# **Deciding Where to Meet for Dinner:**

# **Simple Problems and Joint Intentionality**

Robert Bass

Abstract: Certain apparently simple problems of coming to an agreement are surprisingly difficult to analyze in terms of individually rational behavior with a given set of preferences and beliefs. Though initially the solution appears obvious, the reasoning that would be needed to reach the solution on the part of a pair of rational individuals seems baroque and doubtful. This is used to suggest that a more fruitful tack is to analyze the situation in terms of a kind of joint or shared intentionality. If that can be invoked, what appears simple in practice will turn out to be simple in theory as well.

Keywords: Individual rationality, joint intentionality, agreement, cooperation

## Deciding Where to Meet for Dinner

Jack and Jill are deliberating on where to meet for dinner. Consider the following dialogue snippet:

Jack: How about dinner at eight? At Alice’s?  
Jill: Sure, but not Alice’s. What about Bennie’s?  
Jack: The service was really bad the last time I was there. What about Corrinne’s?  
Jill: That will do.  
Jack: Fine. See you then!

Such conversations occur every day, and the participants do, often enough, actually succeed in meeting for dinner. What is going on? The obvious – and correct – answer is that Jack and Jill are coming to an agreement. But just putting the familiar label, ‘agreement,’ on it is apt to conceal ways in which it is formally puzzling. One might think that something so prosaic and so much a matter of common experience could be readily analyzed in terms of what the participants believe, want and have reason to do. In fact, however, such an account is not readily forthcoming.

I shall go through attempts to provide that analysis at some length, and no doubt, some will wonder why Jack and Jill can’t just agree. Once each knows what the other wants, why is it a problem for them just to agree on Corrinne’s? In a sense, I think that is exactly right. What they *should* do is agree on Corrinne’s. The problem comes in understanding what is meant by saying they should agree on Corrinne’s or, equivalently for my purposes, what they *are* doing in agreeing on Corrinne’s.

The way I will be telling the story and developing the analysis (for some time to come) involves a certain picture of what is involved in Jack and Jill coming to an agreement. I shall call this the individualist picture of agreement.[[1]](#footnote-1) This picture, which is culturally reinforced in various ways, comes very naturally to us – so naturally that we may not see that there is any possible or imaginable alternative. Roughly, Jack and Jill are depicted as each having reasons for action which are some function of their respective preferences and beliefs. Each can engage in various kinds of thinking or deliberation about preferences and beliefs in a way that may result in the recognition of additional reasons for action. There is a sense in which their reasons and deliberations are private. Jack’s reasons for action are entirely a function of his preferences and beliefs; Jill’s, entirely a function of hers. Jack has his reasons and deliberates on their basis; Jill has hers and does likewise. Each engages in this deliberative process and, from time to time, publicly announces provisional conclusions to the other. When these provisional conclusions (a) appear stable, and so not apt to be upset by a further round of deliberation and (b) coincide, we say they have reached an agreement. According to this picture, the upshot of their conversation is that Jill ends up with a reason for going to Corrinne’s and Jack does as well; therefore, they both have a reason to go there. Since their conclusions agree, we say that *they* have agreed.

The point of the extended discussion is to show that the individualist picture is highly problematic. It is not at all clear that, if *this* is what must be involved, Jack and Jill *can* agree to meet at all. More precisely, it is not clear that the two of them can rationally, and without deception or manipulation of the other, agree to meet. Moreover, even if they can somehow come to a real agreement, this is hardly the most difficult case to analyze in terms of the individualist picture. How, for example, is a group of a half-dozen to select a place to have lunch when it will not even be true that everyone in the group is familiar with all the restaurants suggested? Or consider how a committee, an organization or a club comes to a decision on a long-term policy. Larger numbers must be consulted, and among them, there may be differing assessments even of what the likely outcomes are, not to mention differential rankings of the outcomes. If analyzing a two-person case with only a few options reveals unexpected complexity, adding additional parties and additional options may make the problems intractable.[[2]](#footnote-2)

This begins to indicate the shape of a larger argument. I shall not be claiming that it is impossible for some analysis in terms of the individualist picture to be given, either of the case of Jack and Jill or in various more complicated agreements. However, in delineating the various complexities involved in understanding what would have to be true for such an analysis to work, I shall be building a case that, even if such an account is possible in principle, it is not an adequate account of what *we* do in reaching agreement. *We* find it easy to see that Jack and Jill should agree on meeting at Corrinne’s Restaurant. Jack and Jill themselves find it easy to agree on meeting there. If we find the problem easy, while our *theory* of what is going on makes it look difficult, that is reason to question the theory.

