Game Counterpossibles

Felipe Morales Carbonell

There is an ongoing debate about the status of counterpossibles, counterfactuals with impossible antecedents. There are roughly two camps: one defends the view that counterpossibles are vacuously true, while the other defends the view that counterpossibles can be non-vacuously true or false.[1] One of the main motivations for the non-vacuity position is the defense of a series of metaphysical views (Nolan (2014) gives an overview of the many topics in which counterpossibles might play a crucial role). However, metaphysics is contentious enough that the case for counterpossible non-vacuity has remained inconclusive. Recently, some authors who defend non-vacuity have tried the different strategy to show that there are less contentious independent contexts in which it is necessary to distinguish between the truth value or acceptability of counterpossibles. For example, Baron, Colyvan, and Ripley (2017) argue that there can be genuine mathematical counterpossibles, and Tan (2019) argues that there can be genuine counterpossibles in the natural sciences. Without holding an opinion on whether those applications of the strategy work, the main aim of this paper is to examine whether this strategy can pan out in the context of games and play. I will argue that the strategy does indeed work in this context, albeit with significant restrictions. By observing how counterfactuals behave in the context of games, we can get indirect evidence about whether this strategy can be useful more broadly. If we can only account for counterpossibles with restrictions even in the case of games, there might be restrictions also for the use of counterpossibles in different contexts.[2]

1 Some generalities about counterfactuals in games

Reasoning in the context of games is often *explicitly conditional* reasoning. For concreteness, take chess. Planning a move involves reasoning about the consequences of the move: if we move a pawn to a certain position, the king will be exposed; if we castle, the attacker will have to move their knights to a certain area of the board if we want to push from this side; and so on. The conditionals that are evaluated in the context of play encode information about the outcomes of hypothetical scenarios where strategic choices are made, and they can be highly complex: in multi-player games, they are often not only about the direct effects of actions in the game state, but also about the beliefs of other participants of the game about the game itself: 'if I move this piece here, my opponent will think that I plan to do this, so he will…'.

While some of the relevant conditionals in game playing are indicative (as in the examples I just gave), it can be equally common to reason using counterfactuals of the form:

1) If A were to happen, B would happen.

There is a rich literature on counterfactual reasoning in the context of games from the perspective of game theoretical issues (cf. Binmore (1987), Bicchieri (1988), Stalnaker (1996) and Skyrms (1998)).

[1] Lewis (1973), Williamson (2007), Emery and Hill (2016) and Vetter (2016) defend the orthodoxy. Nolan (1997), Kim and Maslen (2006), Brogaard and Salerno (2013), Kment (2006), Priest (2016), Berto et al. (2018), Locke (2019), Tan (2019), and Berto and Jago (2019) defend non-vacuity.

[2] The topic of counterpossibles in games is interesting also from a different perspective. One could defend the non-vacuity of counterpossibles from an anti-realist perspective by treating counterpossible-talk as a kind of game or fiction (cf. Kim and Maslen (2006)). Our discussion here could also bear on the scope of this research direction.

For example, we can describe the prisoner's dilemma in counterfactual terms: in that formulation, it is about what would happen if a number or individuals were made to choose between cooperating or defecting against each other, given certain payoffs. The use of counterfactuals instead of future indicative conditionals seems to provide greater flexibility, since it allows the evaluation of situations that are detached from the current circumstances. Think of the different contexts in which we would use the conditionals 'if Liam doesn't kick the ball to the right, someone else will' and 'if Liam hadn't kicked the ball to the right, someone else would have'. Clearly, there are contexts in which the truth conditions of these conditionals diverge.[3]

Counterfactual reasoning in the context of games can be also *backwards looking*, in case where what we are talking about isn't the outcome of an action, but what explained the action:

2) If A had happened, B would have happened.[4]

For example, given a surprising event in a game, we might reason about what explained it, so that we can then evaluate our future strategy. In some contexts we also might want to evaluate backwards looking counterfactuals for reasons that don't bear on future play at all. I will assume that both forward and backward looking counterfactual can be given a unified account.[5]

2 Counterpossibles in games

Ordinarily, we only consider game counterfactuals with possible antecedents. This is reasonable because we are interested in problems of evaluating courses of action, where the possibility of acting on the information given by those evaluations matters (a different way of putting this point is that when reasoning counterfactually when playing games, we are looking for *guidance*). However, this is not decisive on whether there couldn't be cases where we evaluate counterfactuals with impossible antecedents. In principle, it is possible for judgements about what we ought to do to be independent of judgements about what would happen, even though in many cases it is clear that our judgements about what we ought to do are informed by what we believe would happen.[6] Pushing this line of thought would force us to take a stance on a whole host of difficult issues.[7] Instead, we should examine whether there can be direct counterexamples to the restriction of having possible antecedents.

Consider the following chess board:

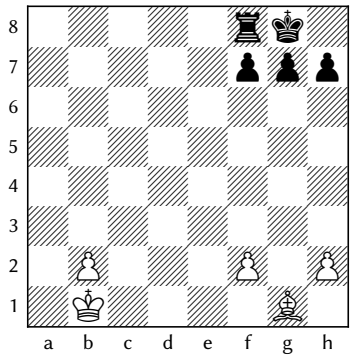
[3]It is important to observe, with Lewis (1973, 4), that there are apparently subjunctive conditionals which have the same truth conditions as indicative conditionals, so that the apparent use of subjunctive conditionals in reasoning does not immediately entail that we are dealing with counterfactual reasoning.

[4]In contexts where one could produce this conditional, it might be possible to also produce the conditional expressions 'if A happened, B must have happened', 'given that A happened, B must have happened'.

[5]Cf. Bennett (2003 ch. 18) for a defense.

[6]Sinnott-Armstrong (1984) raises this possibility about the 'ought implies can' principle, which he takes to be a mere conversational implicature. In his view, while someone literally could be obligated to do something that they wouldn't be able to do, it would be pointless to say that they ought to do it because such saying wouldn't provide advice.

