Game semantics for weak depth-bounded approximations to classical propositional logic

A. Solares-Rojas[1]*, O. Majer[2], and F. E. Miranda-Perea[3]†

[1] Instituto de Ciencias de la Computación, UBA, Argentina
asrojas@dc.uba.ar
[2] Institute of Philosophy, The Czech Academy of Sciences, Czech Republic
ondrej.majer@gmail.com
[3] Departamento de Matemáticas, Facultad de Ciencias, UNAM, México
favio@ciencias.unam.mx

Tractable deductive systems approximating classical propositional logic (CPL) have interest in areas that require models of bounded rationality (see [3]). In a series of papers culminating in [1], *depth-bounded* approximations have been studied, which can be intuitively related to the deduction power of resource-bounded agents. Among these approximations, the so-called *weak* ones are defined in terms of the *depth* of derivations within a KE-KI system, and are decidable in polynomial time whenever their associated depth is suitably parameterized. The 0-depth approximation can't be characterized by a set of finitely valued matrices. So far, two alternative semantics have been given to characterize that basic approximation and, recursively, all of its successors. Namely, a *modular* one and a 3-valued *non-deterministic* one [1, Sec. 1.3, 1.5]. Both are well motivated and intuitive for the 0-depth approximation, though not necessarily for those of greater depth. In this work, we introduce a game semantics which in our opinion provides a more intuitive framework for the whole hierarchy of approximations. Namely, we define a game where *negative* constraints are associated with understanding the informational meaning of the connectives, while resource consumption is transparently modeled by the *expense* of questions that are within a finite number. Although related to standard dialogical accounts [4], our question-answer framework seems more intuitive in the context of the approximations.

Proof-theoretical background We work with system $\mathcal{NT}$ in Table 1, formulated with *signed* formulas of the form $\mathsf{T}A$ or $\mathsf{F}A$, meaning that the agent "holds the information that A is true (respectively, false)" [1, Sec. 2.1]. It results from combining two complete systems for CPL: the refutation system KE and the direct-proof system KI; respectively related to, but more efficient that, Tableaux and truth-tables. Like Natural Deduction, $\mathcal{NT}$ has introduction and elimination rules, though ones that are non-branching and involve only information practically available to the agent and with which she can *operate*. The only branching rule implements the Principle of Bivalence (PB), which allows for the introduction of *hypothetical* information from no premises and thus can be used anywhere in a derivation. An unrestricted application of PB isn't amenable to proof-search, but it can be restricted to applications on the set of *subformulas* of the initial assumptions without affecting completeness. In this regard, the introduction rules can potentially be applied on and on, yielding ever more complex formulas. Yet, they can also be tamed so as to satisfy the subformula property while preserving completeness.

It's in terms of the PB rule that a measure of complexity of derivations is introduced. Namely, a derivation's *depth* is defined as the maximum number of nested applications of PB needed to obtain it. $\mathcal{NT}$ is advantageous over KE or KI alone, since it reduces the number of PB instances required to obtain a derivation and is closer to human reasoning. This generates a

*Funded by CONICET Postdoctoral Fellowships Program.
†Funded by UNAM-DGAPA-PAPIIT grant IN101723.

