

# Agreement as Joint Promise

Hanoch Sheinman

Draft – please do not quote w/o permission

Comments are welcome. Please send to sheinman@rice.edu

I propose a largely individualistic general account of agreements according to which an agreement is a joint action, namely an action the parties perform together as one. The action agreement parties jointly perform is promise. Specifically, each party to an agreement promises the other to do her part of the agreement, neither more nor less. Two promises add up to a joint promise (agreement) just when they interdepend in the right way. On my proposal, the interdependence of agreement promises is largely causal-motivational. Specifically, each party's promise causally depends on that party's earlier belief that (she promises later to do her part of the agreement if and only if the other party promises later to do his), and each party is aware of this causal dependence.

## I. Introduction

When you and I open our umbrellas as it starts raining, we open our umbrellas together; when we tango, we tango together as a group—or *as one*. My general hypothesis is that, whenever you and I enter into an agreement, we do something together as one in the same intuitive sense.

Now we can also carry out an agreement together as one. We don't have to, however. My hypothesis is that, whenever we enter into an agreement, we cannot but do something together as one. And to explain agreements is to explain the entering into—not the carrying out of—agreements. When two agents act in such a way as to warrant the image of their performing some action together as one, philosophers of action say that these agents perform that action *jointly*. My hypothesis then is that *an agreement is a joint action*.

### *Agreement as Joint Promise*

Good, but what is it, exactly, that we do together jointly when we enter into an agreement?

According to one view, agreement parties jointly decide to do their respective parts of the agreement.<sup>1</sup> My own hypothesis is that agreement parties jointly *promise* to do their respective parts of the agreement. By this I mean two things.

(1) Each agreement party promises the other to do her part of the agreement, neither more nor less.

(2) These promises interdepend in such a way that warrants the claim that the parties make these promises jointly—that they promise each other together as one.

Let me briefly elaborate on these two claims in order.

#### *(1) Each Agreement Party Promises to Do Her Part*

Many would probably agree with me that a promise is a three-place commitment:

##### *Assumption 1*

When A promises B to do X, A commits to B to do X.

I also believe that an agreement involves two three-place commitments:

##### *Assumption 2*

When A and B enter into an agreement that A is to do X and B is to do Y, A commits to B to X, and B commits to A to do Y.

While I find Assumption 2 hard to resist, it is not obviously beyond reproach. Suppose we enter into an agreement that I am to wash the dishes and you are to take out the trash. Assumption 2 entails that I commit to you to wash the dishes, just like so. If I fail to wash the dishes, I break our agreement *even if you fail to take up the trash*. This judgment might strike you as peculiar. Surely, so goes the objection, if you do not do your part of the agreement, I cannot be committed to do mine.<sup>2</sup> I am wrong to suppose that our agreement commits me to you

to do my part of the agreement, neither more nor less. It commits me to do *less*, namely to do my part of the agreement *if* you do yours. Generally, so goes the suggestion, when A and B enter into an agreement that A is to do X and B is to do Y, A commits to B to (do X if B does Y) and B commits to A to (do Y if A does X). So there is a fundamental asymmetry between the commitments of promises and agreements: agreement commitments are conditional in a way promissory commitments are not.

But I do not see a reason to deny that my failure to wash the dishes breaks our agreement even when you fail to take out the trash. Just because you break the agreement does not mean I cannot break it, too. Again, just because I have a complaint does not mean that you cannot have a complaint, too. Indeed, the suggestion that neither one of us breaks the agreement seems to me to fail to take our agreement seriously. After all, we could enter into a significantly different agreement, namely the agreement that

I am to (wash the dishes if you take out the trash), and you are to (take out the trash if I wash the dishes).

Instead, we chose to enter into the agreement that

I am to (wash the dishes) and you are to (take out the trash).

It seems clear to me that, insofar as this second agreement is concerned, the case in which I wash the dishes and you take out the trash is preferable to the case in which I fail to wash the dishes and you fail to take out the trash. The first is a case of double success; the second, of double failure.

The view that rejects Assumption 2 in favor of the assumption that agreement commitments are mutually conditional makes it difficult to explain the difference between the case in which A chooses to enter into an agreement that A is to do X and B is to do Y, and the case in which they

choose to enter into the rather different agreement that A is to do (X if B does Y) and B is to do (Y if A does X). If Assumption 2 is right, then A's commitment under the first agreement is to do X, and A's commitment under the second agreement is to do (X if B does Y). These are obviously different commitments. But if agreement commitments are mutually conditional, then A's commitment under the first agreement is to do (X if B does Y), and A's commitment under the second agreement is to do [(X if B does Y) if B does (Y if A does X)]. But are these really two different commitments? I do not think so. Surely, A breaks the first agreement if and only if B does Y and A fails to do X. Now, A breaks the second agreement if and only if A fails to (do X if B does Y) and B (does Y if A does X). But the only case in which this conjunction comes out true is the case in which B does Y and A fails to do X.

In the end, it matters not whether our agreement is that (I am to wash the dishes and you are to take out the trash), or that (I am to wash the dishes if you take out the trash, and you are to take out the trash if I wash the dishes). Our commitments under these two seemingly different agreements are indistinguishable. In each case, each party's commitment is to do her chore if the other does his, a commitment each party breaks if and only if she fails to do her chore and the other does his. Our choice to conditionalize our commitments can make no difference to what we are committed to do under the agreement. This strikes me as counterintuitive.

Underlying Assumption 2 is the methodological assumption that the commitments agreement parties make depend on what *their* agreement requires (given its context), not the general account of agreements. What a particular agreement requires is largely up to the parties. In particular, whether one (or each) party commits to perform some action X, or only some action X if the other party performs another action Y, depends on what the parties choose their agreement to require. In principle, the parties can choose to require one (or each) party to

perform some action X if the other performs some other action Y, in which case the commitment of one (or each) party will be conditional. However, the parties can also choose to require each party to perform some action X, just like so (regardless of what the other does), in which case the commitment of each party will be unconditional. Now what any given agreement actually requires the parties to do might require interpretation and depends on the context as well as the words and deeds of the parties. However, if it is established that you and I have in fact entered into an agreement that I am to wash the dishes and you are to take out the trash, then what my agreement commits me to do is wash the dishes. If I fail to do so, then I break the agreement, whether or not you take out the trash. If we both fail to do our chores, we both break the agreement.

Rather than say more on behalf of that assumption, I would like to ask the reader to acquiesce in it for the duration of the paper.

Assumptions 1 and 2 help to explain why I think an agreement is a joint *promise*. If a promise involves a three-place commitment by the promisor to the promisee to do what she promises to do, and an agreement involves two (or more) three-place commitments by each party to the other to do her part of the agreement, then perhaps we should believe that each agreement involves two (or more) promises by each party to the other to do her part of the agreement.

