Abstract: I show how perceptual and apriori elements of chess knowledge may be cleanly distinguished. Examples suggest that apriori elements may be either demonstrative (i.e., implicitly proof-based) or inductive, and may be socially distributed. The examples illuminate liberalized neo-rationalist accounts of apriori warrant, such as that of Tyler Burge. A further example suggests a complication, however, for Burge's treatment of computer-enabled apriori warrant.

The *Apriori* in Chess

Bernard W. Kobes
Department of Philosophy
Arizona State University

"... one did not have to deal with visible, audible, palpable pieces whose quaint shape and wooden materiality always disturbed him and always seemed to him but the crude, mortal shell of exquisite, invisible chess forces. When playing blind he was able to sense these diverse forces in their original purity. He saw then neither the Knight's carved mane nor the glossy heads of the Pawns — but he felt quite clearly that this or that imaginary square was occupied by a definite, concentrated force, so that he envisioned the movement of a piece as a discharge, a shock, a stroke of lightning — and the whole chess field quivered with tension, and over this tension he was sovereign ..." — Vladimir Nabokov, *The Defense*

"Of chess it has been said that life is not long enough for it. But that is the fault of life, not chess." — William Ewart Napier (attributed)

The game of chess has often been used as a source of analogies in philosophical discussions of language, mind, rules, and normativity. I will argue that chess can be a fruitful source of specific examples of apriori epistemic warrant, examples that illuminate recent liberalized neo-rationalist notions of apriority.[1] A focus on this sharply delimited domain permits a degree of resolution that is missing from the larger, more general analogies.

I will understand by the apriori primarily a type of *epistemic warrant*. A subject S who believes that p does so with apriori warrant just in case S has an epistemic warrant for believing p, the force of which derives from reason or understanding and not from sense experience. On this conception, apriority does not entail infallibility, unrevisability, or certainty. S may have apriori warrant for believing a proposition that later turns out to have been false. There may be cases in which S has both apriori and empirical warrant for her belief that p, and cases in which S's apriori warrant is over-ridden by empirical counter-considerations.

Secondarily, we may say that S's *knowledge* that p is apriori just in case S has apriori warrant adequate for knowledge. That is, given S's apriori warrant, S's belief that p stands in need of no *further warrant* to count as knowledge. (Of course, other conditions for knowledge must also be satisfied.)

Finally, we may say that a *proposition* p is apriori just in case p is *knowable* apriori; more exactly, p is the *kind of proposition* that subjects relevantly like us could, perhaps under severely idealized conditions, know apriori.

The word 'chess' may be construed more or less broadly. On one broad and wholly legitimate usage, 'chess' refers to a game that has evolved historically and may take different forms. Thus we may say that in the old Arab form of chess, *shatranj*, the *firzan* could only move diagonally one square at a time, but that the firzan has now been supplanted by the much more powerful queen; or we may say that *shogi* is Japanese chess. However, 'chess' may also be used more narrowly, such that the pieces and the well-defined rules of the contemporary game so called are essential to it, and even the smallest alteration or evolution of the rules, in a thought experiment or in the real world, would result in a distinct game. We might still call such a variant game 'chess', but given strict usage that would amount to a different word, or a different sense of the word. Analogous points apply to words like ‘queen’. In this paper I adopt the strict and narrow usage.

Full mastery of the concept *chess*, then, requires knowing all the rules of chess. Thus the proposition that *the (chess) queen may be moved any number of squares along the rank, file, or diagonal*, is apriori knowable, since the force of the warrant for believing it may derive wholly from understanding the proposition.[2] Sense perception— e.g., observation of play, listening to instruction, or reading— will be causally and psychologically involved in a subject’s coming to believe the proposition. But given our employment of the strict and narrow concept *chess*, the force of a subject’s warrant for believing the proposition will derive from her rational apprehension of the proposition itself. By contrast, the force of the subject’s warrant for believing the proposition *there is a (chess) queen in front of me* will derive at least partly from her sense perception of the physical token.

