Truth and Cognitive Presumptions

Pavlova A. M. (*Saint Petersburg*)

Introduction

The PhD project 'Truth and Cognitive Presumptions' is aimed at analysing logical features of cognitive presumptions of rational agents studied and modelled in various logical theories, as well as at comparing them from the perspective of a possibility of joint following the wining strategies of interaction. Given that, one can define the following tasks:

1. Show that different logical theories modelling rational agents interactions ground this interaction upon certain hypotheses closely related to interpretations of the central logical notions, i.e. truth (logical value) and deducibility. Those hypotheses are often implicit and one of the aims of the project is to reveal and clarify them;
2. Find out the connection between the formal logical aspects of the above mentioned hypotheses and the way cognitive notions of agents concerning interaction and their presumptions about other agents are based on the logical notions of truth and entailment;
3. Analyse the possibility of interaction between various agents with different cognitive presumptions in the game-theoretical framework.

In order to accomplish the above tasks we shall use different logical theories: game-theoretical approach (mostly represented by the dialogue logic and IF-logic with its game-theoretical semantics), BDI logic (Belief-Desire-Intention logic) and dynamic epistemic logic. Generally speaking the two main approaches will be considered: the game approach towards the notion of truth and entailment and multi-agent systems that formalise cognitive presumptions that help us to build an inference. In the present paper we shall present some of the results that has already been reached in this respect and outline the ongoing work and its difficulties.

To begin with we should outline the different aspects of the issue of the diversity of agents and distinguish it from other logical sources of variety. We shall have a look at the variety from different perspectives, including: the distinction between the diversity of agents and the one of logics, different levels and sources of variety agents.

Variety of Logics vs Variety of Agents

To begin our overview of different agents with regard to their cognitive presumptions we should have a look at the difference between logics and agents. Now it seems obvious for us that there are more that one possible logical system. However, if we switch to logics with more that one agent we get a number of questions that arise with respect to what are the basis for an agent to act, whether he or she has a default logic inside and what does it look like. To approach these questions we should first have a look at different logical systems that tried to model some sort of interaction between agents from the point of view of a game or a dialogue.

Dialogue Logic

First of all we shall have a look at the dialogue logic by P. Lorenzen [19] and K. Loreanz [20] that was later developed by E.Krabbe [15]. Dialogue logic represents an *operationalist approach* to logic and understanding of truth. Here are some properties of Dialogue logic:

1. Proof-theoretical approach
2. Deals with *validity*, or truth in all models (in terms of model-theoretical semantics). A first-order sentence A is *dialogically valid* in the classical (or intuitionistic) sense, if all finished plays belonging to the classical (resp. intuitionistic) dialogue $D(A)$ are *closed*[26].[1]
3. Operational presentation of truth (purely deductive and constructive methods suffice)
4. *Two levels of rules*:
   — Logical rules;
   — Structural rules.
5. In a strict sense the deducibility is defined, not validity[2], though some authors treat it not as logic but as a model [1].
6. *Operationalism*: the meaning of logical constants is defind by an operation we can conduct with it
7. Considering the above the notion of truth itself is *redundant* as strictly speaking there is a deducibility condition, - the existence of the winning strategy for Proponent - not truth condition.

Game-Theoretical Semantics

Now we shall pass to the game-theoretical semantics. *Game-Theoretical Semantics* was invented by J. Hintikka and then developed by G. Sandu as an alternative to model-theoretical semantics of Tarski-type. It is commonly used for IF logic and other branches of logic, like *Diagrammatical logic* invented by Peirce and nowadays studied by Pietarinen [23]. Here are some properties of *Game-Theoretical Semantics*:

1. Semantical (model-theoretical in a sense pointed out earlier) approach. *Game-theoretical* presentation of truth (an alternative to model-theoretical way of setting up semantics);

[1]if there are models at all, as it is worth mentioning that in the dialogue logic account there are no models as special elements of the system. However, one may treat dialogue logic as a model construction game, as the Proponent tries to build a counter-model. See [3]

[2]We should note that J. van Benthem claims that in model construction games to which he attributes Lorenzen's games there is not just one, but double task as both players are in a sense *positive*, as both try to accomplish some ask in a game associated with some formula:

— proof of the formula

— model construction

2. *Truth in a given model*:
   — A is true in M, in symbols $M \models^{+}_{GTS} A$, if there exists a winning strategy for *Eloise* in $G(A, M)$.
   — A is false in M, in symbols $M \models^{-}_{GTS} A$, if there exists a winning strategy for *Abelard* in $G(A, M)$.
3. *Two levels of rules*:
   — *Rules of Semantical Game*, defining roles and possible moves of the players and also the wining function;
   — *Metarules*, defining the basic notions of strategy, winning strategy and truth.
4. Truth conditions are defined[3]
5. There is an objective consisting in defining truth, notably:
   — defining *truth conditions* - an existence of the winning strategy for initial verifier;
   — defining *truth predicate* on the basis of truth conditions in the form of *Skolem functions*.
6. Is bound to the *particular model*
7. Notion of truth is significant as a part of the project of refutation of *Tarski's impossibility result*
8. The aim is to find witness that will satisfy the formula in a model.

Correspondence between Intuitionistic Dialogues with Hypotheses and GTS

After giving a brief presentation of the main features of the systems mentioned above we should say that there is a *correspondence between Intuitionistic Dialogues with Hypotheses and GTS*. In order to establish correspondence between Dialogues and Semantical Games we should transform our initial dialogues in such a way that the will be encoded a model by purely operational and proof-theoretical means, that is invent hypotheses at the initial step, as there is a restriction on moves for Proponent as regards the atomic formulas, i.e. **(SR-5) (Formal use of atomic formulas)**[4][26] The hypotheses should be of the following form:

$$H = p \vee \neg p$$

for all p, being propositional formulas, that occur in the formula A. So the initial Proponent's move is of the following form:

$$? - p \vee \neg p$$

[3]Truth conditions in GTS are expressed in the form of winning strategies and, thus, by formulae in a Skolem form, that is Σ^1_1

[4]P cannot introduce positive literals: any positive literal must be stated by O first. From this it follows that a dialogue about an atom cannot have non-empty plays. Positive literals cannot be attacked.

and the corresponding defence of Opponent is the choice of the *right or left disjuct* which corresponds the model of the Semantical game. Thus, now our dialogue is of the form $D(A, H)^5$.

Having transformed our dialogues in such a way we can now establish the following correspondence[6].

Theorem 1. *Let A be any formula of propositional logic. The following conditions are equivalent:*

i There is a wining strategy for Proponent in $D(A; p_1 \vee \neg p_1, ..., p_n \vee \neg p_n)$*;*
ii For every M*, there is a winning strategy for Eloise in* $G(A, M)$*;*

where $p_1, ..., p_n$ *are the propositional atoms appearing in A. Furthermore, there is an algorithm turning winning strategy of Proponent into a family of winning strategies for Eloise, and vice versa.*

This theorem is successfully proved in [26]. Thus, one can conclude that there is correspondence when we talk about truth in a particular model.

Correspondence between Intuitionistic Dialogues and GTS without a Model

Looking at the result presented above, one can think of trying to model the similar correspondence, however with regard to validity. To do so we shall transform the very semantic games in order to be able to show validity as truth in every model. One should note here that above we established a correspondence between the two systems without implication. The case of implication is not as trivial as one might think as we are dealing with the intuitionistic logic (in connection to this we shall consider an example below).

We shall reorganise the rules of a semantic game as follows:

— In the network of one game there can be played more than one subgames[7], so the players can switch the branches, but except for the players being not able to chose another disjunct after having chosen once (that corresponds to the rule **(SR-1.I) (Intuitionistic round closing rule)**);
— We invent a new structure which is in some way similar to the model that is called the *Commitment store* that tracks all the formulae (which are in fact the subformulae of initial formula in question) which the players committed themselves to, that is which they stated.
— Eloise cannot state literals, so that they will be put in her Commitment store, if there is no such a literal in the Commitment store of the Abelard.

[5]Note that in case of FOL we should use the hypotheses of the following form[26]:

$$\forall x_1 ... \forall x_n (E(x_1 ... x_n \vee \neg E(x_1 ... x_n)$$

[6]Note that the formulae should be in *negation normal form*

[7]That corresponds to the **(SR-2) (Branching rule for plays)** and **(SR-3) (Shifting rule)** in Intuitionistic dialogue.

