ISSA Proceedings 2014 – Logical Validity, Bounded Rationality, And Pragma-Dialectics: Outline Of A Game-Theoretic Naturalization Of Classically-Valid Argumentation

Abstract: This paper outlines how classical propositional logic, particularly the notion of 'obtaining a classically-valid logical proof', can be understood as the outcome of an argumentation-game. We adopt two game-rules from dialogical logic under which obtaining such as proof is a matter of due course, as both rules together guarantee a winning-strategy for one player when logical consequence holds. We then show how these rules can arise from players' preferences, rather than be imposed externally, and can hence count as 'player self-imposable'. Subsequently, this game is shown to comply with the Pragma-dialectical Code of Conduct, while some of the Code's rules become gratuitous as their content arises directly from player's preferences instead. Our discussion is oriented towards future inquiries into how logics other than its classical variant can be similarly "naturalized."
Keywords: game theory, classical logic, proof, proponent, opponent, winning-strategy, pragma-dialectical code of conduct rules.

1. *Introduction*
Viewing logic as one language game among many, Ludwig Wittgenstein had offered an analogy between having a proof and winning a game (Wittgenstein, 1953). The formal details of this analogy have been mostly studied by formal logicians who, in viewing logical proofs as regimented argumentation-procedures, sought to give an argumentative characterization of logic.**[i]** Game-theory in particular became a natural framework to model episodes of natural language argumentation that characterizes logical inference, giving rise to game-theoretic semantics (GTS) (Hintikka & Sandu, 1997) and dialogical logic (DL) (Rahman & Keiff, 2005) as the two main approaches.

GTS and DL partially reduce logic to argumentation-procedures by restricting players' strategies so that games realize the model-checking procedures and proof procedures typical of logical inference. The motivation for such restrictions, however, remains internal**[ii]** to the model, receiving primarily pragmatic justification through successfully recovering logical inference formally from particular constraints on argumentation. This article shows DL-restrictions that are imposed to recover first-order logical consequence from argumentation to be instead forthcoming from preference-profiles of boundedly rational players. Such players, we take it, cannot optimize their strategies because they lack the ability to compute complete representations of a game, while we understand constraints on such a game to be player-self-imposable through strategic reasoning (provably) equivalent to the elimination of dominated strategies.

The following outlines how classical propositional logic, particularly the notion of 'obtaining a classically-valid logical proof', can be understood as the outcome of an argumentation-game (2.1), and introduce two game-rules under which obtaining it is a matter of due course, for both rules together guarantee a winning-strategy (2.2), then raise the claim that the strategies adopted by players in this game are 'player self-imposable', because these same strategies may be inferred from players' preferences by (reasoning employing) a maximin-principle (Sect. 2.3 to 2.5). Subsequently, this game is shown to comply with the Pragma-dialectical Code of Conduct (3.1), but that some among the Code's rules are gratuitous, so to speak, whenever normative content already arises from player's preferences (3.2). Our discussion, in Sect. 4, is oriented towards future inquiries into how logics other than its classical variant might similarly be "naturalized." We close with brief conclusions in Sect. 5.

2. *The game-theoretic apparatus*

To start, we will sketch the elements of an argumentation-game as they appear from a game-theoretic perspective, introducing further relevant notions as we go along.

2.1. *Logic as an argumentative game*

The players' choice of a language, L, is a preliminary step to any language game. Agreement on the language in which the argumentation will be couched determines the actions arguers can take (e.g., how to attack and defend complex sentences; how to assess an atoms' truth value). We restrict L to a propositional language corresponding to a fragment of vernacular English where basic sentences (aka atoms) contain a subject phrase referring to individuals, a verb phrase, and terms referring to individuals, e.g. “The cat is on the mat”; “Alice is taller than Bob.” Complex L-sentences combine atoms through connectives (and, or, if… then…), and locutions equivalent to negation (is not, or it is not the case that), or locutions that combine such complex sentences, collectively called operators.