Recall the view that an agreement is a joint decision. This view cannot avail itself of the explanation I have just offered. For a decision only involves a *two*-place commitment relation: the decider commits herself to do what she decides to do, but she does not commit herself to anyone in particular. On the face of it, we cannot explain the multiple commitments of an agreement as multiple decision commitments, only as multiple promissory commitments. Put

differently, agreement commitments seem *relational* in a way promissory commitments are but decision commitments are not.

A defender of the joint decision account of agreements can object that, while a personal decision does not involve the kind of relational commitment essential to promises, a *joint* decision does. The idea here is that the sheer jointness of joint decision can explain the relational nature of agreement commitments. At work here is the theory that *every* joint action (joint walk, joint dance, joint promise) involves relational commitments: each party commits to the other to do its part of that action. Therefore, joint decision involves relational commitments.<sup>3</sup>

But relational commitments to do what, exactly? Let's accept, for argument's sake, the theory that every joint action involves relational commitments to do one's part of that action. This theory tells us that, if you and I jointly walk, then I am committed to you to do my part of our walking, and you are committed to me to do yours. By the same token, if you and I jointly decide, then I am committed to you to do my part of our *deciding*, and you are committed to me to do yours. This theory gives no reason to suppose that either one of us commits to the other to do her part in *carrying out* the joint decision. This would only follow if you and I jointly carried out the decision, which is hardly entailed by our joint decision.

(Of course, just as my personal decision to do X involves a commitment to do X, our joint decision to do X and Y involves commitments to do X and Y. But the theory under consideration has no tendency to show that *these* commitments—as opposed to our commitments to decide to do X and Y—are relational).

The sheer jointness of a joint decision cannot explain the relational nature of agreement commitments. To explain that, I suggest, we have to suppose that the action agreement parties perform jointly is one that would issue in a relational commitment *even if each party performed*

*it individually*. I therefore suggest we reject the joint decision account of agreements in favor of the joint promise one: When we enter into an agreement that I am to do X and you are to do Y, we jointly promise that I do X and you do Y.

## *(2) Agreement Promises Interdepend*

That was just the first part of my hypothesis that an agreement is a joint promise. The second part is the claim that the promises agreement parties give each other interdepend in such a way that warrants the claim that the parties make these promises *jointly*, that they promise each other *together as one*. We can promise each other to do X and Y, respectively, without entering into an agreement that I am to do X and you are to do Y, because one or both of our promises does not depend on the other in the right way. Our promises add up to an agreement if and only if each depends on the other in the right way. To describe the sense in which agreement promises interdepend is to explain what makes them joint promises. To echo Wittgenstein, interdependence is what's left over if we subtract the fact that each agreement party promises from the fact that they jointly promise each other.

## *Promise, Commitment, Obligation*

Now given Assumption 1, promises are commitments, so the notion that an agreement is a joint promise (interdependent promises) entails that an agreement is a joint commitment (interdependent commitments). Now unless "obligation" is just another name for commitment, this last claim does not entail that an agreement is a joint obligation (interdependent obligations).<sup>4</sup> Here I take no view on whether promissory commitments are obligations. Instead, we can simply *assume* that promissory commitments are obligations. Given this assumption, we can say that an agreement is a joint obligation (interdependent obligations).

### *The Challenge: Explaining Interdependence*

I think it is fair to say that everyone agrees that the commitments or obligations of an agreement interdepend in some explanatorily important way. Nevertheless, many seem to reject the notion that *promissory* commitments or obligations can interdepend in that sense. For example, several legal theorists reject the claim that contractual (i.e. legally binding) agreements can be explained as interdependent promises.<sup>5</sup> The complaint seems to have nothing to do with the fact that contractual agreements are legally binding, and everything with the fact that they are *agreements*. Arguably, if contractual agreements cannot be explained as interdependent (contractual) promises, it would be surprising to learn that noncontractual agreements can be explained as interdependent (noncontractual) promises.

Indeed, skepticism about promissory accounts is not confined to contractual agreements; it extends to agreements generally. The most comprehensive rejection of the notion of agreement as joint promise belongs to Margaret Gilbert. She claims that every account of agreements must meet three criteria of adequacy by explaining how every agreement has the following properties:

- (1) The agreement gives each party an obligation to do her part of the agreement (*obligation* criterion).
- (2) It gives these obligations simultaneously (*simultaneity* criterion).
- (3) These obligations interdependent (*interdependence* criterion).

Gilbert rejects all promissory accounts of agreements because she thinks that no such account can meet all three criteria of adequacy:

no promise-exchange, however complex, is capable of simultaneously delivering unconditional, interdependent obligations to the parties (2006: 220; 2000:61; 1996: 328).

More to the point, Gilbert claims that no promissory account of agreements can meet criteria (1)-(3) because *none can meet criterion (3)*:

no pair of promises can meet the interdependence criterion (1996: 221f); cf. also Mintoff (2004: 50).

Now Gilbert has her own account of the interdependence of agreement obligations, and I would like to accept, for argument's sake, her claim that no promissory account of agreement can explain how agreement obligations interdepend in *that* particular sense. All I wish to take from her critique is the notion that agreement obligations interdepend in *some* explanatorily important sense. Not any sense will do, however. The interdependence has to warrant the claim that the parties *jointly* obligate themselves to each other to do their respective parts of the agreement—that they obligate themselves to each other *together as one*. For the proponent of a *promissory* account of agreements, the challenge takes this form:

Describe an explanatorily important sense in which two promises interdepend just when they constitute an agreement, which interdependence warrants the claim that the parties jointly promise each other.

Assuming each promise is a commitment and an obligation, to meet this challenge is also to explain how agreement commitments and obligations interdepend.

#### *Two Assumptions about Joint Action*

This is an essay in applied shared agency theory. Since I seek to explain agreements as a joint promise, my account will inevitably presuppose some working account of joint action. Here I simply want to make two relevant commitments of that account explicit.

(1) The first assumption is *individualism*, understood as the view that joint action consists of interrelated individual actions of personal agents. No appeal can be made, explicitly or by

implication, to a *sui generis*, irreducibly joint action, agent, or commitment. In this, I am basically following Michael Bratman (1999: Part II).<sup>6</sup>

More specifically, to each joint intentional action there correspond at least two individual intentions and actions. If you and I jointly and intentionally walk, then I intend to walk, and walk, and you intend to walk, and walk. If one of us stops walking, then that is the end of our joint walk. And if one of us stops intending to walk, then that is the end of our joint intentional walk.<sup>7</sup>

On the face of it, the claim that agreement promises interdepend presupposes individualism; it presupposes at least two individual promises that can depend on one another. It is just not clear what could possibly interdepend in an irreducibly joint action.<sup>8</sup>

(2) My second assumption is that, while the parties to a joint action may well commit to do their respective parts of that action, they do not necessarily commit to do anything else. In particular, they do not necessary *promise* to do anything other than their respective parts of the joint action.<sup>9</sup> This assumption means that an agreement is a rather special case of joint action. When the action the parties jointly perform here is promise, the parties promise each other—and so commit—to do something other than their respective parts of the joint action (promising): they promise each other to do their respective parts of the agreement. The promises and their attendant commitments then have nothing to do with the fact that the parties perform a joint action; it has everything to do with the fact that the joint action they perform is promise.