$\dfrac{\mathsf{T}\,A}{\mathsf{T}\,A \vee B}$	$\dfrac{\mathsf{F}\,A \quad \mathsf{F}\,B}{\mathsf{F}\,A \vee B}$	$\dfrac{\mathsf{F}\,A}{\mathsf{F}\,A \wedge B}$	$\dfrac{\mathsf{T}\,A \quad \mathsf{T}\,B}{\mathsf{T}\,A \wedge B}$	$\dfrac{\mathsf{F}\,A}{\mathsf{T}\,A \to B}$	$\dfrac{\mathsf{T}\,B}{\mathsf{T}\,A \to B}$	$\dfrac{\mathsf{T}\,A \quad \mathsf{F}\,B}{\mathsf{F}\,A \to B}$
$\dfrac{\mathsf{T}\,A}{\mathsf{F}\,\neg A}$	$\dfrac{\mathsf{F}\,A}{\mathsf{T}\,\neg A}$	$\dfrac{\mathsf{T}\,A \vee B \quad \mathsf{F}\,A}{\mathsf{T}\,B}$	$\dfrac{\mathsf{F}\,A \vee B}{\mathsf{F}\,A}$	$\dfrac{\mathsf{F}\,A \wedge B \quad \mathsf{T}\,A}{\mathsf{F}\,B}$	$\dfrac{\mathsf{T}\,A \wedge B}{\mathsf{T}\,A}$	$\dfrac{\mathsf{T}\,A \to B \quad \mathsf{T}\,A}{\mathsf{T}\,B}$
$\dfrac{\mathsf{T}\,A \to B \quad \mathsf{F}\,B}{\mathsf{F}\,A}$	$\dfrac{\mathsf{F}\,A \to B}{\mathsf{T}\,A}$	$\dfrac{\mathsf{F}\,A \to B}{\mathsf{F}\,B}$	$\dfrac{\mathsf{T}\,\neg A}{\mathsf{F}\,A}$	$\dfrac{\mathsf{F}\,\neg A}{\mathsf{T}\,A}$	$\dfrac{\mathsf{T}\,A \vee A}{\mathsf{T}\,A}$	$\dfrac{\mathsf{F}\,A \wedge A}{\mathsf{F}\,A}$
			$\dfrac{}{\mathsf{T}A \mid \mathsf{F}A}$			

Table 1: $\mathcal{NT}$ (symmetry of $\vee$ and $\wedge$ is assumed for brevity)

hierarchy of k-depth approximations to CPL, each one tractable whenever the application of PB and the introduction rules are restricted to a suitable subset of formulas as conclusions. Less restrictions on that subset yield deductively more powerful approximations, and tractability crucially depends on appropriate restrictions thereof.

Negative constraints A valuation is a mapping v from any set of formulas Φ to the set of values $\{0, 1, ?\}$, respectively standing for *informational* truth, falsity and indeterminacy. These values are partially ordered by the usual flat relation $\preceq$, defined as $? \preceq x$ and $x \preceq x$ for each $x \in \{0, 1, ?\}$. If v, w are valuations on Φ, then w is a *refinement* of v, if and only if $v(A) \preceq w(A)$ for all $A \in \Phi$. It is *proper*, if there is a $B \in \Phi$ such that $v(B) \prec w(B)$. We introduce concise notation for refinements consisting of changing the value of a single formula: $v^{A:=x}$ is a refinement of v such that $v^{A:=x}(A) = x$ and $v^{A:=x}(B) = v(B)$ for $A, B \in \Phi$, $B \neq A$, $x \in \{0, 1, ?\}$. The classical truth conditions imposed by, e.g. the standard truth-tables, are not suitable for the notion of informational truth. For example, if an agent holds the information that $A \vee B$ is true, she does not necessarily holds the information that A is true or that B is true. Similarly for all cases where the reading of the classical truth-table going from the formula to its components is informationally non-deterministic. This prevents us from giving a direct definition of admissible valuation. Instead, the *negative* constraints expressed in the tables below detect valuations that are *inadmissible* for any agent who 'understands' the informational meaning of the connectives:

A	B	$A \vee B$
0	0	1
1	?	0
?	1	0

A	B	$A \wedge B$
0	?	1
?	0	1
1	1	0

A	B	$A \to B$
1	0	1
0	?	0
?	1	0

A	$\neg A$
1	1
0	0

An agent who understands this meaning can *update* her information state by *uniquely* determining, from her practically available or operational information, the value of formulas that were previously indeterminate. For example, if she holds that $A \wedge B$ is false and B is true, then she can update her state with A being false by complying with the tables for negative constraints, since the refined state with A being true is inadmissible. This 'local' task is computationally and cognitively easy.