Given a language L, a proof demonstrates that a conclusion C follows from a set P of premises. We will here be mostly concerned with the semantic view, where P collects situations where the set’s members are true, and to prove C is to demonstrate that C is true in every situation.**[iii]** In a DL game, the proponent (PRO) is committed that C is true if P is assumed, while the opponent (OPP) is committed that C may be false in at least one case where all members of P are true. In order to prove C, PRO must demonstrate that, once OPP concedes P explicitly, C is conceded implicitly. Players’ legitimate moves are attacks, which ask for explicit commitment to the consequences of a statement, and defenses, which incur commitments. A move’s legitimacy is partly determined by L; both players are allowed the same moves. Independently of the PRO or OPP role, for instance, if player X states “A and B” then player Y can constrain her to commit to A, to commit to B, or to both. If player X states “A or B,” then player Y can only constrain her to commit to (at least) one of the disjuncts, while X retains the option to commit to A, or to B, or to both. In tree form, Table 1 provides the complete set of attacks and defenses. Atoms are noted ‘ψ’ and are indexed by 1 or 2 when these occur in complex sentences; the prefix ‘Y?’ indicates an attack, followed by the specific sentence it targets, where some attacks allow to ask for a commitment that, when relevant, is specified after a forward-slash (‘/’). These rules can be applied systematically to any sentence player X has stated, eventually forcing X to commit to a basic sentence or its negation.

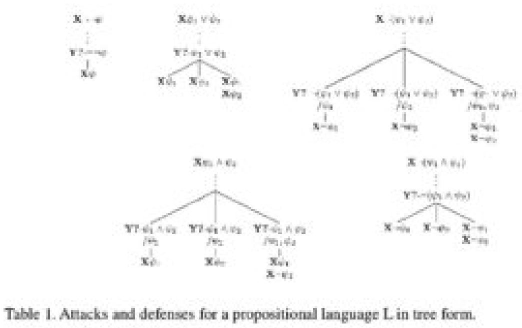

Table 1. Attacks and defenses for a propositional language L in tree

form.

2.2 *Structural rules that guarantee a winning-strategy*

Provided the conclusion, C, is a finite statement, OPP is restricted to a finite number of genuine attacks, i.e., excluding repetitions. As we saw above, asking PRO to commit to complex expressions eventually brings it about that OPP asks PRO to commit to a literal, i.e., an atom, or its negation. (By convention, negated atoms cannot be attacked.) The first of the two structural rules, the *Structural Rule for Literals* (abbreviated SR-L), amounts to PRO having the last say in a play, only if she can, using merely the premises, P, defend C in that play. SR-L restricts only PRO's strategies.

Structural Rule for Literals (SR-L): Unless OPP has previously stated the literal A, PRO cannot defend herself against an attack that requires of her to state A.

The second, the *Structural Rule for Repetitions* (SR-R), prevents delay-tactics. After all, by repeating a genuine attack, one player could keep denying the other player's win, and so (forever) delay reaching the play's end-state.

Structural Rule for Repetitions (SR-R): Should player X have previously attacked statement A of player Y to which player Y has responded, then X cannot repeat this attack.

Together with agreement on the meaning of L's logical terms laid out in the attack and defense-rules in Sect. 2.1, these two rules suffice for representing argumentative games in tree form. The analogy between proving and winning a game thus gains precision. We now turn to the strategic reasoning of the players.

2.3 *Strategy selection*

Game theory generally explains strategy-selection by an inference called 'elimination of dominated strategies'. This inference considers all strategies available to players, ranks these on an outcome-ordered ordinal scale, and eliminates all strategies but that, or those, at the highest rank. (Once eliminated, the succession of moves such strategies consists of is not played.) Leaving implicit those preferences that are instrumental in generating the outcome-ordered strategy ranking has, in our opinion, prevented argumentative approaches to logic from becoming genuinely game-theoretic treatments. As we now argue, these shortcomings prevent DL from describing genuine language games, which thus fails to resonate with its self-professed Wittgensteinian origins. As we also argue, however, DL semantics can be suitably "fixed."**[iv]** The required modifications apply at each of the following steps:

At step (1), each player must be provided with a preference-profile over the game's outcomes. From it, one may infer players' preferences over all possible moves of a play, thus postulating an inference from outcome-preferences to move-preferences. A genuine import from game-theory would otherwise be hard to discern.

