

FIVE STEPS TO RESPONSIBILITY

ELENA LISANYUK*

ABSTRACT

Responsibility has entered the academic discourse of logicians hardly more than few decades ago. I suggest a logical concept of responsibility which employs ideas both from a number of theories belonging to different branches of logic as well from other academic areas. As a comment to this concept, I suggest five steps narrative scenario in order to show how the logical dimension of responsibility emerges from diverse tendencies in logic and other sciences. Here are the five steps briefly stated:

Step 1. Developing modal formalisms capable of evaluative analysis of situations (deontic, epistemic and etc.).

Step 2. Drawing a conceptual borderline between normal and non-normal (weak) logical systems.

Step 3. Using different kinds of models.

Step 4. Agent- and action- friendly turn in logic.

Step 5. Creating formalisms for modeling different types of agency.

An idea advocated here within 5-Steps route to responsibility is that this concept is a complex causal and evaluative (axiological) relation. A logical account may be given for causal and normative aspects of this relation. Unfolding the responsibility back and forth through 5 Steps will result in different concepts. The technicalities are minimized for the sake of keeping the philosophical scope of the paper. For the same reason I also refrain from discussing legal and juridical ramifications of the issue.

KEYWORDS: Responsibility. Agency. Deontic Logic. Norms.

INTRODUCTION. BACK AND FORTH TO RESPONSIBILITY.

The idea of responsibility originates in legal sources and grows out of legal notions of liability and accountability.¹ Despite apparent conceptual closeness to each other the two legal notions mark an important borderline between deontological and causal approaches to responsibility. To say that agent α is accountable for some action or situation α in the face of agent β means that there is an asymmetric relation established between the two agents α and β in which α is the object of an obligation imposed on him by means of the accountability relation and β is the subject or

* Associate Professor, Department of Logic, St Petersburg State University.
E-mail: elenalisanyuk@yandex.ru.

1 RICOEUR P. Le Juste. Edition Esprit, Paris, 1995 (2001). p. 41.

beneficiary to whom the corresponding power, right or privilege belongs. It is in respect to the relation accountability that we find pioneering instances of logical analysis of responsibility in J.Bentham² in the XIX c. and W.N.Hohfeld³ in the beginning of XXc.⁴

Accountability is a kind of social relation between agents; it is usually elaborated in deontological terms and may be further applied to some particular actions and states. Contrary to this, liability relation takes such particular stance and develops the idea of causal connection between agent α and action or situation φ rather than of the accountability relation between agent α and agent β . It stresses the idea of agent's being liable for something rather than agent's being accountable to somebody.

In the view of the two faces of responsibility, liability and accountability, we may choose two different ways of how to pursue it: either we start with causal relations among agents, actions and facts and proceed then further to evaluate these causalities in deontic terms, or, alternatively, we do the other way round and begin with defining social relations among agents deontologically and then try to include also agents and actions into the picture. The two ways amount to the pair of back and forth routes through Steps 1-5. In philosophy, responsibility studies developed according to the both scenarios, and we may easily find deontological⁵ or ontic⁶ approaches to it that apparently point to the forward scenario, as well as functional⁷

2 BENTHAM J. *Of Laws in general*. London, 1970.

3 HOHFELD W.N. *Fundamental Legal Conceptions*. Yale University Press, 1964 (orig. 1923).

4 See substantial outline of both approaches in LINDAHL L. *Position and change*.

5 KANT I. *Kritik der praktischen Vernunft*. Hrsg. von Joachim Kopper, Reclam, Stuttgart 1961. See esp. B.1, Ch.1 §135.

6 INGARDEN R. *Über die Verantwortung. Ihre ontischen Fundamente*. Phillipp Reclam Jun., 1970.

7 HART H.L.A. *Punishment and Responsibility*. Clarendon Press, Oxford, 1968; Lenk,

or agent-wise⁸ positions which lean to the backward one. Perhaps, the imperative conception of responsibility advanced by H. Jonas is the only exception in this aspect, for it explicitly starts by pursuing causal issues and conceptual restrictions preventing human beings from getting the former totally controlled by means of science and proceeds further for the sake of stating the responsibility imperative on ethical grounds.⁹

In both routes of moving through Steps, there is a tension among certain groups of them which marks out the borderline between causal and evaluative aspect of responsibility relation. This tension forces those responsibility theorists who adopt the forward route to keep away from deepening into discussions of individual agents and personal responsibility, whereas those who adhere to the backward line are compelled to handle normative frameworks either as elements of general outer world eventual necessitations imposed on agents or as agent-dependent duties. In logic, responsibility studies start in S.Kanger who has advanced them in a manner which seems to be closer to the forward scenario.¹⁰ However, the backward way tends to become a mainstream approach for the contemporary logical considerations of responsibility. This is manifest in the conceptions of B.Kooi & A.Tamminga¹¹ and especially of M.Mose Bentzen on which I mostly rely in my outline.