Basic game Our semantics is defined in terms of a win-lose perfect information game of two players. An *approximation game* $\mathcal{G}_k(\Gamma, C)$ is given by a set of formulas Γ, a single formula C and parameter $k \in \mathbb{N}$. In the *basic* version of the game, we fix the set of formulas available to the players during the game (the 'game board') to the set of subformulas of $\Gamma \cup \{C\}$, denoted $\mathsf{sub}(\Gamma \cup \{C\})$. This set constitutes the 'smallest' game board. Now, the goal of the first player called Questioner is to show that C follows from the set of initial assumptions Γ, while the goal of the second player, Responder, is the opposite. For any $i \in \mathbb{N}$, a *state* $S_i = (\Gamma_i, v_i, n_i)$

of a game $\mathcal{G}_k(\Gamma, C)$ is given by a set of formulas Γ_i, a valuation $v_i : \Gamma_i \cup \{C\} \to \{0, 1, ?\}$ and a parameter $n_i \leq k$. The valuation v_i represents explicit information currently held by Q, which is updated during the game via admissible refinements. The number n_i is a counter for the number of questions used in the game up to the state S_i. At the beginning of a game, i.e. at the state S_0, v_0 evaluates all the formulas in Γ to 1, while C and all the formulas in $\mathsf{sub}(\Gamma \cup \{C\}) \setminus \Gamma$ are evaluated to $?$. The value of n_0 is set to 0.

Moves The moves in the approximation game consist of admissible (possibly improper) refinements of a current valuation. An *answer* consists of determining the value of some currently indeterminate formula B, a task that is only allowed if there is a unique admissible proper refinement determining B. A *question* is an explicit request for a determinate value of a formula A, currently indeterminate. A question itself does not involve a change of the current valuation. An *inadmissibility detection* means that Q found a currently indeterminate formula such that each of the two proper refinements making it determinate is inadmissible. It intuitively corresponds to detecting that the answers given by R, if any, lead to a situation violating the negative constraints.

Formally, a move is a couple lA, such that A is a formula and $l \in \{0, 1, ?, \curlywedge\}$ is a label. Sequences of moves are called *histories*, so $h = l_1 A_1 \ldots l_m A_m$. We denote by H the set of all histories and by $h \sqsubseteq h'$ the relation 'h is a subsequence of h''. Let $S_i = (\Gamma_i, v_i, n_i)$ be the current state of a game and h_i be the current history, that is the history up to S_i. Then a move lA is *legal* in S_i if and only if it has not been played (i.e. $h_j lA \not\sqsubseteq h_i$ with $j < i$) and:

- (**question**) $l = ?$, $n_i < k$, and *at least one* of the proper refinements $v_i^{A:=1}$ and $v_i^{A:=0}$ is admissible, then the game proceeds to $S_{i+1} = (\Gamma_i, v_i, n_i + 1)$; or
- (**answer**) $l \in \{0, 1\}$ and there is a *unique* admissible proper refinement $v_i^{A:=l}$, then the game proceeds to $S_{i+1} = (\Gamma_i \cup \{A\}, v_{i+1}^{A:=l}, n_i)$; or
- (**inadmissibility**) $l = \curlywedge$ and *neither* proper refinement $v_i^{A:=0}$ nor $v_i^{A:=1}$ is admissible, then the game proceeds to $S_{i+1} = S_i$ (and ends).

Player function The roles of players are not symmetric, the possibilities of R are quite restricted, since he cannot ask questions and can answer only if Q explicitly asks. In contrast, Q can ask questions any time and she can also play answers, which can be seen as replies to an 'implicit' question Q asks to herself, thus updating her explicit information. Q can also detect inadmissibility in the sense mentioned above, using the $\curlywedge$-move. Formally, the only histories which are moves for R are those of the form $h?A$ for some $h \in H$. All the other histories, including the empty one, are moves for Q.