At step (2), the rules SR-L and SR-R are promoted from being reasonable game-rules to the status of player-self-imposed restrictions. Here, one might postulate another inference that derives both rules from players' preference-profiles. But players might as well agree upon these restrictions explicitly, making them non-inferred game-rules that promote players' interests (see the next section).

At step (3), one requires some explanation on how players can each prefer selecting a strategy that, in combination with the other player's strategy, gives rise to a pair – called a 'strategy-profile' – which would be mapped to a semantic proof obtained when implementing mechanical constructions that guarantee this proof to terminate if and only if C follows from P.

As we show in the next section, comparatively weak assumptions suffice to equip players with suitable preferences.

2.4 *Preferences*

Being rather natural ones, our assumptions seemingly describe but mildly idealized arguers. Furthermore, a single inference-principle – called 'Harsanyi-Maximin', introduced below – apparently suffices to let players
(i) individually select move-preferences,
(ii) jointly self-impose the game-restricting rules SR-L and SR-R, and
(iii) jointly select only strategy-profiles that generate the equivalent of systematic tableaux proofs. These assumptions have been formalized in Genot & Jacot (2014). The following provides an informal version:

A1 – *Meaning Coordination*

Players have coordinated on the meaning of logical operators, and have the means for disambiguating non-logical terms.**[v]**

A2 – *Asymmetric Burden*
Players agree that, to win the game, PRO must meet every challenge raised by OPP, and so must win every play; OPP may challenge PRO by raising all alternatives compatible with the common ground, and OPP subsequently wins the game as soon as he has won a play.

A3 – *Comparative Efficiency*
Both players prefer games with fewer to those with more moves.

A4 – *Termination over Frustration*
Both players prefer losing a play, or a game, over playing indefinitely long.

A5 – *Imperfect Foresight*
Both players' ability to anticipate the other's moves is limited.

A6 – *Common Knowledge***[vi]**
Both players know A1 to A5 to be the case.

As sketched in Sect. 2.1, A1 can be satisfied by players agreeing on rules for attacks and defenses for connectives, and by their referencing atoms ostensibly (i.e., pointing to a term's referent).**[vii]** A2 is equivalent to having agreed on semantic consequence,**[viii]** while an explicit notion thereof remains gratuitous as long as it is well-defined how a play is won (which occurs by agreement on L). A3 is immediate whenever playing is costly, for instance time-wise. A4 is reasonable whenever players can contemplate the prospects of winning *future* games, while they might lose the present one. A5 typically holds for boundedly rational self-knowledgeable players unable to grasp the game's full combinatorial structure, and assuming as much of the other player. A6 holds whenever players explicitly agree to A1 to A5, in the sense that each then knows that the other does, too.

As a consequence of assuming bounded rationality, players cannot be meaningfully said to distribute probabilities over alternative courses of the game, and so cannot form rational expectations based on these. They can, however, always apply the rationality principle that Harsanyi (1977) proposed for reasoning in games where (probabilistic) expectations are not well-defined:
Harsanyi Maximin (HM): If player X cannot form rational expectations about the probability that Y will not select the strategy leading to X's least preferred outcome, then X should play the strategy that best responds to Y's most

detrimental strategy for X.**[ix]**

The rationale for HM consists in a simple consequentialist consideration: acting as HM prescribes guarantees minimizing losses that are incurred in worst case scenarios. Hence, for HM to be applied, it must be clear what the most detrimental strategy is. Together with A1 to A4, HM suffices in DL-games to vindicate informal arguments that are typically provided for the collapse of symmetrical options in dialogical games to the asymmetrical rules of semantic tableaux. More importantly, as is shown in the next sub-section, from HM, together with A1 to A4, SR-L and SR-R can be obtained as self-imposed strategic principles. Finally, if players agree to sequentially conduct all plays necessary to demonstrate whether PRO has a winning-strategy, or not, then this sequence simulates a tree proof. As noted above, when L is a first-order language, the possibility of infinite plays arises, and consensus can therefore only be found in the limit, by assuming that infinite plays are won by OPP.