On the face of it, my claim that an agreement is a joint promise presupposes my second assumption; it presupposes that at least some joint actions do not involve promises (to do something other than the joint action). If every joint action involved such promissory

commitments, it is just not clear how we could explain the difference between agreement and, say, tango.

I now turn to develop my general account of agreements by progressively modifying proposals. Following Bratman, I will employ the commonsense causal notion *because of*, treat it as unanalyzed, and occasionally use counterfactuals as a proxy. To say of events X and Y that X occurs because Y does is to say that Y is an efficient cause of X. Roughly (but only roughly): X wouldn't occur if Y didn't.<sup>10</sup>

## II. Conditional Promises

The discussion of shared intention in Velleman (2000: ch. 9 and 242f) suggests to me

### *Agreement as Conditional Promises*

A and B enter at t into an agreement that A is to do X and B is to do Y just when, at t

- (1) A promises B to (do X if B promises A to do Y).
- (2) B promises A to (do Y if A promises B to do X).

This account threatens incoherence. By Assumption 2, each agreement party commits to do *its part* of the agreement (neither more *nor less*). In particular, A commits to (do X). But by Assumption 1, each promisor commits to do what's promised (*neither more nor less*). So condition (1) implies that A merely commits to (do X if B promises A to do Y). This is certainly *less* than the commitment to do X.<sup>11</sup>

A slightly different account (cf. Bratman, 1999: 131) tries to avoid the problem by attributing to each party both a simple *and* a conditional promise.

### *Agreement as Conditional and Unconditional Promises*

A and B enter at t into an agreement that A is to do X and B is to do Y just when at t

- (1) (a) A promises B to do X.

- (b) B promises A to do Y.
- (2) (a) A promises B to (do X because B promises A to do Y).
- (b) B promises A to (do Y because A promises B to do X).

Condition (1) seeks to explain how agreements commit each party to do her part (period), and condition (2) seeks to explain how the promises interdepend. But this is an attempt to have the cake and eat it too: the account requires each party to make incompatible promissory commitments. Condition (1) implies that each party promises (and so commits) to do her part of the agreement, period. Condition (2) implies that each party promises (and so commits) to (do her part because the other promises to do his part). Even granting that one can sensibly promise to perform-an-action-as-a-result-of-this-or-that-efficient-cause, the promissory commitments of conditions (1) and (2) conflict with respect to the case in which the other party does not promise to do his part: as a party's condition-1 promise commits her to do her part, her condition-2 promise commits her not to do it.

The lesson I draw is the need to switch to accounts in which agreement promises are unconditional in that each party promises to do her part of the agreement, period.<sup>12</sup>

### III. Unconditional Promises

Consider now

*Agreement as Strictly Interdependent Promises*

A and B enter at t into an agreement that A is to do X and B is to do Y just when,  
at t

- (1) (a) A promises B to do X.
- (b) B promises A to do Y.
- (2) (a) **Because (1b), (1a).**
- (b) **Because (1a), (1b).**

By changing the scope of condition (2), we have just removed the possibility of conflict with condition (1), which is now the sole requirement as to the content of agreement promises (and commitments). The present account finally gets the relation between the content of the parties' promises and the content of their agreement right. It fails for two independent reasons, however. First, it poses a problematic causal loop in which each party's promise both causes, and is caused by, the other party's promise. This seems to run roughshod over causation's temporal structure (the cause precedes the effect). Under condition (1a), A promises B at *t*. Given causation's temporal structure, condition (2a) implies that B promises A *before t*. Yet under condition (1b), B promises A *at t*. Since A and B promise each other simultaneously, neither can promise *because* the other does, under our causal interpretation of "because."

But even apart from this problem, Agreement as Strictly Interdependent Promises fails to capture the *motivational* aspect of agreement-making. Arguably, when A and B enter into an agreement to do X and Y, respectively, A is motivated to promise B to do X by B's promise to do Y. To use a contemporary philosophical idiom, A *takes* B's promise *as A's reason* to promise. Cf. e.g. Setiya's (2007); Bratman (2007). But we cannot hope to explain A's agreement-making motivation without reference to A's beliefs.

The lesson I draw is the need to switch to accounts in which each party's promise depends on that party's belief about the other party's promise. There are really two complementary ideas here, corresponding to the two problems with Agreement as Strictly Interdependent Promises. The negative idea: To say that A's and B's promises interdepend in a way that warrants the joint promise imagery, we need not say that each promises because the other does, in the causal sense of "because;" we needn't posit a causal loop. It is sufficient to say that each party promises because of some belief she has about the other's promise (again, in the causal sense of

“because”). The positive idea: To say that promises interdepend in a way that warrants the joint promise imagery, we must say that each party promises because of some belief she has about the other party’s promise.

#### IV. The Belief Account

We now modify the efficient cause of each promise in condition (2) to get the

##### *Belief Account*

A and B enter at t into an agreement that A is to do X and B is to do Y just when,  
at t

(1) (a) A promises B to do X.

(b) B promises A to do Y.

(2) (a) **Because of A’s belief that (1b), A promises B to do X.**<sup>13</sup>

(b) **Because of B’s belief that (1a), B promises A to do Y.**

One trouble with the Belief Account is that making one’s belief the efficient cause of one’s promise is not enough to capture the motivational, taking-as-one’s-reason aspect of the interdependence we’re after. One can perform an action because of one’s belief, in the causal sense of “because,” without being aware of this causal fact, as when Freud sweeps his inkstand on the floor because of his belief that this will persuade his sister to buy him a new one. See Velleman (2000: 2-9); cf. also Setiya (2007: 33f). Arguably, Freud’s unconscious yet causally efficacious belief moves him to act without quite motivating him. In any event, Freud fails to perform a full-blooded intentional action. Now agreements and promises are not just any actions; they’re speech-acts. It is a platitude of speech-act theory that such acts must be performed intentionally. At least in the central case, one promises, or enters into an agreement, with the intention to do just that. See e.g. Owens (2006: 59f); cf. also Anscombe (2000/1963: § 47); Setiya (2007: 57).

What's missing, I think, is the agent's second-order belief that she acts because of her causally efficacious or *operative* belief. Hence we need to add condition

**(3) (a) A believes (2a).**

**(b) B believes (2b).**

### *Objection*

This addition raises the psychological worry that, together, conditions (2) and (3) render impossible the unconditional promises condition (1) requires. Cf. Velleman (2001). Suppose

(i) A promises B at t to do X;

(ii) Because A believes that B will promise A at t to do Y, A promises B at t to do X;

(iii) A believes (ii).

