Some Useful Social Metaphors in Logic

Ahti Pietarinen

Abstract— **The purpose of this paper is to register a couple of novelties related to social conceptions and social metaphors in logic. As an overarching methodology, game-theoretic notions will be used. The purpose is not to try to explore how logical methods can be made to apply to actual human behaviour in everyday scientific reasoning and social contexts in science, but rather to (i) see how game-theoretically inclined factors involving sociality can be made to apply to logic, and to (ii) discuss what the nature and current status of these social metaphors in the context of logical investigations are. It is shown that issues such as consistency, team theory and bounded rationality are of increased importance in logic. The general logical framework where these notions are investigated is provided by game-theoretic semantics (GTS) and extension of classical first-order logic known as "independence-friendly" (IF) logic.**

Keywords— **Social metaphors, IF logic, semantic games, consistency, team theory, rationality.**

I. Introduction

THE topic of this paper can best be delineated by a quotation from C.S. Peirce, one of the founders of modern concept of logic. In (Peirce 1931-5, 4.240), he remarked "Formal logic, however, is by no means the whole of logic, or even its principal part. It is hardly to be reckoned as a part of logic proper. Logic has to define its aim; and in doing so is even more dependent upon ethics, or the philosophy of aims, by far, than it is, in the methodeutic branch, upon mathematics."[1] This is why in this paper we do not propose to study logic in its own sake. Instead, the quote is meant to be instructive in pointing out the generality of the overall science of logic beyond the purview of its purely formal or mathematical use. The purpose here is thus to investigate, in the light of recent advancements in the philosophy of logic, a couple of aspects as to what directions these general viewpoints may take. Among these recent advancement is the idea of extending classical concepts of logic to "independence-friendly" ones.

II. Independence-friendly logic and semantic games

In contrast to classical logic, the recently proliferated approach in logical philosophy and philosophy of logic known as "independence-friendly" (IF) logic is meant to create a logic where the flow of semantic information in formulas is no longer perfect. Following [4], this can be generated by a special slash notation. For example, in the sentence $\phi = \forall x(\exists y/x)\, Sxy$, the choice of the value for the universally quantified variable x is not taken reach the point in ϕ where the choice of the value for the existentially quantified variable y is being made. This idea can be made precise in the framework of game-theoretic semantics (GTS) [5], [15], [18], especially in the extensive-form representation of such semantic games that the formulas of IF logic give rise to (for details, see [20]). Without going into the details of this story of how to marry IF logic and GTS, suffice it to note that the idea uses the concept of an information set in a very similar way as in the traditional theory of games. Information sets are in essence equivalence relations between those histories or situations that the player is not supposed to be able to distinguish with respect to his or her current situation in the game (for some inevitable shortcomings of these notions in the traditional theory of games, see [17] though). As far as the basic format of semantic games is concerned (either in association to classical or IF logic), the game rules prescribing the legitimate game rules remain the same. The novelty is that the strategies that the players have an access to are defined on reduced argument sets, as some previous choices may not be visible to later parts of the game. The basic game rules for the first-order logic (again, either IF or classical) are such that conjunction and universal quantifier prompt a move by the player that tries to falsify a formula (called the Falsifier, Nature, Abelard), and disjunction and existential quantifier prompt a move by the player that tries to verify it (the Verifier, Myself, Eloise). Each move reduce the complexity of the formula, and when an atomic formula is reached, if it is true in the given interpretation, the verifying player wins, and if it is false in the interpretation, the falsifying player wins. The truth of the sentence is captured by the notion of a winning strategy: the whole sentence is true if and only if there exists a winning strategy for the verifying player, and likewise, the whole sentence is false if and only if there exists a winning strategy for the falsifying player.

III. Consistency and games

As soon as we have a game-theoretic framework at hand that functions as a semantics for logic, various ways of setting up these games, or choosing between different game characteristics, will inevitably follow. For example, it is known [4] that semantic games for IF logic are not determined, which means that neither of the players will have a winning strategy. A consequence of this to logic is that the law of excluded middle fails. (In another nomenclature, one would say that the logic is partial, that is, contains a truth-value of Undefined.) The important thing to note in non-determinacy is that instead of being classical and behaving contradictorily, the negation sign in IF logic is game-theoretic, defined by a role-switch between the two players.