2.5 *Structural rules as self-imposed constraints*

Formal proofs are given in Genot & Jacot (2014) that HM suffices to (i) collapse the best options for PRO and OPP to tree-building rules, (ii) obtain SR-L and SR-R as self-imposed restrictions, and (iii) lead players to realize proofs. We point readers to this paper for the third claim and will not separately treat the first claim here, either, as particle rules depend on the language L and thus on the pre-play agreement. But the second claim concerns structural rules which are in force in any DL-game, and for any language. How boundedly rational arguers can self-impose the structural rules SR-L and SR-R should therefore be relevant to the reduction of logical reasoning, classical or other, to argumentation. We now sketch how SR-L and SR-R can be justified argumentatively.

As for SR-L, the strongest position for the proponent of a thesis C in a pro and contra argumentation entails the ability to always support C *ex concessis*, i.e., through arguments raised by the opponent. In PRO's case, then, supporting C comes down to supporting those literals for which PRO has incurred commitments as a consequence of upholding a commitment to C vis-à-vis OPP's doubt about C. PRO can maintain the strongest position only if these same literals have previously been stated by OPP. And were PRO about to state a literal A that OPP had not *yet* stated, then PRO's worst case would consists in OPP systematically avoiding to state A. Since, *qua* A4, PRO cannot form a rational expectation as to the probability of OPP avoiding to state A, *qua* HM, PRO should never state literals, unless these had first been stated by OPP.

Turning now to SR-R, consider cases where PRO might want to repeat an attack, because PRO's previous attempt to obtain a suitable literal A from OPP had failed, while PRO could possibly obtain a better response through repetition. PRO's worst case here consists in OPP repeating the response that had already proved non-suitable to PRO. *Qua* HM, PRO should therefore *not* repeat the attack. Doing so would merely extend the play, but bring no further benefits, an option that is ruled out by the preference expressed in A3.

OPP's reasons to enforce the content of SR-R are symmetrical to PRO's reasons, as the only situation where an attack-repetition is plausible is exactly that where PRO has answered all previous attacks. And even here, OPP could at best hope, but not rationally expect that PRO might, upon OPP's repetition of the attack, give responses that PRO cannot defend. The worst case for OPP, then, is that course of the game where PRO selects the same responses that PRO had previously managed to defend. Qua HM, as above, therefore also OPP should not repeat the attack.

3. *Comparison with the code of conduct*

Players' choices with respect to L, and with respect to preferences, may yield argumentation-games that instantiate different systems of logical inference. In particular, starting from an impoverished L, characterizing players' preferences through the assumptions A1 to A6, and using the Harsanyi Maximin principle (HM) suffice in order to obtain classical logic, modulo quantifiers. On these, see Genot & Jacot (2014). Classical logic is therefore said to result from self-imposed restrictions when argumentation is treated as a game that to win presupposes the existence of a winning-strategy, but not knowledge of its existence. This provides a formally precise sense in which logic can in principle emerge from arguers' preferences, thus clarifying the Wittgensteinian analogy mentioned in the introduction.

Were the formal relation between logic and arguer-preferences more fully understood, then one might perhaps obtain one from, and in terms of, the other. Until future research has shown as much, a modest but no less important insight is that classical logic needs no mentioning in normative argumentation-rules for it to nevertheless dictate the game's winner, because the constraints that make classical logic "the ruler" can arise from arguers' preferences, and so need not be explicit.

