semantical conditions restricting obligatory actions only to the good ones and the forbidden actions only to the bad ones. It is also worth noting that each generator (and atom) in $\mathcal{DAL}$ satisfying all the axioms introduced above is interpreted as *Leg* or *Ill* (see figure 3).

Finally, we obtain a structure similar to the one resembling Belnaps's bilattice from section 3.1. To preserve the intuitions of boolean algebra in figure 4 we reversed the order of saturation.

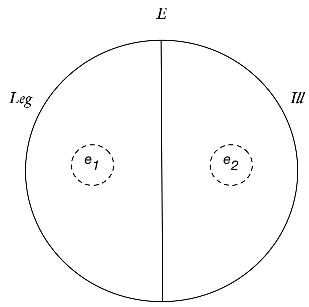


Fig. 3. Model for Kalinowski's deontic logic

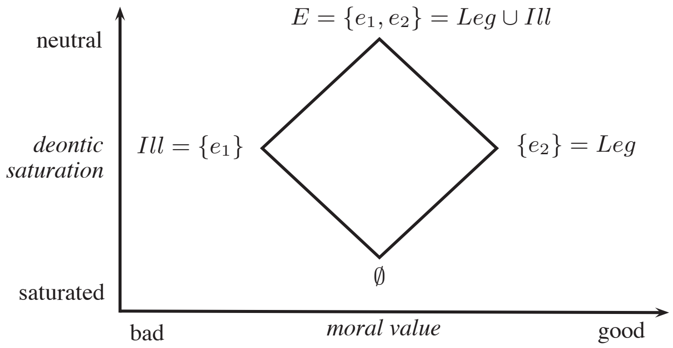


Fig. 4.

It turned out that only the impossible action $\emptyset$ is saturated. Both actions $\emptyset$ and E are neutral.

Since there are only four elements in the algebra we can represent it by the following matrices.

$\sqcap$	Ill	E	$\emptyset$	Leg
Ill	Ill	Ill	$\emptyset$	$\emptyset$
E	Ill	E	$\emptyset$	Leg
$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$
Leg	$\emptyset$	Leg	$\emptyset$	Leg

$\sqcup$	Ill	E	$\emptyset$	Leg
Ill	Ill	E	Ill	E
E	E	E	E	E
$\emptyset$	Ill	E	$\emptyset$	Leg
Leg	E	E	Leg	Leg

a	$\overline{a}$
Ill	Leg
E	$\emptyset$
$\emptyset$	E
Leg	Ill

The only difference between these matrices and 4-valued matrices from section 3.1 lies in the definition of negation. There negation of $\bot$ is $\bot$ and negation of $\top$ is $\top$, here respective values – E and $\emptyset$ interchange.

Kalinowski's weak permission can be defined in the following way:

$$\mathsf{P}_K(\alpha) \triangleq \mathsf{P}(\alpha) \vee \mathsf{N}(\alpha) \tag{28}$$

where

$$\mathsf{N}(\alpha) \triangleq \alpha = 0 \vee \neg(\mathsf{P}(\alpha) \vee \mathsf{F}(\alpha)) \tag{29}$$

Action is (syntactically) neutral (N) iff it is impossible or is neither good nor bad. Neutrality of the impossible action is assumed for the homogeneity reason. The impossible action is the only one which is at the same time good, bad and neutral. The operator defined in such a way satisfies the desired property that α is neutral iff its complement is neutral too:

$$\mathsf{N}(\alpha) \equiv \mathsf{N}(\overline{\alpha}) \tag{30}$$

A weakly permitted action (in Kalinowski's sense) is thus defined as the one that is either good or neutral.

a	$\mathsf{P}_K(a)$	$\mathsf{P}(a)$	$\mathsf{F}(a)$
Ill	0	0	1
E	1	0	0
$\emptyset$	1	0	0
Leg	1	1	0

For P_K the only axiom of Kalinowski's deontic logic K_1 can be proved:

$$\neg\mathsf{P}_K(\overline{\alpha}) \to \mathsf{P}_K(\alpha) \tag{31}$$

One can also check that the following formula is a theorem of extended $\mathcal{DAL}$:

$$\mathsf{P}(\alpha) \equiv \neg\mathsf{P}_K(\overline{\alpha}) \tag{32}$$

Conclusions and further work

We have discussed systems of deontic action logic based on three values of actions: good, bad and neutral. In some of them the neutral value was divided into two. The proposed formalism was useful for discussing different approaches to conflicts of norms or values. Future works will be directed to the application of such logics for modelling agents in the context of norms.

References

1. Arnon Avron and Iddo Lev. Non-deterministic multiple-valued structures. *Journal of Logic and Computation*, 15(3):241–261, June 2005.
2. Nuel Belnap. A useful four-valued logic. In J.M. Dunn and G. Epstein, editors, *Modern uses of multiple-valued logic*, pages 8–37. 1977.
3. Pablo F. Castro and T.S.E. Maibaum. Deontic action logic, atomic boolean algebra and fault-tolerance. *Journal of Applied Logic*, 7(4):441–466, 2009.
4. M. Fisher. A three-valued calculus for deontic logic. *Theoria*, 27:107–118, 1961.
5. J. Kalinowski. Theorie des propositions normativess. *Studia Logica*, 1:147–182, 1953.

6. Jerzy Kalinowski. *La logique des normes*. Presses Universitaires de France, 1972.
7. Andrei Kouznetsov. Quasi-matrix deontic logic. In Alessio Lomuscio and Donald Nute, editors, *Deontic Logic in Computer Science*, volume 3065 of *Lecture Notes in Computer Science*, pages 191–208. Springer Berlin / Heidelberg, 2004.
8. Piotr Kulicki and Robert Trypuz. How to build a deontic action logic. In *Logica 2011*. To appear.
9. Krister Segerberg. A deontic logic of action. *Studia Logica*, 41:269–282, 1982.
10. Robert Trypuz and Piotr Kulicki. A systematics of deontic action logics based on boolean algebra. *Logic and Logical Philosophy*, 18:263–279, 2009.
11. Robert Trypuz and Piotr Kulicki. Towards metalogical systematisation of deontic action logics based on boolean algebra. In *Proc. 10th International Conference Deontic Logic in Computer Science*, volume 6181 of *Lecture Notes in Computer Science*. Springer, 2010.
12. Robert Trypuz and Piotr Kulicki. A norm-giver meets deontic action logic. *Logic and Logical Philosophy*, 20:59–72, 2011.
13. G. H. von Wright. Deontic logic. *Mind*, LX(237):1–15, 1951.