Chess is a finite game, in the sense that only a finite number of positions are possible. Indeed, under certain simplifying assumptions, at most 6,350 moves can be made in a single game.[3] In game-theoretic terms, chess is a *determinate* two-person zero-sum game with perfect information. To say that chess is determinate is to say that in each possible chess position, exactly one of the following three alternatives obtains: either White has a winning strategy, or Black has a winning strategy, or both players have strategies to avoid losing. This idea was made precise by the set-theorist Ernst Zermelo in 1913. In the first formal result of game theory, Zermelo proved that if a player can force a win, then she can do so in a number of moves bounded by the number of possible chess positions.[4]

If White can force checkmate, let us call the position a *win for White*. Let a *position* be any arrangement of the pieces that could be reached using standard rules, and it is specified which side is to play, and there is no possibility of present or future ambiguity as to which moves are legal.[5] The notion of a *win for White* may be defined recursively, as follows:

Base clause:
(1) A position is a *win for White* if Black stands checkmated. (Black stands checkmated iff Black is to move, Black's king is under direct attack, and Black has no legal move.)

Recursion clauses:
(2a) A (non-concluded) position is a *win for White* if White is to move, and some move available to White results in a win for White.
(2b) A (non-concluded) position is a *win for White* if Black is to move, and each move available to Black results in a win for White.

The notion of a *win for Black* is recursively defined analogously.[6] These notions are objective and independent of the psychologies or abilities of the players.

Now it may be apriori demonstrable that a given position is a win for White. Figure 1, for example, is a mate-in-2 composition from the 13th century. The position shown, call it P_1, may be construed as a *chess* position since it does not use pieces whose rules have changed since the 13th century. That *position P_1 is a win for White* has a precise meaning given the recursive definitions, and that *White can force checkmate in two*

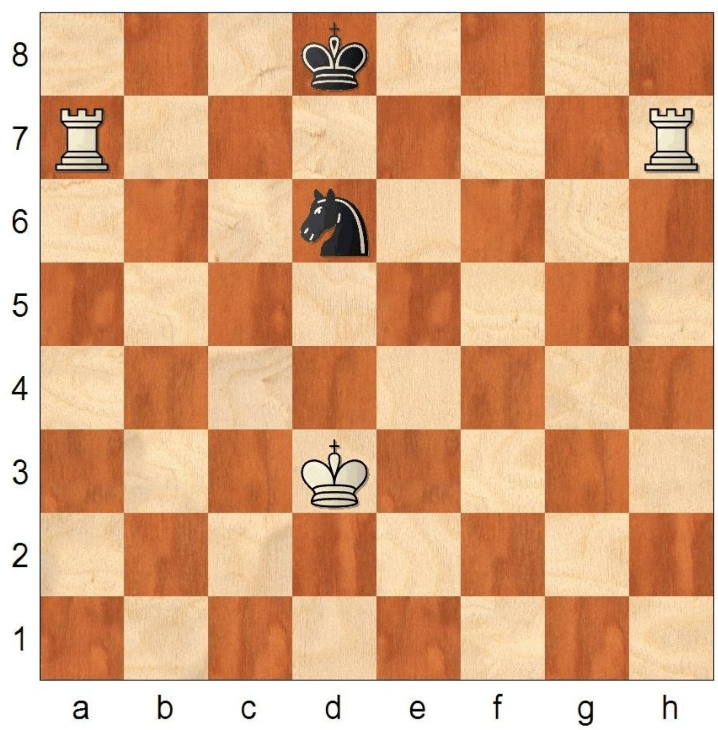

Fig. 1: Position P_1. White to play and mate in two moves. The solution is given in an endnote. From the 13th century *Bonus Socius* manuscript.

moves is apriori demonstrable, as chess-playing readers may wish to verify.[7] Note that clause (2a) is first used, then (2b), then (2a) again, then (1). In general, if a win for White can be demonstrated by a recursive argument the longest path of which uses clause (2a) n times, then the position is a *win for White in (at most) n moves*.

In this and later examples 'P_i' should be understood to abbreviate a structural-descriptive rigid designator of the relevant position.[8] That *the position before me (*or *the position depicted here) is a win for White* is, I assume, partly empirical, because one's warrant for it must derive partly from one's sense perception of the position. But this proposition factors cleanly into an empirical component *the position before me is position P*, and an apriori component *position P is a win for White*.