In the remainder, we argue that reaching a consensus on the kind of logical consequence that shall apply for some argumentation-game, amounts to endorsing a particular specification of the Code of Conduct in the Pragma-dialectical theory (PD), and so may be viewed as a special case thereof. Sect. 3.1 compares the fifteen PD-rules to our structural rules. SR-L and SL-R are said to be specifications of PD-rules whenever the Code does not prevent participants from specifying its content in this way. We moreover discuss the assumptions A1 to A6 vis-à-vis PD's higher-order conditions that are placed on arguers seeking to settle a difference of opinion on the merits, and provide a brief discussion of the HM-principle (Sect. 3.2).

3.1 *Comparison of Structural Rules with PD-rules*
We assume familiarity with the fifteen Pragma-dialectical discussion-rules, aka the Code of Conduct. Its latest version is found in Van Eemeren & Grootendorst (2004), 123-157; Zenker (2007a) compares it to the Code's 1984 version. We refer to the Code's n-th rule as PD-n.

The structural rule for repetitions (SR-R) is a near-verbatim copy of PD-13, serving the same function: preventing delays. In contrast, the content of the structural rule for literals (SR-L) specifies more than one PD-rule. Moreover, some specifications of the Code arising from SR-L do so in combination with the assumptions A1 to A6, as will be discussed further below.

SR-L distributes the proponent and opponent rules, which remain the same throughout the game, thus specifying PD-4. Moreover, SR-L specifies the right to challenge, thus specifying PD-2, assigning it to OPP, and the obligation to respond to a challenge, thus specifying PD-3, assigning it to PRO. This allocation, in turn, implies a corresponding distribution of the burden of proof, regulated likewise through mutual implication in PD-3. Provided that player's agree on the circumstances of winning qua accepting SR-L, this also specifies PD-5, for players now agree on a successful attack, and a successful defense, in this discussion. (Recall that, per SR-R, a successful attack – of the claim that C follows from P – must not have been used already in the same discussion; and that a successful defense of that claim may not recur to material other than that conceded by OPP.)

PD-6 demands that players attack and defend only by argumentation. We do not so much take PD-6 to be specified, but to be implied by SR-L and SR-R. After all, neither SR-L nor SR-R leave room for moves other than argumentative attacks and defenses. Thus, one may not strictly need PD-6 in the sense of a necessary condition for the resolution of a difference on opinion, provided certain preferences. Similarly, for a critical discussion the rules PD-7, PD-8, and PD-9 demand that participant-agreement is reached on a successful attack and defense of a propositional content and of its justificatory potential, and on a conclusive defense. Such definitions are effectively provided by SR-L, along with Asymmetric Burden (A2), to which we return below. Moreover, if we view the defense of a sub-standpoint, regulated in PD-9, to amount to winning a play, as opposed to winning a game, then SR-L and A2 jointly imply the content of PD-9.

PD-10 and PD-11 assign the right to attack and to defend undefended standpoints to the proponent and the opponent, respectively. We had only introduced a single standpoint, expressed as: C follows from P. Therefore, neither PD-10 and PD-11, nor their negations, apply to our argumentation-game; hence these rules cannot be violated, either; a fortiori they cannot be meaningfully called necessary. Having discussed PD-13 above, PD-14, which assigns an obligation to retract upon a conclusive defense, is implied by SR-L. Finally, PD-15 states an unconditional right to demand usage-declaratives. This is either not needed (when stipulating players to assign truth values without analyzing the meaning of literals) or assumption A1 states as much, but also more (see Sect. 3.2). Finally, PD-1, which denies special preparatory conditions on arguers or their arguments, can be viewed as being fulfilled, but has no direct or indirect counterpart in the assumptions A1 to A6.

In sum, the Code of Conduct does not bar logical argumentation from occurring as a result of playing, with suitable preferences, according to PD-rules. This being so is far from incidental, and should rather be viewed as a desired consequence of the PD model. At any rate, our rules and assumptions yield a limiting case of the Code, while it also became clear that the content of PD-rules that regulate agreement on a conclusive attack and defense are not needed as explicit rules. In Sect. 2, players' preferences as to how the game should be played were shown to arise on the assumptions A1 to A6. We now turn to these.