So one can see that Commitment store serves to construct a model, thus the semantic game is transformed from the game of verification into the model-construction game[8]. Let C be a commitment store, then

$$C = C_{\exists} \cup C_{\forall},$$

where $C_{\exists} = p_1, p_2, ..., p_i, ..., p_n$, where p_0 is initial formula and $p_i \in P$, P being a set of Eloise's statements; $C_{\forall} = q_1, q_2, ..., q_l, ..., q_m$, where $q_l \in Q$, Q being a set of Abelard's statements.

As in the initial semantical game the moves of the players were defined by the main logical operator, so we cannot define the winning conditions as it is done it dialogue logic.

Definition 1. There is a closure in the commitment store C of the semantic game $G(A, C)$ iff one and the same atomic formula (a) $a \in C_{\exists}$ and $a \in C_{\forall}$.

Definition 2. In a semantical game with commitment store $G(A, C)$:
— Eloise wins iff there is a closure in the commitment store.
— otherwise Abelard wins.

Let us consider an example. Let $A = (\neg p \wedge q) \vee (p \vee \neg q)$ Consider first a dialogue $D(A)$:

Proponent	**Opponent**
1. $(\neg p \wedge q) \vee (p \vee \neg q)$	
3. $p \vee \neg q$ (2)	2. $? - \vee$ (1)
5. $\neg q$ (4)	4. $? - \vee$ (3)
	6. q (5)

It is easy to notice that Opponent has a winning strategy in the dialogue, due to the **(SR-1.I) (Intuitionistic round closing rule)**, as if it was a classical one, the proponent would have a winning strategy[9].

Let us now consider the corresponding semantic game $G(A, C)$. The commitment store of the game will be as follows:

$C_{\exists}$	$C_{\forall}$
I. $(\neg p \wedge q) \vee (p \vee \neg q)$	
II. $\neg p \wedge q$	
III. $\neg q$	q

There is no closure in the commitment store C, thus Abelard has a winning strategy in the game $G(A, C)$.

[8]However, if one may see the same is may be done with the dialogue, as we can also associate a model with it, but the method we chose in the previous section still reserves the very nature of the dialogue in the sense that it uses deduction to check the formula.

[9]One should mention that as it is a finite game with *zero sum*, according to the *Gale-Stewart theorem* such dialogues are *determined games*

One can easily see that there is a strict correspondence between the table of the dialogue $D(A)$ and the commitment store of the semantic game $G(A, C)$. Thus it may be argued that given a winning strategy for Proponent (or Opponent) in the intuitionistic dialogue $D(A)$, we can get a winning strategy for Eloise (or Abelard respectively) following a finite algorithm. Thus, we can define the validity (truth in all models) within $G(A, C)$.

We have just seen that we can represent game-theoretical truth in dialogue logic and vice versa. However, what does it tell us about the status of agent in those games? As for dialogue logic, one can see that the structural rules influence the type of validity (or truth) that we finally get in the system. That means that both players should have one and the same idea of truth and inference rules as a precondition. Thus it does not allow agents to have different ideas about basic logical notions, on the contrary, we get those ideas by constructing the game. Some scholars claimed that dialogue logic is just another variant of Beth analytic tableaux. However, the variant of the dialogue with hypotheses gives a new incentive to the understanding of the dialogue logic as a game. We can look at those hypotheses from two different perspectives:

— the hypotheses represent a *common knowledge* (cf. a set K_c-set[10], which *'is the common state of knowledge of those present at the disputation'* [5, 4], that we find in the disputations *de obligationibus* and their formalisation by C. Dutilh Novaes);
— the hypotheses represent *personal beliefs* of the *Opponent*. Then we can view the dialogue as a procedure checking the *compatibility* of the opponents beliefs with the formula in question, when the *Proponent* tries to build a corresponding model.