H. Macht und Machbarkeit der Technik. Stuttgart, 1994. S. 37 und v.

- 8 ARENDT H. Personal responsibility under dictatorship \ Kohn J. (ed.) Responsibility and Judgement. Schocken Books, 2003 (1964). pp. 17-48.
- 9 JONAS H. Das Prinzip Verantwortung. Versuch einer Ethik für Technologische Zivilization. Frankfurt am Main: Surhkamp 1984. S.35-40.
- 10 KANGER S. Law and Logic \ Holstrom-Hintikka G., Lindstrom S., Sliwinski R. (eds.) Collected papers of Stig Kanger with essays on his life and work. V. 1. Dordrecht, Kluwer, 2001. Pp. 146-169.
- 11 KOOI B., TAMMINGA A. Conflicting obligations in multi-agent deontic logic. \ Goble L., Meyer J.-J.C. (eds.) DEON 2006, LNAI 4048, Springer Verlag, Berlin – Heidelberg, 2006. pp. 175-186.

Here I consider a backward view from Step 5 to Step 1, and place evaluative aspect of responsibility relation on top of its causal aspect which is often viewed in terms of determinism or indeterminism.

Despite the fact that Steps remain substantially the same in the route of defining responsibility relation by moving hither and thither, the resulting definitions are conceptually different. Let us have a brief look on why this is so.

ACCOUNTABILITY, LIABILITY AND AGENTS. AROUND STEPS 4 AND 5.

The accountability relation is elaborated in the line of normative approaches in the framework of which the evaluative aspect of responsibility relation is investigated. Deontologically understood, responsibility relation may be outlined in a remarkable diversity of social or axiological terms with the help of which one can pursue legal obligations, moral postulates, socially rooted rules of conduct and personal duties. There is a great many of philosophers and legal theorists who explicitly point to the deontological character of social relations: Aristotle and Plato, St Augustine and Thomas Aquinas, to mention just some of most influential ancient anticipators of the idea. Deontological character of social relations in a community is an important prerequisite for evaluative aspect of responsibility relation, for it informs us of two crucial issues regarding to this relation. One of them is that there are social facts that tend to be viewed in axiological or normative terms to somewhat larger extent than natural facts which seem to be less amenable in this sense. Another issue is that the distinction between natural and social facts is not impermeable, for in many aspects it is a matter of conceptual standpoint and philosophical justification. To put it otherwise, attributing values to certain kinds of facts and discriminating between social and natural facts are two conceptually different

endeavors.¹² One can refrain from causal aspects as relative to natural facts and identify responsibility relation in deontological terms as a kind social relation between two agents. This way explores the accountability. In this case the study starts with Step 1 and moves further up to Step 5. Step 1 with its deontic and epistemic logical frameworks will suffice to give a general formal view for such approach. This view abstracts from the distinction between action and situation based norms and demonstrates the accountability relation in terms of the stronger \square -type modalities, such as obligations and prohibitions, whereas permission as the weaker \diamond -type modality appears as logical dual for the stronger modalities. The best example for such approach is the standard deontic logic (SDL) of von Wright-type.¹³ In order to develop the idea of diverse kinds of the accountability relation such as right, obligation, privilege and etc. Step 3 provides us with game-theoretical and dynamic frameworks.

Here I place evaluative aspect of responsibility relation which is usually outlined in axiological terms on top of its causal aspect which is normally viewed in terms of determinism or indeterminism. Consequently, responsibility is identified first as liability and this is outlined in causal terms which are ranged from eventual causation as the weakest up to guilt as the strongest. On the basis of this, actions and their results at question are then investigated in axiological terms as praiseworthy or blameworthy, and this manifests that such evaluation is always relative to certain normative order rooted in a social community. The idea to take the evaluative aspect of responsibility relation as distinct of its causal aspect constitutes a considerable part of Step 5. However, the idea to place the former on top the latter

12 Cf. ALCHOURRON C.E., BULYGIN E. Normative knowledge and truth \ Philosophical Analysis in Latin America \ eds. Gracia J.J. et al. Dordrecht, 198. Pp. 25-45.

13 WRIGHT G.H.VON. Norm and Action. London, 1963.

unveils the communicative character of responsibility relation when applied in the framework of social practices.

THE STRUCTURE OF RESPONSIBILITY RELATION. WHY THESE 5 STEPS?

Responsibility is a multi-sort complex relation among the following statements:

- (0) that an action φ has in fact occurred;
- (1) that there is an agent α who performed φ intentionally;
- (2) that agent α has been able to perform φ ;
- (3) that φ performed by α is praiseworthy (blameworthy);
- (4) that some obligations are imposed on α because of statements (1) – (3).

The statements (1)-(5) play different roles in defining responsibility relation. (0) implies either (1) or (2) or both and depending on that we arrive at the subsequent concepts of responsibility in terms of liability and guilt. On the basis of these, (3) is applied on top of the corresponding causal relation either as it is, and thus leads to a default kind of agent's social responsibility which may be described as absolute or moral duty, or, alternatively, it supports (4) and provides us with a variety of particular legal or moral obligations.

Now I briefly comment on these statements taken first separately and then in groups in order to show those groups of them for which a logical relation can be. In doing so, I aim to contrast the handful of logically accessible relations among the statements with two counterpart groups of the statements: those for which no logical notion may be given and those for which we are still awaiting for a logical model to be suggested.

To give a logical definition of a relation means to identify this relation in terms of logical consequence and truth values. Step 1 suggests looking for a logical account of evaluative modalities and deontic logic seems to be an obvious candidate for this.

Instead, Step 5 advances focusing on actions and there types in regards with causal connections among event, action and agent.