In solving this puzzle, typically White will literally see with her eyes (or feel with her fingers, if she is blind), or visually (tactually) imagine, that after White's key move, then for each move of Black's knight or king, White's rooks can safely sweep through all squares accessible to Black's king. White *visually notices* these features of the position. But the role of sense perception is to *bring to mind* and to *causally enable a cognitive grasp of* a sequence of chess positions and their relations to each other. The subject's *epistemic warrant for believing* that those relations obtain is distinguishable, and derives wholly from her rational understanding of the relevant positions grasped, and of the nature of chess. The demonstration from recursive clauses, described above, makes explicit the logic of her rational understanding.

Figure 2, with Black to play, is also known to be a win for White. It is *metaphysically possible* for a person to carry out a recursive demonstration of this fact, although doing so would require severe (and perhaps nomologically impossible) idealization of time, space, and memory. That *position P_2 is a win for White* is therefore an apriori proposition, since it is *knowable* apriori. But all of this is *in fact* known via computer search, which reveals that White requires an astonishing 517 moves to convert to a familiar winning sub-endgame! [9] Best play for both sides is assumed; White aims for the fastest conversion, while Black aims for the slowest conversion. This sort of computer endgame has been called "chess with God"; the win is an extremely lengthy, beautiful, and surpassingly strange sequence of maneuvers in which no pattern or progress can be discerned even by the strongest human grandmasters.

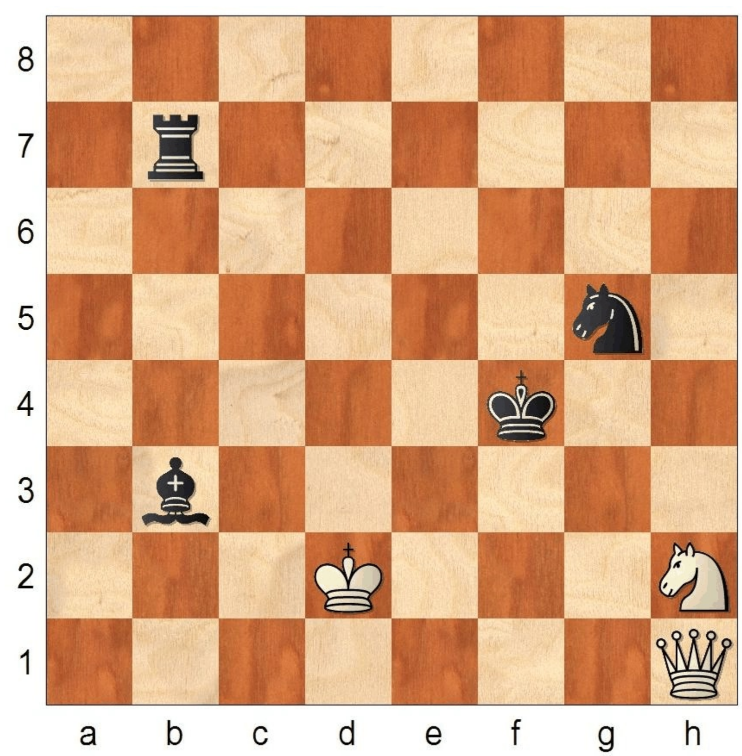

Fig. 2: Position P_2, Black to play, White wins. From item #316 at http://www.xs4all.nl/~timkr/chess2/diary.htm. Computer search by Marc Bourzutschky and Yakov Konoval, May 2006, shows that White requires 517 moves to convert to a familiar winning sub-endgame. Best play for both sides is assumed. White aims for the fastest conversion, Black aims for the slowest conversion.

Tyler Burge has argued that reliance on computers in mathematical demonstration does not by itself preclude the demonstration's being apriori.[10] Burge's argument is complex, but the key idea is that if one can use the computer source as an amplification of one's own rational powers, and one is able to appreciate the source's powers of inference "from the inside", then reliance on the computer source is compatible with one's warrant being apriori. One must, on Burge's view, be able to incorporate the source's ability into one's own point of view. "One's empirical activity comes to be submerged into one's knowing how to use the computer as an extension of one's own rationality."[11] Perception of the computer's activity then plays an *enabling* role rather than a *justificatory* role.