As for the game-theoretical semantics approach we shall see that that tree-form of a game is used for epistemic games with different kinds of information. However, what is important is that neither dialogue logic, game-theoretical semantics does not allow the agents to have different default logical systems and, thus, notions of truth and entailment. That is so due to the fact that instead of inference rules we have structural rules in those systems that define the meaning of the logical connective, operator or quantifier according to moves that players can make, given the formula with the connective in question as the main one. However, that does not mean that we cannot have this property in a system. Imagine the situation where different agent have different ideals about what validity means, for example one agent might believe in classical logic, but the other one prefers the intuitionistic notion of validity. That is possible in the system of formal dialectics proposed by Ch. Hamblin [9]. However, his system gives rise to a eat number of questions as it has no semantics and the rules are not specified in an accurate way.

[10]The signature used here was proposed by Catarina Dutilh Novaes, for instance in [5]

Diversity of agents and some questions about their interaction

As we have already seen that different logical systems give us different understanding of truth and entailment, thus, different types of rationality. We claim that the same principle is applicable when we talk about diversity of agents. What we mean here is that the approach itself to some extent defines the way in which we shall look at the diversity of agents, its sources and the way to model them. Thus we want to emphasise that cognitive presumptions can be seen from different perspectives that can be divided into the following groups:

I Epistemic and doxastic;
II Intentions of the agents in interaction (modelled in BDI logic);
III Logical systems that lie behind the resoning of different agents that is represented in different understanding of truth and entailment.

Let us have a look at some of those general groups.

Basic agent diversity classification Given the epistemic and doxastic approach, we can define different types of agents studying their abilities to reason. Following [16] we define different types of agent on the basis of certain parameters. The following diversity if applicable to the agents from the viewpoint of the *epistemic logic* and *imperfect information games.*

a *inferential power*, i.e. the ability to make all necessary proof steps. The feature is question is usually associated with the disputable formula of logical omniscience: $K(\varphi \to \psi) \to (K\varphi \to K\psi)$;
b *introspection* i.e. the ability to make statements about one's own epistemic condition. The formulae encoding this feature ot devoid of uncertainty: (i) $K\varphi \to KK\varphi$, (ii) $\neg K\varphi \to K\neg K\varphi$;
c *observation*, i.e. different powers for observing events (limitations on the observation capacity usually give rise to so-called 'imperfect information');
d *memory*, i.e. limitations on the storing of the number of the last events observed;
e *revision policies*, i.e. different types of revision, from the conservative to the radical one[11].

When we say that the diversity arises from the variation on the parameters mentioned above, we mean that there are perfect agents with the respect to those variables, for instance, agent with *Perfect Recall*, i.e. having ideal memory; or *omniscient agent*, that is, the one who for whom the logical omniscience formula is satisfied. One can see that all those types of diversity can be dealt with given various modifications of the dynamic epistemic logic, except for [e] as *belief revision theory* is needed there.

[11] One should have noticed that this type does not belong to the epistemic logic but to the doxastic logic. However, it can also be analysed from the game-theoretical point of view, for instance, using *defective information games* [14].

Various epistemic games When we talk about games and information that the players possess we can also single out several types of games based of a variation of some of the following parameters with respect to the information about:

— the structure of the game;
— history of the current run of the game (that might include previous states $s, t, ...$ and actions $a, b, ...$);
— the current state s the player is in (relevant for the perfect/imperfect information games distinction);
— the future course of rent run of the game (usually is assumed as not available, though, we can introduce the notion of *decided games* for those where the agents possess this sort of information).

As the terminology is concerned, we use the one presented in [14] as follows:

Definition 3 (**Perfect Information Game**)**.** A *perfect infromation game* is a game where each player knows:

— what the structure of the game is like;
— what has happened up to the current state

That means that the players know everything about the game and its run, however, they have no access to the information concerning the future actions of the other players.

Definition 4 (**Imperfect Information Games**)**.** An *imperfect information game* is a game where:

— each player knows the structure of the game;
— a player might be uncertain of which moves were taken up to the current state, thus he might be in a situation where he considers more than one state as a possible candidate for the current state.

In [14] it is proposed to consider all perfect information games as a subclass of the imperfect information games due to structure of the *tree-form* semantics for both of those types. In any case both imperfect and perfect information games belong to the type of *complete information games*.

Definition 5 (**Complete Information Games**)**.** A *complete information game* is the a game where the structure of the game is known by *each* player.

Definition 6 (**Incomplete Information Games**)**.** An *incomplete information game* is a game in which the structure of the game *need not* be known by every player.

