On the Logical Approach to the Rationality of an Intelligent Agent

Maria A. Mikheyenkova [1,2]

[1] Federal Research Centre "Computer Science and Control", RAS, Vavilova str., 40, Moscow, 119333, Russia
[2] Russian State University for the Humanities, Miusskaya sq. 6, Moscow, 125993, Russia

Abstract

The paper deals with the formalization of an artificial agent activity using the representation of the agent's actions by logical means. The proposed approach characterizes the rationality of the intelligent (cognitive) agents' activity from the logical consistency point of view. The dependence of rationality on the chosen logical semantics is shown. The presentation of rationality based on an argued choice of actions using the logic of argumentation is also considered.

Keywords

Intelligent agent, multi-agent systems, rational activity, three-valued logics, argumentation

1. Introduction

From the beginning of Artificial Intelligence (AI)'s development, the notion of "an intelligent agent" is significant for the research area [1]. Moreover, they often see AI as exactly and only a science on agents perceiving the environment and affecting it through executive mechanisms [2]. Agents include a wide range of objects – reactive objects, real time planners, decision-making systems, deep self-learning systems, etc. Thus, one can see that the notion of "an agent" is very vague [3] and may vary from individuals to software.

Theoretically, an intelligent agent must possess a wide range of capabilities such as powers of action, communication, and interaction; reactivity; obligations; intentional features; goal setting; reasoning; etc. [4]. Each of these capabilities is an object of a separate research[2], so it is natural to consider some of them individually. Modelling agents' actions, including with methods of logics, is important here (see [7]). They believe that agents engaged into resolving problems are rational, which implies that the agents take the best (according to one or the other criteria) decision.

One can reduce general understanding of rational behavior to the search of optimum relation between the goal of an action, on one hand, and the available knowledge, objective possibilities, and chosen instruments, on the other hand. Contemporary studies on Artificial Intelligence describing behavior of artificial agents and their groups (Multi-Agent Systems, MAS) also accept this all-purpose conception [2].

The development of decision-making theory naturally led to engaging issues related to cognitive mechanisms of human rationality into its scope [8]. It is the awareness of need to secure analysis of rational choice with formal instruments that made researchers turn to ample opportunities of AI methods that, admittedly, make a significant contribution to the development of cognitive studies [9]. Methods of logics for cognitive modelling of agents and multi-agent systems play an important role here [10].

Russian Conference on Artificial Intelligence (RCAI-2021), October 11–16, 2021, Taganrog, Russia
EMAIL: m.mikheyenkova@yandex.ru
ORCID: 0000-0001-9645-6974
© 2021 Copyright for this paper by its authors.
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

[2] For example, [5] is devoted to the formal representation of the "cognitive agent", the use of a "reasoning agent" for the development of the Semantic Web is discussed in [6], etc.

CEUR-WS.org/Vol-3044/paper06.pdf

2. Logical consistency as the basis of rationality

Parallels between social networks and Multi-Agent Systems are a source of reciprocal beneficiation for social science and AI researches [3]. Thus, sociologists see MASs as an instrument for modelling social communities. AI in turn develops instruments for imitating and enhancing human intelligent capabilities, including those related to essential aspects of social activity. Accordingly, such instruments can constitute basis for formalizing behavior of artificial agents and their groups. For example, in [11] one finds an approach allowing extending formal methods of analyzing public opinion to Multi-Agent Systems. There are approaches to modelling agents' actions (largely based upon action theory accepted in social science [12]) considering directed action, individual and group, within certain social group.

Let's consider an algebraic model of a multi-agent system MAS = (Ag, ACT, F, L) [13]. Here $Ag = \{C_1, \dots, C_m\}$ is the set of agents; $ACT = \{p_1, \dots, p_n\}$ is the set of agents' actions in the MAS, F: $Ag \to 2^{ACT}$; $ACT_{C_j} = F_j$ is the set of actions of agent $C_j \in Ag$, L is the subset of extended set ACT', which describes the action of the entire MAS.