Statement (0) can easily be expressed by means of propositional variables. I will refer to it also as event-statement or fact-statement. Statements (1) and (2) are based on event-statements, but operate over them in a special way. They are commented in the following two sections. Statements (3) and (4) presuppose that there exists a normative system which is used to evaluate the results of previously considered statements (1)-(2). Consequently, in order to discuss statements (3)-(4) deontic modalities are introduced with the help of which different kinds of causal responsibility relations are then defined.

ACTIONS AND STRATEGIC AGENCY. STEP 5 AND AROUND STEP 2.

Let us adopt basic J. Habermas-style distinction between two kinds of actions: strategic and communicative. Strategic actions are causal or can be viewed as causal, whereas communicative actions in many aspects exhibit dynamic nature in what concerns their basic objectives. In other words, communicative and creative actions have their ultimate objectives in the performing of an action itself and in most cases the objectives of such actions unfold in the course of the action. Contrary to that, strategic action is performed for the sake of its goal which can be identified is distinct from the action. Creative actions have dynamic character, for their objectives can undergo constitutive change while being performed. We may model strategic actions by means of utilitarian or consequentialist frameworks which have been fruitfully studied in game theory.¹⁴

14 BENTZEN, M. MOSE. *Stit, Iit, and Deontic Logic for Action Types*. PhD Dissertation. Roskilde University, 2009; KOOI B., TAMMINGA A. *Conflicting obligations in multi-agent deontic logic*. \ Goble L., Meyer J.-J.C. (eds.) *DEON 2006*, LNAI 4048, Springer Verlag, Berlin – Heidelberg, 2006. pp. 175-186.

Such models form up Step 3 and are referred to here. However, communicative actions and creative agency are poorly amenable to being researched in terms of strategic action.

Strategic actions can be viewed as normal or non-normal strategic actions. To say that agent α has performed a strategic action ϕ means that α has formulated a goal and gets it accomplished by performing ϕ . This kind of strategic action presupposes that agent α is both rational and reasonable. A person is regarded rational when he is able to outline a coherent line of behavior; we also call a rational person reasonable whenever he is assumed to be capable of means-ends considerations. Being rational and reasonable also implies agent's free will which is a natural prerequisite for strategic agency. Such free-will-based actions are called here normal strategic actions, or C-strategic actions. This is how normal strategic action are symbolized with the help of stit-frameworks¹⁵:

(C-strategic action) $[\alpha \text{ cstit }] \phi$

stit- operators are abbreviated from 'sees to it that' and describe the idea of strategic agency which is meant to include both cognitive planning and physical performance of action ϕ .

Strategic planning can be interpreted in two philosophically distinct ways: in terms of freedom of will and in terms of both freedom of will and freedom of choice. Agent's freedom of choice points to deliberative character of action whereas freedom of will is manifest of the agent being rational. Deliberative action is called non-normal strategic action here and is expressed by means of special dstit-operator which says what 'agent α deliberately sees to it that' ϕ :

(D-strategic action) $[\alpha \text{ dstit }] \phi$

15 For, perhaps, the best account of stit-frameworks in deontic logic and references see HORTY J. Agency and Deontic Logic.

Logical distinction between C- and D-strategic actions may be seen as parallel to philosophical distinction between acting according to freedom of will and that of freedom of will and freedom of choice altogether. The idea of distinguishing actions into normal and non-normal comes from the corresponding classification of modal systems in logic according to their inferential and semantic properties as expressed by the postulates the system at issue validates. In terms of modal systems C-strategic actions are viewed with help of S5-like normal modal system in which a number of important postulates hold. Let us briefly look at those pertinent to responsibility relation:

- T $[\alpha \text{ cstit } \phi] \supset \phi$.
 B. $\phi \supset [\alpha \text{ cstit } \neg[\alpha \text{ cstit } \neg\phi]]$.
 RE. $\frac{\phi \equiv \psi}{[\alpha \text{ cstit } \phi] \equiv [\alpha \text{ cstit } \psi]}$
 N. $[\alpha \text{ cstit } T]$,

Postulate T underlies the causal connection between agent's α strategic behavior and the action ϕ performed. T-postulate may be also understood as effectiveness postulate. B-postulate insists on the agent's freedom of will which provides the coherency of his strategic performance. J.Horty observes the importance of RE-postulate for responsibility studies, for it makes it perfectly clear that strategic agency is a kind of both cognitive and cause-related project, and for this reason, if two actions ψ and ϕ are identical, then agent α is regarded responsible for action ψ inasmuch as for action ϕ of which he may well be unaware of. N – postulate points to the necessary and consistent character of agent's rational will, and taken together with B-postulate draws up the borderline between normal and non-normal strategic action, or C- and D-strategic actions, according to the symbolism employed. In case of deliberative strategic action N – and B-postulates do not hold. Instead, we have

- ~~N.~~ $\neg[\alpha \text{ cstit } T]$.

We will also need a weaker dual of strategic stit-operator in order to talk about allowing in the sense of weakly preserving a situation as is it or as it goes by itself:

(allowing) $\langle \alpha \text{ cstit } \varphi \rangle =_{\text{Df}} \neg[\alpha \text{ cstit } \neg\varphi]$

The third distinction over actions is that between positive and negative ones. This distinction enables us to distinguish acting from preventing or doing from refraining respectively.