Put yourself in A's shoes and ask whether your unconditional promising here feels any different from your conditionally promising, namely promising B to (do X if B promises A at t to do Y). If these logically different promises are indiscernible from a first-person perspective, we are back in Agreement as Conditional Promises.

In my intuition, this is a logical distinction with a psychological difference. Something very much like the distinction I'm invoking is commonly acknowledged with respect to intention.<sup>14</sup> I believe that my intention to jump in front of the steering wheel is caused by my beliefs that the starter will be in functioning condition, the battery will not be empty, and gravity will apply. These beliefs are at the background of my intention; intuitively, they do not affect what I intend to do, which is simply to drive off. Now if the car does not start, my beliefs will change, and I will no longer have that intention. But this is phenomenologically different from my intending, when I suspect the battery will be empty, to drive because (or only if) it will not be empty.

I think the same principle applies to my intention to promise—and so to my (intentional) promising. I might believe that, because of my belief that you will promise me at  $t$  to do  $Y$ , I will promise you at  $t$  to do  $X$ . Still what I intentionally promise you is simply to do  $X$ .

### *Matching Cross Promises*

Let us say that the promises  $A$  and  $B$  make when (a)  $A$  promises  $B$  to do  $X$ , (b)  $B$  promises  $A$  to do  $Y$ , but (c)  $A$  and  $B$  do *not* enter into an agreement that  $A$  is to do  $X$  and  $B$  is to do  $Y$  are *matching cross promises*. By describing how agreement promises interdepend, we seek to exclude matching cross promises. The Belief Account begins to do so. Consider an easy case.

*Consecutive Promises.*  $A$  to  $B$  at  $t_1$ : “I promise to do  $X$ .”  $B$  to  $A$  at  $t_2$ : “I promise to do  $Y$ .”

Intuitively,  $A$  promises  $B$  at  $t_1$ , thereby committing herself to do  $X$  later, and  $B$  promises  $A$  at  $t_2$ , thereby committing himself to do  $Y$  later. Intuitively, they do not enter into an agreement. And indeed, the case violates condition (1), which requires the parties to promise simultaneously. The case also violates condition (2). It is not true that  $A$  promises because she believes  $B$  will: at no point before  $A$  promises does she have that belief. Trivially, she would promise apart from it.

We can also have simultaneous matching cross promises, however. Consider

*Simultaneous Promises.* Unbeknownst to  $B$ ,  $A$  emails  $B$  at  $t_1$ : “I promise you my car.” At the same time, and unbeknownst to  $A$ ,  $B$  emails  $A$ : “I promise you \$5k.”  $A$  and  $B$  read these messages at  $t_2$ .

Arguably, promises are made when the promisor communicates her promissory intention to the promisee (when the promisee is led to believe he is being promised). Arguably, then,  $A$  and  $B$  promise each other at  $t_2$ , when they read each other’s messages. They simultaneously promise each other, moreover, exactly what they would simultaneously promise each other at  $t_2$  if they entered into a sale of car agreement at  $t_2$ :  $A$  promises to transfer  $B$  her car, period;  $B$  promises to

pay A \$5k, period. So this time condition (1) is met. Condition (2) remains unmet, however. Take condition (2a). It is true that, by t2, A believes that B presently promises to pay A \$5k (she presently reads B's email), but that belief comes too late to *cause* A to promise at t2. At no point before t2 does A believe that B will promise at t2. Trivially, A would promise at t2 even if she didn't believe that.

### *Accepted Offer Agreements*

Unfortunately, the Belief Account excludes some perfectly good agreements. Consider

*Accepted Offer (Conversation)*. A to B at t1: "I offer you my car for \$5k." B to A at t2: "Accepted."

Intuitively, A and B enter at t2 into a sale of car agreement, yet condition (2) goes unmet. But first let me explain why condition (1) is in fact met here. Appearances to the contrary notwithstanding, accepted offer agreements involve two simultaneous promises.

A's offer can be said to include her *conditional* promise, namely her promise to (transfer the car to B if B promises to pay A \$5k). B's acceptance can be said to include an unconditional promise, namely his promise to pay A \$5k, which promise makes the condition of A's conditional promise obtain, turning it into an unconditional promise to transfer B the car. Hence at t2, B promises A the car (just like that) *and* A promises B the money (just like so). Cf. Bach (1995) for similar analysis.

The bad news for the Belief Account is that it cannot explain the interdependence of the promises in such accepted offer agreements; condition (2) remains unmet. Properly understood, condition (2a) implies that, A's pre-t2 belief that B will promise (at t2) is the efficient cause of A's t2 promise. Alas, at no point before t2 does A believe that B will promise (at t2). It is only

at t2—when B accepts A’s offer—that A learns about B’s promise. *That* belief comes too late to cause A to promise at t2.<sup>15</sup>

The Belief Account’s failure to explain the interdependence of agreement promises becomes clearer still where the offeree learns about his own acceptance after the fact. Consider

*Accepted Offer (Email)*. At t1, A emails B: “I offer you my car for \$5k.” At t2, B reads A’s message. At t3, B emails A: “Accepted.” At t4, A reads B’s message.

Intuitively, A and B enter at t4 into a sale of car agreement, but condition (2) goes unmet. As before, let me first explain why condition (1) is met here. At t1, A sends B an offer message, but A does not quite offer anything until she communicates her message to B at t2 (when B reads it). At t3, B sends A an acceptance message, but B does not quite accept anything until B communicates his message to A at t4 (when A reads it). As before, we can think of A’s offer as a conditional promise, namely a promise to (transfer the car to B if B promises A \$5k), and of B’s acceptance as an unconditional promise, namely a promise to pay A \$5k, which renders A’s promise unconditional by making its condition obtain. That is why we can plausibly say that, at t4, B promises A the money, just like that, and A promises B the car, just like so.

As before, however, condition (2) remains unmet. Properly understood, condition (2b) requires that, B’s pre-t4 belief that A will promise (at t4) is the efficient cause of B’s t4 promise. But for all we know, it is not until *after* t4 that B learns that A has read his acceptance message! (That’s right: B can accept A’s offer before he believes he does, in much the same way B can hit his shooting target before retrieving it). In any event, B’s t4 promise is not caused by his earlier belief. Roughly, B would promise the money at t4 even if he did not believe earlier that A will promise the car.

The Belief Account is too demanding. Often, at least one agreement party does not believe the other will promise until they both do. That party cannot promise because of his belief that the other will. What to do?