# Deciding Where to Go – Without Communication

To begin to make it clear that it is not easy to understand what Jack and Jill are doing in agreeing where to meet for dinner, suppress for the time being the knowledge that they have talked and agreed. Suppose just that they want to meet for dinner, want to have dinner (at one of three restaurants) whether or not they meet, and that both know how both rank the various possible outcomes. (Perhaps they agree to meet for dinner but neglect to decide where.) Suppose also that each is rational in the sense that neither will select a lower-ranked over a higher-ranked option nor will either strictly prefer[[3]](#footnote-3) one of two equally ranked options to the other. And suppose that both know that both are rational in this sense. It is worth attending explicitly to several features that either are or may be present, but are not apparent on the surface, in order to bring out the difficulties involved.

First, both Jack and Jill have a definite (and strict) preference ordering over the three restaurants mentioned. Jack prefers Alice’s to Corrinne’s to Bennie’s, while Jill prefers Bennie’s to Corrinne’s to Alice’s. Second, each would prefer meeting in a given restaurant over going to that restaurant but not meeting (because the other is elsewhere). Third, because the restaurant at which they agree to meet (Corrinne’s) is not the first choice of either, neither has available (in the absence of agreement – whether agreement makes a difference will be considered later) a dominance argument for being there.[[4]](#footnote-4)

We can represent the options that face them in this way.[[5]](#footnote-5)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Jill | | | | |
| Jack |  | Alice’s | Bennie’s | Corrinne’s |
| Alice’s | 6, 2 | 4, 4 | 4, 3 |
| Bennie’s | 1, 1 | 2, 6 | 1, 3 |
| Corrinne’s | 3, 1 | 3, 4 | 5, 5 |

Table I: Meeting for Dinner

If we assume in addition that the restaurants are close together, so that moving from one to another is a negligible inconvenience,[[6]](#footnote-6) then only two of the possible outcomes of their joint actions, represented by the two shaded boxes, are stable. If Jack is at Alice’s while Jill is at Bennie’s, Jack cannot do better by moving to a different restaurant. He would rather be at Alice’s without Jill than at Bennie’s with her or at Corrinne’s without her. Similarly, Jill cannot do better than to remain at Bennie’s. She would rather be at Bennie’s without Jack than at Alice’s with him or at Corrinne’s without him. For like reasons, if they were both at Corrinne’s, neither could do better, acting alone. This is to say, in game-theoretic parlance, that these two outcomes are both equilibria. Neither can be unilaterally improved upon by either Jack or Jill. Plainly, though both are equilibria, one is superior to the other. If Jack and Jill both end up at Corrinne’s, each will be better off in terms of his or her own preference ranking than if Jack ends up at Alice’s and Jill ends up at Bennie’s.

None of the other outcomes is an equilibrium. If Jack somehow finds himself at Bennie’s (with or without Jill) or at Corrinne’s (without Jill), he can improve by going to Alice’s. And if Jill somehow finds herself at Alice’s (with or without Jack) or at Corrinne’s (without Jack), she can improve by moving to Bennie’s. Note that this means that if they do not happen both to go to Corrinne’s first, they will, either immediately or by moving, end up with Jack at Alice’s and Jill at Bennie’s. That is, they will end up in the inferior equilibrium position.

The question to ask here is whether there is any reason for each, and therefore for both, to go to Corrinne’s first. In other words, is it just a matter of what they *happen* to do if they both go first to Corrinne’s? The first thing to say is that there doesn’t appear to be any reason for either of them *not* to go to Corrinne’s first. As was pointed out earlier, neither has a dominance argument in favor of some particular restaurant. But there is also no dominance argument, for either of them *against* going to Corrinne’s first. Jack has a dominance argument against going to Bennie’s and Jill has one against going to Alice’s, but that leaves either Alice’s or Corrinne’s as live options for Jack and either Bennie’s or Corrinne’s as live options for Jill.