On the basis of the latter two distinctions let us now see how causation, instrumental and strategic actions and their negative counterparts are expressed with the help of stit-operators. Causation is expressed immediately by means of ‘allowing’, for it is manifest of the eventual character of such agency when compared to the strategic character as expressed by the stronger stit-operators:

(allowing) $\langle \alpha \text{ cstit } \varphi \rangle$

For example, because of the street dirt on his shoes a terrorist allows the floor to be muddy in the room where he holds the hostages. The terrorist purposed to hold the hostages and he did not intend to stain the floor, but eventually did the latter too.

Another kind of action is an instrumental action. The account of instrumental action rests on the idea of performing φ as means to accomplish ψ , provided the two are distinct from each other:

(instrumental action) $[\alpha \text{ cstit } (\varphi \supset \psi)]$

For example, a terrorist holds hostages men, women and children altogether in order to reserve a room for negotiation maneuver *by means* of releasing women and children when needed. He himself does organize neither negotiation nor a room for maneuver; instead, he performs the action of holding women and children to the effect of possible supporting the negotiation endeavor.

The negative counterpart line for these kinds of actions consists of preventing and refraining. Let us identify preventing here and leave refraining to be dealt with in the next section.

Preventing is a deliberative strategic action, for it presupposes agent's α awareness of both positive and negative scenarios for the action ϕ at question and secures him choosing the latter thus either refraining from the former or allowing it not to occur. This contention suggests that preventing is a non-normal strategic action and it should be symbolized with *dstit*-operator and not with *stit*-operator:

(preventing) $[\alpha \text{ dstit } \neg\phi]$

As it becomes clear with this symbolism, preventing is just a kind of strategic action, but nevertheless it gets specially identified as exactly preventing because it conceptually marks out the distinction between positive and negative kind of strategic agency and does so not only at the propositional level, as both negative causation or instrumental action do. Consequently, the strategic character of preventing implies that that neither causation nor instrumental action may have their negative counterparts. Can we interpret preventing as C-strategic action? Perhaps, many theorists would reply positively, but I would rather go *negative* here and argue that negativeness at the action-level is something different from propositional negation.

Now we may conclude that both kinds of strategic action, normal and non-normal, are genuinely strategic, whereas causation and instrumental actions may be called quasi-strategic. Conceptually the distinction between allowing as the weakest instance of agentive action, instrumental action and both kinds of strategic action lies in the course of agentive intentions and abilities to which we now turn.¹⁶

16 During recent decades, intentions have been intensively and fruitfully studied in philosophy, logic and linguistics. For the two most influential approaches in logic see MEYER J.-J., VELTMAN F. Intelligent agents and common-sense reasoning. \

ABILITIES AND INTENTIONS. AROUND STEPS 2 AND 3.

Statement (2) has to do with two kinds of logical modalities: ability and intentions. Both kinds of modalities belong to formalisms developed in Step 4 and both of them are agent-friendly in the sense that they employ a notion of agent, though they do it in a different way. Apart from this, ability modal formalisms differ from that of intentions in two important aspects. One of them concerns the logical properties of the two, the other one has to do with the contemporary state of research at issue. Let us first take a glance at their logical properties.

Intensive study of ability modalities goes back to the last two decades of the XXc.¹⁷ and the results achieved in the course of it can be recapitulated in two basic lines. Parallel to what has been said above regarding C- and D- strategic actions, one can think of abilities as of kinds of \Box -like and \Diamond - like modalities in the sense described by K-system which is the weakest of normal modal systems. Alternatively, one can interpret abilities as non-normal irreflexive modalities and thus take them in the sense of E-system or yet weaker ones. In both cases abilities may be formally modeled by means of possible world's structures where actions are matched with their outcomes. Moving in this way we get \Box -like strong and \Diamond - like weak ability modalities by minimal and maximal amount of matches correspondingly. Another way of introducing ability modalities employs stit-frameworks.

Stronger A - ability is understood as 'agent α will see to it that' φ :

(will-ability) $A[\alpha \text{ cstit}] \varphi$

BLACKBURN P., VAN BENTHEM J., WOLTER F. (eds.) Handbook of Modal Logic. Elsevier, 2007. Pp. 991-1029 and RAO, A.S., GEORGEFF M.P. Modelling rational agents within a bdi-architecture \ HUHNS M.N., SINGH M.P. (eds) Readings in Agents. Morgan Kaufmann, 1994. Pp. 317-328.

17 See HORTY J. Agency and Deontic Logic. Oxford UP, pp. 2-24.

Weaker E - ability is taken to say what ‘agent α *might* see to it that’ φ :

(might-ability) $E[\alpha \text{ cstit}] \varphi$

Intention modality is a modality of a different sort than that of abilities. Agent’s intention to do something is closely related to his cognitively understood intellectual abilities and his freedom of will. In order to formulate an intention for action agent has to create a strategic outline of the intended action thus choosing¹⁸ a line of behavior out of those available to him according to the action outline.

Then, on the basis of the choice he has to evaluate the possible outcomes of the action according to different scenarios of performing it. Finally, he has to pick out the scenarios with the best outcomes as regards to the intended goal. Briefly stated, intention is the last stage of this agent’s action strategic preview and it is a function over the agent’s choices; intention function merges the agentive choices with the best outcomes and is normally modeled with the help of utilitarian frameworks. Any strategic action implies that it is intended action, but the contrary does not hold. Agent may have intentions which are waiting to be realized as well as those that are never realized:

$[\alpha \text{ cstit}] \varphi \supset [\alpha \text{ iit}] \varphi$

$[\alpha \text{ dstit}] \varphi \supset [\alpha \text{ iit}] \varphi$

iit -operator is a normal D45 operator and it validates standard normality postulates.

