IS AN AGREEMENT AN EXCHANGE OF INTENTIONS?

BY
JOE MINTOFF

Abstract: Margaret Gilbert has argued that an agreement is not exchange of *promises*, since no such exchange plays all the roles she claims are distinctive of agreements. After briefly discussing the notion of intention and the principles governing intentions, I argue that a certain type of exchange of intentions — in which one person forms a conditional intention to act if the other does, and the other forms an unconditional intention to act on the presumption that the first will do what they have said — plays all these roles, and so conclude that an agreement is in fact an exchange of *intentions*.

Is an agreement an exchange of promises?¹ Margaret Gilbert thinks not (Gilbert, 1993). After specifying three criteria any agreement must satisfy, and an exhaustive examination of the particular exchange of promises an agreement might be, she concludes that an agreement is not such an exchange. At the end of her paper (p. 646), she suggests that an agreement is instead a collective or joint decision. But if it is a joint decision, then might it not instead be an exchange of individual decisions? In this paper, I ask whether an agreement is an exchange of decisions or, as I shall put it, of intentions.

An answer to this question would provide not only understanding of an everyday phenomenon, but would also be directly relevant to the fate of social contract theories. A central idea of such theories is that social norms, whether moral or political, are the result of an agreement between pre-social individuals — that is, individuals who are not at the start committed to such norms. There is a familiar problem with such theories: if social norms are the outcome of an agreement, and an agreement is some type of exchange of promises, then what reason do pre-social individuals have to keep their promises, and thus what reason do they have to follow the relevant social norms? The norm that one should keep one's promises, after all, looks like a social norm. But if Gilbert is right, there is a further problem: if social norms are some type of agreement, but agreements are a type of collective decision, then

¹ Thanks are due to the members of the Philosophy Department at Bowling Green State University and the Philosophy Program, RSSS, at the Australian National University, for their helpful comments on earlier drafts. Thanks are due as well to Margaret Gilbert for discussion and comments, and to the University of Newcastle, Australia, for sabbatical leave allowing work on this paper, and the Australian National University for financial assistance during this time.

what sense is there at all to the idea of an agreement between *pre-social* individuals? On this view, agreements are already a form of sociality, and thus are simply unavailable to pre-social individuals. If the argument in this paper is correct, both of these questions have answers.

My argumentative strategy has two stages.² First, I shall specify the roles distinctive of agreements: roles such that it would be plausible to identify agreements with whatever plays those roles. For the sake of argument, I shall take as given (with some re-interpretation, and one addition) the normative roles which Gilbert herself claims to be distinctive of agreements: that they directly generate reasons to perform, that these reasons are generated simultaneously, that they lapse with the default of one party, and that a later agreement can rescind an earlier one. Second, I shall identify some fact or web of facts which play those roles. I start by reviewing Gilbert's claim that various exchanges of promises cannot play these roles, and therefore are not agreements. I continue by arguing that instead a certain kind of exchange of intentions – one in which the first person forms an intention to act conditional on the other's acting, and the second forms an unconditional intention to act on the presumption that the first will do what they have said – *would* play those roles. And this, I shall claim, supports my conjecture that an agreement can be identified with such an intention-exchange.

1. Agreements

The first stage of my argument is to specify the roles distinctive of agreements, using Gilbert's discussion as a base. Suppose, then, we come across the following interaction (p. 627, modified). Rita: "Someone must walk Fido this afternoon, and Tibbles has to be groomed." Peter: "If you'll walk Fido, I'll groom Tibbles." Rita: "OK, then I'll walk Fido." Assuming this really is an agreement (rather than, for example, play-acting), Gilbert claims that it satisfies a number of criteria.

The first condition is the obligation criterion: "the agreement directly generates the relevant performance obligation for each of the parties" (p. 630). An agreement generates for each participant an obligation to perform some act, which Gilbert dubs the *performance obligation*. For example, Peter gains the performance obligation to groom Tibbles, and Rita the obligation to walk Fido. Furthermore – this will be important later on – intuitively these are the only performance obligations this particular agreement generates.

This criterion implies that each party to an agreement is assigned some action *of their own* as their part of the agreement.³ Any agreement between Peter and Rita generates an obligation for each; but *Peter* has an obligation only if

² This way of expressing the strategy owes much to Bratman, 1999, p. 142ff.

³ Thanks to Rob Bass for raising the issues discussed in this paragraph.

there is something *he* is obliged to do, and this is so only if there is something it makes sense to say that *he* does. But we must be careful how we interpret this. For what if Peter and Rita had instead agreed to *walk together*? Then, since walking together is a *joint* activity, we may wonder what Peter or Rita could be *individually* obliged to do. The answer, it seems to me, is straightforward (though otherwise uninformative): Peter is obliged *to walk with Rita*, and she to walk with him. Although the obligation criterion implies that each party to an agreement is assigned an *individual* action (that is, an action which only they may be said to do), it does not imply that the action assigned to them is an *independent* action (that is, an action which they can do without anyone else). While Peter's grooming Tibbles is individual (*he* is grooming Tibbles) and independent (he can do it without Rita), his walking with Rita is individual (*he* is walking with Rita) but nevertheless dependent (he can't do it without Rita). More precisely, then, the obligation criterion implies that each party to an agreement (joint or otherwise) is assigned some individual (but not necessarily independent) action as their part of the agreement.⁴

The second condition Gilbert claims needs to be met by any adequate account of an agreement is the simultaneity criterion: "[t]he obligations that derive from [our sample agreements] are surely simultaneous. That is, neither party has one until the other does" (p. 630). Although Peter speaks before Rita, that does not mean that he receives his performance obligation before she does. It is only once Rita has spoken that each of them, at the same time, gains their appropriate performance obligation. As Gilbert points out, the ability to produce simultaneous obligations is clearly a very useful one, because in many cases "a given party will have no desire to take on some particular obligation before the other party has taken on a corresponding obligation" (p. 630).

The simultaneity condition, though, does not deal very well with e-mail (or any mode of communication which is not simultaneous).⁵ For suppose Peter and Rita's agreement was conducted over e-mail, while they are at their respective work places and, for the most part, focusing on their work. When does Peter get his obligation to groom Tibbles? Certainly not when he sends: "I'll groom Tibbles if you walk Fido" — he does not at that time know what Rita's response will be. Nor is it when Rita sends: "OK, then I'll walk Fido" — Peter *still* doesn't know what Rita's response is, since it hasn't arrived or he hasn't checked his email yet. Most likely, it is when he *reads* this email. But if the simultaneity criterion were correct, then this would be the exact time, and no earlier, Rita gets her obligation to walk Fido. But how can that be? Rita does not know when Peter reads his email, and if Peter was under no obligation at the time Rita sent her email, because he didn't know she had, then surely Rita is

⁴ The issue remains, of course, of how we are to understand joint action, though this is not the issue for the present paper. For some details, see Gilbert, 1990, and Bratman, 1992b.