### *A Modest Response*

The discussion in Bratman (1999: 154-60; cf. also Black, 2004: 83) suggests an easy way out. While we cannot require each party to believe, before the agreement is concluded, that the other will promise, we can require each party to believe that the other will *probably* promise. Under “normal communicative conditions,” agreement parties are in a position *reliably to predict* each other’s promises. It is this prediction that causes and motivates each to promise the other. Under normal communicative conditions, B in Accepted Offer (Email) reasonably predicts, at t3, that his acceptance message will reach A’s inbox and be read at t4, and so that B will probably accept A’s offer. Arguably, B would not send A the message—and so would not accept A’s offer—apart from that prediction.

The response is too modest to handle the problem. Focus on A for a moment. It is simply not clear that A believes, at any point prior to her t4 promise, that B will probably promise her the money. For all we know, when A sends her offer message, she reasonably believes that B is much more likely to reject her offer than to accept it. And she might well persist in this belief until she learns about B’s highly unlikely acceptance. Clearly, the sheer fact that B’s acceptance was unlikely does not prevent it from turning A’s offer into an agreement. By the time A learns about B’s acceptance, both parties already promise, by which point the beliefs she forms are too late to cause her to promise.<sup>16</sup>

The lesson I draw is the need to switch to accounts in which the operative belief is *(bi)conditional*.

## V. The Biconditional Belief Account

Consider the

### *Biconditional Belief Account*

A and B enter at t into an agreement that A is to do X and B is to do Y just when, at t

- (1) (a) A promises B to do X.  
(b) B promises A to do Y.
- (2) (a) **Because of A's belief that (A promises B to (do X) iff B promises A to do Y), A promises B to do X.**<sup>17</sup>  
(b) **Because of B's belief that (A promises B to (do X) iff B promises A to do Y), B promises A to do Y.**
- (3) (a) A believes (2a).  
(b) B believes (2b).

Each party's operative belief is not that the other will, or will probably, promise at t to do his part; it is rather this:

- (i) I will not promise at t to do my part if he will not, at the same time, promise to do his;
- (ii) he will not promise at t to do his part if I will not, at the same time, promise to do mine.

Here is the key idea. Each party views her promissory commitment as an indispensable part of a package of which the other party's promissory commitment is another indispensable part, and her viewing her promissory commitment in this way is part of what causes and motivates her to make it. Each agreement party takes the fact that the promises of the agreement will stand or fall together as her reason for promising. From the perspective of each party, the promises and commitments of the agreement are essentially reciprocal.<sup>18</sup> Cf. also Bach (1995: 666).

### *Overcoming Intransitivity*

In the previous section I tried to show how condition (1) is met in accepted offer agreements.<sup>19</sup> What remains to be explained is how the new condition (2) is met in such agreements, how the same biconditional operative belief causes the offeror and the offeree to promise. This explanation faces a problem, which is clearest with respect to condition (2a). Suppose we can show that A's biconditional belief causes her to offer B to enter into the sale of car agreement. Since causation is intransitive, it does not follow that the same belief causes A to *promise* B the car. And yet that is exactly what we need to show.

To solve the problem I suggest we relax our definition of "because of." So far, we have taken it to require efficient causation. But ordinary causation-talk is not restricted to strict causation; it typically extends to *chains* of causation. Following Lewis (1986: ch. 21), we can say that event Z *causally depends* on event X just when there is a chain of strict causal relations leading from X to Z, say when X causes Y and Y causes Z. We now redefine "because of" in terms of causal dependence, which is transitive. To say of events X and Y that X occurs because Y does is to say that X *causally depends* on Y.

We are now in a position to explain how the new condition (2) is met in accepted offer agreements. Consider Accepted Offer (Conversation), and take condition (2a) first. As A offers, she understands that, if B accepts her offer, it will transform into two simultaneous promises, including in particular B's promise to pay A \$5k, and that, if B does not accept her offer, it will *not* transform into any promise, including in particular A's promise to transfer B the car. This twofold belief causes and motivates A's offer. Roughly, if A believed she will be the only one promising at t<sub>2</sub>, she would not bother making her offer, which is partly designed to *exclude* this possibility. And if A believed that B will be the only one promising at t<sub>2</sub>, she would also not

bother making her offer, which is partly designed to *ensure* that B will promise. Since A's offer is clearly an efficient cause of A's promise (B would not have accepted, and so A would not have promised, if A did not offer), a causal chain leads from A's operative belief to A's promise.

Now take condition (2b). As soon as B hears A's offer, he understands that, if he accepts it, thereby promising to pay \$5k, A will simultaneously promise to transfer the car, and otherwise she will not. This twofold belief causes and motivates B to accept (promise). Roughly, if B believed that he will be the only one promising at t2, he would not bother giving his acceptance, which is partly designed to *exclude* this possibility. And if he believed that A will be the only one promising at t2, he would also not bother making his acceptance, which is partly designed to *ensure* that A will promise. A similar story can be told about Accepted Offer (Email).

### *Objection*

It will be objected that this explanation depends on the assumption that the offeror A forms her operative belief (that she promises at t iff the offeree B promises at t) just *before* she offers. Yet A might form that belief *as* she offers. That belief would come too late to cause the offer.

There are two complementary ways of acknowledging the truth in the objection but insisting that the promises of accepted offer agreements interdepend in the way condition (2) requires. The first response emphasizes the fact that offering (or promising) is made up of more basic actions. When A offers in Accepted Offer (Email), for example, she first *writes* an offer message, *sends* it, and *communicates* it to B. It is true that the offeror's biconditional belief need not precede every action the offeror performs in offering (or promising). It must still precede at least one such action, however. In Accepted Offer (Email), for example, A does not necessarily form her operative belief before *writing* her offer message, but she certainly *sends* that message

only after she forms that belief. But A's sending the offer message causes B's acceptance—and A's promise. Hence a causal chain leads from A's belief to her promise.

The complementary response points to the simple fact that, like any action, offering and promising are extended in time.<sup>20</sup> When we say that A offers, accepts, or promises at t, this is just shorthand for saying that A *completes* her offering, accepting, or promising action at t. It is true that A's operative belief does not necessarily precede her entire offering or promising action, and so cannot cause her to *begin* to offer or promise. However, it must still cause and motivate her offering and promising action by *sustaining* it once begun. The offeror can kick off the agreement-making process without first acquiring the operative belief, but neither she nor the offeror can continue and complete that process without that belief's causal-motivational contribution.

## VI. Agreement as Joint Promise

I tend to think the Biconditional Belief Account is sufficiently weak to include all agreements, but I suspect it is not sufficiently strong to exclude some matching cross promises. Consider

*Bogus Reports.* A receives a convincing report that B is about to email A: "I promise to pay you \$5k," which message A will receive by t2. B receives a convincing report that A is about to email B: "I promise to give you my car," which message B will receive by t2. These reports prompt their recipients to reciprocate. At t1, A (B) emails B (A) the message B (A) expects to read at t2. The parties receive and read the messages at t2. The reports were bogus: at the time they were sent and received, neither party intended to promise the other.