In chess endgame search, the computer's fundamental algorithm is not difficult to take on board; at bottom and conceptually, it is only a few steps removed from our discussion above of the mate in two. Indeed, Burge holds that our warrant for *position* P_2 *is a win for White* is apriori.[12] Yet the strikingly *alien* quality of the resulting endgame maneuvers would seem to be in tension with using the computer as "an extension of one's own rationality". A human player is used to seeing a kind of logic at a larger scale in the theater of operations; for example, larger goals will typically parse into intelligible sub-goals ("first drive the opposing king into a corner of the same color as your bishop, then bring your knight to attack squares adjacent to the corner"). No such progression or logic is discernible when viewing chess endgames played "by God".[13]

Our registering the computer's play as *alien*, relative to familiar patterns of human endgame play, is in tension with our incorporating the computer's powers into our own point of view. It makes salient our epistemic reliance on the well-functioning of a physical device. I will assume, therefore, that our actual warrant for *position* P_2 *is a win for White* is partly empirical, and our actual knowledge of it is aposteriori. However, the proposition itself remains apriori, because given that it is *actually true*, we easily rationally apprehend the metaphysical possibility of our carrying out an apriori demonstration of it.

The point here is not fundamentally about iteration; if the proposition is apriori, then the proposition that it's apriori is also apriori— since it's metaphysically possible, under severe idealization, for us to demonstrate apriori that the proposition is apriori. The point is rather that it's the *proposition* that is apriori, while our *actual warrant* for it, and our *actual knowledge* of it, are partly empirical. But striking iterations do follow: our actual warrant for, and our actual knowledge of, the complex proposition *that position* P_2 *is a win for White is an apriori (true) proposition* turn out to be partly empirical.

Figure 3 shows the final position in the shortest decided serious game ever played by strong players in a tournament.[14] The game went: 1. d4 Nf6 2. Bg5 c6 3. e3 Qa5 check 4. resigns. White's third move was a blunder, and he resigned because he knew that the position was obviously and objectively a win for Black, and that Black was a sufficiently strong player that he could realize the objective win without difficulty. I want to consider the proposition that *position* P_3 *is a win for Black.*

The "hypermodern" chess theorist Gyula Breyer famously quipped, "After 1.e4 White's game is in its last throes." (A theme of hypermodern theory was that centrally placed pawns can become vulnerable

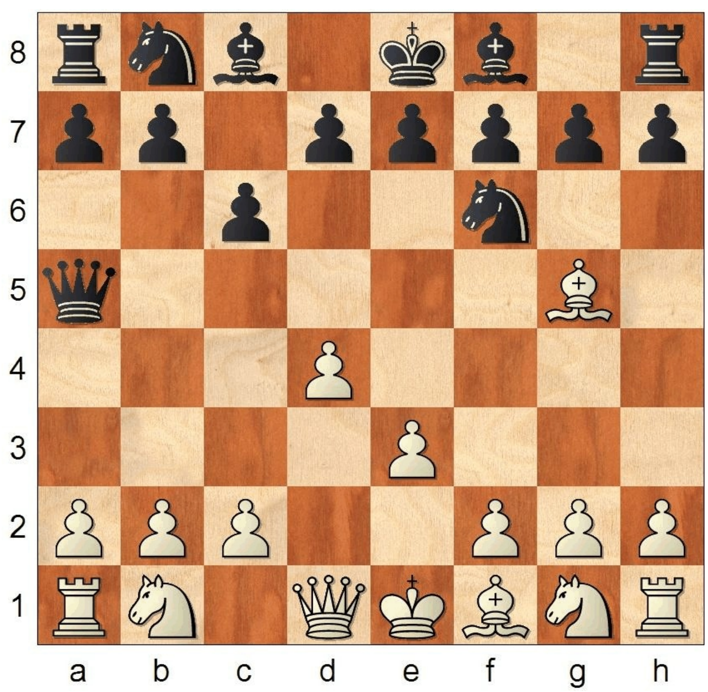

Fig. 3: Position P3, White to play. Djordjevic (White) – Kovacevic (Black), Bela Crkva (1984). This is the shortest decided tournament game between strong players. White is lost, and resigned.

targets for the opposition's long-range pieces.) On a literal construal Breyer's statement was certainly and obviously false— its status as a *bon mot* depends on that. I want to also consider the related proposition that *the opening position in chess is <u>not</u> a win for Black*.