⁵ Thanks to Marv Belzer for raising this general issue.

under no obligation at the time Peter reads her email, similarly because she doesn't know he has.

This problem can be fixed. The motivation for the simultaneity condition, recall, was that "a given party will have no desire to take on some particular obligation before the other party has taken on a corresponding obligation" (p. 630). Now the best way, it seems to me, of respecting such desires when we are talking about email is to say that Rita gets her obligation when she sends her email "OK, then I'll walk Fido", and that Peter gets his when he reads it. This respects the letter of Peter's desire not to be obligated before Rita, since on this view he is obligated *afterwards*. And it also respects the spirit (but not the letter) of Rita's desire not to be obligated before Peter, for even though Rita knows at the time she sends her e-mail that Peter is *not* at that time obliged (since she knows he hasn't read it yet), she does know at that time that he *will* become obliged, because (presumably) she knows he *will* read it sometime, even though she does not know exactly when. More precisely, then, a given party will have no desire to take on an obligation while there is no *assurance* that the other party has taken on, or will take on, a similar obligation, though this assurance may well be received at different times for different parties. Re-drafting the simultaneity criterion—I will persist with the name—we get: an obligation that derives from our sample agreements is simultaneous with the assurance that the other has, or will have, one. That is, neither party has an obligation until they are assured that the other has, or will have, one.

The third condition Gilbert lays down is the interdependence criterion: "if one party defaults on his performance obligation, the other ceases to have his original performance obligation" (p. 632). If Rita defaults by not taking Fido for a walk, then Peter ceases to have his performance obligation to groom Tibbles. Gilbert points out that this is not to say that Peter might not have some other obligation to groom Tibbles—particularly if Tibbles is in dire need of some grooming—but only to say that the performance obligation generated by his agreement with Rita no longer exists. So Peter may still have an obligation to groom Tibbles, but not because he agreed with Rita to do so.

An important implication of the interdependence and obligation criteria is that if it is settled right from the start that one party to an "agreement" will default, then that interaction is no agreement. To see why, consider the following three forms the interaction between Peter and Rita might take. First, it might be that Rita is to act first and Peter second. As we have seen, if Rita does not walk Fido by the allotted time, then Peter will cease to have an obligation to groom Tibbles. This is a straightforward application of the interdependence criterion. Second, it might be that Peter is to act first and Rita second. Nevertheless, if it becomes settled before Peter is to act that Rita will not walk Fido (she gets on a plane thus rendering it impossible, or she simply declares an intention not to do so, etc.), then again Peter will cease to have his performance obligation. Although this is not a case in which one party has already defaulted, it is clearly in the spirit of our common understanding, and Gilbert's own understanding, of agreements (p. 630). Generalizing, we may say that, according to the interdependence criterion, the time at which Peter would cease

to have an obligation to groom Tibbles is the time at which it is settled that Rita will not walk Fido. Third, however, it might be that it is settled right from the start, as Rita is uttering the words "OK, then I'll walk Fido", that she will not do so. The interdependence criterion therefore implies that the interaction between Peter and Rita fails to generate a performance obligation for Peter, since he had no such obligation to start with, and even if he had, Rita's being settled on not doing her part would have meant it ceased to exist. And thus, in turn, the obligation criterion implies that the interaction between Peter and Rita is not an agreement, since it generates no performance obligations. An implication of the way Gilbert understands agreements, then, is that an insincere agreement is no agreement at all.

Closely related to the interdependence criterion is another Gilbert mentions in passing, but does not explicitly set down. In the course of her discussion of the interdependence criterion, she makes the point that: "an unrescinded agreement generates its obligations for the parties until either these obligations have been fulfilled or one party has defaulted. The concurrence of the parties is required in order to rescind the agreement" (p. 631). This passage suggests there are three conditions under which the original performance obligations will cease to exist. First, when each is fulfilled, about which we will say nothing. Second, when one person defaults, a condition covered by the interdependence criterion. And third, when the agreement is rescinded, a condition for which I suggest we introduce the rescission criterion: if neither party defaults on their agreement, then the concurrence of both parties is necessary and sufficient to rescind the performance obligations generated by the agreement. Although Gilbert herself does not claim concurrence is sufficient for rescission, our common understanding is that what two people have put together does not require anyone else to put asunder.

We have now discussed a number of problems and implications of the criteria Gilbert assumes – and which I shall assume – any adequate account of agreements will explain; but there is one problem remaining. As we have seen, Gilbert states her criteria in terms of performance *obligations*: that is, obligations to do what one has said one will do. She spends little time discussing the nature of these obligations, though at one point she does say that "we speak of agreements as immediately generating *obligations* because a particular *tie* or *bond* between persons is involved. There is a sense in which these are not *moral* obligations" (p. 632, fn. 13, emphasis in original). The problem is that of explaining what type of obligations these are if not moral obligations. Now rather than pursuing this – an issue Gilbert takes as not directly relevant to her concerns – I will focus solely on whether or not an agreement gives us what I will call performance *reasons*: that is, reasons to do what one has said one will do. Making this change, the four criteria become:

(OC) Obligation Criterion (for reasons): The agreement directly generates the relevant performance reason for each of the parties.

(SC) Simultaneity Criterion (for reasons): A reason that derives from an agreement is simultaneous with the assurance that the other has, or will have,

one; that is, neither party has a reason until they are assured that the other has, or will have, one.

(IC) Interdependence Criterion (for reasons): If one party defaults on the agreement, the other ceases to have his original performance reason.

(RC) Rescission Criterion (for reasons): If neither party defaults on their agreement, then the concurrence of both parties is necessary and sufficient to rescind the performance reasons generated by the agreement.

Construing the relevant criteria in these terms has two advantages: instead of talking in terms of the unclear notion of a performance obligation, we can conduct our discussion in terms of the comparatively clearer notion of a reason for action; and such an approach is more congenial to the social contract theorist, who wants a theory of agreement which does not employ moral or quasi-moral notions from the very start. I shall, then, take these criteria to be distinctive of agreements, in that it would be plausible to identify agreements with whatever happened to play those roles.

2. Promise Exchanges

The second stage of my argument is concerned to identify some fact or web of facts which play those roles. I start by reviewing very briefly why Gilbert claims that an exchange of promises cannot play these roles.