3.2 *The assumptions A1 to A6, and the HM rationality-principle*

Immediately above, *Meaning Coordination* (A1) was seen to be slightly stronger than PD-15, for A1 assumes players to coordinate successfully, while the Code merely reserves the right to demand usage declaratives, without stipulating semantic success. *Asymmetric Burden* (A2) amounts to a definition of winning a play, and thus the game, for both PRO and OPP. It hence specifies PD-7 to PD-9, along with both of our structural rules, as discussed above. *Comparative Efficiency* (A3) spells out an assumption that seemingly fails to correspond to any PD-rule, but neither is A3 in violation of the Code. The same holds for the remaining three assumptions: *Termination over Frustration* (A4), *Imperfect Foresight* (A5), and *Common Knowledge* (A6). As stated, A4 characterizes a preference of players to rather seek playing the argumentation-game, while the constraint A5 mirrors players' cognitive limitation, of which A6 says that players know it. All assumptions are compatible with the Code.

Further, in PD, so-called higher-order conditions spell out additional features on arguers, for instance, their willingness to settle a dispute. See Zenker (2007b)for a non-exhaustive list of such conditions. We find it plausible to view A4 to A6 as higher-order conditions that describe what one might reasonably expect on behalf of boundedly rational players and their cognitive states. Also, endorsing HM as a rationality principle may be understood as a higher order condition. As we saw, HM ensures that, if an argumentation-game has a winning-strategy, then PRO or OPP will find it. Recognition of HM, or a principle similar to it, bars player X from assuming that Y plays anything but that strategy, or those strategies, on which Y eventually wins the game, if Y could win, and *vice versa*. Therefore, as HM states, the best response to any such Y-strategy is for player X to pursue a strategy that does not in principle fail to reach the same goal, so both players are kept from playing in ways that lead nowhere near the desired result anytime soon.

While HM amounts to a generalized form of pessimism, nothing in the Code keeps HM from applying to players or to their game. For idealized arguers – idealized with respect to possessing sophisticated game-theoretic knowledge – HM is clearly a reasonable choice. But we cannot find that HM would even be questionable for boundedly rational arguers. After all, when properly understood, the content of HM is hardly more complex than the final sentence of the previous paragraph. Put differently, failure to understand, or to endorse, HM would arise from cognitive, emotional, or perhaps ecological boundaries outside the normal range of boundedly rational agents. All the same, HM remains a genuinely game-theoretic principle of rational interaction. Its acceptance by players, as a rationality principle, cannot be motivated other than by explicitly viewing argumentation as a game whose outcome depends on the way in which a strategy-profile, i.e., the particular pair of strategies chosen by X and Y, generates the game's outcome.

4. *Discussion*
The Code of Conduct provided by the Pragma-dialectical theory (PD) normatively governs attacks and defenses of a standpoint in a merit-based critical discussion aimed at a resolution of a difference of opinion, or consensus, where arguers assume the dialectical roles of proponent and opponent. This framework was seen to be consonant with attempts at capturing logic as formal argumentation, understood as a Wittgensteinian language game, as currently implemented in dialogical logic (DL) and game-theoretic semantics (GTS). All three approaches view natural language argumentation as an interactive process between a proponent, who states and argumentatively supports a thesis, and an opponent attacking it.

Logic is regularly equated with the rules one *should* apply to implement logical reasoning, thereby deriving a valid consequence from the premises; DL and GTS make no exception to this, as both represent logic in a game by imposing logical rules onto its players. Equating logic with its rules, however, is in conflict with the view ascribed to Wittgenstein, above: what matters in a language game are not the rules, but the players' goals and preferences. For players who self-regulate their argumentative conduct, the status of logical rules was consequently seen to be demoted to that of a description, useful for instance when instructing newcomers pursuing the same goals. Wittgenstein's view being in principle vindicated by the theory of games that DL and GTS build on, players can therefore dispense with such rules altogether, at least as primitive notions. Embracing this demotion of logical rules brings DL and GTS closer to their professed philosophical and methodological sources. So far, however, both DL and GTS do not yet characterize players who meaningfully *prefer* arguing logically, as opposed to being forced to do so.