We can thus simplify the representation of their choices about where first to go and the associated outcomes in this way[[7]](#footnote-7):

|  |  |  |  |
| --- | --- | --- | --- |
| Jill | | | |
| Jack |  | Bennie’s | Corrinne’s |
| Alice’s | 4, 4 | 4, 3 |
| Corrinne’s | 3, 4 | 5, 5 |

Table II: Meeting for Dinner, Simplified

Again, there are two equilibria, one superior to the other, for both Jack and Jill. If they both go first to Corrinne’s, then they will remain there, since neither can do better. Suppose, however, that Jack goes first to Alice’s and reasons in the following way:

Jill must be either at Corrinne’s or Bennie’s. If she’s at Corrinne’s, then I can do better, than by staying here, if I go to Corrinne’s, too. Of course, she might be at Bennie’s. Then, I’d do better to stay put, since neither Corrinne’s without her nor Bennie’s with her is as good as being at Alice’s without her. But since I don’t know, why not take a chance? I’ll check out Corrinne’s.

Meanwhile, Jill did indeed decide to go to Corrinne’s first. But when she arrived there and found Jack absent, she moved to Bennie’s. Had Jack foreseen this, he wouldn’t have bothered to check whether Jill was at Corrinne’s. He would have realized that, even if she did go there first, she wouldn’t stay unless he was there, too. Of course, a similar argument works for Jill. If she doesn’t go to Corrinne’s first, she will have no reason to check to see if Jack did. Once again, unless they both go to Corrinne’s first, they will end up in the inferior equilibrium position.

Now, one might think – indeed, I am inclined to think – that it is just obvious that they should both go to Corrinne’s first. However obvious it appears, it is not clear from the situation as described.

Fairly abstractly, the point can be put this way. There are two principal solution concepts,[[8]](#footnote-8) both mentioned earlier, that are recognized within game theory, dominance and equilibrium. There is a dominance argument in favor of an option when, among the possible outcomes of an action, at least one of the outcomes is ranked higher than the outcome or outcomes that would occur if some different option were selected, and no outcome is ranked lower than any corresponding alternative outcome or outcomes.[[9]](#footnote-9) For equilibrium, a set of options selected by the different players is in equilibrium when it is not possible for any player to unilaterally improve by selecting a different option, given that other players do not alter their play.[[10]](#footnote-10) Improvements, if any are possible, require the coordinated action of two or more players. Neither of these solution concepts, however, applies to the case of Jack and Jill. Neither has a dominant option, and though it is true that there is an equilibrium,[[11]](#footnote-11) there is more than one, so recommending equilibrium play does not determinately pick out one option as superior for either of them.

So much is orthodox game theory. McClennen’s heterodoxy also deserves brief mention. In his view, dominance and equilibrium reasoning need to be supplemented, and for certain cases replaced, by considerations of Pareto-superiority. One outcome is Pareto-superior to another when it is better for at least one of the parties and not worse for the other. I shall briefly indicate why I do not think the addition of Pareto-superiority as a solution concept helps in the present case. The key point is that we are dealing with a single-shot case, one in which Jack and Jill are meeting (or not) for dinner on a single occasion. What is wanted is an explanation of why it is rational for them to meet at Corrinne’s *this time*, without appeal to the effects of this case on others. Consider the following simple game:

|  |  |  |  |
| --- | --- | --- | --- |
| Jill | | | |
| Jack |  | A | B |
| A | 1, 1 | 0, 1 |
| B | 1, 0 | 0, 0 |

Table III: Pareto-Superiority

Here, neither Jack nor Jill has a dominant choice and all four of the possible outcomes are the results of equilibrium sets of choices. Nonetheless, one of the outcomes, represented by the north-west box, is Pareto-superior to all the others, so this seems an ideal testing ground for the plausibility of Pareto-superiority as a solution concept. The problem derives from the fact that whether Jack gets the outcome he prefers depends entirely on what Jill does, and *vice versa*. If Jill chooses A, it makes no difference to Jack whether he chooses A as well. If it is given that Jill chooses A, then Jack gets an outcome that he ranks identically, whether he chooses A or B. Similarly, if Jill chooses B, Jack gets an outcome ranked the same (but worse than if Jill had chosen A), whether he selects A or B. Of course, Jill is in the symmetrical position. The outcome she gets depends entirely upon Jack’s choice.