In examining various attempts to model agreements as exchanges of promises, Gilbert makes two assumptions about the relationship between promises and obligations. The first is the unstated assumption that every promise generates an obligation, which Gilbert also calls a performance obligation. This is uncontroversial, for it is very difficult to see how an agreement, construed as a promise exchange, would generate its own type of performance obligation without promises themselves generating performance obligations. The second assumption is that: "a given promise stands until its performance obligation has either been fulfilled or violated, unless it is rescinded. Such rescission requires the concurrence of the promisee" (p. 636). This is more controversial, for it may seem that if Peter and Rita exchange promises, it is unfair to insist that Peter is still obliged even if Rita does not do her part. However, whether or not the relevant promises are part of an exchange, it is important to remember that all promises generate obligations to some specific person, and that an obligation to a specific person can be rescinded only by that person.

The second assumption is crucial in Gilbert's demonstration that various forms of promise exchange fail to meet the interdependence criterion. She examines twelve attempts to model agreements as promise-exchanges, and shows that they all fail to satisfy one or other criterion. However, to illustrate this point we need be concerned only with the two models which turn out to be closest in structure to the suggestion I shall make below.

Promise-exchange 7 (p. 641). Peter: "I promise to groom Tibbles, if you walk Fido." Rita walks Fido. This fails to meet the obligation criterion because

Rita fails to incur any performance obligation at all, and, since Peter does end up with a performance obligation, it trivially also fails to meet the simultaneity condition. While this type of exchange presumably happens all the time, it is not really an agreement. Rita needs to make a promise.

Promise-exchange 8 (p. 642). Peter: "I promise to groom Tibbles, if you promise to walk Fido." Rita: "I promise to walk Fido." This fails as well. The chief problem in this case is with the interdependence criterion. For once Peter and Rita have their respective unconditional performance obligations, then they will be released from those obligations not by the failure of the other to do their part, but only by the concurrence of the other. This is, in fact, the central problem with the idea that agreements are an exchange of promises: "[i]t seems that enough has already been said to show that no exchange of promises can model our example agreement, for surely none can meet the interdependence criterion. Once I have a performance obligation through a promise, no one else can eliminate it by breaking some other promise. So no pair of promises can meet the interdependence criterion" (p. 637). After an exhaustive examination of ten other promise-exchanges, Gilbert concludes that none satisfy the relevant criteria.

But if an agreement is not an exchange of promises, might it not instead be an exchange of individual intentions? I shall continue the paper by arguing that a certain kind of exchange of intentions — one in which the first person forms an intention to act conditional on the other's acting, and the second forms an unconditional intention to act on the belief that the first will do what they have said — would play the normative roles distinctive of agreements. In particular, I shall: say something about the notion of intention, and the relevant principles governing that notion; explain in some detail a simple form of intention exchange, and examine the extent to which it plays the normative roles of an agreement; and finally introduce a more complex form of intention exchange, of wider applicability than the simpler ones.

3. Intentions

First, then, to say something about the notion of intention and the principles governing that notion. Of central interest are so-called future directed intentions, intentions held at one time to do something at a later time.

Broadly speaking, there are two types of account of intentions extant in the literature. So-called reductive accounts define an intention as identical to some combination of beliefs, desires, or preferences. For example, a person intends to perform some action A when: they believe their doing A is likely, and they desire to A more than they desire to do anything else (Audi, 1973); or, they believe they will A because their desire to A will so motivate them (Davis, 1984). By contrast, so-called functionalist accounts define an intention to be that state — whatever it is — which plays various causal roles claimed to be peculiar to intention. Thus, for example, functionalist accounts have treated intentions as: reason-centered and conduct-controlling attitudes, subject to constraints of consistency and means-end coherence (Bratman, 1987, pp. 15ff.); initiating,

sustaining, guiding and monitoring motivators of intentional action, which can also be elements of coordinative plans, and which can act as prompters and appropriate terminators of practical reasoning (Mele, 1989, pp. 20-24). Note that, although these two types of accounts are sometimes taken to be exclusive of one another, it is in principle possible that the combination of beliefs, desires or preferences favored by the reductionist plays all the roles favored by the functionalist.

It is not my purpose in this paper to argue for one type of account over the other. Rather, we need only introduce a minimal account of intention, motivate some plausible principles governing that notion, and argue that those principles or ones like them are likely to be acceptable to either a functionalist or reductionist understanding of intentions.

The motivation for a minimal account of intention lies in the fact that there are often strong pragmatic reasons for deliberation about whether to A to have been completed significantly before the time to A (Bratman, 1987, pp. 2ff.). First, our capacity for deliberation is limited. Deliberation takes time and resources, and there are obvious limits to the amount of time and resources we have to expend just before we need to act. Second, separating deliberation and action allows intra- and inter-personal coordination and cooperation. As we will see, if Rita believes that Peter intends to groom Tibbles if she will walk Fido, then (in certain circumstances) she can use that belief to ground an expectation that this is what he will actually do, which will enable her to make a decision of her own. However, since there is no causal action at a temporal distance, there must be *something* about the agent explaining how the earlier completed deliberation leads to the later action. Intentions, it seems plausible to suggest, fill exactly this role. An intention to A may be defined as that state – whatever it is – whose role is to causally link completed deliberation regarding A with one's doing A later on.

The central principle governing intentions relates their formation, reconsideration, and implementation. Intuitively, if Rita rationally forms an intention to walk Fido later (for whatever reason) and rationally does not reconsider it in the meantime, then it would be rational for her later to act as she intends. In general:

(1) *The Intention Principle*. If it is rational for you to form the intention to A at t_1 on the basis of deliberation at t_0 ($t_0 < t_1$) about whether to A at t_1 , and you do, and it is rational from t_0 to t_1 for you not to reconsider this intention, and you do not, and you are able to A at t_1 , then it is rational for you to A at t_1 .

We saw above that promises satisfy the conditions they do (namely, generating obligations, and being rescindable only by the promisee) whether or not they occur in the context of a promise exchange. Similarly, we need to note now that intentions satisfy the conditions they do (the Intention Principle, and those below) whether or not they occur in an intention exchange.

Both the functionalist and reductionist should be happy with something like the Intention Principle. The functionalist's commitment to this principle –

presuming that Bratman fairly represents the functionalists – is exemplified by the conjunction of his *Historical Principle of nondeliberative rationality for the basic case* (Bratman, 1987, p. 80) and his *Intention-Action Principle* (Bratman, 1987, p. 55), though the Intention Principle differs from the first in that it applies to more cases than what Bratman calls the 'basic case', and from the second in that the hypothesis that the intended action needs to have been performed has been replaced by the hypothesis that the agent is able to perform it. The reductionist's commitment follows from two facts. The first fact is that beliefs and desires are themselves subject to principles structurally similar to the Intention Principle, in that a rationally formed belief or desire at t_0 rationally not reconsidered between t_0 and t_1 is rational to hold at t_1 . And so – presuming that Audi fairly represents the reductionists – if an intention to A simply is a belief that doing A is likely and desire to do A greater than a desire to do anything else, then any rationally formed intention to A at t_1 rationally not reconsidered between t_0 and t_1 will also be rational to have at t_1 . The second fact is that any adequate reductive account of intentions will imply that an intention to A at t_1 rationally held at t_1 is rational to act on at t_1 . This will follow since, on Audi's view, it is rational at t_1 that the desire to A at t_1 is greater than a desire to do anything else, and since, in general, it is rational to do what it is rational to desire most to do. The Intention Principle should be common ground.