I claim: A promises B the car, and B promises A \$5k, at t2, but they do not enter into an agreement. Yet surprisingly, conditions (2)-(3) are met here. When A emails B at t1, A reliably predicts that B will shortly send her a message, which will, as soon as A reads it, transform into a promise to pay her \$5k. But A also reliably predicts that her own message will, as soon as B

reads it, transform into a promise to transfer B the car. So when A emails B, A believes that she will promise B the car at  $t_2$  iff B will promise her the money at  $t_2$ . And this belief seems to cause and motivate A to send the message, and so to promise: condition (2a) is met. Since the case is symmetric, so is condition (2b). Condition (3) also seems to be met. Each party is aware of the causal-motivational role her biconditional belief plays in the giving of her own promise. (Nothing Freudian is going on).

What to do? I suggest we leave the causal condition (2) intact and strengthen the cognitive condition (3). Hence my considered account:

*Agreement as Joint Promise*

A and B enter at  $t$  into an agreement that A is to do X and B is to do Y just when, at  $t$

- (1) (a) A promises B to do X.
- (b) B promises A to do Y.
- (2) (a) Because of A's belief that (A promises B at  $t$  to (do X) iff B promises A at  $t$  to do Y), A promises B to do X.
- (b) Because of B's belief that (A promises B at  $t$  to (do X) iff B promises A at  $t$  to do Y), B promises A to do Y.
- (3) (a) A believes (2).**
- (b) B believes (2).**

The new condition (3) requires each party to be aware not only of the causal-motivational role her biconditional belief plays in her own promising, but also of the causal-motivational role the other party's biconditional belief plays in *his* promising. In Bogus Reports, each party is aware of the report available to her, but neither is aware of the report available to the other. Suffice it to consider condition (3a). At no point does A become aware of B's belief that A will promise the car just when B will promise the money. For all A knows, B's promise to pay has nothing to do with anything B believes about A's promise. This is why their simultaneous matching promises still *cross* each other.

### *Agreement Promises, Commitments, and Obligations*

I believe that Agreement as Joint Promise gives credible content to, and so tends to confirm, the hypotheses I put forward in the introduction. In the first place, it tends to vindicate the notion of agreement as joint promise. Since promises are commitments, it also tends to vindicate the notion of agreement as joint commitment. Finally, on our simplifying assumption that promissory commitments are obligations, Agreement as Joint Promise also tends to confirm the notion of agreement as joint obligation. Agreement as Joint Promise tends to confirm these hypothesis by describing a causal-motivational sense in which the promises, commitments, and obligations of an agreement interdepend, which interdependence warrants the claims that the parties promise, commit, and obligate each other together as one.

### *Extending the Account*

Philosophers of action say that you and I *share an intention* whenever we intend together as one. As we have noted, promises and agreements are *speech-acts* that must be performed intentionally. In the central case, I promise, or enter into an agreement, with the intention to do just that. Agreement as Joint Promise seems to me to suggest the following account.

#### *Shared Intention to Enter into an Agreement*

A and B share at t an intention to enter into an agreement that A is to do X and B is to do Y just when, at t

- (1) (a) A intends to promise B to do X.  
(b) B intends to promise A to do Y.
- (2) (a) Because of A's belief that (A promises B at t to (do X) iff B promises A at t to do Y), A intends to promise B to do X.  
(b) Because of B's belief that (A promises B at t to (do X) iff B promises A at t to do Y), B intends to promise A to do Y.
- (3) (a) A believes (2).  
(b) B believes (2).

When the parties enter into an agreement at  $t$ , they must share at  $t$  an intention-in-action to enter into an agreement (promise). But the parties can also share, at some earlier time  $t-1$ , a prior intention to enter into an agreement (promise) later, which they may or may not carry out.

## VII. Agreement and Shared Understanding

It may be objected that, while I set out to explain the interdependence of agreement promises, my account really explains something else, to wit, the dependence of each party's promise on that party's belief. The answer to *this* objection is simple. I have treated the expression "the interdependence of agreement promises" as a label for the intuitive image of two people promising each other together as one (as opposed to giving each other cross promises). This is just an image in need for explanation. Why assume that the best explanation requires the promises to be caused by, or causally depend on, each other? Indeed what I have been trying to show is precisely that promises that meet the conditions of Agreement as Joint Promise interdepend in some explanatorily important sense. This is not the strict causal sense in which each party's promise depends on that party's belief. But that does not matter, provided only that it is strong enough to cash out the joint promise imagery.

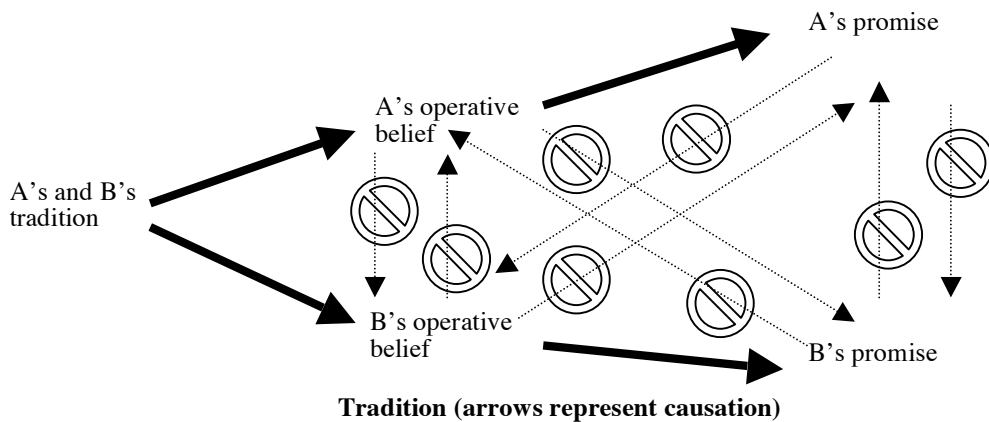
But is it strong enough? We can reinterpret the objection to claim that the sort of interdependence Agreement as Joint Promise describes is too weak to demarcate agreements from matching cross promises. A striking feature of Agreement as Joint Promise is that it admits agreements in which there is neither causation nor causal dependence between the constitutive promises. I have steered clear of strict causal interdependence to avoid the specter of causal loops. But according to the objection, this yields unacceptable results. Consider

*Tradition.* A and B have a tradition whereby every year on March the fourth A emails B: "I promise you my car," and B replies: "I promise to pay you \$X" (the

car's market price). Each year, the parties read the messages on March the fifth. This year is no different, except the market price of A's car is \$5k.