There is a dilemma here for McClennen’s theory, at least as it applies to the single-shot case. The way the outcomes are ranked is supposed to encode everything that matters to the agents, so Pareto-superiority should make no difference to the outcome rankings. But then, Jack and Jill should each be indifferent between the choice of A and B. Considerations of Pareto-superiority will not affect their choices. To illustrate, suppose that, if Jill selcts A, then Jack gets a monetary prize of ten dollars, whether he selects A or B, and if Jill selects B, he receives no prize – again, whichever of his options he selects. If all that Jack cares about is the monetary prize, then he should be indifferent as to whether to choose A or B.

It appears that the only way that it can be rational for Jack (or for Jill) to strictly prefer A over B would be if the Pareto-superiority of the outcome in which they both select A affected the way that each ranked A *vis-à-vis* B. If it does, we get a game-representation that looks like this (where *p* represents some positive value assigned to the outcome in virtue of its being Pareto-superior):

|  |  |  |  |
| --- | --- | --- | --- |
| Jill | | | |
| Jack |  | A | B |
| A | 1 + *p*, 1 + *p* | 0, 1 |
| B | 1, 0 | 0, 0 |

Table IV: Pareto-Superiority Affecting Outcome-Rankings

Here, choosing A is dominant for both, so it may appear that we have what we want. What is not clear, though, is why rational players would alter the way that they rank outcomes on the basis of Pareto-superiority. McClennen has an answer to this that runs approximately as follows: *Agents* who differentially rank outcomes (in this kind of case) on the basis of Pareto-superiority will do better than agents who don’t. They will always coordinate on the Pareto-superior outcome when they encounter one another and will be more likely to coordinate on the Pareto-superior outcome when they encounter agents who attach no special weight to Pareto-superiority. This is to say that what is being compared is the expected returns to different *strategies* for such interactions, strategies that bear on the average return from multiple encounters. But, if so – if that is the rationale for attending to Pareto-superiority – the argument does not address the question of the rationality of the choice *for the single-shot case*.[[12]](#footnote-12)

Going through the restaurant-choice case a bit less abstractly, suppose Jill is thinking through her options in advance and considers whether she has a (decisive) reason to go to Corrinne’s first. Based on the way she ranks the options, she realizes that she does if and only if Jack will be there. So she turns to considering whether Jack will be there. Since she knows him to be rational, she will expect him definitely to be at Corrinne’s just in case he has a decisive reason to be there.[[13]](#footnote-13) However, since their situations and preferences are symmetrical, she will realize that he will have a decisive reason to be at Corrinne’s if and only if he thinks that *she* has a decisive reason to be there. Now, assuming that each has correct beliefs about what the other believes, we’ve come full circle: Jill has a decisive reason to be at Corrinne’s if and only if Jack does, but Jack has a decisive reason if and only if Jill does. So, Jill has a decisive reason to be at Corrinne’s just in case she has a decisive reason to be there. Similarly, Jack has a decisive reason to be at Corrinne’s just in case *he* has a decisive reason to be there.

*These* conclusions might have been reached more quickly, just as a matter of logic, and of course are entirely consistent with neither having a decisive reason for being at Corrinne’s.[[14]](#footnote-14) Parenthetically, we can ask if this is the end of the story. Perhaps not. We can also reach a more qualified conclusion along the following lines: *If* either has a decisive reason for going first to Corrinne’s, it is not adequately captured or expressed by their preference rankings, common beliefs or common rationality in the senses indicated earlier. Since their preference rankings are given, and their common beliefs are correct, this means that if they have a decisive reason, it is the conception of rationality, the conditions of which they satisfy, that stands in need of revision. For the present, I leave that as a suggestion to which we can return later.

# Deciding Where to Go – With Communication

Now, with a better sense of the shape of the problem, let us return to consider the issue with which we began – how, given that Jack and Jill have the beliefs and preferences described, they can reach an agreement. In short, what headway can we make if, to the initial conditions pertaining to their beliefs, preferences and rationality, we add that they can communicate?