Yet if this much is the case, why is not the law of non-contradiction invalidated? The reason is that the games have previously been assumed to be *strictly competitive*,

Author address: Department of Philosophy, University of Helsinki, P.O. Box 9, FIN-00014 University of Helsinki. E-mail: pietarin@cc.helsinki.fi.

[1]The reference CP is to [13] by a volume and paragraph number.

that is, both players cannot come out as the winners. This means that in any semantic game, there can be no winning strategies for both players. However, as noted this holds only if the class of games is restricted to contain only those games that are strictly competitive. Yet it is perfectly feasible to relax this assumption and take some semantic games to be *non-strictly competitive*. This now means that the following no longer holds: If there exists a strategy f that is winning for the Verifier then there does not exist a strategy g that is winning for the Falsifier, and if there exists a strategy g that is winning for the Falsifier then there does not exist a strategy f that is winning for the Verifier. If the game is not strictly competitive, call it non-strictly competitive. In non-strictly competitive games, it may happen that there exist a winning strategy for both players. To implement this, one can for instance stipulate that there may be some terminal histories Z that are winning for both players. This can be denoted by a game-theoretic devise of using the payoff function $u_i(h)$ that outputs the matrix $(1, 1)$ for a history $h \in K$, in addition to the zero-sum matrices for some other payoffs. Consequently, given a literal p, it will be interpreted so that it has both the truth-value True and the truth-value False, and hence will have a truth-value of Over-defined.

The abundance of the classes of non-strictly competitive games in the theory of games supports the fact that in logic where game-theoretic concepts are quite commonplace, one should not rule such classes of games out offhand. Just as physical instances of games can be used as evidence for the claim that games in logic may encompass imperfect information and failures of the law of excluded middle, the other basic class of games in logic may spring into existence in a comparable manner.[2] The other effect is that non-strict games may be determined even if their strict counterparts are not. Non-determinacy can in that case be restored only by introducing partial models into the language (see [16]).

By the same token, it should be mentioned that some systems of paraconsistent logic admit that non-atomic but inconsistent formulas are trivial, which means that anything can be derived from them and thus non-triviality holds only for atomic inconsistencies [22]. The game-theoretic perspective does not endorse this, as games can transmit inconsistencies from non-constant-sum payoffs of atomic formulas to complex ones via suitable winning strategies. In this case the existence of imperfect information may nonetheless affect this transmission.

Related to this is the so-called Jaśkowski's problem [8], which says that any logic claiming the name of a paraconsistent one needs to satisfy three conditions. First, such a logic has to come with a one-place operator leading to a decent paraconsistent system (that is, to an inconsistent but non-explosive one). Second, its negation must be strong enough to be called negation. Thirdly, its semantics needs to be well motivated.

This problem has proved elusive thus far, as the received paraconsistent logics fall short of having a satisfactory semantic explication. Now the previous remarks are calculated to provide pointers to a comprehensive answer to this problem, however. For GTS is a well-motivated and systematic semantic method for all kinds of logics which admit of coherent definitions of game-theoretic rules, many of the logics being "non-classical" ones. What is more, the previous attempts to solve this problem by paraconsistent systems in the literature remain to be based on some negative criteria — for example, they describe principles that must be rejected, such as ex falso, consistency, or triviality. Yet the game-theoretically defined negation is a genuine negation, what can be observed for example from the relation it has with negative constructions and negative operators in natural language [16].

IV. Team Theory: toward decentralised processing in logic

As soon as we allow unrestricted notation in representing various ways of expressing variable dependencies and semantic information flow within formulas, IF logic becomes equipped with a way of capturing the phenomenon of forgetting information — or the phenomenon of imperfect recall, as game theorists would perhaps prefer to say. Without going into the details of this notion and its consequences here (see [15]), it can be observed that imperfect recall already follows from the definition of semantic games of imperfect information where there may be independencies between existentially or between universally quantified variables themselves. What this means is that players would then forget some information previously presented to them. But this can be accounted for by viewing players as "teams" of players or "multiple-selves" of a single player, where it is new members of a team or new "selves" who become responsible for the individual decisions in the course of the game as the need arises. Indeed, the team approach is by far the most common and natural way of capturing the game-theoretic notion of forgetting, and is spontaneously resorted to in a number of game and decision-theoretic problems ([14], [19]).