We will be interested only in some elements of such a model, namely: the actual activity of agents (without taking into account their interaction) in some En environment, the features of which significantly determine this activity. Let's digress for now also from the sources of motivation that guide the action of agents. The presence of the agent's intentional characteristics (see [4]) allows to diversify its actions in different environments (situations of action), dividing them into permissible, forbidden and indefinite (nonsense) for each of the dynamically changing environments. Accordingly, three-valued logics with truth values $\nu \in \{0, \frac{1}{2}, 1\}$ can be used to formalize the action. The possible semantics of the truth value $\frac{1}{2}$ is considered in [14]: strong nonsense (mathematical) in Bochvar logic B_3, weak nonsense (linguistic) in Ebbinghaus logic E_3, uncertainty (unknown, true or false) in the version of Lukasiewicz logic L'_3 proposed by V.K. Finn. If $\frac{1}{2}$ is interpreted as a strong nonsense, then a complex statement $\varphi(p)$ containing the occurrence of a nonsense atomic statement p (with truth value $\frac{1}{2}$) is also nonsense. When interpreting $\frac{1}{2}$ as weak nonsense, a complex statement $\varphi(p_1, \dots, p_n)$ is nonsense if and only if all atomic statements $p_1, \dots, p_n$ included in φ are nonsense. Interpretation of $\frac{1}{2}$ as uncertainty means that the statement p is either true or false, but its evaluation is unknown.

In all three logics – B_3, E_3 и L'_3 – logical connectives negation ~, conjunction and disjunction on {0,1} are defined as in two-valued logic, and $\sim \frac{1}{2} = \frac{1}{2}$ (negation of nonsense is nonsense, negation of uncertainty is uncertainty). Disjunction and conjunction for logics B_3, E_3 и L'_3 are defined with idempotence safekeeping (see Table 1 and Table 2 below). The definition of formulae is standard.

Table 1

Disjunction truth tables

B_3				E_3				L'_3			
$\dot\cup$	0	$\frac{1}{2}$	1	$\dot\vee$	0	$\frac{1}{2}$	1	$\vee$	0	$\frac{1}{2}$	1
0	0	$\frac{1}{2}$	1	0	0	0	1	0	0	$\frac{1}{2}$	1
$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	0	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
1	1	$\frac{1}{2}$	1	1	1	1	1	1	1	1	1

To represent the possible actions of the agent, we will use the logical connectives J_ν introduced by D.A. Bochvar (named afterwards Rosser-Turquette J-operator), $\nu \in \{0, \frac{1}{2}, 1\}$, $J_\nu p = \begin{cases} 1, if\ v[p] = \nu \\ 0, if\ v[p] \neq \nu \end{cases}$, p is a propositional variable, v is the valuation function. Thus, J_ν-operators correspond to characteristic functions that recognize truth values $\nu \in \{0, \frac{1}{2}, 1\}$. Then $J_1 p$ means that the agent performs action p, $J_0 p$ – the agent refrains from performing action p (for example, the action is prohibited in this situation/environment), $J_{\frac{1}{2}}p$ – the execution of the action is undefined (nonsense). The complete set of actions of the agent X_j from the set of actions $ACT = \{p_1, \ldots, p_n\}$ is represented as $[\varphi_j] = \{J_{v_1^{(j)}}p_1, \ldots, J_{v_n^{(j)}}p_n\}$, where $v_i^{(j)} \in \{0, \frac{1}{2}, 1\}$, $i = 1, \ldots, n$; $j = 1, \ldots, r$; $r = |Ag|$.

External control influence can impose limitations on the agents' activities, representing them as dependencies between the performance of certain actions. Let us represent these dependencies in the form of a consistent set of formulae $\Sigma = \{\psi_1, \ldots, \psi_s\}$ of the corresponding three-valued logic (B_3, E_3 or L'_3), and the conjunction $\bar{\psi} = \psi_1\&\ldots\&\psi_s$ is not a tautology (the symbol & here is conditional, the conjunction is determined by the corresponding truth table of the chosen logic, Table 2).

Table 2

Conjunction truth tables

B₃				E₃				L′₃			
⩀	0	$\frac{1}{2}$	1	⩀	0	$\frac{1}{2}$	1	&	0	$\frac{1}{2}$	1
0	0	$\frac{1}{2}$	0	0	0	$\frac{1}{2}$	0	0	0	0	0
$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{1}{2}$
1	0	$\frac{1}{2}$	1	1	0	$\frac{1}{2}$	1	1	0	$\frac{1}{2}$	1

Consistency of Σ can be established by the method of analytic tableaux for the logics B_3, E_3 and L'_3, where the so-called designated formulae $J_\nu\varphi$ are used, φ is an un designated formula, $\nu \in \{0, \frac{1}{2}, 1\}$[3]. Designated formulae $J_\nu\varphi$, $J_\mu\varphi$ ($\nu \neq \mu$) are the contrary pairs. The corresponding inference α-rules (rules of conjunctive type), β-rules (rules of disjunctive type) and special χ-rules are formulated. Analytic tableau $\mathcal{T}_\Sigma$ for the set of formulae $\Sigma = \{\psi_1, \ldots, \psi_s\}$ is an analytic tableau such that the root of its inference tree is a sequence of formulae $\psi_1, \ldots, \psi_s$.