[7]Cf. the debate between Streumer (2007) and Heuer (2010) on whether there can be reasons to do or try the impossible.



Two questions: a) how can we proceed from this position to a win? b) how did we arrive at the current position? I will leave the former aside. The second question is characteristic of retrograde (or "retro") analysis.[8] During retro analysis, one can make counterfactual judgments like

3) If white were in this position, the bishop in g1 would have, at some point in the course of the game, moved from h2 or along the a7–g1 diagonal.[9]

Now, it turns out that it is impossible to arrive at this position during actual play (that is, in a game that begins from the standard starting position). I will call this type of impossibility *chess-impossibility*.[10] Chess-impossible positions are also called illegal (cf. the FIDE's laws of chess (2018, 3.10.3)).[11] The relevant modality pertains to *states* of a game, although there is closely related modality that pertains to *game-plays* (which in the case of chess are sequences of states, but in other cases can be processes in a richer sense). In this paper I will limit myself to the discussion of chess-impossibility in the stative sense only, but I think that much of what I will say here can be applied to the process conception of possibility. In any case, there should be a closely related counterfactual that says:

4) If the board had come to be in this position, the bishop in g1 would have, at some point in the course of the game, moved from h2 or along the a7–g1 diagonal.

The antecedent of the counterfactual is chess-impossible, so we classify the counterfactual as *chess-counterpossible*.

How to evaluate counterfactuals like (4)? It is usually recognized that counterfactual evaluation is always performed against the background of some body of relevant assumptions. This body of

[8] Smullyan (1979) is the most accessible introduction to retrograde analysis.

[9] It is plausible that a seasoned player or problem solver would recognize the impossibility of the board *visually* or *imaginatively* instead of relying on explicit counterfactual reasoning. Cf. the psychological literature on chess cognition tracing back to De Groot (n.d.) and Chase and Simon (1973).

[10] For a more formal treatment of chess-possibility, see the appendix.

[11] Dawson and Hundsdorfer (1915) make an interesting distinction between impossibility and illegality:

> We use with forethought the word *illegal* to define any condition which could not arise in actual play. The word *impossible* is often used in the same sense, but it is not satisfactory, and we shall not use it. There is no such thing as an impossible position, provided you have enough chess-men in your box to draw on. The word always provoked Sam Loyd. "Impossible?" he would say, "you say these men could not have got into such a position! Why, they *are* in that position; I put them there myself!" To this no answer can be made. (9)

Dawson and Hundsdorfer's point is that all chess diagrams are *constructible*, whereas not all of those possible diagrams are legal or could happen in actual play. The size of the possibility spaces is vastly different: roughly speaking, there are 10^{71} possible diagrams ($((6x2) + 1)^{64}$), and while the number of possible legal positions is an open question, it is widely believed to be within the 10^{40} to 10^{50} range. Cf. Steinerberger (2015).

Sam Loyd (1841–1911) was a well-known chess problem composer and puzzle creator. For a very interesting overview of his position on the significance of chess impossibilities, which is more nuanced than Dawson and Hundsdorfer report, see White (n.d., 444–54).

assumptions is fixed by the context, which in turn is fixed by the task at hand. In the case of (4) we should presumably include assumptions about the rules of chess in this background; for example, that the bishop moves diagonally an arbitrary number of free spaces, and that the starting position of white's bishops is `c1` and `f1`. Since we know that bishops move diagonally and that the bishop in `g1` could not have started in its current position (call it P), the last move of the piece could have only initiated at either a position in the diagonals `a7-g1` or `h2-g1` (this is something that we can deduce or imagine). It is at this point that we could judge that if P were to happen, it would have followed such move (since they are the only apparently possible moves); that is, we could be disposed to accept (4). However, in both diagonals there are pawns blocking the bishop, which we also know couldn't have moved from their initial positions (pawns do not move backwards). So the bishop couldn't have arrived at `g1` from either direction, since they are blocked, and the position is impossible.[12] Here we can suspend judgement on whether this means that we should reject our initial acceptance of the counterfactual; in any case, the standard semantics gives the verdict that the counterfactual is vacuously true.[13]

Once we reach an impossibility like this, we might be interested in evaluating whether there are changes to the setup that would make it possible (for example, we might realize that the type of play that would follow from an illegal position is interesting in a way that we judge should be allowed). Since the impossibility follows from rules about the movement of the chess pieces, and more precisely of a subset of those pieces, one might want to exchange those rules for more suitable ones. This immediately puts us in the position to consider rules that would deliver situations which are strictly speaking impossible in the relevant sense (chess-impossible in the case of (5)). There may be several viable variations of the set of rules that would yield the wanted result.[14] Consider the following counterfactual:

5) If bishops in chess jumped over pieces of their own color once, the bishop in `g1` would have moved from `e3`.

Again, the antecedent of this counterfactual is chess-impossible since the bishop in chess does not jump over pieces of their own color. Conditional (5) codifies a change to the rules that would allow a bishop in the `a7-g1` diagonal to reach `g1` (since the diagonal `h6-c1` is empty, we can allow free movement for the bishop from its original position to `g1`). This might suggest that we should accept (5) as true in a way that doesn't follow automatically from the orthodox vacuity assumption.[15]

[12] It is plausible that a seasoned player or problem solver would recognize the impossibility of the board *visually* or *imaginatively* instead of relying on explicit counterfactual reasoning. Cf. the psychological literature on chess cognition tracing back to De Groot (n.d.) and Chase and Simon (1973).

[13] On the supposition that the antecedent is indeed impossible; otherwise we have reason to think that in the closest worlds, whenever the antecedent is true, the consequent is false, so the counterfactual evaluates as false. Suppose that we rejected (4), and moved on to judge that it is false; our options would be either a) to reject the orthodoxy about counterfactuals, or b) to reject the classification of (4) as a counterfactual conditional. Lewis (1973, 24) already considers the possibility that so-called counterpossibles might be *sui generis*, but dismisses it without much comment.