For the sake of simplicity, assume that whatever Jack and Jill say to one another is an honest expression of their intentions or beliefs, that one who intends to do something will actually do it, and that both are aware of these facts.

Granting this much will get Jack and Jill through two and a half or three lines of the dialogue given earlier. Remember, that went:

Jack: How about dinner at eight? At Alice’s?  
Jill: Sure, but not Alice’s. What about Bennie’s?  
Jack: The service was really bad the last time I was there ….

In terms of their respective preference rankings, Jack can propose Alice’s and Jill can propose Bennie’s; also, Jack can reject Bennie’s and Jill can reject Alice’s as unacceptable. Each is expressing a conditional intention to go to a favorite restaurant (if the other will be there) and an unconditional intention[[15]](#footnote-15) not to go to a disfavored restaurant (whether or not the other will be there). If what has been stated or proposed so far between them fully captured the relevant preferences, then they would not meet for dinner. They have not yet reached an agreement.

The dialogue continues with Jack proposing that they go to Corrinne’s. Note that, though it may be strongly suggested, Jack has not yet said that he will be at Corrinne’s. He has only asked the question, “What about Corrinne’s?” It appears that that, too, is possible given the way he ranks the various possible outcomes. There is only one outcome he ranks more highly than dinner with Jill at Corrinne’s, namely, dinner with Jill at Alice’s, but he knows that isn’t possible since Jill won’t agree to it. In asking the question, he has expressed only a *conditional* intention to be there, provided that Jill will be. No *unconditional* intention to be at a particular restaurant to the exclusion of the other two[[16]](#footnote-16) has yet been expressed by either of them.

Suppose, though it is no more than suggested by the dialogue, that each has now managed to convey to the other his or her complete preference ranking of the possibilities. They both know what Table I and Table II above would look like. Since (a) their positions are symmetrical and (b) neither has yet expressed an unconditional intention to be in a particular place (much less, have both of them), no agreement on where to meet for dinner has yet been reached. What is needed is some kind of symmetry-breaking. One or the other of them must express an unconditional intention to be at one of the three restaurants.

That is what appears to occur when, in reply to Jack’s proposal that they meet at Corrinne’s, Jill says, “That will do.” What is going on at this point? How does Jill move from a conditional intention to be at one of the restaurants (neither has, prior to this point, expressed any more than this) to saying, unconditionally, that she will be at Corrinne’s? There is a deeper question here: *Is* that what Jill is doing? Is she saying, unconditionally, that she will be at Corrinne’s? She is not saying it in so many words, and the grammar of her actual reply is not sufficient to make it determinate that that is what she means. Surely, she is expressing *some* unconditional intention or other and that the object of the intention – the state of affairs that would, if it were actual, be the fulfillment or carrying out of that intention – is in some sense satisfactory, but it is not clear that the object of the intention is that she be at Corrinne’s. For the present, let us suppose that she is indeed expressing an unconditional intention to be at Corrinne’s. An alternative interpretation will be examined later.

From our earlier characterization, it follows that if Jill says that she unconditionally intends to be at Corrinne’s, she does so intend, and that if she intends to be, she will be. Looking ahead, we can see that once Jack hears her saying that she unconditionally intends to be at Corrinne’s, he will, since he is aware that she is honestly expressing her beliefs and intentions and will do whatever she intends, have sufficient reason to go to Corrinne’s rather than anywhere else. The condition of his conditional intention to go to Corrinne’s will be satisfied. Since it is, he will have a dominance argument in favor of going to Corrinne’s: Once it is given that Jill will be there, he can do better in terms of his preferences by going to Corrinne’s than by going anywhere else.[[17]](#footnote-17)

The crucial question, then, is how Jill arrives at or forms the unconditional intention to be at Corrinne’s. In one sense, it is plain that she can do so. She just *does*. That leaves open, however, the question whether she is *rational* to do so. Here, the problem is that, *until* she forms the intention to go to Corrinne’s, she is in exactly the same situation as in the earlier case in which she was fully informed about the preferences of both herself and of Jack but was unable to communicate with him. In that case, though she *could* go (first) to Corrinne’s, she had no decisive reason to. In fact, she could not have a decisive reason to go to Corrinne’s unless Jack did, and Jack could not have a decisive reason unless she did.