End of the game Winning conditions for Q are simple, either the valuation of the conclusion C is set to 1 in some move (either by Q or by R) or she detects an inadmissibility. R wins if Q cannot move any more, which includes the case when she has spent all her k questions. In contrast, setting the conclusion false does not suffice, for Q might still spot inadmissibility on a formula in the remaining part of the game. Formally, h_i is a terminal history and S_i is a terminal state of the game $\mathcal{G}_k(\Gamma, C)$ iff:

- (**conclusion true**) $h_i = h_{i-1} 1C$ and consequently $v_i(C) = 1$;
- (**inadmissibility detected**) $h_i = h_{i-1} \curlywedge A$, and thus neither proper refinement with $v_i(A) = 0$ nor $v_i(A) = 1$ of v_i is admissible;

- (**no moves**) $h_i \neq h_{i-1}?A, n_i = k$, and there is no $B \in \Gamma_i$ for which there is a *unique* proper admissible refinement $v_i^{B:=l}$, $l \in \{0,1\}$.

There are no special *procedural* rules in the basic version of the game. Q starts the game and then plays answers and questions in an arbitrary order as long as she can, i.e. until the end of the game is reached. R moves only when asked a question. No player can repeat her/his moves according to the definition of a legal move.

Correspondence theorem We show that here is a k-depth proof of $\mathsf{T}\,C$ from $\mathsf{T}\,\Gamma$ in $\mathcal{NT}$ over the sub-bounded ('analytic') *search space* [1, 2.1] if and only if there is a winning strategy for Q in the basic game $\mathcal{G}_k(\Gamma, C)$.

Adequacy and intuitiveness Some natural *liberalizations* of the basic version of the game are, for example: (i) setting a wider 'board game', in particular, allowing *questions* from a superset of $\mathsf{sub}(\Gamma \cup \{C\})$; (ii) allowing for controlled move repetition. By contrast, some natural *restrictions* are, for instance: (i) Q asks *only* when she cannot answer herself (corresponding to pushing PB as down as possible); (ii) questions restricted to atomic formulas. In any case, a remarkable intuition is that questions are a resource worth keeping! Actually, it is exactly when Q exploits the information she holds as much as possible that the answers of R correspond only to the introduction of information that was not even implicitly contained in the information held by Q.

Thoughtful questions might guarantee a win, while hasty questions correspond to a wasteful play. Accordingly, question selection is not a trivial task, whose difficulty increases proportionally with the number of questions needed and the freedom on the subset from which these are selected. The levels of the hierarchy of approximations, under proof-theoretical restrictions, can intuitively be associated with increasingly better questioners, in terms of their connective-meaning mastery and their 'ingenuity' when selecting questions, under suitable playing freedom. More liberalized settings are naturally associated with more competent questioners, and thus with more efficient playing in that the number of questions needed can dramatically decrease.

Strategies by the Questioner arise naturally when balancing her question 'budget' with playing freedom and competence thereof. These strategies intuitively correspond to different *procedures* when implementing the background proof-theory.

Finally, we envisage *extensions* of our semantics to non-classical depth-bounded approximations, such as FDE [2] and IPL [5], by modifying our game in the spirit of dialogics.

References

[1] M. D'Agostino, D. Gabbay, C. Larese, and S. Modgil. *Depth-bounded Reasoning. Volume 1: Classical Propositional Logic.* College Publications, 2024.

[2] M. D'Agostino and A. Solares-Rojas. Tractable depth-bounded approximations to FDE and its satellites. *Journal of Logic and Computation*, 34(5):815–855, 2024.

[3] G. Lakemeyer and H. Levesque. A first-order logic of limited belief based on possible worlds. In *Proceedings of the 17th International Conference on Principles of Knowledge Representation and Reasoning*, volume 17, pages 624–635. IJCAI Organization, 9 2020.

[4] B. Mendonça. Dialogue games and deductive information: A dialogical account of the concept of virtual information. *Synthese*, 202(3):73, 2023.

[5] A. Solares-Rojas, P. Baldi, and R. O. Rodriguez. Free-variable KE tableaux for IPL. Manuscript submitted for publication, 2024.