[14] We could come up with rules by transposing and varying the movesets of the relevant pieces (e.g., having the pawns move backwards or sideways) or of other pieces (e.g., having the bishop move like the knight). We could also come up with entirely new move ideas; for example, having the bishop wrap around the board (so that it could continue from the diagonal `a3-f8` into the diagonal `g1-h2`, for example), which no other piece does.

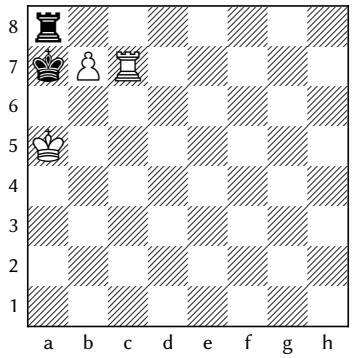
[15] Does this count against the orthodoxy? We can see the orthodox view as giving explanations for why counterfactuals are true. Does the view have to be committed to those explanations being the only possible explanations? Perhaps some true counterpossibles are overdetermined as true: vacuously and non-vacuously (this could be spelled out in terms of counterpossibles possibly having multiple truthmakers, cf. Armstrong (2004, 21)). For a view that went in this direction, it would be more important to establish the possibility of false counterpossibles, since the orthodoxy does not have resources for handling them.

3 Defending the legitimacy of game counterpossibles

There are several ways to handle counterfactuals like these. In this section, I will address several of them and argue that there are reasons to think that to handle games counterfactuals we have to be able to account for non-vacuous counterpossibles.

Perhaps the non-vacuity intuition could be explained by an ambiguity in the description of the evaluation of the counterfactual. Call the variation of chess that has the modification to the rules that we just described, chess*. Since P is a possible state of chess*, we can say that the chess-impossibilities of the antecedents of (4) and (5) are chess*-possibilities, so that the judgements about the chess-impossibilities' non-vacuity is simply a reflection of the judgements about the chess*-possibilities' non-vacuity (this sort of strategy is used often by defenders of orthodoxy). But then, there should be a worry that the reasonableness of counterpossible-talk relies on changing the subject, and thus on a form of modal illusion. The idea would be that in cases like these, our acceptance of the counterfactuals would rely on our acceptance of counterpart counterfactuals about similar things which are nonetheless strictly speaking different from the ones we are taking the counterfactuals to be about. If bishops in chess jumped over pieces of their own color once, it wouldn't be chess anymore.[16] While this might work to dismiss counterpossible-talk as misguided in a range of cases, it seems that this strategy of ambiguity elimination cannot be applied so clearly in many game cases.

Take for an example the following chess problem.[17] Suppose that the board is as follows, and it is white's turn:



Can white win in one move? According to the current rules of chess, this is impossible (white cannot capture the king or put it in check in one move). However, consider:

6) If white were to promote the pawn in b7 to a black knight in b8, white would win in one move.

This counterfactual seems true for non-vacuous reasons (if the white pawn is promoted to a black knight in b8, the black king is in check from the white rook, and can only move to a6 and b6 where it can be captured by the white king). However, according to the current rules for chess, the move described by the antecedent of (6) is disallowed because one can only promote a pawn to a piece of its same color. So the antecedent of (6) seems to be chess-impossible, and we should treat (6) as a

[16] Cf. Kripke (1981, 113–14): "[…] could *this* table have been made from a completely *different* block of wood, or even of water hardened into ice[…]? […] thought we can imagine making a table out of another block of wood or even from ice, identical in appearance to this one, and though we could have put it in this very position in the room, it seems to me that this is *not* to imagine *this* table as made of wood or ice, but rather is to imagine another table, *resembling* this one in all external details, made of another block of wood, or even of ice." Cf. Yablo (1993) for a different account of modal error, and Yablo (2000) for criticism of so-called "textbook Kripkeanism". Cf. also Byrne (2007) and Stoljar (2006) on 'proposition confusion'.

[17] I take this example from Smullyan (1979, 77). He also comments on a position where it is not obvious whether it is possible to castle, that raises similar worries.

chess-counterpossible. As we sketched above, one could say that the move is possible for the game (call it chess**) with the less restrictive rule for promotion where there is no restriction about the color of the promoted pieces, and then explain the non-vacuity intuition by reference to the intuition about the chess**-possibility. However, it is not obvious that chess** is *not* chess. It would seem odd to say that after that restriction was put in place, the original game ceased to be and it was replaced by a different game (from the perspective of people endorsing the unrestricted rule, chess would become something else, but from the perspective of people endorsing the restricted rule, something turned into chess); rather, it is more natural to say that chess *itself* changed. The problem is not only theoretical, since the historical rules of chess actually changed in order to prevent this sort of situation.[18] The difficulties here lie in the identity conditions for the referent of the term 'chess', and these difficulties ramify in various directions.

One immediate response might be to notice that the term 'chess' can be equivocal between a broad sense and a narrow sense. In the broad sense, when we talk about 'chess', we talk about a class of games that share similarities (in the structure of the board used, the type of pieces, the rules, the goals). In this case we also talk of chess *variants*.[19] In the narrow sense, when we talk about 'chess' we refer to a specific instance of chess in the broad sense. However, what precisely, is that type? The precise reference of the term 'chess' when used in the narrow sense will vary from context to context: a person talking about chess in a narrow sense now could be talking about a different thing than what a person talking about chess in a narrow sense a hundred years ago would be talking about. We can expect the issues involved in fixing the reference of the term in a given context to be similar to those that solving our main problem requires, so the distinction between broad and narrow senses of the term is not sufficient.[20]

A more promising observation is that not all rules for chess will have the same role in fixing the reference of the term. While the game supervenes on the rules, the identity of the game might not supervene on the set of all the rules that apply to it. If so, then varying certain rules will not yield chess impossible situations, and consequently counterfactuals that involve variations to those rules will not be counterpossibles. One way to implement this strategy is to deploy Searle (1969)'s distinction between *constitutive* and *regulative* rules. The former "create or define new forms of behavior", while the latter "regulate antecedently or independently existing forms of behavior" (ibid, pp. 33–34). The type of rule will determine the modal character of facts about the bindingness of the rules (that is, whether they are necessarily or contingently binding). Someone who adopted this strategy could rely on something like the following plausible sounding principles:

Constitutive Necessity For some A regulated by a set of rules R, if some r in R is a constitutive rule for A, it is necessary that instances of A must obey r (where the inner necessity, which is deontic, is different from the outer necessity).