Now we are equipped enough as to consider refraining of which there exists two versions parallel to the distinction between normal and non-normal strategic actions.

(C-refraining) $E[\alpha \text{ cstit}] \varphi \wedge \langle \alpha \text{ cstit} \neg \varphi \rangle$

18 This notion of choice in what concerns intention is not to be confused with term ‘choice’ in ‘freedom of choice’.

(D- refraining) $E [\alpha \text{ dstit } \varphi] \wedge \neg [\alpha \text{ cstit } \varphi]$.

Crucial aspect in distinguishing normal C-refraining from non-normal D-refraining is the agent's freedom of choice. In both cases agent α might have done φ in the respective strategic way. However, in case of D-refraining agent α not only declines performing φ , but also shows no preference or care in what regards both occurring and non-occurring of φ thus ipso facto allowing either of the two to happen. Unlike this, C-refraining suggests that agent α knowingly allows happening exactly $\neg\varphi$ but not the contrary. Leibniz's *Fatum Mahometanum* and *Fatum Stoicum* are suitable examples of these kinds of refraining,¹⁹ respectively, though, perhaps, in Leibniz's version the abilities should refer rather to God than to the agent.

MODELS FOR ACTIONS. STEP 3.

In modeling diverse entities like facts, actions, intentions, agency and etc., logicians widely use different frameworks aiming to grasp their formal properties by means of ascribing semantic values to entities analysed. Models provide abstract patterns of real world viewed from a certain angle and in this way supply inferential formalisms with the systems of logical values. Selecting particular model for defining a logical relation between some speculative entities constitutes a substantial part of logical heuristics and creativity.

There is a number of mainstream types of such models, for instance, possible worlds style or game-theoretical style, to recall just two most influential. Possible worlds, or model theoretical semantics appear in two basic versions: standard relativistic which fits *relatively* well with normal modal systems closed under logical consequence, and neighborhood frameworks which are specially designed to express their weaker formal properties such

19 LEIBNIZ G.W.F. *Theodicy: Essays on the Goodness of God, the Freedom of Man, and the Origin of Evil*. Open Court Publishing, 1985. P.43.

as closedness under equivalence instead of that of consequence, among others.²⁰

Responsibility of an agent is a relation sensitive not only to the formal properties of connections between facts, actions and agents, which can be modeled in game-theoretical style,²¹ but also to the effectiveness of agent's activities. In the view of imposing on agents some obligations resulting from agents' previous actions, the outcomes hit in fact have to be estimated in their relation to the range of possible outcomes of the action at issue. This can be done in two ways, by calculating probabilities or utilities, and either of the two can be outlined in relation to two core elements of action, in agent-wise or event-wise perspective. It is worth observing here that adopting either of the two perspectives is far from being a matter of philosophical taste, for it will provide us with different conclusions regarding investigating responsibility. Agent-wise perspective proceeds backwards along the Steps, but event-wise perspective takes the evaluative vein and moves forward from Step 1 to Step 5.

In calculating probabilities or utilities, outcomes of agentive actions play crucial role. By means of intentions agent picks out a number of best outcomes in the line of behavior he has chosen before out of those available to him. Consequently, we have to differentiate the measure of his responsibility in case he achieves his best outcome, his second best and etc. best one, or fails to hit any of them. This is done by means of employing utility or consequentialist models which evaluate the outcomes. In the view of these game theory inspired approaches we distinguish agent's α responsibility for performing φ from

20 For the substantial exposition of these kinds of logical models see CHELLAS B. *Modal Logic. An Introduction*. Cambridge 1980., GABBAY D. *Investigation in Modal and Tense Logic with Application to Problems in Philosophy and Linguistics*. Dordrecht, 1976.

21 Cf. HINTIKKA, JAAKKO AND GABRIEL SANDU. "Game-theoretical semantics," in JOHAN VAN BENTHEM AND ALICE TER MEULEN (eds.), *Handbook of Logic and Language*, Amsterdam: Elsevier, 1997. pp. 361–410.

attempting φ or risking φ respectively. For example, Hero succeeds in releasing most of the hostages kept by Terrorist, but some of the hostages die. We say that Hero is causally responsible for risking those hostages' lives which have been lost dead as well as for attempting to save all the hostages alive. Depending on how the circumstances are evaluated in causal terms we also may say that Terrorist is guilty of attempt of killing the hostages, whereas Hero is liable for the deaths of some of the hostages. Let us put these notions in more exact terms.

CAUSAL RESPONSIBILITY. STATEMENTS (1) AND (2).

The distinctions drawn among the kinds of actions provide us with a firmer ground for formulating of a number of notions of causal responsibility.²² We call agent α positively responsible for φ [α dstit] φ , if φ has been performed by agent α , or simply – there is event φ , and φ has been a deliberative strategic action of α .

(positively responsible) $\varphi \wedge [\alpha \text{ dstit}] \varphi$

Being positively responsible φ implies that agent performed the action to the effect that hits minimal best or minimal positive outcome.