The objection I have in mind claims that (1) the conditions of Agreement as Joint Promise are all met in Tradition, but (2) the parties do not enter into an agreement. I think claim (1) is correct. Condition (1) of Agreement as Joint Promise is met because on March the fifth we have two simple simultaneous promises: A promises B the car; B promises A \$5k. Condition (2) is also met. The parties' tradition is a form of shared agency, namely *shared understanding*—the parties believe together as one. Given that understanding, just before A clicks Send, A believes that B will promise her \$5k, and she will promise B the car, on March the fifth. So A believes that she will promise B the car just when B will promise her \$5k. Since this belief causes A to click Send, and since A's doing so causes her to promise, there is a causal chain leading from A's belief to her promise. We can tell an exactly similar story about B. Condition (3) is also met: given their tradition, A and B are aware of the causal-motivational structure of their promising actions.

There is therefore no escape from assessing claim (2). What prevents the simultaneous promises here from being an agreement, so goes the objection, is that no attitude or action of any party causally depends on any attitude or action *of the other*. To see this, suffice it to show how A's agency is causally independent of B's. Neither A's *promise*, nor A's *intention* to promise, nor A's operative *belief* causally depends, directly or indirectly, on B's simultaneous promise, intention to promise, or operative belief. Instead, what plays a causal role in producing A's operative belief is the parties' *tradition* (shared understanding).



We should bite the bullet and accept that the promises in Tradition constitute an agreement, after all. Notice that this hardly treats shared understandings as agreements; like promises, agreements are essentially intentional in a way shared understandings are not. The agreement or promises in Tradition are every bit as intentional as in all our previous examples. To be sure, the case is somewhat unusual; in the more common case, the operative belief of at least one party causally depends on the other's agency. (In accepted offer agreements, for example, the offeree's operative belief causally depends on the offer). However, when the shared understanding between the parties is robust enough to meet the conditions of Agreement as Joint Promise, I think the image of promises *crossing* each other in the dark is out of place; the appropriate image is that of parties promising together as one.

## VIII. Agential Control

It may be objected, next, that my explanation of accepted offer agreements is inadequate.

Consider Gilbert's leading example of everyday agreements:

*Peter and Rita.* Peter to Rita: "Why don't you walk Fido, and I'll groom Tibbles?" After a few minutes Rita replies: "Fine."

Among the promissory proposals Gilbert explicitly rejects is one that looks very much like an application of Agreement as Joint Promise. According to this proposal, the parties here exchange simple (“unconditional”) promises to walk (groom)—as opposed to conditional promises to walk (groom) *if* the other grooms (walks). I think this proposal is on target: to the extent that we understand the agreement here as an agreement that Rita is to walk Fido and Peter is to groom Tibbles, I think we should say that Rita promises Peter to (walk Fido), and Peter promises Rita to (groom Tibbles). Gilbert’s argument against this proposal is simple:<sup>21</sup>

(1) The agreement in Peter and Rita generates Rita’s walking obligation, and Peter’s grooming obligation, *simultaneously*. While Peter *speaks* first, his statement does not obligate him. He does not get an obligation to groom before she gets an obligation to walk.

(2) The proposal suggests that Peter promises, and gets his obligation, *before* Rita.

I accept (1) but reject (2). Claim (2) tacitly assumes that, if we think of the agreement in Peter and Rita as an exchange of simple promises (to walk, to groom), we must infer that Peter promises, and obligates himself, minutes before Rita does. Rita promises when she says “fine,” but she is the only one promising at that time. This assumption is spurious. More to the point, it is incompatible with the analysis of accepted offer agreements put forward in Sections IV-V. On this analysis, Peter’s promise-making agency only begins when he makes his statement. His statement must be construed not as an unconditional promise to do his part of the agreement, but as an offer to enter into an agreement. That offer contains his merely *conditional* promise to groom Tibbles if Rita promises to walk Fido. It is only when Rita accepts Peter’s offer—thereby giving him an unconditional promise to walk, and making the condition of his earlier promise obtain—that Peter gives Rita an unconditional promise to groom. When Rita accepts Peter’s offer, Rita promises Peter to walk Fido, just like that, and Peter promises Rita to groom Tibbles, just like so.

But just how plausible is this? After all, Gilbert's assumption that Peter does not promise when Rita speaks appears to be supported by the commonsense

*Control*

What one does (or intends) is up to one.

If Peter promises when Rita accepts his offer, Peter's promise is outside his control; it is in Rita's control, for if she does not accept, he does not promise. (Similarly, if Peter intends to promise when Rita accepts his offer, his intention is outside his control; it is in Rita's control).

This objection from control equivocates. On a biting interpretation, action (or intention) requires *exclusive* agential control. This interpretation would prevent Peter from promising, but would also make Control unacceptably demanding. If Peter fires a bullet through a black shoebox, intending to destroy whatever Rita has placed there, and the bullet destroys Tibbles, then I suspect Peter intentionally destroys Tibbles, whose placing in the box was outside his control. Rita makes it the case that Peter intentionally destroys Tibbles.<sup>22</sup> Since what we do or intend is hostage to contingencies beyond our control—including other peoples' actions and intentions—there need be no mystery about Peter's promising (or intending to promise) when Rita accepts his offer. Peter can (intentionally) promise to groom Tibbles without exclusive control just as he can (intentionally) destroy Tibbles without such control. Odd as this might sound at first, Rita can make it the case that Peter promises, and intends to promise, Rita.

On the more plausible interpretation of Control, the agent must have *substantial* control over what she does or intends. But Peter has substantial control over his promissory intention and action. After all, he only loses as much agential control over his promise as he voluntarily surrenders to Rita in his offer. True, by accepting the offer, Rita makes it the case that Peter promises her. But without Peter's offer, Rita would have nothing to accept. Peter would not

promise *via* Rita if he did not make her an offer first (and refrained from revoking it before she accepted). So when Rita accepts, Peter can hardly disavow his promise by denying substantial agential control.

## IX. Conclusion

Recall Gilbert's three criteria of adequacy on theories of agreements from the introduction. Let's assume, for argument's sake, that Agreement as Joint Promise fails to meet these criteria. I wish to close by reconfiguring these criteria and saying how Agreement as Joint Promise meets them.

(At this point I am retracting my working assumption that agreement commitments are obligations. Those who think agreement commitments are obligations are free to substitute obligation-talk for my commitment-talk).

### (1) *Commitment*

An agreement commits each party to the other to do her part of the agreement.