Similar to the case of imperfect information games *complete information games* are believed to be a special case of *incomplete information games* due to the same reason.

Given these definitions, we can see that when it comes to the *observation* capacity of an agent, we can divide it two deparate characteristics of an agent:

1. observation of the previous states and actions as was noted above that is expressed by equivalence relation $\sim_i$, where i is an agent, between states $s \sim_i t$ or actions $a \sim_i b$;
2. awareness of an agent about the structure of the game.

Additional types of agent variety We can also add two more basic sources of agent diversity that seems to cause big problems when it comes to model construction:

— In the framework of epistemic and doxastic logic we can also distinguish agent on the basis of their *access to the other agent's knowledge or believes* (including agent's awareness about the structure of the game, which is modelled in [14])

— As we have also mentioned before we can imagine a situation when different agents have not only different knowledge and beliefs, but also different notions of truth and entailment as basic, i.e. their default logical systems differ. However, it is possible for a dialogue or a game where actions are not determined (and are not strictly defined by the logical connectives). This kind of diversity is the most problematic with respect to the agents interaction. The reason is that we need at least one winning strategy for effective communication or at least the notion of strategy to characterise the interaction. Still it would not lack some interest if we tried to think of the possible ways of representing a possibility of having different presumptions about truth and entailment in semantics and syntax, however, it is yet unclear which approach suits this goal better.

So far we have considered the diversity of cognitive positions of rational agents, thus there is still remaining two important tasks:

— to formally represent different types of agents on the level of syntax and semantics (a large part of this work has been done in the epistemic and doxastic logics, for instance by [4], [14], [16], [17], and in the BDI approach [21] which should also be included into our discourse);

— to find and model a possibility of effective interaction of agents.

Definition 7 (**Effective Interaction**)**.** By the *effective interaction* we understand here a question of existence of at least one winning strategy for agents.

It is important to find out whether there is a connection between the existence of a winning strategy and establishing of some sort of concord between agents with respect to semantics and procedure.

Conclusion

In conclusion we would like to emphasise that the question of diversity in multi-agent systems can be approached in two different ways:

— It might be seen and, thus, modelled as a diversity between systems within which all the agents have equal abilities and ideas about logical notions, though, their actual knowledge and beliefs may differ;

— We can represent it as a diversity between different agents that interact (and not only co-exist) inside one system. This understanding of diversity is of the greatest interest to us.

All sources of diversity can be generally divided into three groups, two of which we have already discussed here, whereas the third one should be added yet:

— Epistemic conditions: knowledge, beliefs and awareness;
— Intentions and preferences;
— Logical notions of truth and entailment (thus, logical systems used as basic)

Given those major groups of sources for diversity it is interesting to see that in the former two groups we have at least some universal rules for the system which are determined by the logic that is shared by all the agents, whereas for the case of diversity based on different notions of truth it is not so obvious what might be the universally working there, except for some rules that limit and determine the procedure. That is something that should also be investigated in the course of our future work on the project.

Finally, it is worth mentioning that the research presupposes not only detailed analysis of the existing approaches but it is intended to propose a model for the interaction of agents given the variety of cognitive presumptions of agents.

Bibliography

[1] Alama J. and Uckelman S. L.: Lorenzen Dialogue Games as Logical Semantics (2011)

[2] Barth E. M.: From Axiom to Dialogue. A philosophical study of logics and argumentation, Walter de Gruyter, Berlin, New York (1982)

[3] van Benthem J.: Logical Construction Games. *Acta Philosophica Fennica*

[4] van Benthem J, Liu F.: Diversity of Logical Agents in Games, *Philosophia Scientiæ,* 8 (2), 2004, 165–181. 78, T. Aho and A-V Pietarinen, eds., Truth and Games, essays in honour of Gabriel Sandu, pp. 123-138 (2006)

[5] Dutilh Novaes C.: Medieval Obligationes as Logical Games of Consistency Maintenance. *Synthese 145(3)*, pp. 371-395 (2005)

[6] Eklund M.: On How Logic Became First-Order, *Nordic Journal of Philosophical Logic, Vol. 1, No. 2*, Scandinavian University Press, pp.147-167 (1996)