Why D-strategic action and not altogether C- and D-strategic actions matter regarding agent positively responsible? The answer is that to say that somebody is responsible is to make a statement of type (3). For this reason, it is worth considering the issue of agent's freedom to act and outlining the question of agentive freedom in agent-wise manner which amounts to seeing the agent acting on the basis of either free will or free will and free choice altogether. Demonstrating the latter, for instance, by means of pointing to the fact that there had been an alternative

22 See MOSE BENTZEN. OP.CIT. P. 42-45.

choice for the agent would already suffice to prove the strategic and free character of the agency at issue. Justifying agent's relying on free will only seems to be a far more sophisticated endeavor. The question of agent's positive responsibility is the point where the communicative aspect of imposing responsibility originates.²³

Agent α is liable for φ if φ has actually happened, the agent had been able to prevent φ , but he didn't:

$$\text{(liable)} \quad \varphi \wedge E [\alpha \text{ dstit}] \neg\varphi$$

This symbolism provides us with a notion of liability in a weak or minimal sense as regards to causal connections between the agent, his action and the fact. This is the case when agent did not intent for φ , but at the same time he did nothing to prevent φ despite his being able to. To continue the Hero-Terrorist example, Hero is liable for the hostages lost dead, for he could have foreseen some of Terrorist's tactics, say, that Terrorist would keep the hostages in two groups located in two separate places, but Hero did not.

The notion of being liable for risking expresses a somewhat stronger causal connection among the three elements:

$$\text{(liable for risking)} \quad \neg\varphi \wedge \langle \alpha \text{ cstit } \varphi \rangle \wedge E [\alpha \text{ dstit}] \neg\varphi$$

Liability for risking is a typical case when φ did not happen, but still the agent could have made both φ and $\neg\varphi$, depending on his choice of the respective line of conduct. Thus, the agent did not enhance $\neg\varphi$, but could have prevented it. In the above example Hero is apparently liable for risking the hostages' lives.

On order to say that an agent is guilty for something we have to consider agentive intentions in regards to the action and its effect at issue. Common sense considerations prompt that

23 See LISANYUK E. Responsibility, rationality and power. In DOUDNIK S.I., OSIPOV I.D. (eds) Constitutional state and individual responsibility. St Petersburg: St Petersburg Philosophical Society, 2011. Pp. 61-77. (In Russian).

being guilty for ϕ means being responsible for ϕ and intended to do ϕ . Agent's being positively responsible for ϕ implies his being intended to do ϕ , as it has been observed above. We also say that being positively responsible is a necessary condition for being guilty. An agent α is guilty for ϕ , if he is liable for ϕ and he intended to do ϕ :

(positively guilty) $\phi \wedge E [\alpha \text{ dstit}] \neg\phi \wedge [\alpha \text{ iit}] \phi$

This symbolism provides us with the notion of an agent's positive guilt for conscious doing something. In order to construct a negative counterpart notion of agentive guilt we now combine the notion of agentive liability for risking in which the effect of the action at issue shows up with a negative outcome with the agent's intention. This is how we arrive at the notion of being guilty for attempt:

(guilty of attempt) $\neg\phi \wedge E [\alpha \text{ dstit}] \neg\phi \wedge [\alpha \text{ iit}] \phi \wedge \langle \alpha \text{ cstit } \phi \rangle$

In our example Terrorist is apparently guilty of attempt as regards to the hostages' death, for he has intended to hold them what could have lead him to killing them. Alternatively, had not Hero saved some of the hostages they may have been all lost dead because of Terrorist's intention and action which both allowed them all being lost.

In order to sum up the section about the kinds of causal responsibility, let us observe two important issues. We measure causal engagement of an agent in the event produced by means of distinguishing two lines of analysis. One concerns agentive reasons to act, from allowing up to strategic performing with the help of abilities and intentions. The other one is about the effectiveness of the agent's activities which is outlined on the basis of formal frameworks for modeling situations, actions, agentive choices and outcomes. Combining the two lines provides us with either probabilistic, or utilitarian, or yet two-dimensional picture of cognitive and causal aspects of agency. However, despite the

fact that we call agents liable, responsible or guilty in relation to this picture, we have established analytic idea of responsibility only, without evaluating agent's causal and cognitive behavior in axiological terms. In order to do so there should be a system of corresponding values.

DEONTIC LOGIC IN AGENT-WISE PERSPECTIVE. STEP 1.

The idea to study normative aspects of human behavior originates in antiquity and continues in two lines which here have been called causal and evaluative. Up to recent decades the latter stance has been far more investigated. Its most influential and well explored contributors are Leibniz to whom we owe the idea of deontic-alethic parallelism, J.Bentham whose accountability analysis is often called Bentham tradition in logic as distinct from that of Leibniz's, and last but not least is G.von Wright whose pioneering contribution to the evaluative line launches its rapid and fruitful development in the direction of pursuing logic of norms and normative systems.²⁴

In Leibniz-von Wright line of deontic logic norms are expressed with the help of deontic characters O – obligations, F – prohibitions and P – permission by means of applying them to norms' contents expressed by propositional variables. So we get $O\phi$, $F\phi$ and $P\phi$, as obligation to do ϕ , obligation to refrain from doing ϕ and permission to perform ϕ respectively. This is how absolute deontic logic starts and continues further to identify the logical relations among norms or sentences expressing them. In its dyadic or relative version deontic logic also takes agents and situations as correlates for norms' applications.

The conception of norm underlying the development of deontic logic in this perspective advances norms as situation-

24 For a concise overview of the contemporary state of deontic logic MCNAMARA P. Deontic Logic. In: ZALTA P. (ed.) *Stanford Encyclopedia of philosophy*. 2011.

related prescriptions and views them in an objectivistic way. Norms understood in this way provide individual agents with absolute or relative objectives for their social conduct and these norms may be further reformulated as rules applying to certain situations.²⁵ Agents are seen as ultimate norms' addressees to whom the norms are directed and whose social behavior they are meant to govern.