I hold that *position* P_3 *is a win for Black*, and *the opening position in chess is <u>not</u> a win for Black*, are both apriori propositions. For it is metaphysically possible that a person carries out recursive demonstrations of them. The metaphysical possibility here assumes severe, and nomologically impossible, idealization of life, memory, space, and time. Yet if the propositions are *actually true*, then we can easily rationally apprehend that apriori demonstrations of them are metaphysically possible. And the propositions *are* actually true. But what is the nature of our actual warrant for believing this?

In each example, no human being or computer has ever demonstrated it, and no human being nomologically could, because of limitations of life, memory, space, and time. Even a computer-assisted demonstration is nomologically impossible for *the opening position in chess is <u>not</u> a win for Black*. (Perhaps in the case of *position* P_3 *is a win for Black* a computer-assisted demonstration is nomologically possible.) So our actual warrant is plainly *inductive* rather than deductive, although serious chess players would take both propositions to be obviously and certainly true— so much so that in chess literature they would normally be simply presupposed rather than articulated. There is some temptation to suppose that such an induction must be at least partly empirical, because it depends on a vast accumulation of *experience*, primarily over-the-board experience of the material advantage of a bishop in the first example, and familiarity with chess opening theory in the second example. But I want to resist this idea. I hold that our actual warrant for believing both propositions is *inductive*, *apriori*, and *socially distributed*.

The body of our communal familiarity with positions of material advantage, and with chess opening theory, can be construed as a large, socially distributed probabilistic search for evidence of objective positional properties like *win for Black*. Of course, chess is a zero-sum game, so within the context of a particular chess struggle White's and Black's interests are strictly opposed. But from our present larger perspective White and Black can be seen as *cooperating* in a socially distributed inductive search for evidence of objective chess truths. In this cooperative inquiry, perception of the pieces is psychologically and pragmatically necessary, but the force of the warrant for our inductive conclusions derives wholly from reason and understanding of the positions and of their relations to each other.

We can imagine a *single mind* carrying out this probabilistic exploration of the space of recursive possibilities, searching for evidence of the objective condition *win for Black*. The nature and outlines of such quasi-mathematical inquiry are familiar from smaller-scale examples of a single mind searching lines in a chess puzzle or endgame, deciding first that it is *probably* a *win for White* (say), and later finding a conclusive demonstration that it is. There is nothing in this conception resembling the alien quality of the computer proof that P_2 is a win for White. We can assimilate the socially distributed inquiry to the point of view of a single, recognizably human thinker. The resources epistemically relied on are those of reason and understanding of abstract matters, and not those of perception of concrete particulars, though perception supplies us with the positions to think about.

So we should think of the induction as apriori and socially distributed. That is the nature of our actual warrant for believing, and our actual knowledge that, *position P_3 is a win for Black*, and that *the opening position in chess is <u>not</u> a win for Black*. These propositions, I think, are *doubly* apriori: first in that they are actually known with *inductive* apriori warrant, and second in that it's metaphysically possible for them to be known with *deductive* apriori warrant. (But my actual warrant for the latter claim is apriori inductive.)[15]

Chess has become a test bed for theories of expertise in cognitive science, because chess skill can be easily measured and subjected to lab experiments.[16] Similarly chess can serve as a test bed for articulating the concept of, and theories about, apriority. The "clean" chess examples here adduced bring out unexpected permutations implicit in the liberalized neo-rationalist treatment of apriority. Empirical and apriori elements can be distinguished with a precision that has not been previously achieved. The apriori elements may be either demonstrative or inductive, and may be socially distributed. I conjecture that these ideas are applicable to messier domains in which apriority is said to figure, e.g., to our understanding of mathematics, and to our understanding of language and of linguistically-mediated social institutions.