$$\left.\begin{matrix}\psi_1\\ \vdots\\ \psi_s\end{matrix}\right\}\mathcal{T}_\Sigma$$

It is easy to show that $\mathcal{T}_\Sigma$ is equivalent to the analytic tableau $\mathcal{T}_{\bar{\psi}}$ with the root $\bar{\psi}$. For a consistent set Σ, the analytic tableau $\mathcal{T}_\Sigma$ ($\mathcal{T}_{\bar{\psi}}$) is not closed. By the completeness theorem of the analytic tableaux method for B_3, E_3 and L'_3 (see [16]) logics, if $\bar{\psi}$ is a tautology, then $\bar{\psi}$ is provable in B_3, E_3 and L'_3, respectively, i.e., $\mathcal{T}_{J_0\bar{\psi}}$ и $\mathcal{T}_{J_{\frac{1}{2}}\bar{\psi}}$ are closed.

Let's assume that the agent's actions in a certain environment (situation) are characterized by the complete set described above, then the actual activity ACT'_{X_j} of the agent $X_j \in Ag$ can be represented by the maximal conjunction $\varphi_j \eqcirc J_{v_1^{(j)}}p_1\&\ldots\&J_{v_n^{(j)}}p_n$ ($\eqcirc$ is graphical equality, & is a symbol for the conjunction of logics B_3, E_3 or L'_3) of atomic actions $J_{v_i^{(j)}}p_i$, $\overline{ACT} = \{ACT'_{X_1}, \ldots, ACT'_{X_r}\}$, $r = |Ag|$. This conjunction is determined by analogy with the maximal conjunction of two-valued logic, i.e., $J_{v_k^{(j)}}p_k$ for each p_k ($k=1, \ldots, n$) includes into φ_j without repetitions, and $J_{v_k^{(j)}}p_k, J_{v_i^{(j)}}p_k$, $\nu_k^{(j)} \neq \nu_i^{(j)}$, are not included in φ_j together.

Let *consis*(Σ) to be the consistency meta-predicate of the set of formulae Σ. Then it is possible to determine whether the activity of the agent $ACT'_{X_j} = \varphi_j$ is in contradiction with Σ, i.e. does *consis*($\Sigma \cup \{ACT'_{X_j}\}$) hold. For this, we'll also use the analytic tableaux method. Let's construct the set of analytic tableaux $\bar{\mathcal{T}} = \{\mathcal{T}_{\Sigma\cup\{ACT'_X\}} \mid ACT'_X \in \overline{ACT}\}$, choose only such $ACT'_{X_{j_1}}, \ldots, ACT'_{X_{j_k}}$, where *consis*($\Sigma \cup \{ACT'_{X_{j_l}}\}$) holds, i.e. analytic tableau $\mathcal{T}_{\Sigma\cup\{ACT'_{X_{j_l}}\}}$ with the root $\Sigma \cup \{ACT'_{X_{j_l}}\}$ is not closed, $l = 1, \ldots, k$. We'll call agents whose activities do not contradict Σ rational. The set of rational agents is

---

[3] Compare with the method of analytic tableaux for J_m-logics ($m \geq 3$) [15], which are an extension of two-valued logics.

$Ag^* = \{X \mid (consis(\Sigma \cup \{ACT'_X\}) \& (X \in Ag)\}$, $Ag^* \subseteq Ag$. $\mathcal{T}^* = \{\mathcal{T}_{\Sigma \cup \{ACT'_X\}} \mid X \in Ag^*\}$, $\mathcal{T}^* \subseteq \bar{\mathcal{T}}$. Note that α-, β- and χ-rules for the logics B_3, E_3 и L'_3 are formulated differently. Accordingly, checking whether $consis(\Sigma \cup \{ACT'_{X_j}\})$ holds or not is dependent on the chosen logic, i.e. from the semantics of truth values.

The identification of rational agents can be useful for semi-autonomous agents' control, allowing you to block the activities of non-rational agents in the environment *En*. If the environment is transformed – as a result of the agents' actions or under the influence of external control action – non-rational agents can turn out to be rational, and vice versa. Let's consider the sequence of environment changes $En_1, \ldots, En_s$. The set of possible action $ACT = \{p_1, \ldots, p_n\}$ is assumed to be general for all En_q, $q = 1, \ldots, s$; at the same time, in each environment, all actions are not necessarily implemented. Accordingly, for En_q Σ_q is given ($q = 1, \ldots, s$).