Perhaps, now that she *can* communicate with Jack, she can reason in this way:

If I announce that I intend to be at Corrinne’s, Jack will believe me. If he believes I intend to be at Corrinne’s, he will believe that I will be there. If he believes I will be there, then he will be there. If he will be there, then I have a decisive reason to be there, too. So, I will announce my intention to be there.[[18]](#footnote-18)

Does this get what we want? I think it is not clear that it does. Notice that the initial clause begins, “if I announce that I intend” and the conclusion is “I will announce my intention.” On the face of it, this is an argument in favor of making an announcement, not in favor of having an intention. Does Jill just announce that she intends or does she also intend?

Now, this question is not about whether she *will* intend to be at Corrinne’s once she has made the announcement. Once she does make the announcement, she will have a decisive reason to be there and, knowing this, will intend to be there. However, for the announcement to be honest, it must be true, *when it is made*, that she has the intention. Given the assumptions about her honesty with which we have been working, she will not announce that she has an intention unless she really does. This implies that if she does announce an unconditional intention to go to Corrinne’s, she must already have formed the intention to do so or else she must be forming the intention by making the announcement. For reasons already canvassed, she can’t have formed the intention earlier. Can she be forming the intention in or by making the announcement?

If she is forming the intention in or by making the announcement, then the announcement would have something of the character of a performative. Just as saying, “I promise,” in appropriate circumstances, is not just describing an occasion of promise-making but *is* the making of a promise, so, according to this suggestion, when Jill says, “I intend,” she is not just reporting an intention, but *is*, by that action, intending. The announcement would be either identical to or would partially constitute the intention.

To see if this adequately describes what is going on, we need to look more closely. There is, on one hand, Jill’s overt behavior, uttering the words, “I intend to go to Corrinne’s.” There is, on the other hand, the situation in which she utters the words, including any reasons she has for the utterance. By itself, making the utterance is surely not sufficient for her to be intending to go to Corrinne’s. The words might be uttered in a play, as a joke or even as a lie. If saying the words constitutes her intention to go to Corrinne’s, it is saying them honestly or sincerely in an appropriate situation that does so. Since we are assuming her honesty, in virtue of what about her situation is it appropriate for her to utter the words? Only one answer is plausible: The utterance is appropriate, if it is, because of the expected consequences of that very utterance. When Jill says that she intends to be at Corrinne’s, that will bring it about that Jack is there, and therefore will bring it about that being there also is her best option.

To my mind, this has a suspicious air of boot-strapping. The problem is not that there is some general objection to acting on the basis of expected consequences. That is just ordinary, garden-variety, means-end reasoning. Rather, the problem is that, in ordinary means-end reasoning, the end must be assumed to be *already* aimed at to support the claim that the means are appropriate. If Jill does not already intend, unconditionally, to be at Corrinne’s – she has, instead, only conditional intentions to be there – it is not transparent what she is doing. How can she be adopting a means – making the announcement – to an end that she does not have unless she adopts the means?

*Perhaps* this is possible. We are in difficult territory here, and I am uncertain whether it is reasonable to rule something like this out. It could more comfortably be admitted if there were analogous, but uncontroversial, cases of intention formation having the same structure. I do not know of any such cases. But even if parallels cannot be found, it should not be a decisive consideration against an analysis for the case analyzed to turn out to be *sui generis*. Still, it ­*is* puzzling and I think sufficiently so that we are warranted in trying to find an alternative account.[[19]](#footnote-19)

# Sharing Reasons

Consider a different approach. The problems so far have centered around the fact that Jack makes decisions only for himself and Jill only for herself. Each can take into account what the other may be expected to do under various conditions and adjust decisions accordingly, but each treats the other’s behavior, actual or expected, simply as a parameter relevant to his or her own decision-making.

Suppose instead that Jack and Jill can somehow appoint an agent to make a decision for both of them together. What would they want that agent to be like? First, he should have no stake of his own in where or whether they meet for dinner. That is, he will take into account their preferences but will not favor or be biased toward one over the other, nor will he be concerned to prevent them from satisfying their preferences. Second, he can in fact make a decision for both of them. Once he has reached a decision, he, so to speak, waves a magic wand, and Jack and Jill appear at the appointed restaurant or restaurants. Third, he shouldn’t be able to go *too* badly wrong, so the scope of his ability to reach a decision for them should be limited. He can make a decision for both, but can’t make them stay where they have been placed if either would rather move. Fourth, his decision should be such that it will not be vetoed by either in fact deciding to move.