Regulative Contingency For some A regulated by a set of rules R, if some r in R is a regulative rule for A, it is contingent that instances of A must obey r (that is, it is possible that instances of A

[18] A late 19th century rulebook, Steinitz (1889), states the following promotion rule: "A Pawn is 'queened' when it has reached the last square of a file on which it is advancing, or when it captures a hostile piece on the eight row. It may then be exchanged for a Queen or Rook or a Bishop or Knight. Thus, a player may have two or more Queens, Rooks, Bishops or Knights on the board at the same time, or he may refuse promotion to his Pawn." (p. xxiv). Note also that the pawn is not obliged to promote.

[19] Pritchard (2007) gives a compendium of chess variants, counting more than 1600 games. To those we should add variants that have been created only for the construction of problems. In his introduction, John Beasley counts as a variant 'any game [...] related to, derived from, or inspired by chess' (ibid, p. 13), which probably includes too much, but he also holds the opinion that strictly speaking 'true' chess games keep the goal of the game to capture the 'king' piece, and distinguishes these from other games that change the goal but keep the pieces, and from games that call themselves 'chess' but hold no resemblance from it whatsoever.

[20] There can be a range of senses between the broadest and narrowest. When I talk about the narrow sense, because of the contextual sensitivity I already mentioned, I mean the variably narrow sense that is sufficient to determine legality for positions.

must obey r and it is possible that instances of A are not obliged to obey r; again, the inner necessity is not the dual of the outer possibilities).

Given these, we could say that counterfactuals about the application of different constitutive rules in the context of a practice A are A-counterpossibles, while counterfactuals about the application of different regulative rules have possible antecedents. Rules about the starting positions and basic movement of chess pieces seem like good examples of constitutive rules; then, we should count (4) and (5) as chess-counterpossibles. Whether we should count (6) as chess-counterpossible depends on whether we count the promotion rule as regulative or constitutive rule. If we count the rule as regulative, we should say that (6) is an ordinary counterfactual. However, there is a problem with treating the promotion rule as merely regulative. If we didn't have the promotion rule, a whole class of possible chess games would be excluded.[21] While the rule was adopted independently of the basic rules about the movement of the pieces (so it obviously didn't contribute to the *creation* of chess playing), it nevertheless *defines* what chess games are possible, and how they will pan out. This suggests that we should treat any rule that affects the possibility-space of chess (defined in this case as the set of possible positions) as a constitutive rule.[22] But if so, we cannot rely on the distinction to dismiss the legitimacy of chess-counterpossibles; on the contrary, the problem itself might turn out to be about what are the constitutive rules of the game, or what rules can play a constitutive role for chess. I should make it clear that my point here isn't that the application of the distinction couldn't work in any case; in effect, it might be useful to handle counterfactuals about rules like those of refereeing and tournament play (if applicable), since rules like those seem to be correctly characterized as regulative. It is not correct to say that by refereeing being done in one way or another, the game that is being played is different in one case or another. It is also incorrect to say that amateurs and professional players play different games. For those cases, the distinction between regulative and constitutive rules might be useful, with the appropriate revisions.[23]

Perhaps, then, we should treat *some* game counterfactuals as genuine counterpossibles (those we cannot rule out by the simple application of the broadness and constitutivity criteria), and try to account for their non-vacuity in a less indirect way.[24] For this, I think we should consider the functions that these counterpossibles could play in their contexts of use. As we pointed out above, ordinary counterfactuals often arise in play because they are needed for planning and strategic thinking. On the contrary, we don't expect counterfactuals like those in our examples to arise during normal play (and to be taken as true or false), except in cases where our reasoning about play is faulty (for example, to evaluate (6) as true and to move according to the antecedent because one wants to arrive at the consequent position would be a misplay). Rather, we expect these counterfactuals to be evaluated in contexts where play is not the point. In the case of chess counterpossibles there seem to

[21] It might also mean that the game rules would give no direction about what to do when pawns moved to the opposite end of the board, which would make the pawns unique. However, historically the promotion rule was not universal.

[22] In turn, this seems to indicate that Constitutive Necessity is incorrect at least for games, since for at least some rules that could play a constitutive role, there is a possibility where instances of the game must obey the rule, without it being necessary that instances of the game must obey it. This is compatible with it being necessary that the game obeys some of the rules that can constitute it. Alternatively, but at a greater theoretical cost, we could keep Constitutive Necessity but allow for incompatible constitutive rules to apply to instances of games, in which case we would also need to add a pragmatic story about why contextually certain rules are salient instead of others (cf. footnote 6 above). To make the incompatibility more palatable, we could adopt a logic where it can happen that 'A is true', 'B is true', but 'A and B is not true' (cf. Lewis (1983, 277) observations on the 'method of union' for truth in fiction).

[23] How much weight should we give to these intuitions about what counts or not as the same game? Couldn't it be that the ordinary conception of games is incoherent, or that alternative conceptions are at least equally good? While these are definite possibilities, in the case of games any potential mismatch between their nature and ordinary talk about them must be treated with care, because the constitution of games is given by the practices of people who engage in them, including our talk about them. So while these intuitions are not infallible, the objection has less bite than usual. However, this line of defense of intuitions doesn't necessarily generalize to cases unlike games.

[24] There might be other ways to dismiss game counterpossibles as non-genuine that I haven't considered. Here I am only claiming that the lines of attack above are not sufficient to dismiss them.

be two contexts where counterpossibles might arise.