Alternative way to talk about norms is to view them as agent-wise regulations in the sense that norms are taken not only to govern agents' action but also are necessarily constituted by agents. According to that, norms are considered in an indeterministic way and are viewed as part of outer world in which the agent chooses his line of behavior. Perhaps, the first agent-wise perspective for studying norms has been advanced by J.Bentham, but conceptually this tradition starts in a somewhat confused contribution of E.Mally, who has introduced his ill-formed logical system called *Deontik* in 1926.²⁶ There two contemporary versions of agent-wise indeterministic tradition of deontic logic. One has Scandinavian origin and its landmarks are contributions by S.Kanger, L.Lindahl and K.Segerberg; another owes much to N.Belnap and J.Horty and has been inspired by temporal logic of A.Prior. stit-operators introduced above belong to the latter version.

According to the agent-wise idea of deontic logic deontic operators may be viewed as norm imposing functions over agent's choices. There are several ways how such function can be introduced depending on the formal framework which is employed at issue. Here we will rely on the utilitarian framework suggested by M.Mose Bentzen, but take it in a less cumbersome way. He defines such agent-wise norms as certain optimal

25 Perhaps, the best example of such scholarly application may be found in ALCHOURRON C.E., BULYGIN E. Normative systems.

26 LISANYUK E. Deontic 'cocktail' according to E.Mally's receipt. In: Logical Investigations 19 (2013). Moscow-St Petersburg, 2013 P.5-27.

strategies formulated and picked out by the agent for the sake of arriving at the best available outcome. It is clear that agent-wise norms are genuinely individual, for they originate in the agent's free will and do so due to agent's rational choice to obey or disobey them. The conceptual idea behind the agent-wise norm is that norms effectively regulate human social practice just in case the humans as ultimate norms addressee regard them as necessary constraints for themselves.²⁷

Here I use term 'duty' to refer to agent-wise □-type positive norm and contrast it with term obligation to which standard deontic O-operator corresponds.

(agentive duty) Shall [α cstit] ϕ

Shall – operator is a normal modal operator over actions (ought-to-do operator) which validates K and some other normality postulates.

In conclusion to this brief story of how different approaches to deontic study of norms emerge, it is worth observing that the idea of a conception of norm that would be capable of giving a descriptive account of how agent actually behave in the view of diverse normative systems existing as a part of the social reality, has been to some extent anticipated in the paraconsistent approach to normative systems.²⁸

RESPONSIBILITY EVALUATED. CONTINUING WITH STEP 1.

Let us now impose some evaluations onto the ideas of causal responsibility with the help of the deontic operators. An

27 LISANYUK E., ANTONOV M. Deontic logic and the theory of normative systems. A foreword. In: Lisanyuk E. (ed) "Normative systems' and other works in legal philosophy and logic of norms. St Petersburg Univ.Press, St Petersburg, 2013. P. 4-53. (In Russian).

28 PUGA, L. Z.; DA COSTA, N. C. A., VERNENGO, R. J. Normative logics, morality and law. Experts systems in law. A. Martino (ed.). Elsevier Sc. P., 1992.

agent is regarded blameworthy of ϕ , if he is guilty of ϕ and his duty was to prevent ϕ :

(blameworthiness and guilt) $\phi \wedge E [\alpha \text{ dstit}] \neg\phi \wedge [\alpha \text{ iit}] \phi$
 \wedge Shall $[\alpha \text{ cstit}] \phi$

An agent is blameworthy of attempt to ϕ , if in fact ϕ did not happen, but the agent intended to ϕ and allowed ϕ to happen despite his duty to prevent it:

(blameworthiness for attempt) $\neg\phi \wedge E [\alpha \text{ dstit}] \neg\phi \wedge$
 $[\alpha \text{ iit}] \phi \wedge \langle \alpha \text{ cstit } \phi \rangle \wedge$ Shall $[\alpha \text{ cstit}] \phi$

Yet weaker blame on an agent is that of neglect causation, when ϕ occurs, but the agent had no intention to ϕ , however his duty was to prevent it and he could have done so:

(blameworthiness of neglect) $\phi \wedge E [\alpha \text{ dstit}] \neg\phi \wedge \neg [\alpha \text{ iit}]$
 $\phi \wedge$ Shall $[\alpha \text{ cstit}] \neg\phi$

A blame of risking is imposed on an agent, if ϕ actually did not happen, but it might have happened should the agent have made another choice. At the same time the agent did not intend it that ϕ , but his duty was to prevent ϕ and he was able to do so:

(blameworthiness of risk) $\neg\phi \wedge E [\alpha \text{ dstit}] \neg\phi \wedge \neg [\alpha \text{ iit}]$
 $\phi \wedge$ Shall $[\alpha \text{ cstit}] \neg\phi$

It is quite easy to find legal notions of culpability corresponding to these of blameworthiness. Let us now observe that all notions of blameworthiness are relative to agentive duties. Blameworthiness may be also introduced on the basis of obligations and it is precisely the way how it is being done in many cases. However, viewing blameworthiness in this deterministic perspective inevitably leads us to lean towards evaluative stance according to which not only evaluative aspects of responsibility are imposed but the causal ones are also seen as imposed rather than investigated.