Note that the case $\Sigma_1 = \Sigma_2 = \ldots = \Sigma_s$ is of no interest, since it initially fixes the sets of rational and non-rational agents. Let for some m $\Sigma_{m-1} \cap \Sigma_m = \varnothing$. In this case, the identification of rational agents for En_m occurs anew in accordance with the procedure described above, although for some l and m ($l \neq m$) it is possible $Ag_l^* = Ag_m^*$.

If $\Sigma_{m-1} \subset \Sigma_m$, to identify Ag_m^*, it is enough to check whether the activity of rational agents Ag_{m-1}^* of the En_{m-1} environment does not contradict the new dependencies, i.e. $Ag_m^* = \{X \mid (consis((\Sigma_m \setminus \Sigma_{m-1}) \cup \{ACT'_X\}) \& (X \in Ag_{m-1}^*)\}$.

In the case $\Sigma_m \subset \Sigma_{m-1}$ some non-rational agents from the set $(Ag \setminus Ag_{m-1}^*)$ may turn out to be rational, i.e. $Ag_m^* = Ag_{m-1}^* \cup \{X \mid (consis(\Sigma_m \cup \{ACT'_X\}) \& (X \in (Ag \setminus Ag_{m-1}^*))\}$.

The case $((\Sigma_{m-1} \cap \Sigma_m \neq \varnothing) \& \neg(\Sigma_{m-1} \subset \Sigma_m) \& \neg(\Sigma_m \subset \Sigma_{m-1}))$ does not allow us to reduce the procedure for dividing the set of agents into rational and non-rational.

3. Rationality and argumentation

The development of logical theories of argumentation [17, 18, 19] finds practical application in decision-making theory, conflict analysis, knowledge representation in intelligent systems, and multi-agent systems engineering. The actions of intentional agents can be based on widely understood argumentation, which can be (conditional) beliefs, motivations, obligations, etc. Argued decision making (choice of actions), not reducible to deductive reasoning, is supposed to be rational [20].

Following the logic of argumentation proposed in [21], let A be the set of arguments (argumentation base) regarding the acceptance or non-acceptance of certain statements, that is, performance or non-performance of some actions from $ACT = \{p_1, \ldots, p_n\}$ of a multi-agent system by an intentional agent. Note that the argumentation base A is considered as common to all agents – it can be, for example, the union of argumentation bases of all agents. Let's determine functions $g^+(p_i)$ and $g^-(p_i)$: $g^\sigma: ACT \rightarrow 2^A$, where $\sigma \in \{+, -\}$. These functions give a set of arguments "for" and a set of arguments "against", respectively:

$g^+: ACT \rightarrow 2^A$, $g^+(p_i) \subseteq A$, $i = 1, \ldots, n$.

$g^-: ACT \rightarrow 2^A$, $g^-(p_i) \subseteq A$, $i = 1, \ldots, n$.

A pair of functions $g+$, g^- will be called normal, if for all $p_i \in ACT$ $g+(p_i) \cap g^-(p_i) = \varnothing$, $i = 1, \ldots, n$.

Permissible actions from $ACT = \{p_1, \ldots, p_n\}$ take the truth value «1», forbidden – «–1», undefined – «τ». Let us define the argumentation semantics of the three-valued logic A_3 by analogy with the semantics of the four-valued logic A_4 from [21].

Atomic valuations for truth values $V = \{1, -1, \tau\}$ are defined as follows:

$v[p_i] = 1 \leftrightarrow g+(p_i) \neq \varnothing, g^-(p_i) = \varnothing$;

$v[p_i] = -1 \leftrightarrow g+(p_i) = \varnothing, g^-(p_i) \neq \varnothing$;

$v[p_i] = \tau \leftrightarrow (g+(p_i) \neq \varnothing, g^-(p_i) \neq \varnothing)$ or $(g+(p_i) = g^-(p_i) = \varnothing)$;

($i = 1, \ldots, n$).

Of course, each agent X_j has its own set of argument functions $g^+_{X_j}$, $g^-_{X_j}$, $g^\sigma_{X_j} = \{g^\sigma_{X_j}(p_1), \ldots, g^\sigma_{X_j}(p_n)\}$, $\sigma \in \{+, -\}$. For the agent X_j to be rational, it is necessary (but not enough) to satisfy the condition $\forall p_i (g^+_{X_j}(p_i) \cap g^-_{X_j}(p_i) = \varnothing)$, $i = 1, \ldots, n$.