The first case is that of retro problems. While they can be interesting from the perspective of endgame analysis (and thus implicitly from the perspective of chess possibility), they exist independently of play.[25] Chess problem solving exists outside the institution of play that extends to tournament play, and consequently has entirely different criteria of fairness, and depending on the setup, of what counts as an admissible solution.[26] This might suggest that the notion of impossibility in use here differs systematically from the notion of impossibility in use during play. However, in the case of retro problems with impossible setups, the relevant notion of impossibility is often the regular one: the point of the problems is to explain the illegality of the positions, which is not always obvious. Backtracking to a move and position that couldn't have happened, we reason about intermediate steps that also couldn't have happened. It seems to me that the more flexible way to do this is by allowing counterpossible reasoning.

The second case is the evaluation of rules; for example, when faced with issues that require a decision on how to implement a rule (due to ambiguity in the rule, or because the rule doesn't handle corner cases). This could be observed above in the case of the restricted and unrestricted promotion rules. In regular play, finding oneself in an illegal position indicates that someone made a mistake or cheated; finding oneself in an ambiguous situation, on the other hand, forces an examination of the rules, and of the consequences of possible changes to the rules. In those cases we want to distinguish between game-impossible scenarios, so we need a way to hold the relevant counterfactuals as true or false. Counterpossible reasoning could be used here.

In games like chess the practices that can allow for counterpossible reasoning and playing are relatively independent. However, this is a contingent feature of these practices. Peter Suber's 'nomic' game illustrates how both practices can be fully integrated.[27] In nomic, each 'move' can consist in the modification of the game's rules. A nomic game starts with a minimal set of rules about how the players should proceed, and specifies how rule changes can be incorporated (by default there is a 'democratic' mechanism where a player proposes rules and the other players either accept or reject the proposal). Given these facts about the game, what can be nomic-possible and nomic-impossible is much less clear than in the case of chess possibility and impossibility (with suitable changes, the sphere of possibilities at any stage can grow and shrink widely). While one could say that everything is nomic-possible and nothing is nomic-impossible, these are *not* the notions of possibility and impossibility that would be used in counterfactual strategic reasoning during actual nomic play, which would be the proper counterparts of the notion of chess-possibility and chess-impossibility that we examined earlier. Thus, there might be a need for the evaluation of counterfactuals about genuine nomic-impossibilities. Admittedly, one could adopt the possibilist view according to which everything whatsoever is nomic possible, and supplement it with a pragmatic account that filters out irrelevancies. However, given the context sensitivity of counterfactuals, this might underutilize the resources that the context provides to determine their semantic content.[28] While we still get a liberal account of entertainable 'situations' or 'worlds' (that includes impossibilities *stricto sensu*), we have an 'inner' notion of possibility that we can then use to pragmatically rule out irrelevancies in context.

My proposal to understand counterpossible talk in the context of games (and perhaps more generally) can be sketched as follows. Games of the type we have discussed here supervene on

[25]The problem literature precedes the existence of modern chess, with many medieval examples. It is worth mentioning that in some cases problems were embedded in games of gambling (cf. Murray (1913, II, ch. VII)).

[26]Cf. White (n.d., 449) on the construction of problems with illegal positions: "If you want to use an extra officer or two, why not do so? There is nothing morally wrong about it. Your result will be distasteful to many solvers; but it will do them no harm." White, of course, assumes the modern practice of treating problems as intellectual exercises, while historically this was not always the case (see footnote 25).

[27]Suber (1990). For a multi-player chess variant of nomic, see Howe (2000).

[28]The case of nomic is important because it puts pressure on the idea that we could understand the possibility of non-vacuous counterpossible-talk in terms of a sharp distinction between object languages and meta-languages for games (where counterpossibles are vacuous at the object level and possibly non-vacuous at the meta-level).

rules.[29] If you change the rules too much, you start playing a different or divergent game. But before that happens, you will have potential variations that still count as the same game as we have been playing all along. What counts as merely a variation and what counts as a divergent game depends on criteria which are given in the context, and which are themselves subject to revision. In practice, surrounding or embedded in the game proper there is always a meta-game (or a collection of meta-games) that deals with managing revisions of this sort. It seems like counterpossible-talk can play a crucial role here, because it offers a way to express and discuss the consequences of adopting variant rules while keeping the distinction between variants and divergencies using a constant modal conceptual framework. Chess-impossibility stands in a relation to chess-possibility that chess*-possibility does not stand in relation to chess-possibility. While counterpossibles are context sensitive, they don't shift the modal framework in use implicitly.[30] If they did, they would be pointless in many cases, since they would change the subject too radically. Even when they don't change the subject, they can still be pointless in cases where the task at hand is to evaluate courses of actions, since it is doubtful that they could be of direct use for guidance.[31] This is why we don't find them in play. Instead of giving a pragmatic account of the acceptability of counterpossibles, we should also be able to give a pragmatic account of the restrictions that we make in ordinary contexts to counterfactuals with possible antecedents (in which case instead of having a restricted default semantics which is pragmatically extended, we have a liberal default semantics which is pragmatically restricted).[32]

4 Divide and conquer, or normativist subsumption?

It can be useful to contrast the current proposal in its general form with two recent views: Vetter's (2016) 'divide and conquer' strategy, and Locke's (2019) normativist account.

Vetter's (2016) aim is to defend the orthodoxy about counterpossibles using what she calls a 'divide and conquer' strategy, by distinguishing between cases where counterpossibles should be vacuous, and cases where they might not be. Arguably, the current proposal shares this 'divide and conquer' structure, although it draws the division between admissible and inadmissible cases differently.

The crux of Vetter's argument lies on the distinction she makes between epistemic and circumstantial modality, which she contrasts as follows:

> [...] circumstantial modality concerns the objects, properties, and relations that a given claim is *about*, not [like in the epistemic case] any representational or cognitive features of the terms we use to refer to them. (p. 2698)

With this distinction in hand, she proceeds to argue that non-vacuous seeming counterfactuals are always epistemic. The reason for this is that they would give rise to referential opacity, which gives evidence for an epistemic reading. This suggests that in the cases of seemingly non-vacuous game counterpossibles we have considered, the non-vacuity intuition can be explained away by proposing

[29] Cf. Kreider (2011) for discussion of the relation between rules and games, and Ridge (2019) for an overview of the philosophical literature on the nature of games.