This principle raises three general questions.

One question raised by the Intention Principle is: under what conditions is it rational to *form* an intention to act, or, as we may put it, to decide to act? The answer to this question is controversial amongst those interested in social contract theories of morals or politics. The orthodox view is that the rationality of an intention to A is determined solely by the question of whether *doing* A would serve one's desires, according to one's beliefs: "[t]he reasons an agent has for intending to do something are basically of the same sort as the reasons an agent has for acting intentionally; they consist of both desires (and other pro attitudes) and beliefs. If someone intends to polish his right shoe, it must be because there is some value he wants to promote by polishing his right shoe" (Davidson, 1985, p. 213). An alternative view is that the rationality of a decision is also partly determined by the question of whether *intending to* A would serve one's desires, according to one's beliefs: "I shall then say that it is rational to choose an action if and only if choosing the action is maximally conducive to the supremely rational aim. So according to Theory U, it is rational for an agent to choose an action if and only if choosing it is maximally conducive to bringing about those outcomes that would maximize her utility" (Gauthier, 1989, p. 184). Fortunately, we do not need to take a stand on this issue, and it will be sufficient to make two assumptions. First, that if Peter's saying "If you'll walk Fido then I'll groom Tibbles" expresses a conditional intention, then it is consistent with what Rita knows about him that this intention is rational. Second, that if Rita's saying "OK, then I'll walk Fido" expresses an intention, then it is consistent with what Peter knows about Rita that this intention is rational.

A second question raised by the Intention Principle is: under what conditions is it rational to *reconsider* an intention? Now, as Bratman points out, the reconsideration of an intention comes in at least three different varieties (Bratman, 1987, pp. 60ff). Reconsideration may be *deliberative*, in that it is the result of deliberating about whether to reconsider, and so presumably the result of considering the reasons for and against reconsidering. It may be *non-reflective*, in that it is the result of certain habits governing when one reconsiders. Or reconsideration may be *policy-based*, in that it is the result of a previously adopted policy to reconsider in such-and-such a circumstance. In this paper, I shall focus only on deliberative reconsideration, and thus on the question: under what conditions is there a reason to reconsider an intention? Furthermore, I shall focus only on cases in which one gets reason to reconsider because of some change in the *beliefs* one based the original decision on, rather than some change in *desires*.

Consider Peter again, but this time outside of the context of his agreement with Rita. As he gets home from work, he can see without Rita's telling him that Tibbles badly needs grooming, although (partly because he has left it so long) doing so would take up a significant amount of time. That Tibbles badly needs grooming is a reason for Peter to decide to do so; though the fact that grooming Tibbles will take a significant amount of time is a reason for him to decide not to do so. On the basis of his deliberations, including these considerations, he decides – quite rationally, let us suppose – to groom Tibbles after dinner tonight. Now these beliefs may or may not persist. But if they do persist, say at least until dinner time, then it is plausible to suppose that they will not at that later time be a reason for Peter to reconsider his intention; though if they do not persist (if, for example, it turns out that grooming Tibbles will take longer than Peter originally thought) then he may have reason to reconsider. This suggests the following principle:

(2) *The Reconsideration Principle*. If you rationally decide at t_0 to A at t_2 because of the reason that p_1 in favor, which overrides the reason that p_2 against, and you continuously intend between t_0 and t_1 ($\forall t_2$) to A at t_2 , then (i) if you continuously rationally believe between t_0 and t_1 that p_1 and that p_2 , then your belief at t_1 that p_2 is no reason for you to reconsider your intention at t_1 to A at t_2 , and (ii) if you lose either belief, then this fact is a reason, other things being equal, for you to reconsider your intention at t_1 to A at t_2 .⁶

⁶ Here is a possible objection to Reconsideration Principle, based on Kavka, 1983. Suppose an eccentric tycoon offers you a million dollars to decide at midnight to drink a toxin tomorrow afternoon. He assures you (and you believe him) that he can tell whether you really have decided, that you get the money tomorrow morning even if you don't end up drinking but so long as you had decided, and that the toxin will only make you sick for a day. Then, it seems possible that (i) you rationally decide at midnight to drink the toxin because of the million dollars, and overriding the fact that after you have been paid it is better for you to change your mind; (ii) you in addition continuously believe these facts to be the case, and yet (iii) after you have been paid, the fact

The key idea underlying this principle is that, since the relevant belief was not at the time of deliberation a reason for you to have made a different decision, then if it persists it will not later be a reason for you to reconsider that very decision, though if it does not persist you will, other things being equal, have reason to reconsider your decision.⁷

But when are "other things being equal"? It is fairly clear, at least to me, that when the beliefs which one based one's decision on remains exactly the same, then they will *never* provide any reason to reconsider; but if some belief does change that does not mean one *always* has reason to reconsider. For reconsideration takes time, and the degree to which the relevant belief has changed might not be very great, or that belief may have played only a minor role in your original deliberations, etc. Suppose, for example, Peter comes to believe that grooming Tibbles will take about 10 minutes longer than he originally believed. If so, he need have no reason to reconsider. If forced to provide at least a rough account of the conditions under which "other things are

that after you have been paid it is better for you to change your mind now is a good reason for you to reconsider your intention to drink.

The appropriate response to this objection depends the correct answer to the following question (which we briefly canvassed above): under what conditions is it rational to form an intention to act? On the one hand, if you believe the orthodox theory – that only considerations relating to the desirability of *the intended action* are relevant in deliberation – then the fact that you get paid one million dollars to intend to drink the toxin does not make the intention formed on this basis rational on the basis of deliberation, and so (i) above is false. On the other hand, if you believe the alternative view – that considerations relating to the desirability of *the intention itself* are also relevant in deliberation – then the fact it would be useful to lose the intention after you have been paid is not a reason to reconsider, and so (ii) above is false. If such pragmatic considerations really were acceptable in deliberation, then they should obey all the constraints that other, more familiar, considerations obey, including the Reconsideration Principle.