The method of analytic tableaux for logics JA_4 and JA_5 (four- and five-valued logics with argumentation semantics) has been formulated in [20]. It is easily transformed for three-valued logic JA_3 with argumentation semantics.

Unary logical connectives J_1, J_{-1}, J_τ and binary logical connectives &, ∨, →; are used here t, f – (external) truth values of two-valued logic “true” and “false”, respectively. $J_\nu p = \begin{cases} t, \ if\ v[p] = \nu \\ f, \ if\ v[p] \neq \nu \end{cases}$, $v[p]$ is the valuation function, $\nu \in \{1, -1, \tau\}$. Accordingly, $v[J_1 p_i] = t \leftrightarrow g^+(p_i) \neq \varnothing$, $g^-(p_i) = \varnothing$ and so on. The analytic tableaux are built by use of designated formulae $t\varphi$ and $f\varphi$, t and f are signs for φ. Designated formulae $t\varphi$ and $f\varphi$ are contrary pairs so as undesignated formulae $J_\nu p$, $J_\mu p$, $\nu \neq \mu$, ν, $\mu \in \{1, -1, \tau\}$.

As for the three-valued logics B_3, E_3, and L'_3 considered above, the corresponding α-rules (rules of conjunctive type), β-rules (rules of disjunctive type) and special χ-rules are formulated. Accordingly, the set of rational agents is determined by means of JA_3 logic in accordance with the procedure described above.

Let $\overline{ACT} = \{\varphi_j \mid \varphi_j \eqcirc J_{\nu_1^{(j)}} p_1 \& \ldots \& J_{\nu_n^{(j)}} p_n, \nu_i^{(j)} \in \{1, -1, \tau\}, i = 1, \ldots, n; j = 1, \ldots, r; r = |Ag|\}$ be the set of agents’ activities.

Here $J_1 p$ means that the agent has arguments for performing the action p in the environment and there are no arguments against, $J_{-1} p$ – there are arguments for refusing the action p and there are no arguments for its performing, $J_\tau p$ – the execution of the action is undefined due to the absence of arguments or the presence of both arguments “for” and “against”.

Let’s give a simple example of rational agents’ identification.

Let $Ag = \{C_1, C_2, C_3\}$ to be the set of agents, $ACT = \{p_1, p_2, p_3\}$ to be the set of agents’ actions. The actions of agents C_1, C_2, C_3 in a certain environment are represented by the sets $[\varphi_1] = \{J_1 p_1, J_1 p_2, J_\tau p_3\}$, $[\varphi_2] = \{J_1 p_1, J_1 p_2, J_1 p_3\}$, $[\varphi_3] = \{J_1 p_1, J_1 p_2, J_0 p_3\}$, respectively. Then ACT'_{C_j} for agent $C_j \in Ag$ is represented by maximal conjunction $\varphi_j \eqcirc J_{\nu_1^{(j)}} p_1 \& J_{\nu_2^{(j)}} p_2 \& J_{\nu_3^{(j)}} p_3$, $j = 1, 2, 3$; $\overline{ACT} = \{ACT'_{C_1}, ACT'_{C_2}, ACT'_{C_3}\}$. Limitations $\Sigma = \{J_1 p_1 \rightarrow (J_1 p_2 \vee J_1 p_3)\}$ are imposed on the actions of agents in the environment. To identify non-rational agents whose activities contradict the imposed restrictions and should be blocked, we will construct analytic tableaux $\mathcal{T}_{\Sigma \cup \{ACT'_{C_i}\}}$, $i = 1, 2, 3$.

For this we need the corresponding

α-rules $\frac{tJ_\nu p}{J_\nu p}$ ($\nu \in \{1, -1, \tau\}$) and $\frac{t(\varphi \& \psi)}{t\varphi, t\psi}$, β-rules $\frac{t(\varphi \rightarrow \psi)}{f\varphi | t\psi}$ and $\frac{t(\varphi \vee \psi)}{t\varphi | t\psi}$, χ-rule $\frac{fJ_1 p}{J_{-1} p \mid J_\tau p}$

of the method of analytic tableaux of JA_3 logic. Here φ, ψ denotes arbitrary formulae of JA_3 logic, p is propositional variable.