Now, if Jack and Jill could appoint an agent with those properties, *his* decision would be simple. Only two decisions would not be vetoed, namely, placing both at Corrinne’s and placing Jack at Alice’s and Jill at Bennie’s. Of those two, Corrinne’s is better for both Jack and Jill, so that’s what the agent would choose.

Minus the metaphor of appointing a third-party equipped with a magic wand, I think this is close to what is actually going on. It is not that Jack and Jill are each reasoning about what to do and, fortunately, coming up with concordant decisions. Rather, they are reasoning *together* to reach a single, shared decision, one that is not just some compound of their separate decisions but *their* decision.

What Jack and Jill are doing is – in this case, for limited and temporary purposes – constituting themselves as a single decision-making unit. When Jill says “That will do,” she is taking it to be the case that the two of them are willing to so constitute themselves and is *announcing* the decision that is reasonable from that joint or shared perspective.[[20]](#footnote-20) She is indeed expressing an unconditional intention to be at Corrinne’s and also that being there, with Jack, is satisfactory, but it is not an intention or satisfaction simply from her own perspective; it is the intention and satisfaction of the decision-making unit that she and Jack are constituting themselves as.

1. As I am using the term here, I am not addressing moral or political individualism. The individualism of the individualist picture is a matter of the explanatory concepts used or invoked in accounting for agreement and is most closely akin to theses of methodological individualism in the social sciences. Some extremely interesting discussion of methodological individualism can be found in Nozick 199?. [↑](#footnote-ref-1)
2. It might be that we just can’t – rationally, honestly and non-manipulatively – reach agreements, but that we do reach agreements anyhow because we are not clever enough to see this. I take it that this is a supposition of last resort, adoption of which is to be indefinitely postponed. [↑](#footnote-ref-2)
3. An option, *x*, is strictly preferred to another option, *y*, when it is regarded as being definitely better than *y*. It is weakly preferred when it is regarded as being at least as good as *y*. [↑](#footnote-ref-3)
4. Suppose that Corrinne’s was Jack’s first choice, provided that Jill was also there, and just as good as any other if Jill were not there. Then, Jack would have available to him a dominance argument for going to Corrinne’s. He could do as well as with any other option by going there and might do better (if Jill was there, too). Then, if Jill would rather be at Corrinne’s with Jack than elsewhere without him, Jill, if she knows that going to Corrinne’s is Jack’s dominant choice, would have a reason to be at Corrinne’s, too. [↑](#footnote-ref-4)
5. The numbers represent ordinal rankings, with larger numbers representing higher rankings. The first number in each pair represents the way Jack ranks the outcomes, while the second represents the way that Jill ranks the outcomes. The particular numbers used do not license any inferences as to *how much* one outcome is preferred to another. Jack, e.g., may regard his highest-ranked outcome as much better than his second highest-ranked, but also regard his second highest-ranked as only slightly better than his third choice, though the difference between the numbers used to represent them is the same in both cases. To make comparisons of that sort, cardinal rankings would be needed. [↑](#footnote-ref-5)
6. Also assume that in moving between restaurants, they do not run into one another. The current discussion is meant to abstract from any additional considerations that might be brought into play if they can communicate. [↑](#footnote-ref-6)
7. Remember that the numbers represent ordinal rankings. [↑](#footnote-ref-7)
8. A solution concept for a game recommends some determinate option from among those available to the players. [↑](#footnote-ref-8)
9. Technically, this is *non-strict dominance*. An option is *strictly dominant* as well when every outcome is ranked more highly than its corresponding alternative or alternatives. All strictly dominant options are also non-strictly dominant but not *vice versa*. Strict dominance captures the idea that if some option is better, no matter what others do, then it is, simply, better. Non-strict dominance captures the idea that if some option does at least as well as all other options, no matter what others do, and does better than other options on some assumptions about what others do, then it is, simply, better. [↑](#footnote-ref-9)