[7] Eklund M. and Kolak D.: Is Hintikka's Logic First-Order?, Synthese 131(3), pp. 371-388 (2002)

[8] Gödel K.: Uber die Vollstandigkeit des Logikkalkuls (1929) (On the completeness of the calculus of logic (1929)). In Kurt Godel *Collected Works* Volume I, Publications 1929-1936, Oxford University Press, pp. 61-101 (1986)

[9] Hamblin Ch.: Fallacies, London. (1970)

[10] Hintikka J.: The Principles of Mathematics Revisited, Cambridge University Press, Cambridge (1996)

[11] Hintikka J.: Post-Tarskian Truth, Synthese 126. Kluwer Academic Publishers. Printed in the Netherlands, pp. 17-36 (2001)

[12] Hintikka J. and Sandu G.: What is Logic? In D. Jacquette (ed.), *Handbook of the Philosophy of Logic*, Elsevier, Amsterdam, pp. 13-40 (2007)

[13] Hodges W.: Compositional Semantics for a Langauge of Imperfect Information, *Logic Journal of the IGPL*, 5, pp. 539–563 (1997)

[14] Hötte T.: A Model for Epistemic Games. Marter thesis (2004)

[15] Krabbe E. C. W.: Dialogue logic. *Handbook of the History of Logic*, Volume 7, pp. 665-704 (2006)

[16] Liu F.: Diversity of Agents and their Interaction (2007)

[17] Liu F.: Changing for the Better: Preference Dynamics and Agent Diversity. PhD thesis (2008)

[18] Lorenz K.: Arithmetik und Logik als Spiele. PhD Dissertation, Christian-Albrechts-Üniversität, Kiel. Selections reprinted in [20] (1961)

[19] Lorenzen P.: Constructive philosophy (1987)

[20] Lorenzen P. and Lorenz K.: Dialogische Logik. Wissenschaftlische Buchgesellschaft, Darmstadt (1978)

[21] [2] Meyer J.-J. and Veltman F. (2007). Intelligent agents and common-sense reasoning, in P. Blackburn, J. van Benthem and F. Wolter (eds), Handbook of Modal Logic, Elsevier, pp. 991–1029.

[22] Pietarinen Aht.: Varieties of IFing (2001)

[23] Pietarinen Aht.: Existential Graphs: What a Diagrammatic Logic of Cognition Might Look Like, 'History and Philosophy of Logic' 32 (3):265-281 (2011)

[24] Rahman Sh.: Remarks on Dialogical Meaning: A Case Study, pp. 1-30 (2011)

[25] Rahman Sh., Ruckert H. ≪Preface≫, Synthese 127: 1–6, 2001.

[26] Rahman Sh., Tulenheimo T.: From Games to Dialogues and Back: Towards a General Frame for Validity. In Ondrej Majer, Ahti-Veikko Pietarinen and Tero Tulenheimo (editors), *Games: Unifying logic, Language and Philosophy.* Springer (2009)

[27] [3] Rao A. S. and Georgeff, M. P., 1991, "Modeling rational agents within a BDI-architecture", San Francisco: Morgan Kaufmann, pp. 473–484.

[28] Reed Ch and Walton D.: Argumentation Schemes in Dialogue, In H.V. Hansen, et. al. (Eds.), *Dissensus and the Search for Common Ground*, CD-ROM (pp. 1-11), Windsor, ON: OSSA (2007)

[29] de Rouihan P., Bozon S.: The Truth of IF: Has Hintikka Really Exorcised Tarski's Curse? *The Philosophy of Jaakko Hintikka*, The Library of Liveing Philosophers, v. XXX, pp. 683-705 (2006)

[30] Stegmüller W.: Remarks on the completeness of logical systems relative to the validity-concepts of P. Lorenzen and K. Lorenz. Notre Dame J. Formal Logic Volume 5, Number 2, pp. 81-112 (1964)

[31] Uckelman S. L., Johnston Sp.: A Simple Semantics for Aristotelian Apodeictic Syllogistics, pp. 454-469

[32] Weingartner P.: Basic Questions on Truth. Kluwer Academic Publishers (2000)

[33] Wells S. and Reed Ch.: Formal Dialectic Specification, *Argumentation in Multi-Agent Systems First International Workshop*, pp. 31-43 (2004)