The analytic tableau for $\mathcal{T}_{\Sigma \cup \{ACT'_{C_1}\}}$ is presented below:

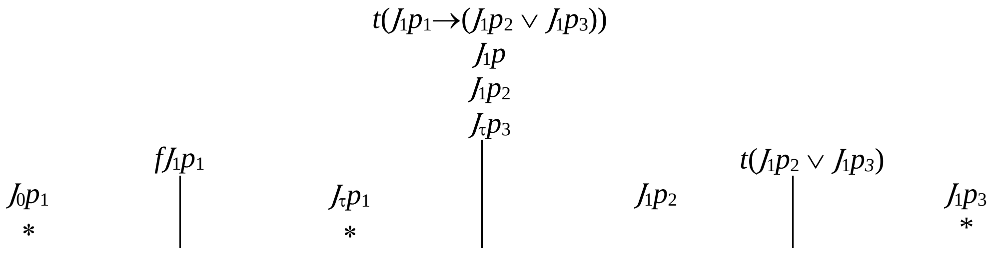

Figure 1: The example of the analytic tableau for JA_3

The tableau is not closed, there is an open branch, therefore, the activity of agent C_1 does not contradict Σ, C_1 is rational agent, $C_1 \in Ag^*$.

Analogously, the tableaux $\mathcal{T}_{\Sigma \cup \{ACT'_{C_2}\}}$ and $\mathcal{T}_{\Sigma \cup \{ACT'_{C_3}\}}$ have open branches, hence $Ag^* = Ag$, all agents can act according to their instructions.

However, as mentioned above, the analysis of agents' rationality is closely related to the interpretation of actions evaluation, on which the choice of the logic depends. The possible semantics of truth values $\nu \in \{0, \frac{1}{2}, 1\}$ are defined in the section 2.

Let us show what the result of the given example will be if we choose the logic B_3. Then the environment limitations have the form $\Sigma = \{p_1 \rightarrow (p_2 \Cup p_3)\}$, $\overline{ACT} = \{ACT'_{C_1}, ACT'_{C_2}, ACT'_{C_3}\}$, and $ACT'_{C_1} \eqcirc J_1 p_1 \Cap J_1 p_2 \Cap J_{\frac{1}{2}} p_3$, $ACT'_{C_2} \eqcirc J_1 p_1 \Cap J_1 p_2 \Cap J_1 p_3$, $ACT'_{C_3} \eqcirc J_1 p_1 \Cap J_1 p_2 \Cap J_0 p_3$.

To construct the tables $\mathcal{T}_{\Sigma \cup \{ACT'_{C_i}\}}$, $i = 1, 2, 3$, use the corresponding α-rule $\frac{J_1(\varphi \Cap \psi)}{J_1\varphi, J_1\psi}$ and corresponding χ-rules $\frac{J_1(\varphi \rightarrow \psi)}{J_0\varphi \,\middle|\, J_{\frac{1}{2}}\varphi \,\middle|\, J_1\psi}$ and $\frac{J_1(\varphi \Cup \psi)}{J_1\varphi, J_0\psi \mid J_1\varphi, J_1\psi \mid J_0\varphi, J_1\psi}$ of the method of analytic tableaux for B_3 logic. Here φ, ψ are formulae of B_3 logic.

The analytic tableau $\mathcal{T}_{\Sigma \cup \{ACT'_{C_1}\}}$ has the following form:

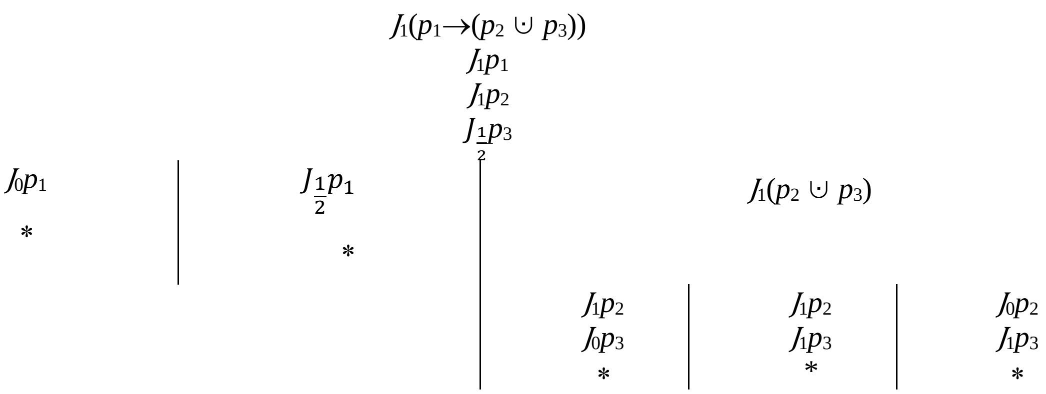

Figure 2: The example of analytic tableau for B_3

The tableau is closed, there are no open branches, the agent's activity contradicts Σ, $C_1 \notin Ag^*$.