We have indicated how to tell a different story: take a fragment of natural language (restricted to noun phrases, verb phrases, and any complements needed) no more expressive than a formal propositional language; then understand logical argumentation taking place between a proponent and an opponent as the outcome of a particular type of argumentation-game; finally, provide sufficient conditions under which players' preferences and abilities restrict their argumentative moves to logically valid inferences. In this way, enforcing the consensus through the imposition of logical rules becomes superfluous, for logical rules now emerge from a game where well-defined preferences are ascribed to players who achieve meaning-coordination. Importantly, our assumptions about players' preferences and abilities were said to characterize *boundedly rational* agents, thus remaining much closer to human reasoners than to the ideal reasoners typically assumed in DL and GTS approaches.

Comparing what such assumptions induce with the Pragma-dialectical Code of Conduct, we observed a similarity between natural-language argumentation and logical argumentation that is far from incidental. Some of our assumptions on players' abilities and preferences were seen to be specifications of the Code's rules, or its higher order conditions, while assumptions that remained unrelated to the Code did not violate its normative content. Hence, logical argumentation can arise within the Pragma-dialectical framework for a critical discussion among boundedly rational players without assuming prior knowledge of, or explicit agreement on, the norms of logic. This being as it should be, we hope to have made understandable how logic can systematically emerge from natural language argumentative practice.

5. *Conclusion*

While our story here had ended with classical propositional logic, the main result presented in the present paper has been successfully extended to full classical first-order logic (Genot & Jacot, 2014). Consistent with the conjecture that a similar story can be told for logic's ontogenesis, only a natural language and boundedly rational players were taken to be necessary to make a first step towards a naturalization of logic. To carry this naturalization-attempt further, future research should be conducted in a theoretical and in an empirical manner. Similar argumentative accounts of logic should be extended to non-classical logics, by considering richer natural language fragments, for instance, as well as different goals and preferences. Moreover, assumptions that constrain players' preferences and abilities should be validated, e.g., in focus interviews, but also through systematic experimental work.

Acknowledgements

For comments and useful criticism, we thank the audience at the 2014 ISSA conference, 1-4 July, 2014, Amsterdam, as well as Jeroen Smid, John Woods, Maurice Finocchiaro, and Erik Krabbe. The authors gratefully acknowledge funding from the Swedish Research Council (VR) and the Lund University Information Quality Research Group (LUIQ) led by Erik J. Olsson. For an extended manuscript that significantly extends Sect. 2 and 4 of the present paper, please contact the first author.

NOTES

i. Logic had thus returned to its origins in argumentation, if one views Aristotelian logic to emerge from the argumentative practices in the Academy and the Lyceum (Robinson, 1971), being proceeded by the Socratic elenchus, among others. For a brief historical overview, see Dipert, Hintikka, & Spade (2014), and Hintikka (1996).

ii. Our use of 'internal' and 'external' breaks with standard game-theory where preferences are part of the definition of a game set-up, and in this sense internal to the game, while restrictions imposed on players' strategies to guarantee a proof are called externalities whenever being independent of such preferences.

iii. Viewed syntactically, P is a set of grammatically well-formed L-sentences, so to prove C is to demonstrate that, using only the grammatical rules of L, C can be obtained by a transformation and a combination of P-members.