Notes

1. See, for example, T. Burge, "Content Preservation", *Philosophical Review*, Vol. 102, No. 4 (October 1993), pp. 457-488; T. Burge, "Computer Proof, Apriori Knowledge, and Other Minds", *Philosophical Perspective 12: Language, Mind, and Ontology* (Blackwell Publishing, 1998), pp. 1-37; L. Bonjour, "A Rationalist Manifesto", *Canadian Journal of Philosophy*, Supplementary Volume 18 (1992), pp. 53-88; L. Bonjour, *In Defense of Pure Reason: A Rationalist Account of A Priori Justification* (Cambridge University Press, 1998).

2. The example is over-simplified in that it omits reference to obstacles and captures.

3. See http://www-history.mcs.st-andrews.ac.uk/Projects/MacQuarrie/Chapters/Ch4.html for the calculation, which assumes that at least one of the two players claims a draw when permitted to do so by the "50-move draw" rule. The claim that only a finite number of moves are possible in a single game is stronger than the claim that there are only a finite

number of chess positions, and is not required for Zermelo's result. Nor is it strictly required for my own arguments, though I assume it as a convenient simplification.

4. Ernst Zermelo, "On an Application of Set Theory to the Theory of the Game of Chess" (1913). See Ulrich Schwalbe and Paul Walker, "Zermelo and the Early History of Game Theory", *Games and Economic Behavior* 34 (2001), pp. 123-137; also available on-line at http://www.econ.canterbury.ac.nz/personal_pages/paul_walker/pubs/**zermelo**-geb.pdf Zermelo did not assume that the number of moves in a game must be finite. Schwalbe and Walker describe common misconceptions of what Zermelo proved. They also append the full translated text of Zermelo's article.

5. Thus information relevant to castling, *en passant* captures, draw by repetition, and the 50-move draw rule is specified as part of the position.

6. These recursive definitions are original with me, though implicit in Zermelo's different and more complex formalism. Certain rules governing draws and resignation have not been taken into account. The simplification does not matter for purposes of this paper.

7. Key: 1. Rh7-g7. For each of Black's knight or king moves, some White rook move delivers a back-rank mate.

8. There are standard notations for describing chess positions; one widely used system is Forsyth-Edwards Notation (FEN). Thus 'P_i' could be taken to abbreviate a FEN description of the position, which would be a structural-descriptive rigid designator. Likewise, for example, the arabic numeral '1,234' rigidly designates a number by describing it as a certain sum of powers of 10.

9. From http://www.xs4all.nl/~timkr/chess2/diary.htm, item # 316. Black to play, White wins. Computer search by Marc Bourzutschky and Yakov Konoval, May 2006, shows that White requires 517 moves to convert to a winning sub-endgame. Best play for both sides is assumed. White aims for the fastest conversion, Black aims for the slowest conversion. At this link, see also items ## 311, 298, 294, and 282.

10. Burge, "Computer Proof, Apriori Knowledge, and Other Minds", op. cit.

11. Burge, "Computer Proof, Apriori Knowledge, and Other Minds", p. 31.

12. Burge, personal communication.

13. As Tim Krabbé put it, "It's like walking a treadmill in the gym. All you ever see is the walls of the place, and suddenly you're in Kathmandu." Op. cit. item # 298.

14. Djordjevic - Kovacevic, Bela Crkva, 1984. Remarkably, this sequence of moves was subsequently played in a different serious tournament. See Tim Krabbé's discussion at: http://www.xs4all.nl/~timkr/records/records.htm#Shortest%20game

15. Burge, "Computer Proof ...", op. cit., p. 8 and fn. 14, cites the following as instances of *inductive* mathematical knowledge: Newton's knowledge of elementary truths of calculus, our knowledge of the consistency of arithmetic, Zermelo's belief in the axiom of choice, and possibly also new axioms in descriptive set theory and Church's thesis. These are instances of mathematical knowledge in the absence of proof. They are not self-evident. They do not rely on sense experience or on natural science. Rather, Burge argues, mathematical fruitfulness is key; he cites supporting considerations from I. Lakatos and from G. Pólya.

16. Cf. Philip E. Ross, The Expert Mind, *Scientific American*, August 2006, pp. 64-71.