[30] This assumes a more or less traditional contextualist view. Ludlow (2014) offers a more dynamic view along similar lines, where the meaning of terms can change between and within conversations (cf. chapter 5 specially on how he addresses troubles for his account). Like in the current proposal, Ludlow emphasizes the practices of negotiation of meaning and concepts. Unlike in the current proposal, in Ludlow's proposal the negotiation is purely metalinguistic, while I think it might have to do with the open-endedness of the referents of terms as well (in the case of games, the open-endedness of our conceptualizations is grounded on the open-endedness nature of games).

[31] However, see Heuer (2010). In any case, counterpossible talk in the uses I describe here could be indirectly of use because in some cases guidance requires changes to the operative modal framework.

[32] We don't need to choose: my point is that we have both strategies available instead of just the first.

epistemic readings for the counterfactuals. Note that Vetter's view is not that counterpossibles are always vacuous, but that circumstantial counterpossibles always are.

However, this does not seem plausible in the case of the counterfactuals that we have considered. They are explicitly not about the representational features of games, or of our epistemic situation relative to them. They are about what would happen or would have happened in the context of games. This is a circumstantial subject matter, and the corresponding modalities should be correspondingly circumstantial.[33]

Locke (2019) offers a theory of counterpossibles that applies a more general modal normativist framework to the case of counterpossibles. Modal normativism is the view that the primary function of modal claims is expressive or normative.[34] The basic idea is that modal claims, in Brandom's turn of phrase, 'make explicit' the rules of use of our terms. Thomasson (2007) describes modal normativism about metaphysical necessity as the view that modal claims about necessity

> serve the *prescriptive* function of expressing semantic rules for the terms used in them, or their consequences, while remaining in the object language. (p. 136)

The last point is an important similarity between the modal normativist view and the current proposal. As I said before, if we are to accept seemingly non-vacuous counterpossibles, we should be careful not to change the subject. The goal of having modal language belong to the object language is precisely to avoid this issue. Consequently, modal normativist views do not have the problem that I raised for Vetter's account concerning the subject matter of counterpossibles.

Locke states normativism about counterpossibles as follows:

> [...] metaphysical counterpossibles function to illustrate or express changes, or consequences of changes, to the actual constitutive rules that govern language use while remaining in the object language where terms are used rather than mentioned. (p. 8)

This follows the constraint we raised before that if there are genuine game counterpossibles, at least some (if not all) of those should relate to constitutive rules. A further similarity between Locke's view and the current proposal is the way Locke deals with the problem of changing the subject:

> I claim that, since object language claims about metaphysical necessities and possibilities illustrate the actual rules or permissions that govern the use of modal vocabulary, object language claims about non-trivial metaphysical impossibilities illustrate non-trivial changes in those rules and permissions. In the right context, claims about non-trivial metaphysical impossibilities are an important object language resource for "mis-using" language without being subject to rebuke or interpreted as incompetent, e.g. in the case of a charitable philosophical dispute. This is because small, relevant changes in the actual rules that govern the use of some expression neither result in a radically different expression nor do they result in a complete change of subject. (p. 11)

The current proposal manages to tell roughly the same story without having modal language as a whole play a normative or expressive function. Perhaps representational language is normative or expressive, but that is an even greater departure from orthodoxy that we are not forced to make just for the sake of being able to handle counterpossibles. This aspect of the normativist proposal is underplayed by Locke because of his underlying commitment to normativism about modality in general, but in the present context the issue is more pronounced. Furthermore, modal normativism depends on having an account of the adequacy of the constitutive rules of language use (thus,

[33] Locke (2019) raises the same criticism about the scope of Vetter's strategy, giving as a counterexample the counterfactual 'if a steel Penrose triangle were placed in a 4000 deg. F oven, it would melt.'

[34] Cf. Brandom (2008) and Thomasson (2007, 2013)

Thomasson (2007, 138) says that normativism 'requires that we first accept that our terms *have* rules of use'). That makes the possibility of contexts where counterpossibles are used to discuss potential revisions to those very same rules a bit awkward; this seems to be the reason why, in the (2007) paper, she claims that under normativism there are substantive limitations about what kind of revisionary projects can be undertaken. In recent work (2017), Thomasson introduces the idea that metalinguistic negotiations might have non-semantic consequences, which allows for more revisionary projects; Locke (2019) adopts this solution. The solution in the current proposal is that the appropriateness of counterpossibles depends on the features of the local context, not of global standards of use (of course, the local context might in turn refer back to broader standards). This means that disputes about counterpossibles might not necessarily be resolved definitely through conceptual analysis, like Locke suggests (p. 20); indeed, they might only be resolved temporarily or not at all.

Appendix: chess-possibility

Semi-formally, a board b is chess-possible iff it can be reached in any number of steps by the application of chess-rules R, from a starting board s.

A *diagram* is a sentence describing the complete state of a board (essentially the information encoded in a FEN string). We will work in a language with variables for diagrams ($p_1 \dots p_n$), two constants: i for the current diagram and s for the starting diagram, and three modal operators: $\Diamond_{\rightarrow}$, $\Diamond_{\leftarrow}$ and $\Diamond_s$ that build sentences out of sentences. The informal interpretation of these operators is 'it is possible to advance to position…', 'it is possible to have come from position…' and 'it is chess possible that…', respectively. We also have the usual negation and the connectives $\wedge$, $\vee$, and $\supset$.