Contrary to blameworthiness, praiseworthiness may be both duty- and obligation-relative.

(positive praiseworthiness) $\phi \wedge [\alpha \text{ cstit}] \phi \wedge$ Shall $[\alpha \text{ cstit}] \phi$

(negative praiseworthiness) $O\phi \wedge E [\alpha \text{ dstit}] \neg\phi \wedge [\alpha \text{ iit}] \phi \wedge \langle \alpha \text{ cstit } \phi \rangle$

The idea of being negatively praiseworthy suggests that in cases when there is an obligation to do ϕ , but in fact ϕ did not happen, this not occurring of ϕ comes as a result of other causes and not as a result of the agent's actions who actually intended to ϕ . Negative praiseworthiness may have many faces depending on how duties and obligations are sorted out at issue by an agent. For example, obeying an obligation may be viewed as omissible by the agent for some reason, but he nevertheless stays loyal to the obligation instead of active refrain from obeying it. Negative praiseworthiness may be summarized as that of attempt by means of the above introduced notion of liability of risking:

(praiseworthiness of attempt) $O\phi \wedge E [\alpha \text{ dstit}] \neg\phi$

The notions of negative praiseworthiness give us a notable chance to see how obligations go across and even incompatible with duties and in what causal and evaluative terms this quite familiar everyday instance can be pursued. Another remarkable issue about both praiseworthiness and blameworthiness is that adding them on top of the causal determination of actions provides us with a normatively evaluated responsibility relation but gives it in a descriptive way. This is the result available at the backward route to responsibility and unavailable at the forward route where it always shows up as normative.

CONCLUSION

The 5-step concept of responsibility is a kind of logical high-tech endeavor, for it employs a number of notions and means from different areas. It has been demonstrated here that depending of the forward or, alternatively, the backward route through these 5 Steps one arrives at different concepts of responsibility. The backward scenario considered here provides the descriptive concept in which normative evaluation appears

on top of the causal investigation. The logical concept of responsibility outlined here applies to strategic agency and leaves out that of communicative and creative.

REFERENCES

ALCHOURRON C.E., BULYGIN E. *Normative knowledge and truth* \ Philosophical Analysis in Latin America \ eds. Gracia J.J. et al. Dordrecht, 198. Pp. 25-45.

ALCHOURRON C.E., BULYGIN E. *Normative systems*. Springer, Berlin-Wien, 1971.

BENTHAM J. *Of Laws in General*. London, 1970.

BENTZEN, M. M. *Stit, Iit, and Deontic Logic for Action Types*. PhD Dissertation. Roskilde University, 2009.

CHELLAS B. *Modal Logic. An Introduction*. Cambridge 1980.

GABBAY D. *Investigation in Modal and Tense Logic with Application to Problems in Philosophy and Linguistics*. Dordrecht, 1976.

HINTIKKA, J. AND G. SANDU. "Game-theoretical semantics," in Johan van Benthem and Alice ter Meulen (eds.), *Handbook of Logic and Language*, Amsterdam: Elsevier, 1997. pp. 361–410.

HOHFELD W.N. *Fundamental Legal Conceptions*. Yale University Press, 1964 (orig. 1923).

HORTY J. *Agency and Deontic Logic*. Oxford University Press, 2001.

KANGER S. LAW AND LOGIC \ Holstrom-Hintikka G., Lindstrom S., Sliwinski R. (eds.) *Collected papers of Stig Kanger with essays on his life and work*. V. 1. Dordrecht, Kluwer, 2001. Pp. 146-169.

KOOI B., TAMMINGA A. *Conflicting obligations in multi-agent deontic logic*. \ Goble L., Meyer J.-J.C. (eds.) DEON 2006, LNAI 4048, Springer Verlag, Berlin – Heidelberg, 2006. pp. 175-186.

LEIBNIZ G.W.F. *Theodicy: Essays on the Goodness of God, the Freedom of Man, and the Origin of Evil*. Open Court Publishing, 1985.

LINDAHL L. *Position and change: A study in law and logic*. Synthese Library vol. 112. Uppsala 1977.

LISANYUK E. Responsibility, rationality and power. In Doudnik S.I., Osipov I.D. (eds) *Constitutional state and individual responsibility*. St Petersburg: St Petersburg Philosophical Society, 2011. Pp. 61-77. (In Russian).

LISANYUK E., ANTONOV M. Deontic logic and the theory of normative systems. A foreword. In: Lisanyuk E. (ed) "*Normative systems' and other works in legal philosophy and logic of norms*". St Petersburg Univ.Press, St Petersburg, 2013. P. 4-53. (In Russian)

MCNAMARA P. Deontic Logic. In: Zalta P. (ed.) *Stanford Encyclopedia of philosophy*. 2011.

MEYER J.-J., VELTMAN F. Intelligent agents and common-sense reasoning. \ Blackburn P., van Benthem J., Wolter F. (eds.) *Handbook of Modal Logic*. Elsevier, 2007. Pp. 991-1029.

RAO, A.S., GEORGEFF M.P. Modelling rational agents within a bdi-architecture \ Huhns M.N., Singh M.P. (eds). *Readings in Agents*. Morgan Kaufmann, 1994. Pp. 317-328.

RICOEUR P. *Le Juste*. Edition Esprit, Paris, 1995 (2001).