⁷ The Reconsideration Principle is consistent with both functionalist and reductionist accounts of intention. On Bratman's view, deliberative reconsideration is itself a deliberative intentional act, and so its rationality will simply fall under a more general account of the rationality of deliberative action. Since, presumably, reasons to *act* satisfy a principle structurally similar to the Reconsideration Principle, then so in particular will reasons to reconsider. And, on Audi's view, reconsideration of an intention will be identical to reconsideration of some desire-belief complex, and since, presumably, reasons to *believe* and reasons to *desire* also satisfy a principle structurally similar to the Reconsideration Principle, then again so will reasons to reconsider. I do not have the space here to justify my assertion that reasons for actions, beliefs and desires do indeed satisfy structurally similar principles.

equal", I would be tempted by the suggestion that a change of belief gives you a reason to reconsider when it seems that you would likely have made a different decision had that change of belief occurred before you made the decision.⁸ Since it may seem to Peter that the extra ten minutes would not have made a difference to his original decision, this change in his beliefs need not be a reason for him to reconsider that decision.

A third question raised by the Intention Principle is: if it is rational to A at some time in virtue of intending to A at that time then what *reason* will one have to A? The answer to this question relates to the fact that combinations of intentions and beliefs can together provide reasons for action. Consider Rita now, again outside of the context of her interaction with Peter. Suppose, for example, that she is on her way home from work, and wants to decide what to do when she gets home. On reflection, her reasons for walking Fido may seem no more or less weighty than those in favor of doing something else. Still, she wants to decide. As it happens, she decides – albeit arbitrarily, and without Peter's telling her – to walk Fido. But where? Now suppose that she quite rationally believes that in order for her to walk Fido, she must (or it is most desirable to) do so in the park. That's the answer. She ought to walk Fido in the park. But if this is what she ought to do, there must now be a reason for her to do this. And this reason could come only from the combination of her intention and belief together, since her obligation to walk Fido in the park arose no sooner and no later than when she decided to walk Fido. Generalizing,

(3) *The Reasons Principle*. If you rationally intend to A and rationally believe that in order to A one must, or it is most desirable to, do B, then this intention-belief pair gives you a reason to B.⁹

Thus, an intention to walk Fido gives Rita a reason to do anything she believes she must do, or it is most desirable to do, in order to walk Fido.¹⁰

⁸ It is important to put some *vague* qualification such as "it *seems* that you would *likely* have made a different decision" because, before you have reconsidered, your information will only be sketchy about whether or not the relevant consideration is important enough. We must demand lower standards of a reason to reconsider an intention to A than of a reason to intend to A or reason not to intend to A. The latter reasons will often only become clearer after you have begun reconsidering.

⁹ I say that this intention-belief pair "gives" one a reason to B, meaning only that if one has this intention-belief pair then there will be some reason, closely related to this intention-belief pair, to B. I put it this way to leave open the question – unimportant in the context of this paper – of whether the intention-belief pair is itself the reason to B, or the conjunction of the objects of this intention-belief pair are the reason to B, or whether something else "closely related" to the pair is the reason.

The Reasons Principle has two noteworthy implications. The first is that if I rationally intend *if p to A* and rationally believe that p, then this intention-belief pair gives me a reason to A, for a belief that p is a reason to believe that in order to A if p one must A – for example, if Peter believes that Rita *has* walked Fido, then this belief will be a reason for him also to believe that in order to do what he agreed, namely groom-Tibbles-if-Rita-walks-Fido, he must groom Tibbles. The second implication is that if I rationally intend to A and rationally believe that A'ing *just consists* in performing the sequence B1, B2, ..., Bn of actions, then this intention-belief pair gives me a reason to perform the sequence of actions B1, B2, ..., Bn. Thus, since Rita intends to walk Fido in the Park, and since she (correctly) believes that doing this just consists in getting the leash, calling Fido, putting the leash on Fido, etc., then her intention and belief give her a reason to perform this sequence of actions.

4. *Intention Exchanges (I)*

So much for intentions and the principles governing them; I have said that an agreement is an exchange of intentions, and now the time has come to make this idea more precise. I start in this section with a simple model of such an exchange, based on strong background assumptions, and finish in the next section with weaker assumptions but a more complex model.

The simple model is the analogue to Gilbert's *Promise-exchange* 8 – Peter: "I promise to groom Tibbles, if you promise to walk Fido." Rita: "I promise to walk Fido." – but avoids its problems precisely because intentions, with their different rescission conditions, replace promises. In particular, in an agreement, one person forms an intention to act conditional on the other's acting, and the other forms an unconditional intention to act on the belief that the first will do what they have said. Thus, I hypothesize that:

Intention-Exchange 1: It is common knowledge that: in saying "If you will walk Fido, I will groom Tibbles", Peter expresses a decision (and so an intention) that

¹⁰ The Reasons Principle is consistent with both functionalist and reductionist accounts of intention. On Bratman's view, "intentions provide special kinds of reasons – *framework reasons* – whose role is to help determine the relevance and admissibility of options. These reasons do not compete with desire-belief reasons, but rather structure the process of weighing such reasons" (Bratman, 1987, p. 34, emphasis in original). Thus an intention to A, and desire-belief deliberation which establishes that B is the most desirable means to A, together provide one with reason to take B. And since, in general, the combination of a desire to A greater than a desire to do anything else and a belief that to do so one must (or it is most desirable to) do such-and-such gives one a reason to do such-and-such, then anything identical to such a pair – that is, intentions, on the Audi's view – will automatically satisfy the Reasons Principle.

if Rita will walk Fido then he will groom Tibbles; in saying "Then I'll walk Fido", Rita expresses a decision (and so an intention) that she will walk Fido, made because she believes that Peter will do as he intends.

This initial model is accompanied by the following background assumptions, which are common knowledge: (B1) Peter and Rita are *practically rational*, in that each decides (and so intends) what they rationally ought to decide; (B2) Peter and Rita are *epistemically rational*, in that each believes what they rationally ought to believe; and, specifically, (B3) each rationally ought to believe that each can be relied upon, in that each ought to believe that if Peter rationally intends to groom Tibbles if Rita will walk Fido then he will do so, and that if Rita rationally intends to walk Fido then she will do so. In short, the initial model assumes it is common knowledge that each is rational and reliable.

Given these background assumptions, it follows that this intention exchange also involves an interrelated set of beliefs. It is common knowledge that the response to Rita's initial statement is that: (1) Peter says "If you'll walk Fido, I'll groom Tibbles." Rita and Peter draw two conclusions from this statement. First, by hypothesis, this statement expresses an intention that if Rita will walk Fido then he will groom Tibbles; and, since Peter's decision is rational, this intention will be as well. Second, since Peter can be relied upon, Rita ought to believe that Peter will do as he intends; and, since Rita is epistemically rational (and in any case by hypothesis), she has this belief. In sum, it is common knowledge that:

- (2a) Peter rationally intends that if Rita will walk Fido, he will groom Tibbles;
- (2b) Rita rationally believes that if she will walk Fido, Peter will groom Tibbles.