A *chess-frame* is a quintuple $\langle W, s, R_{\leftarrow}, R_{\rightarrow} \rangle$, where W is a set of possible (constructible) boards, s is a selected member of W that represents the starting position, $R_{\leftarrow}$ is a binary relation over W, and $R_{\rightarrow}$ is another binary relation over W. We use two binary relations instead of one because we want to track more perspicuously (1) what moves can be made legally from a position (this is what $R_{\rightarrow}$ tracks) and (2) what moves could have been made legally to arrive at a position (this is what $R_{\leftarrow}$ tracks), and some moves in chess are not reversible (the pawns can only move forward). $R_{\rightarrow}$ and $R_{\leftarrow}$ can be understood as the converse of each other, so that $R_{\rightarrow}ab \leftrightarrow R_{\leftarrow}ba$.[35] It is worth noting that neither relation is reflexive (it is not possible to make a move that doesn't change the state of the board), but both relations are transitive (if it is possible to arrive from one direction at a position A from a position B, and it is possible to arrive from the same direction at a position B from a position C, it is possible to arrive from the same direction to A from C). We extend frames with a function I that assigns a unique diagram to every board in W to obtain a *chess-model* (a different way to present this would be to make boards themselves diagrams, and to let diagrams represent themselves).

We define a valuation V_M for a model M as a function that assigns truth values (0 or 1) to each well-formed formula (wff) to each member of W as follows, where δ is any diagram, ϕ and ψ are any wffs, and w is any board:

$$
\begin{aligned}
V_M(\delta, w) &= 1 \text{ iff } \delta = I(w) \\
V_M(\neg\phi, w) &= 1 \text{ iff } V_M(\phi, w) = 0 \\
V_M(\phi \supset \psi, w) &= 1 \text{ iff } V_M(\phi, w) = 0 \text{ or } V_M(\psi, w) = 1 \\
V_M(\Diamond_{\rightarrow}\phi, w) &= 1 \text{ iff for some } w' \in W \text{ with } R_{\rightarrow}ww', V(\phi, w') = 1 \\
V_M(\Diamond_{\leftarrow}\phi, w) &= 1 \text{ iff for some } w' \in W \text{ with } R_{\leftarrow}ww', V(\Diamond_{\rightarrow}\phi, w') = 1
\end{aligned}
$$

[35] We implicitly assume that we track information about the players and the turns (for example, to prevent white to move twice in a row, or–in some variants–to allow for such things). A different approach would be to have one the frames be a triple $\langle W, s, R \rangle$ where R is a set of binary relations over W where each represents a possible move according to a rule.

$V_M(\Diamond_s\phi, w) = 1$ iff $\phi = I(s)$ or $V(\Diamond_\rightarrow\phi, s) = 1$
or if there is a world w' where $V(\phi, w') = 1$ and $V(\Diamond_\leftarrow I(s), w') = 1$

For $\wedge$ and $\vee$, the truth conditions can be derived as usual from the conditions for $\neg$ and $\supset$. For the three modal operators, there is a derived notion of necessity that is their dual.[36]

Given the assumption that $R_\leftarrow$ and $R_\rightarrow$ are transitive, we have as axioms schemes:

- $\Diamond_\leftarrow\Diamond_\leftarrow\phi \supset \Diamond_\leftarrow\phi$
- $\Diamond_\rightarrow\Diamond_\rightarrow\phi \supset \Diamond_\rightarrow\phi$.

Some theorems:

- $(\Diamond_s p_1 \wedge \Diamond_\leftarrow p_1) \supset \Diamond_s i$. If we could have arrived at the current state from a chess possible position, the current position is chess possible.
- $\Diamond_\leftarrow s \supset \Diamond_s i$. A special case of the previous. If we could have arrived at the current position from the starting position, the current position is chess possible.
- $(\Diamond_\leftarrow s \wedge \Diamond_\rightarrow p2) \supset \Diamond_s p_2$. If we could have arrived at the current position from the starting position, and we can arrive legally at a position p_2, p_2 is chess possible.

$(\Diamond_s\phi \wedge \Diamond_\rightarrow\phi) \supset \Diamond_s i$ is a non-theorem: there can be positions that can move towards chess-possible positions that couldn't have come from the standard position.

On the other hand, if we can advance to an impossible position, the current position is impossible: $(\neg\Diamond_s\phi \wedge \Diamond_\rightarrow\phi) \supset \neg\Diamond_s i$. In this model, some impossible positions share with the starting position the property of being terminal nodes: there is no position that they could have come from. But it is clear that in many cases we want to reason about illegal positions that derive from legal positions through misplay. To model this, we should introduce additional accessibility relations that models transitions from positions through mistakes (forwards and backwards, like above). In the system extended in this way we can reason backwards from impossible positions to positions that caused the illegality.

References

Armstrong, David M. 2004. *Truth and Truthmakers*. Cambridge: Cambridge University PRess.

Baron, Sam, Mark Colyvan, and David Ripley. 2017. "How Mathematics Can Make a Difference." *Philosopher's Imprint* 17 (3).

Bennett, Jonathan. 2003. *A Philosophical Guide to Conditionals*. Oxford: Oxford University Press.

Berto, Francesco, Rohan French, Graham Priest, and David Ripley. 2018. "Williamson on Counterpossibles." *Journal of Philosophical Logic* 47 (4): 693–713.

Berto, Francesco, and Mark Jago. 2019. *Impossible Worlds*. Oxford University Press.

Bicchieri, Cristina. 1988. "Strategic Behavior and Counterfactuals." *Synthese* 76: 135–69.

Binmore, Ken. 1987. "Modeling Rational Players I." *Economics & Philosophy* 3 (2): 179–214.

Brandom, Robert. 2008. *Between Saying & Doing: Towards an Analytic Pragmatism*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199542871.001.0001.

Brogaard, Berit, and Joe Salerno. 2013. "Remarks on Counterpossibles." *Synthese* 190: 639–60.

Byrne, Alex. 2007. "Possibility and Imagination." *Philosophical Perspectives* 21 (1): 125–44.