It needs to be observed that there is a tradition in the game-theoretic literature known as *team theory* ([6], [9], [12], [26]). For one thing, team theory is applicable to many multi-agent problems, although it seems by and large to have been neglected by the multi-agent community thus far (but see [1]). A *team* T is a finite set of non-coordinating players $i = \{1 \ldots n\}$ who have identical payoffs $u_i(h)$ but who act individually. The Verifier team and the Falsifier team have a finite number of individual members. Teams are thus groups of individuals with a common goal but individual information, knowledge and actions.[3] The central result of team theory — which underscores the central place game-like conceptualisations and strategic interactions have in the theory — says that all solutions of two-

[2]See [27] for some vibrant stories concerning the role of non-zero-sum games in our societies, science and everyday life.

[3]Thus, contrary to what has been suggested in [2], coalitional games, as they assume coordination, do not provide proper models for understanding IF logics and imperfect recall in the associated games. Indeed, coalition games have not in fact been considered in relation to imperfect recall in the game-theoretic literature.

person zero-sum games still hold for games played by teams [7]. Admittedly, team theory is a fairly heterogeneous field that aims to bring together various theoretical approaches, such as decision and systems theory, operations research, dynamic games, search and coordination, and parallel processing. But this is also the case in the research on multi-agent systems.

The key observation as regard to logic is that having multiple players in IF logic sets the semantic games here broadly in line with team theory, which sees teams as groups of agents with identical interests but individual actions and individual information. Under this conception, strategies are still be based on previous information within a game, but not on the information other members of the team might have had. If we moreover take these games to be strictly competitive, it follows that the basic solution concepts, namely the existence of winning strategies, are formed in games played by teams precisely as if there were just two players. It needs to be remarked that it is an open question whether these results can be applied also to non-strictly competitive games.

Furthermore, in IF logic the members of a team are not allowed to communicate with one another because this would destroy the team's ability, when viewed as one player, to genuinely forget something. The members of the same team all receive the payoff $u_i(h)$ when the outcome of a play is resolved. In addition, the information for individual team members remains persistent although the teams, when viewed as single players, do not forget information. Hence, whenever a move associated with the team of Verifiers or the team of Falsifiers is regarded as independent of the move made by the member of the same team, we are able to capture that by using a new member who makes the new move in question.

For some of the consequences and concrete examples of team actions in IF logic, see [15]. Just to mention a few, let us remark that team games do not presuppose that every logical component is assigned a distinct member. Only in the case of imperfect recall, a new member will be produced to account for the information loss. The game still contains just two players who, upon reassessing their plans and actions when moving from one information set to another, can control their behaviour at future information sets. Therefore, semantic games for IF logic very rarely form structures that can be termed *agent normal forms*. By this it is meant such games where each information set is assigned a distinct player. This observation in fact runs counter to the arguments given in Rubinstein [19, p. 78]. An example of an imperfect recall game where two consecutive moves are made by the same player is provided by the formula with three existential quantifiers and two slashes: $\exists x(\exists y/x)(\exists z/x)\, Sxyz$.

The team approach to logic advances the view that at the level of individual players the semantic information is persistent and the players do not forget information, but that the two principle players are seen to exhibit imperfect recall. One can think of an implicit map from the "information set" $S = \bigcup_i S_i^j$ containing all the information sets of the respective player to the information sets of the members of a team; in this way the principle "Verifier" and the principle "Falsifier" can coordinate the individual agents. From a slightly different perspective, one can think of players as playing the *roles* of all of the members, one at the time. When a subformula has the first component associated with a member of either of the teams, the player in question assumes the role of a single member. As it happens, she or he is seen to forget information, since the players are not, during a particular turn, allowed to use the information available to the other members of the team. Admittedly, viewing teams as single players usually gives away the coordination aspect of the game and introduces some excess strategies.

Furthermore, there are some important links between this team-theoretic kind of imperfect information and the NP-complete problems, implied by the so-called Tsitsiklis–Athans theorem. This theorem says that the problem of finding a team strategy that guarantees a certain minimum for finite 2×2 tables includes the Hamiltonian circuit problem and is hence NP-complete [25]. And as it is known [3], the Hamiltonian circuit problem is expressible in IF first-order logic.