iv. A tentative explanation why this option had not been considered much earlier, crucially in the Erlangen school (see Krabbe (2006)), is that the requisite reasoning had (falsely) been viewed to demand of players abilities that are equivalent to mathematical induction. After all, in logic and proof theory, it is mathematical induction that is normally used to reason about logical proofs (aka meta-logic or meta-mathematics). However, mathematical induction is here required only to prove that a given proof strategy will be successful, but is not required to implement a proof strategy. So a game-theoretic approach to logic could well have internalized reasoning-about-proofs within a given proof, and thus strictly subordinate logical reasoning to meta-logical, or meta-.

v. Non-logical terms comprise noun-phrases, verb-phrases, etc.; disambiguating these is understood to be part of linguistic competence.

vi. Unlike A1 to A5, which are both necessary conditions to obtain proofs from games, A6 is sufficient but not necessary. Also a weaker assumption may do, such as a belief in the other player's rationality (see Genot & Jacot (2014)).

vii. Genot & Jacot (2014) use pointing to abstract representations such as vertices and edges of a graph to disambiguate atoms, where a vertex represents an individual, and a labeled path of length n represents an n-ary predicate. The representations are motivated cognitively, as they share properties of perceptual representation.

viii. Agreement to consider some, but not all possibilities compatible with P yields a non-monotonic logic where, once drawn, a previously agreed-upon conclusion can nevertheless be retracted if this agreement is subsequently revised, for instance upon taking into consideration additional possibilities, including counterexamples formerly disregarded. Such agreement is independent of the player's agreement on L, and so depends on their preferences.

ix. The most detrimental strategy, aka the worst case, for X is not always the best case for Y. In our games, the worst case for either player is to be denied victory in a play through the other player's use of a delaying tactic. But this tactic is never the best one for any player using it. After all, the outcome of the game would be unnecessarily delayed, so both players would incur a loss, and so both players' preferences (as expressed in A3 and A4) would be satisfied to a lesser degree.

References

Dipert, D., Hintikka, J., & Spade, P. V. (2014). *History of logic.* Encyclopedia Britannica.

Genot, E., & Jacot, J. (2014). Semantic game for first-order entailment with algorithmic players. *Proceedings of the eleventh conference on logic and the foundations of games and decision theory* (LOFT'14). LOFT.

Harsanyi, J. (1977). *Rational behavior and bargaining equilibrium in games and social situations*. Cambridge: Cambridge University Press.

Hintikka, J. (1996). Knowledge Acknowledged: Knowledge of Propositions vs. Knowledge of Objects. *Philosophy and Phenomenological Research,* 56(2), pp. 251-275.

Hintikka, J., & Sandu, G. (1997). Game-theoretical semantics. In J. van Benthem, & A. ter Meulen (Eds.), *Handbook of logic and language* (pp. 361-410). Amsterdam: Elzevier.

Krabbe, E. (2006). Logic and games. In F.H. van Eemeren, P. Houtlosser, & A. Van Rees (Eds.), *Considering pragma-dialectics: A festschrift for Frans H. van Eemeren on the occasion of his 60th birthday* (pp. 185-198). Mahwah, NJ: Lawrence Erlbaum.

Rahman, S., & Keiff, L. (2005). On how to be a dialogician. In D. Vanderveken (Ed.), *Logic, thought and action*. Dordrecht: Springer.

Van Eemeren, F., & Grootendorst, R. (2004). *A systematic theory of argumentation: the pragma-dialectical approach.* Cambridge UK: Cambridge

University Press.
Wittgenstein, L. (1953). *Philosophical investigations*. Oxford: Blackwell.
Zenker, F. (2007a). Changes in conduct-rules and ten commandments: pragma-dialectics 1984 vs. 2004. In F.H. van Eemeren (Ed.), *Proceedings of the international society for the study of argumentation* (ISSA, Amsterdam, June 2006) (pp. 1581-1589). Amsterdam: SicSat.
Zenker, F. (2007b). Pragma-dialectic's necessary conditions for a critical discussion. In H. Hansen et al. (Eds). *Proceedings of the 7th Int. conference of the Ontario society for the study of argumentation* (OSSA), (pp. 1-15). Windsor ON.