It is easy to verify that analytic tables $\mathcal{T}_{\Sigma \cup \{ACT'_{C_2}\}}$ and $\mathcal{T}_{\Sigma \cup \{ACT'_{C_3}\}}$ for C_2 and C_3, respectively, are not closed, i.e., $Ag^* = \{C_2, C_3\}$, $Ag^* \subset Ag$.

It can be seen from the examples given that the formation of groups of rational agents under the environment limitation depends on the semantics of actions' valuations, in accordance with which the logical apparatus is chosen.

JA_3 logic formal means also allows an alternative procedure to be used to identify rational agents. Let, as above, $\overline{ACT} = \{\varphi_j \mid \varphi_j \eqcirc J_{\nu_1^{(j)}} p_1 \& \ldots \& J_{\nu_n^{(j)}} p_n,\ \nu_i^{(j)} \in \{1, -1, \tau\},\ i = 1, \ldots, n;\ j = 1, \ldots, r;\ r = |Ag|\}$. Let $F_{Ag} = (\varphi_1 \vee \ldots \vee \varphi_r)$, $\varphi_j \in \overline{ACT}$, $j = 1, \ldots, r$.

F_{Ag} is a perfect DNF, that can be transformed to reduced DNF using the generalized Quine algorithm modified for JA_3 logic. Thus, the axioms of generalized gluing and absorption are formulated, respectively, as follows (here C, C_1, C_2, C_3 are maximal conjunctions of the JA_3 logic, p is a variable):

(*a*) $(J_1 p \& C_1) \vee (J_{-1} p \& C_2) \vee (J_\tau p \& C_3) \leftrightarrow (J_1 p \& C_1) \vee (J_{-1} p \& C_2) \vee (J_\tau p \& C_3) \vee (C_1 \& C_2 \& C_3)$;

(*b*) $J_\nu p \vee (J_\nu p \& C) \leftrightarrow J_\nu p$.[4]

Applying successively (*a*) and (*b*) to the $F_{Ag} = (\varphi_1 \vee \ldots \vee \varphi_r)$ until their applicability stops, we obtain reduced DNF $(\chi_1 \vee \ldots \vee \chi_h)$ with corresponding implicants set $\{\chi_1, \ldots, \chi_h\}$. We assign to each implicant χ_t $(t = 1, \ldots, h)$ such a set Ag_t of agents $X \in Ag$ that their activity $ACT'_X = \varphi$ is covered by the implicant χ_t, $Ag_t = \{X \mid \chi_t \sqsubset ACT'_X\}$.

---

[4] Note that generalized gluing and absorption can also be formulated for logics E_3 and L'_3, but the axiom of absorption does not hold in B_3.

Let's construct the set of analytic tableaux $\tilde{\mathcal{T}} = \left\{\mathcal{T}_{\Sigma\cup\{\chi_1\}}, \ldots, \mathcal{T}_{\Sigma\cup\{\chi_h\}}\right\}$ and choose such $\chi_{j_1}, \ldots, \chi_{j_m}$, where $\neg consis(\Sigma\cup\{\chi_{j_l}\})$ holds, i.e., the analytical table $\mathcal{T}_{\Sigma\cup\left\{\chi_{j_l}\right\}}$ with the root $\Sigma\cup\{\chi_{j_l}\}$ is closed, $l = 1, \ldots, m$. Then Ag_{j_l} agents are not rational, and their activities in *En* environment should be blocked.

It should be noted that the choice of one of the two described procedures for constructing a set of rational agents depends on their comparative efficiency, which is determined separately in each case.

4. Conclusion

One often sees agent approach in AI as a universal one. However, diversity of conceptions of essential features, attributes, powers of the agents often makes this approach speculative. Gradual progression from Intelligent Systems to Cognitive Systems and further, to Intelligent Robots, appears promising. Intelligent Robots are a type of Intelligent Agents imitating and reinforcing certain intelligent capabilities that designate phenomenology of natural intelligence [22, pp. 99–121].

Clarifying the notion of rationality by methods of logics contributes to exercising by an agent of one of the core powers of natural intelligence – the power of argued decision-making. This power allows us to talk about a rational choice of action. Owing to the suggested methods, we can distinguish agents whose actions are adequate to the features of an environment, and avoid use of actions inadequate to it. Further development of these methods may result practical for describing, researching and understanding of Multi-Agent Systems as well as of social systems and society.

5. References

[1] M.J. Wooldridge, A Brief History of Artificial Intelligence, Flatiron Books, New York, 2021.