Some further evidence for the usefulness of the team perspective in logic has been provided in [10] and [23], albeit somewhat indirectly. They show that games of imperfect recall should use strategies that would be more appropriate than just the traditional mixed ones, proposing for instance team-maxmin strategy profiles. This need is demonstrated to arise for a game of one team playing off against a single player, which in IF logic can be taken to correspond to the semantic game for weak equivalence, that is, for equivalence with respect to the truth of the formulas in a model, or with respect to the falsity of the formulas in a model, but not both.

The logical correlate to the kind of behaviour of teams presented above can also be predicted to have scores of potential applications in system and organisation theory, as well as in distributed computing tasks, which constantly are in the need of both logics and useful notions of teams or groups of agents in system-theoretic and distributed concept modelling.

V. Bounded rationality and the role of strategic concepts in logic

Introduced by Herb Simon in the 1950s and widely researched in economics and game theory since, the concept of bounded rationality in logical contexts usually refers to logical omniscience that may lurk behind the logics of knowledge and belief based on the traditional relational structures of possible worlds. But this is only one way of looking at bounded rationality and its edifications. The theory of semantic games, and especially the extensive forms of these games, would already provide us with a rich framework in which to model and reason about aspects of agent's restricted rationality, including restricted access to information, reduced sets of strategics, and so forth. This is not far removed from some of the recent branches of game theory such as evolutionary game theory, which extend strategic conceptualisations to the realm of non-rational objects (while intelligence in some form is still required). This being the case, it is safe to assume that the notion of strategies make perfect sense also in other inanimate decision tasks, such as semantic games and computational. An approach advocating thus can already be found in C.S. Peirce's philosophy: "Signs require at least two Quasi-minds; a Quasi-utterer and a Quasi-interpreter, and although these two are at one (i.e., are one mind) in the sign itself, they must nevertheless be distinct. In the Sign they are, so to say, welded. Accordingly, it is not merely a fact of human Psychology, but a necessity of Logic, that every logical evolution of thought should be dialogic" (CP 4.551).

There are many other areas within the broad field of multi-agent systems where various strategic conceptualisations either have turned out to be or at least should be recognised as important. For instance, in heterogeneous agent societies [24], despite concerning groups of agents, such strategies are used that are individualistic rather than collective. Because they are not coalitional, heterogeneous societies would thus fall quite naturally within the realm of team theory and its non-collective view of strategies. By the same token, the field known as preference-based non-monotonic reasoning is also one where strategic concerns are receiving increasing attention [11]. It therefore seems vital that such concerns be extended to the general field of non-monotonic reasoning and non-monotonic logic, which to date have been based on ad hoc definitions of non-monotonic entailment relations.

These short remarks are also calculated to pave a way for an account of mitigating logical omniscience in epistemic logic and related disciplines — the problem of how to prevent an indefinite production of logical consequences about what is known: By such semantic games applied to the modalities of epistemic logics where the games are non-strictly competitive, one can readily see how there can be inconsistent worlds, namely worlds that are epistemically possible in the sense that a player can pick them from the semantic structure of all worlds in a game, but which nevertheless are not logically possible. Namely, they contain contradictory sentences in the sense of there existing winning strategies for both of the players. In this sense non-strictly competitive games can be seen as a vindication of one instance of the general concept of bounded rationality in the semantic decisions problems for logical meaning and interpretation.

VI. Conclusions and further research

This paper needs to be regarded as both general and specialised in content and purpose. The general aim is to discuss to what extent social concepts and metaphors may permeate logic. The specialised purpose is to suggest the theory of IF logic and the associated semantic games as illustrative examples where such metaphors rear their heads through some specific game-theoretic terminology. Inevitably, the research toward a marriage of these general and special aims is in its early life.

Application-wise, the potentials for the kind of general logical and game-theoretic systems suggested above have to be confined just to a list of a couple of promising suggestions. They include, but are not limited to: Inconsistent information modelling and reasoning in knowledge-based systems, conflict resolution and negotiation games under uncertainty, routing problems in communicating networks, and the representation of knowledge in message-passing multi-agent systems and parallel processing.

References

[1] Bacharach, M., 2001. Superagency: beyond individualistic game theory, manuscript.

[2] van Benthem, J. 2002. Hintikka self-applied, to appear in H. Lewis (ed.), *Library of Living Philosophers: Jaakko Hintikka*. Available electronically at `http://turing.wins.uva.nl/~johan/H-H.ps`.