It is also common knowledge that the response to Peter's statement is that (3) Rita says: "Then I'll walk Fido." Two more conclusions follow. First, by hypothesis, Rita's statement expresses an intention that she will walk Fido, made because she believes that Peter will act as he intends; and since Rita's decision is rational, then this intention is rational as well. Second, since Rita can be relied upon, Peter ought to believe that Rita will do as she intends; and, since Peter is epistemically rational, he will indeed have this belief. In sum, it is common knowledge that:

- (4a) Rita rationally intends that she will walk Fido;
- (4b) Peter rationally believes that Rita will walk Fido.

It follows, given the above background assumptions, that this interlocking set of intentions and beliefs plays the normative roles which Gilbert herself identified as distinctive of agreement.

First, the account above satisfies the obligation criterion, understood as the claim that the agreement directly generates the relevant performance reason for each of the parties. Consider Peter's case first. As we have seen, by speaking as he does, Peter expresses a rational conditional intention to groom Tibbles if

Rita will walk Fido, and hearing Rita as he does, he comes rationally to believe that Rita will walk Fido; the first noteworthy implication of the Reasons Principle is that this intention-belief pair, since rational, gives Peter a reason to groom Tibbles. Consider Rita's case next. As we have seen, by speaking as she does, Rita expresses a rational unconditional intention to walk Fido; the second noteworthy implication of the Reasons Principle is that this intention, since rational, gives Rita a reason to do whatever she rationally believes walking Fido consists in. So, even if it may not strictly speaking follow that the agreement gives Rita a performance reason to *walk Fido*, it does however give her a performance reason to *get the leash, call Fido, put the leash on Fido, walk him out of the yard, etc., etc.* This means that Rita has all the reason that Peter needs.

Gilbert could respond to this proposal by claiming that too many reasons to act have been generated. To see why I say this, let us examine an objection she has to the following proposed exchange of promises: *Promise-exchange 5* (p. 639). Peter: "I promise this: to groom Tibbles, if you walk Fido." Rita: "I promise this: to walk Fido, if you groom Tibbles." In Gilbert's view, this exchange has serious problems: "In the first place, given that the promises are made one after the other, Peter has an obligation through the promise before Rita does. It is true that his obligation is only to perform a certain act if a certain condition holds. He has, nonetheless, a genuine obligation, from which Rita alone can release him. After he has spoken, she could refuse to make any promise and hence avoid becoming obliged to Peter in any way" (p. 639). In light of this comment, it is easy to see what Gilbert might say to my proposal. Given that Peter forms his intention before Rita forms hers, he ends up with a reason to perform before Rita does. After he has first spoken, Rita can refuse to make any agreement at all, thus leaving only Peter with a reason to perform. My account generates too many performance reasons.

In my view, this is simply not a problem. For in everyday agreements, we distinguish between the *offer* of an agreement by one person and its *acceptance* by the other. Peter's conditional intention is clearly part of his offer of an agreement, an offer which Rita may or may not accept. But this fact hardly counts as an objection to a theory of agreement. In the first place, there is no great harm if Peter's offer remains on the table even though Rita has not (yet) accepted it. For until she does, Peter will not be specifically motivated to do anything at all with regard to the animals. True, he will be motivated to avoid doing things inconsistent with his grooming Tibbles should Rita finally accept the agreement, and this might restrict some of the things he might like to do, but he may well judge that this cost is worth it while there is some chance of coming to an agreement. There are, however, limits. In the second place, after a significant period of Rita's not accepting the offer, it would be perfectly possible and rational for Peter to withdraw the offer, depending on his reasons for making the offer in the first place. Though it is not explicit in our presentation of the case to this point, presumably Peter arrived at his conditional intention because he preferred that he grooms Tibbles and Rita walks Fido to what would happen if he and Rita didn't agree, *and because he believed that if he made this offer then the animals would get the attention they need.* However, if after a significant

period of time Rita still refuses to accept the offer, then Peter should come to realize that his belief in the efficacy of the offer was mistaken, and thus (according to the Reconsideration Principle) that he has reason to reconsider his intention to groom Tibbles if Rita walks Fido, and this may well lead to his withdrawing his offer. In short, far from this being an objection, this feature of the account is a point in its favor.

Second, this account satisfies the simultaneity criterion – that a reason deriving from an agreement is simultaneous with the assurance that the other has, or will have, one; or, in other words, neither party has a reason until they are assured that the other has, or will have, one. Peter gets his reason to groom Tibbles when he comes to believe that Rita will walk Fido. Now given the way the discussion between them has gone, Peter came to this belief because he believed Rita intended to do so. Thus at the time he receives his performance reason Peter is assured that Rita already has one. Rita, by contrast, gets her performance reason before Peter gets his, in particular when she says "Then I'll walk Fido", thereby expressing an intention to so do. Nevertheless, when she says this, Rita is in a position to reason as we did above, that if she says such a thing then Peter will come to believe that she intends to walk Fido and thus to believe that she will do so, and this will mean he will have a performance reason of his own. Thus at the time Rita receives her performance reason she is assured that Peter will have one. Note that this conclusion matches precisely the suggestion above regarding an email exchange between Peter and Rita, that Rita gets her performance obligation when she sends her email and Peter his when he reads that email.

Third, this account also satisfies interdependence criterion – that if one party defaults on the agreement the other ceases to have his original performance reason. Peter's is the simpler case: recall that his reason to groom Tibbles was provided by the combination of his conditional intention to do so if Rita walked Fido and his belief *that Rita will walk Fido*. But if Peter comes to believe that Rita will not walk Fido, then this reason no longer exists. Rita's case is slightly more complex. Recall that she decided to walk Fido because she believed that Peter will act as he intends, or in other words, because she believed the truth of a certain conditional – to wit, *that if she will walk Fido then he will groom Tibbles*. Now Rita still intends to walk Fido, and so believes she will do so (since she can rely on herself), which means that if she comes to believe that Peter will not groom Tibbles, then she will come to believe that the relevant conditional is false. Since she may well not have agreed had she believed right from the start that this conditional was not true, then (again by the Reconsideration Principle) it follows that Rita will have reason to reconsider her intention to walk Fido.