[36] There are four types of possibility in the model: a) the combinatorial possibility of diagrams, which is assumed for W in the frames, b) the forward looking possibility $\Diamond_\rightarrow$, c) the backwards looking possibility $\Diamond_\leftarrow$, and c) the composite $\Diamond_s$, which is what we call chess-possibility properly speaking. Because of this, the model includes worlds which are constructible and sharply distinguishable, but impossible in a definite sense, without a need to mark those explicitly.

Chase, William G., and Herbert A. Simon. 1973. "Perception in Chess." *Cognitive Psychology* 4: 55–81.
Dawson, T. R., and W. Hundsdorfer. 1915. *Retrograde Analysis: A Study*. Edited by Alain C. White. Leeds: Whitehead; Miller Printers.
De Groot, Adriaan D. n.d. *Thought and Choice in Chess*. Amsterdam University Press.
Emery, Nina, and Cristopher Hill. 2016. "Impossible Worlds and Metaphysical Explanation: Comments on Kment's 'Modality and Explanatory Reasoning'." *Analysis* 77 (1): 134–48.
FIDE. 2018. "FIDE Handbook." In. https://handbook.fide.com/chapter/E012018.
Heuer, Ulrike. 2010. "Reasons and Impossibility." *Philosophical Studies* 147 (2): 235–46.
Howe, David. 2000. "Nomic Chess." https://www.chessvariants.com/multiplayer.dir/nomicchess.html.
Kim, Seahwa, and Cei Maslen. 2006. "Counterfactuals as Short Stories." *Philosophical Studies* 129: 81–117.
Kment, Boris. 2006. "Counterfactuals and the Analysis of Necessity." *Philosophical Perspectives* 20: 237–302.
Kreider, A. J. 2011. "Game-Playing Without Rule-Following." *Journal of the Philosophy of Sport* 38 (1): 55–73.
Kripke, Saul. 1981. *Naming and Necessity*. Blackwell.
Lewis, David. 1983. "Postcripts to 'Truth in Fiction'." In *Philosophical Papers*, I:276–80. Oxford University Press.
Lewis, David K. 1973. *Counterfactuals*. Oxford: Blackwell.
Locke, Theodore D. 2019. "Counterpossibles for Modal Normativists." *Synthese*. https://doi.org/https://doi.org/10.1007/s11229-019-02103-1.
Ludlow, Peter. 2014. *Living Words: Meaning Underdetermination and the Dynamic Lexicon*. Oxford: Oxford University Press.
Murray, H. J. R. 1913. *A History of Chess*. London: Oxford University Press.
Nolan, Daniel. 1997. "Impossible Worlds: A Modest Approach." *Notre Dame Journal of Formal Logic* 38: 535–72.
Nolan, Daniel. 2014. "Hyperintensional Metaphysics." *Philosophical Studies* 171 (1): 149–60. https://doi.org/10.1007/s11098-013-0251-2.
Priest, Graham. 2016. "Thinking the Impossible." *Philosophical Studies* 173 (10): 2649–62.
Pritchard, David Brine. 2007. *The Classified Encyclopedia of Chess Variants*. Edited by John D. Beasley. 2nd ed. Harpenden: John Beasley.
Ridge, Michael. 2019. "Play and Games: An Opinionated Introduction." *Philosophy Compass*. https://doi.org/10.1111/phc3.12573.
Searle, John. 1969. *Speech Acts*. Cambridge: Cambridge University Press.
Sinnott-Armstrong, Walter. 1984. " 'Ought' Conversationally Implies 'Can'." *The Philosophical Review* XCIII (2): 249–61.
Skyrms, Brian. 1998. "Subjunctive Conditionals and Revealed Preference." *Philsophy of Science* 65 (4): 545–74.
Smullyan, Raymond. 1979. *The Chess Mysteries of Sherlock Holmes*. New York: Alfread A. Knopf.
Stalnaker, Robert. 1996. "Knowledge, Belief and Counterfactual Reasoning in Games." *Economics and Philosophy* 12: 133–63.
Steinerberger, Stefan. 2015. "On the Number of Positions in Chess Without Promotion." *International Journal of Game Theory* 44: 761–67. https://doi.org/https://doi.org/10.1007/s00182-014-0453-7.
Steinitz, William. 1889. *The Modern Chess Instructor*. New York: G.P. Putnam & sons.
Stoljar, Daniel. 2006. *Ignorance and Imagination: The Epistemic Origin of the Problem of Consciousness*. Oxford University Press.
Streumer, Bart. 2007. "Reasons and Impossibility." *Philosophical Studies* 136 (3): 351–84.
Suber, Peter. 1990. *The Paradox of Self-Amendment: A Study of Law, Logic, Omnipotence, and Change*. Berne: Peter Lang.

Tan, Peter. 2019. "Counterpossible Non-Vacuity in Scientific Practice." *Journal of Philosophy* 116 (1): 32–60. https://doi.org/10.5840/jphil201911612.
Thomasson, Amie. 2007. "Modal Normativism and the Methods of Metaphysics." *Philosophical Topics* 35 (1/2): 135–60.
Thomasson, Amie. 2013. "Norms and Necessity." *The Southern Journal of Philosophy* 51 (2): 143–60.
Thomasson, Amie. 2017. "Metaphysical Disputes and Metalinguistic Negotiation." *Analytic Philosophy* 58 (1): 1–28.
Vetter, Barbara. 2016. "Counterpossibles (Not Only) for Dispositionalists." *Philosophical Studies* 173 (10): 2681–2700.
White, Alain C. n.d. *Sam Loyd and His Chess Problems*. New York: Dover Publications.
Williamson, Timothy. 2007. *The Philosophy of Philosophy*. Oxford: Blackwell Publishing. https://doi.org/10.1002/9780470696675.
Yablo, Stephen. 1993. "Is Conceivability a Guide to Possibility?" *Philosophy and Phenomenological Research* 53: 1–42.
Yablo, Stephen. 2000. "Texbook Kripkeanism and the Open Texture of Concepts." *Pacific Philosophical Quarterly* 81 (1): 98–122.