[2] S. Russell, P. Norvig, Artificial Intelligence: A Modern Approach (4th ed.). – NJ: Pearson, 2020.

[3] Wooldridge M.J. An Introduction to Multiagent Systems, 2nd ed., John Willey&Sons, Chichester, United Kingdom, 2009.

[4] V.B. Tarassov, Ot Mnogoagentnykh System k Intellektual'nym Organizatsiyam: Filosofia, Psikhologiya, Informatika [From Multi-Agent Systems to Intelligent Organizations: Philosophy, Psychology, Informatics], URSS, Moscow, 2002.

[5] O.M. Anshakov, T. Gergely, Cognitive reasoning. A formal approach, Springer, Berlin, Heidelberg, 2010.

[6] C. Walton, Agency and the Semantic Web, Oxford University Press, NY, 2006.

[7] R. Calegari, G. Ciatto, V. Mascardi, A. Omicini, Logic-based technologies for multi-agent systems: a systematic literature review, Auton Agent Multi-Agent Syst 35(1) (2021). doi: 10.1007/s10458-020-09478-3.

[8] D. Kahneman, Thinking, fast and slow, Farrar, Straus and Giroux, NY, 2011.

[9] M. A. Boden, AI. Its nature and future, Oxford University Press, Oxford, 2016.

[10] W. van der Hoek, & M. Wooldridge, Logics for Multiagent Systems, AI Magazine, 33(3) (2012), 92–105. doi: 10.1609/aimag.v33i3.2427.

[11] V.K. Finn, M.A. Mikheyenkova, Ob odnom podkhode k raspoznavaniyu ratsional'nosti v kollektivakh agentov [On the Approach to Rationality Recognition in Agents Systems], Iskusstvennyj Intellekt i prinyatie reshenij [Artificial Intelligence and Decision Making], 3 (2010) 20–32.

[12] T. Parsons On theory and metatheory, Humboldt Journal of Social Relations, 7(1) (Fall/Winter 1979/1980), 5–16.

[13] H.-D. Burkhard, Liveness and Fairness Properties in Multi-Agent Systems, in: Proceedings of the 13th IJCAI, volume 1, Chambery France, 1993, pp. 325–330.

[14] V.K. Finn, O.M. Anshakov, R.Sh. Grigolia, M.I. Zabezhailo, Many-valued Logics as Fragments of Formalized Semantics, Acta Philosophica Fennica, 35 (1982) 239–273.

[15] V.K. Finn, M.A. Mikheyenkova, Plausible Reasoning for the Problems of Cognitive Sociology, Logic and Logical Philosophy, 20 (2011) 113–139.

[16] D.A. Bochvar, V.K. Finn, Nekotorye dopolneniya k stat’yam o mnogoznachnykh logikakh [Some Additions to Articles on Many-Valued Logics], in: V.K. Finn (Ed.), Mnogoznachnye logiki i ikh primeneniya [Many-Valued Logics and their Use], volume 1: Logicheskie ischisleniya, algebry i funktsional’nye svoistva [Logical Calculi, Algebras and Functional Features], URSS, Moscow, 2020, pp. 111–171.
[17] I. Rahwan, G.R. Simari (Eds.), Argumentation in artificial intelligence, Springer, Dordrecht, Heidelberg, London, New York, 2009.
[18] V.N. Vagin Znaniya i ubezhdeniya v intellektual’nom analize dannykh [Knowledge and Believe in Knowledge Discovery], Fizmatlit, Moscow, 2019.
[19] M. Dastani, H. Dong, L van der Torre (Eds.), Logic and Argumentation, volume 12061 of Lecture Notes in Artificial Intelligence, Springer, Switzerland, 2020. Proceedings. doi: 10.1007/978-3-030-44638-3.
[20] V.K. Finn, M.A. Mikheyenkova, O logicheskikh sredstvakh conceptualiztsii analiza mnenij [On Logical Means of Opinion Analysis Conceptualizing], in: V.K. Finn (Ed.), Mnogoznachnye logiki i ikh primeneniya [Many-Valued Logics and their Use], volume 2: Logiki v sistemakh iskusstvennogo intellekta [Logics in Artificial Intelligence Systems], URSS, Moscow, 2020, pp. 152–199.
[21] V.K. Finn Standartnye i nestandartnye logiki argumentatsii [Standard and Non-Standard Logics of Argumentation], ibid., pp. 59–91.
[22] V.K. Finn, Exact Epistemology and Artificial Intelligence, Autom. Doc. Math. Linguist., 54(3), 2020 140–173. doi: 10.3103/S0005105520030073.