However, it seems that the idea an agreement is an intention exchange fails to meet the final, rescission criterion – that if neither party defaults on their agreement then the concurrence of both parties is necessary and sufficient to rescind the performance reasons generated by the agreement. For the very feature of intentions which makes an intention-exchange satisfy the interdependence criterion is the feature which makes it fail the rescission

criterion. An intention exchange satisfies the interdependence criterion because a person's performance reason depends on their beliefs, and not on their having agreed to rescind the original agreement. But it fails the rescission criterion because the pair of reasons generated by an intention-exchange can be rescinded by the change of belief of one of the parties, independently of the other's concurring. By contrast, the very feature of promises which makes a promise-exchange *fail* the interdependence criterion is the feature which makes it *satisfy* the rescission criterion. A promise-exchange fails the interdependence criterion, as we saw, because the persistence of one person's obligation depends on the other party's not *releasing* them, rather than on the other's party's not *defaulting*. But it satisfies the rescission criterion, because the pair of obligations generated by a promise-exchange will be rescinded when and only when each party releases the other from their obligation, which is to say when and only when each party concurs.

5. *Intention Exchanges (II)*

The simple model would be the correct understanding of agreements, if not for two problems: as we have just seen, it fails to satisfy the rescission criterion; and the background assumption that each ought to believe the other is reliable seems (at least to me) too strong. Because of these problems, it would be implausible to identify an agreement with an interaction of the form of *Intention-Exchange 1*, since such an interaction does not play all the roles agreements play (given its failure to meet the rescission criterion), and since such an interaction is not likely to occur very often (given the strength of the background assumption). These problems can be rectified, however.

Consider first the problem created by the background assumption that the parties to an agreement *ought to believe* each is reliable. We can weaken this assumption, however, by talking about what Rita and Peter can *presume*, rather than believe, about each other. The reason for this has to do with the general point that the inputs to deliberation include not only beliefs but, sometimes, presumptions.¹¹ For example, when I was deliberating about how long to spend in the US during my summer break from teaching, I needed to come up with an estimate of everyday living expenses. To do so, I reviewed the financial records of my previous trip to the US, and presumed for the sake of my deliberations that the expenses of my forthcoming trip would be no more than 15% greater. I deliberated on the basis of this presumption, and other beliefs, and came to the conclusion that I could afford a two-month trip to the US. Reflection on this example suggests the ways in which presumptions are, and are not, like beliefs.

¹¹ The idea that rational deliberation may take as input not only beliefs but also presumptions — Bratman refers to this as the "acceptance" of propositions — is defended in greater detail in Bratman, 1992a.

On the one hand, they are not like beliefs in their "input" conditions: it may be rational to form a presumption that *p* even though one has no good evidence that *p* (or, in the case of "simplifying assumptions", even though one has good evidence that not-*p*). Even though I did not *believe* that my expenses would be no greater than this (I had no evidence one way or the other), such a presumption was nevertheless rational. First, if (as I suspected) I would be able to afford a two-month trip even if expenses were 15% higher, then I would be in a position to make a decision without going through the tedium of discovering *exactly* what my expenses would be. And, second, since my presumption is not fixed in concrete, I would have lost it had I come to see that it was obviously mistaken. This presumption was rational because I managed to cut down on the costs of deliberation, but at minimal risk of making and persisting with a mistake.

On the other hand, presumptions are like beliefs in their "output" conditions: a presumption that *p* behaves just like a belief that *p* in practical and theoretical reasoning. After all, there is no point in making presumptions in order to cut down one's deliberative load unless they really can play the same type of role that beliefs play in deliberation. This means, in particular, that presumptions satisfy both the Reconsideration and Reasons Principles.

We may now replace the background assumption that the parties to an agreement ought to believe each is reliable, with the following significantly weaker assumption:

(B3') Peter and Rita initially presume that for each of them: (i) if it is rational for them to decide initially on their allotted action, then it will continue to be rational for them not to reconsider that decision, and (ii) if it is rational for them to decide initially on their allotted action, then they will continue to be able to perform that action.

There are two things to note about this presumption. First, it has as a corollary that each will rationally presume each is reliable. For the Intention Principle (I), the above presumption (B3'), the assumption that each is practically and epistemically rational (B1, B2) together imply that (B3*) each rationally presumes initially that each can be relied upon, in the sense that if they rationally decide initially on their allotted action, then they will indeed perform that action. In general, making this type of presumption allows a person to ground expectations of the behavior of other people, and thus to facilitate coordination and cooperation between people. Second, it is rational for each initially so to presume. Initially, neither will have any reason to believe that they will come to have reason to reconsider, or to believe that they will come to be unable to perform the allotted action. For if, initially, they did have such reason, then it would not be rational for them initially to decide upon their allotted action—it is irrational to decide to do something you believe you will have reason to reconsider, or which you cannot do. But why presume that they will *continue* not to have reason to reconsider, and *continue* to be able to perform their allotted action? The short answer is that this type of presumption is useful,

since, as we have just seen, it facilitates coordination and cooperation between people. And so long as one remembers that these presumptions are not fixed in concrete, but can be revoked should contrary evidence come one's way, then there should be no objection to them. So Peter and Rita rationally presume each is reliable.

Consider now the problem generated by the fact that it seems impossible for Peter and Rita to rescind their agreement. The problem arises, recall, because the lapsing of a person's reason to perform depends on their own discovery of a failed belief, and not on their having agreed to rescind the original agreement. The problem can be solved, however, by the simple expedient of explicitly making a person's performance reason depend on their not both having agreed to rescind the original agreement. Suppose, then, we come across the following interaction, some time after the original interaction. Rita: "I know we agreed that I would walk Fido and you groom Tibbles, but something has come up." Peter: "I'll forget it, if you forget it." Rita: "OK, then I'll forget it." This seems to be an agreement to rescind their original agreement. The question for someone who wants to reduce agreements to intention-exchanges is how to understand this without presupposing the concept of an agreement.

Here is a suggestion. Suppose there is an utterance — "Blah!", for example, or more realistically "Forget Peter's grooming Tibbles and Rita's walking Fido" — such that (i) in saying "If you will walk Fido, I will groom Tibbles", Peter expresses a decision that if Rita will walk Fido *and if not both have made this utterance*, then he will groom Tibbles; and (ii) that in saying "Then I'll walk Fido", Rita expresses a decision that *if not both have made this utterance*, she will walk Fido, and (iii) all this is common knowledge. In short, it is common knowledge that, in saying what they do, they explicitly make the performance of their allotted task dependent on their not both have made a given utterance (which implies that that utterance *thereby* comes to mean what is usually meant by "Forget Peter's grooming Tibbles and Rita's walking Fido"). If Rita and Peter do this, then they can rescind their agreement simply by both making this utterance; and they can both end up making this utterance simply by agreeing to do so. Of course, there may well be more than one utterance meeting conditions (i) to (iii). In that case, we need instead to suppose that each makes the performance of their allotted task dependent on their not both having uttered *any one of* the given utterances.