10. If an option is dominant for one player, there is at least one response to it on the part of any other player or players such that the set of all options selected is in equilibrium. On the other hand, if there is only a single equilibrium, then one of the options must be dominant for at least one player. However, this breaks down if there is more than one equilibrium. Thus, the availability of a dominant option for any player implies the existence of an equilibrium (that includes that dominant option), but the existence of an equilibrium does not imply the existence of a dominant option for any player. [↑](#footnote-ref-10)
11. If cardinal rankings are admitted (see note 14, *infra*), it can be proved that there is always an equilibrium, at least in mixed strategies. [↑](#footnote-ref-11)
12. I think that McClennen would, perhaps reluctantly, agree with this. [↑](#footnote-ref-12)
13. He may be there without having a decisive reason to be since there is no dominance argument against his going there first. [↑](#footnote-ref-13)
14. Another possibility deserves brief mention. If we had cardinal rather than merely ordinal rankings for the outcomes, then there would be a mixed strategy equilibrium as well. Jack and Jill would each be able to calculate a probability mix over the options available to them such that neither could unilaterally do better in terms of expected cardinal pay-offs. This moves the discussion to a different level where what is compared is not outcomes but probabilistically expected outcomes. Even so, the mixed strategy equilibrium will not be the only one. For example, the selection of Corrinne’s first (with one hundred percent probability) is also a strategy in equilibrium, and one with a higher expected pay-off than any mixed strategy equilibrium. Even more importantly, the mixed strategy does nothing to respond to the intuition that Jack and Jill *do* have a reason for selecting Corrinne’s over either Alice’s or Bennie’s. [↑](#footnote-ref-14)
15. A note on conditional and unconditional intentions: In a conditional intention, the condition is part of its content; in an unconditional intention, it is not. However, to say that an agent has an unconditional intention does not mean that she will execute it, no matter what, or even that she will execute it, provided only that it is in her power to do so. Intentions are formed against an assumed background that consists, roughly, of what is expected whether or not the intention is carried out. If the actual background conditions turn out to be significantly and relevantly different from those assumed, that may constitute a defeating or undermining condition for the execution of the intention. It probably does not matter much for present purposes whether we say that, in such eventualities, the agent is not rational to act on the intention or is rational to alter the intention (rather than carry it out). [↑](#footnote-ref-15)
16. One might say that Jack, by expressing an unconditional intention not to go to Bennie’s, has given voice to an unconditional intention to go to either Alice’s or Corrinne’s and that Jill, by expressing her unconditional opposition to Alice’s, has given voice to an unconditional intention to go to either Bennie’s or Corrinne’s. Those unconditional intentions, plainly, are not enough to solve the problem. [↑](#footnote-ref-16)
17. It is an interesting question whether, once all this has transpired, they have reached an *agreement*. I think they have not, but I will not pursue the matter. [↑](#footnote-ref-17)
18. This way of putting matters owes much to discussion with Joe Mintoff, for which I am grateful. However, I would not want to saddle him with responsibility for it. He might well take exception to the formulation or to my treatment of it. [↑](#footnote-ref-18)
19. Here’s a different tack: Suppose that what Jill unconditionally intends or aims at is not, in the first place, that she be at Corrinne’s, but that she achieve the highest-ranked outcome she can. Since that is her objective, she can see that intending and announcing that she will be at Corrinne’s is sufficient for achieving it (it is not necessary because she and Jack might both go to Corrinne’s first without her having announced any intention), and that none of her other options is sufficient, she should so intend and announce. The intention and announcement are justified as means to something else she unconditionally intends.

    I think that this is unsatisfactory for a different reason – one, however, that might well infect the other proposal, too. We began by talking about how Jack and Jill could reach an agreement. What seems to happen here has more the character of manipulation. Jill has figured out where Jack’s buttons are, and she pushes them. [↑](#footnote-ref-19)
20. She may be *mistaken* in taking that to be the case. Then, Jack, by his reply, could indicate to her that he was not yet willing to take the decision announced to be *their* decision. However, saying that she could be mistaken is not to say that she could not be correct. And if she *is* correct, they are reaching a shared decision. [↑](#footnote-ref-20)