[3] Hella, L. and Sandu, G., 1995. Partially ordered connectives and finite graphs, in M. Krynicki, M. Mostowski, and L.W. Szczerba (eds), *Quantifiers: Logics, Models and Computation* 2, Dordrecht: Kluwer, 79–88.

[4] Hintikka, J., 1996. *The Principles of Mathematics Revisited*, New York: Cambridge University Press.

[5] Hintikka, J. and Sandu, G., 1997. Game-theoretical semantics, in J. van Benthem, and A. ter Meulen (eds), *Handbook of Logic and Language*, Amsterdam: Elsevier, 361–410.

[6] Ho, Y.C. and Chu, K.C., 1972. Team decision theory and information structures in optimal control problems I., *IEEE Transactions on Automatic Control* 17, 15–22.

[7] Ho, Y.C. and Sun, F.K., 1974. Value of information in two-team zero-sum problems, *Journal of Optimization Theory and Applications* 14, 557–571.

[8] Jaśkowski, S., 1948. Rachunek zdańdla systemòw dedukcyjnych sprzecznych, *Studia Societatis Scientiarum Toruensis* 5, 55–77. (Translated in English as: Propositional Calculus for Contradictory Deductive Systems, *Studia Logica* 24, 1969, 143–157.)

[9] Kim, K.H. and Roush, F.W., 1987. *Team Theory*, New York: Ellis Horwood.

[10] Koller, D. and Megiddo, N., 1992. The complexity of two-person zero-sum games in extensive form, *Games and Economic Behavior* 4, 528–552.

[11] Lin, F. and Shoham, Y., 1989. Argument systems —- A uniform basis for non-monotonic reasoning, in *Proceedings of the 1st International Conference on Principles of Knowledge Representation and Reasoning*, 245–255.

[12] Marschak, J. and Radner, R., 1972. *Economic Theory of Teams*, Yale University Press, New Haven.

[13] Peirce, C.S., 1931-5. *Collected Papers* 1–6, C. Hartshorne, and P. Weiss (eds), Cambridge, Mass.: Harvard University Press.

[14] Piccione, M. and Rubinstein, A., 1997. On the interpretation of decision problems with imperfect recall, *Games and Economic Behavior* 20, 3–24.

[15] Pietarinen, A., 2001. Varieties of IFing, in Pauly, M., and Sandu, G. (eds) *Proceedings of the ESSLLI 2001 Workshop on Logic and Games*, Helsinki: University of Helsinki.

[16] Pietarinen, A., 2002a. A note on the structural notion of information in extensive-form games, to appear in *Quality & Quantity*.

[17] Pietarinen, A., 2002b. Logic and coherence in the light of competitive games, to appear in *Logique et Analyse*.

[18] Pietarinen, A. and Sandu, G., 1999. Games in philosophical logic, *Nordic Journal of Philosophical Logic* 4, 143–173.

[19] Rubinstein, A., 1998. *Modeling Bounded Rationality*, Cambridge, Mass.: MIT Press.

[20] Sandu, G. and Pietarinen, A., 2001. Partiality and games: propositional logic, *Logic Journal of the IGPL* 9, 107–127.

[21] Sandu, G. and Pietarinen, A., 2002. Informationally independent connectives, to appear in I. van Loon, G. Mints and R. Muskens (eds), *Logic, Language and Computation* 9, Stanford: CSLI Publications.

[22] Sette, A.M., 1973. On the propositional calculus P1, *Mathematica Japonae* 16, 173–180.

[23] von Stenger, B. and Koller, D., 1997. Team-maxmin equilibria, *Games and Economic Behavior* 21, 309–321.

[24] Subrahmanian, V.S. et al., 2000. *Heterogeneous Agent Systems*, Cambridge, Mass.: MIT Press.

[25] Tsitsiklis J.N. and Athans, M., 1984. On the complexity of decentralized decision making and detection problems, *IEEE Transactions on Automatic Control* 30, 440–446.

[26] Witsenhausen, H.S., 1968. A counterexample in stochastic optimum control, *Siam Journal of Control* 6, 131–147.

[27] Wright, R., 2000. *Nonzero: The Logic of Human Destiny*, New York: Pantheon Books.