We may now define the more complex model which avoids the problems of the simpler one. I hypothesize that:

Intention-Exchange 2: There is a set U of utterances (so-called *rescission utterances*) such that it is common knowledge that: in saying "If you will walk Fido, I will groom Tibbles", Peter expresses a decision (and so an intention) that if Rita will walk Fido and if not both have made a rescission utterance, then he will groom Tibbles; in saying "Then I'll walk Fido", Rita expresses a decision (and so an intention) that if not both have made a rescission utterance, she will walk Fido, made because she presumes that Peter will do as he intends.

This more complex model is accompanied by the following background assumptions, which are common knowledge. As before, (B1) each is practically rational and (B2) epistemically rational. From above, (B3*) each rationally presumes that each can be relied upon. And, in addition, (B4*) each ought to believe neither has made the rescission utterance.

Again, given these background assumptions, it follows that this intention exchange also involves an interrelated set of beliefs. The argument will be similar though more complex than the one above. The following reasoning is common knowledge:

(1) Peter says "If you'll walk Fido, I'll groom Tibbles".

Hence

(2a) Peter rationally intends that if Rita will walk Fido and if not both have made a rescission utterance, he will groom Tibbles; *by* hypothesis (that he intends this), and B1 (practical rationality).

(2b) Rita rationally presumes that if she will walk Fido and if not both have made a rescission utterance, Peter will groom Tibbles; *from* (2a), *by* B3* (presumption of reliability).

But

(3) Each rationally believes that not both have made a rescission utterance; *by* B4* (each ought to believe neither has) and B2 (epistemic rationality).

Hence

(4a) Peter rationally intends that if Rita will walk Fido, he will groom Tibbles; *from* (2a) and (3), *by* the first noteworthy implication of the Reasons Principle.

(4b) Rita rationally presumes that if she will walk Fido, Peter will groom Tibbles; *from* (2b) and (3), *by* Modus Ponens for presumptions.

Now

(5) Rita says: "Then I'll walk Fido"; *by* hypothesis on the presumption in (2b)

Hence

(6a) Rita rationally intends that if not both have made a rescission utterance, she will walk Fido; *by* hypothesis (that she intends this), and B1 (practical rationality).

(6b) Peter rationally presumes that if not both have made a rescission utterance, Rita will walk Fido; *from* (6a), *by* B3* (presumption of reliability).

Hence

(7a) Rita rationally intends that she will walk Fido; *from* (6a) and (3), *by* the first noteworthy implication of the Reasons Principle.

(7b) Peter rationally presumes that Rita will walk Fido; *from* (6b) and (3), *by* Modus Ponens for presumptions.

It again follows, given the above background assumptions, that this second interlocking set of intentions and beliefs plays all the normative roles distinctive of agreement.

This is straightforward for the obligation, simultaneity and interdependence criteria, and we do not need to repeat here the argument provided in the previous section. We need only note the form which the relevant performance

reasons take in this case. In particular, the performance reason Peter has to groom Tibbles is provided by (2a) the intention that if Rita will walk Fido and if not both have made a rescission utterance, then he will groom Tibbles, (3) the belief that not both have made a rescission utterance, and (7b) the presumption that Rita will walk Fido. And the performance reason Rita has to walk Fido is provided by (3) the belief that not both have made a rescission utterance, and (6a) the intention that if not both have made a rescission utterance, she will walk Fido.

This account now satisfies the rescission criteria – that if neither party defaults on their agreement, then the concurrence of both parties – which is to say, their both making a rescission utterance – is necessary and sufficient to rescind the performance reasons generated by the agreement. Consider, for example, Peter's performance reason to groom Tibbles. If both parties concur, then Peter will come to believe that each has made a rescission utterance, and so his performance reason simply disappears. So concurrence is sufficient. And it is necessary as well. For Peter loses his performance reason only if he loses either the intention, belief, or presumption, but, given that he has not and will not default on the agreement, he will not have lost his intention, and, given that Rita has not defaulted on the agreement, he will not have lost the presumption. Therefore, Peter will lose his performance reason only if he loses his belief that not both have made a rescission utterance, and that will be only if both *have* made this utterance. A similar argument – though based on statements (3) and (6a) – demonstrates a similar conclusion applies to Rita's performance reason to walk Fido.

6. Conclusion

In sum, my argument is this. The distinctive roles played by an agreement are specified by the obligation, simultaneity, interdependence, and recession criteria. In particular, an interaction of the form of *Intention-Exchange 2* plays each of those roles. And so, an agreement can be identified with such an intention-exchange. Gilbert was right when she claimed that an agreement is not an exchange of promises. However, that claim is perfectly consistent with its being an exchange of intentions, and, if it is, then there should be no temptation to think that agreements between pre-social individuals are impossible.

School of Liberal Arts
University of Newcastle, Australia

NOTES

REFERENCES

- Audi, Robert. (1973): "Intending," *Journal of Philosophy*, 70, pp. 387-403.
- Bratman, Michael. (1987). *Intentions, Plans and Practical Reason*, (Cambridge, Ma.: Harvard University Press).
- Bratman, Michael. (1992a). "Practical Reasoning and Acceptance in a Context," *Mind*, 101, pp. 1-15.
- Bratman, Michael. (1992b). "Shared Cooperative Activity," *Philosophical Review*, 101, pp. 327-341.
- Bratman, Michael. (1999). "I Intend that We J," in his *Faces of Intention: Selected Essays on Intention and Agency*, (New York: Cambridge University Press).
- Davidson, Donald. (1985). "Replies", in B. Vermazen & M. Hintikka (eds.), *Essays on Davidson*, (Oxford: Clarendon Press).
- Davis, Wayne. (1984). "A Causal Theory of Intending," *American Philosophical Quarterly*, 21, pp. 43-54.
- Gauthier, David. (1989). "In the Neighborhood of the Newcomb Predictor: Reflections on Rationality," *Proceedings of the Aristotelian Society*, 89, pp. 179-194.
- Gilbert, Margaret. (1990). "Walking Together: A Paradigmatic Social Phenomenon," *Midwest Studies in Philosophy*, 15, pp. 1-14.
- Gilbert, Margaret. (1993). "Is an Agreement an Exchange of Promises?" *Journal of Philosophy*, 90, pp. 627-649.
- Kavka, Gregory. (1983). "The Toxin Puzzle," *Analysis*, 43, pp. 33-36.
- Mele, Alfred. (1989). "Intention, Belief, and Intentional Action," *American Philosophical Quarterly*, 26, pp. 19-30.