The commitments of an agreement are unconditional in the sense that they commit each party to do whatever the agreement itself requires her to do, neither more nor less. However, this must not be taken suggest that agreement commitments are unconditional in every possible way. In particular, the parties are largely free to make one (or each) of their commitments conditional in the sense that the relevant party is committed to perform some action if the other party performs another action. If the agreement (properly interpreted in context) requires that

A is to do X and B is to do Y

then A gets an unconditional commitment, namely to do A, and B gets an unconditional commitment, namely to do Y. If the agreement (properly interpreted in context) requires that

A is to do X and B is to do (Y if A does X)

then A gets an unconditional commitment, namely to do X, and B gets a conditional commitment, namely to (do Y if A does X). And if the agreement (properly interpreted in context) requires that

A is to do (X if B does Y) and B is to do (Y if A does X)

then A gets a conditional commitment, namely to (do X if B does Y), and B gets a conditional commitment, namely to (do Y if A does X).<sup>23</sup>

Agreement as Joint Promise meets this criterion. It claims that each party promises the other to do its part of the agreement. And given our assumption that promises commit the promisor to do what's promised (exactly), each party commits to the other to do her part of the agreement (exactly).

## (2) *Simultaneity*

The parties to an agreement commit themselves simultaneously.

There is a single point in time at which the agreement-making process concludes. This is the point at which each party commits herself (and so become committed) to the other. As before, it is important to emphasize that this does not mean that agreement commitments are simultaneous in every possible way. In particular, the parties are largely free to make their commitments non-simultaneous in the sense that they are commitments to perform consecutive actions. If the agreement (properly interpreted in context) requires that

A is to do X at t and B is to do Y at t

then A gets a commitment to do X at t, and B gets a commitment to do Y at t. But if the agreement (properly interpreted in context) requires that

A is to do X at t and B is to do Y at t+1

then A get a commitment to do X at t, and B gets a commitment to do Y at t+1 (although they both get these commitments at t-1).

Agreement as Joint Promise meets this criterion. When one party initiates the process, that party does not make a simple promise first; at most, it makes an offer, which will issue in a simple promise if and when accepted. That is when both commitments are made.

The third criterion of adequacy on accounts of agreements should be familiar by now.

### *(3) Interdependence*

The commitments of an agreement interdepend in such a way that warrants the claim that the parties commit to each other jointly, together as one.

The key idea here is that an agreement is a joint commitment; to explain how this is so just is to explain how the commitments of an agreement interdepend. So understood, there is nothing in this criterion to suggest that agreement commitments are mutually conditional in the sense that each party commits to do her part if the other does his.

Agreement as Joint Promise meets this criterion. The commitments of an agreement are promises, and agreement promises interdepend in a causal-motivational way. Each party's promise causally depends on that party's biconditional belief that (she promises just when the other does), and each party is aware of this fact. These causal-motivational and cognitive conditions flesh out an explanatorily important sense in which agreement promissory commitments interdepend.

## References

- Anscombe, Elizabeth. 2000/1963. *Intention*. Cambridge, MA: Harvard UP.
- Bach, Kent. 1995. "Terms of Agreement," *Ethics* 105: 604-12.
- Black, Oliver. 2003. "Independent Promises and the Rescission of Contracts," *Legal Studies* 23: 555-65.
- . 2004. "Agreements, Undertakings, and Practical Reason," *Legal Theory* 10: 77-95.
- Bratman, Michael. 1999. *Faces of Intention*. Cambridge: Cambridge UP.
- . 2007. *Structures of Agency*. New York: Oxford UP.
- de Moor, Anne. 1987. "Are Contracts Promises?" in *Oxford Essays in Jurisprudence: Third Series*, eds. J. Eekelaar and J. Bell. Oxford: Oxford UP.
- Gilbert, Margaret. 1996. *Living Together*. Lanham, MD: Rowman & Littlefield Publishers.
- . 2000. *Sociality and Responsibility*. Lanham, MD: Rowman & Littlefield Publishers.
- . 2006. *A Theory of Political Obligation*. Oxford: Oxford UP.
- . 2009. "Three Dogmas about Promising." Draft on file with the author.
- Lewis, David. 1986. *Philosophical Papers II*. Oxford: Oxford UP.
- Mintoff, Joe. 2004. "Is an Agreement an Exchange of Intentions?" *Pacific Philosophical Quarterly* 85: 44-67.
- Owens, David. 2006. "A Simple Theory of Promises," *Philosophical Review* 115: 51-77.
- Penner, James. 1996. "Voluntary Obligations and the Scope of the Law of Contract," *Legal Theory* 2: 325-57.
- Smith, Stephen. 2004. *Contract Theory*. Oxford: Oxford UP.
- Setiya, Kieran. 2007. *Reason without Rationalism*. Princeton: Princeton UP.
- Velleman, David. 2000. *The Possibility of Practical Reasoning*. Oxford: Oxford UP.
- . 2001. Book review, *Philosophical Quarterly* 51: 119-21.

## Notes

<sup>1</sup> Gilbert (2009), (2006: ch. 10), (1996: chh. 12, 15), (2000: 27f, 59ff).

<sup>2</sup> I do not wish to attribute this objection to anyone. But cf. e.g. Gilbert (1996: 317): “if one party defaults on his performance obligation, the other ceases to have his original performance obligation.” Cf. also Gilbert (2006: 215-19), Mintoff (2004), and Black (2004).

<sup>3</sup> Cf. Gilbert’s claim that joint action consists of *joint commitment*, and “obligations with corresponding entitlements inhere in any joint commitment” (2000: 25, 54-9). “Whenever we understand a shared intention [and so a joint intentional action] to be present we understand the noted obligations and entitlements to be present: the connection is conceptual” (2000: 17).

<sup>4</sup> The term “obligation” is typically reserved to commitments that render contrary action wrongful or blameworthy in some way. Not all commitments seem to be obligations in this sense. We have already noted that a decision commits the decider to do what is decided, but it is far from clear that her failure to do so is wrongful or blameworthy in any way.

<sup>5</sup> Cf. e.g. Smith (2004: 180f): “contracts are agreements rather than promises;” Penner (1996: 326): “the notion of *agreement*, rather than *promise*, is the kind of voluntary undertaking in light of which contracts must be understood;” de Moore (1987: 122): “a foundation of the law of contract must radically abandon the notion of promise.” Cf. also Black (2004, 2003).

<sup>6</sup> Cf. Gilbert’s rejection of “singularism,” according to which joint action is to be explained in terms of “individual human person’s beliefs, desires, goals, commitments, and so on.”

According to Gilbert, joint action is a matter of “joint commitment,” which “stands outside” the “singularist conceptual scheme” (2006: 125f).

<sup>7</sup> Cf. Gilbert rejection of “personal-intention based” accounts: “there could be a shared intention to do such-and-such though none of the participants personally intends to conform their behavior to the shared intention” (2000: 18).

<sup>8</sup> Cf. Gilbert’s claim that joint commitment “does not have parts” (2000: 158). What could possibly interdepend in a partless commitment?

<sup>9</sup> Incidentally, I believe that the parties to a joint action do not necessarily promise anything at all—not even to do their respective parts of the joint action.

<sup>10</sup> This definition will have to be relaxed somewhat in Section V.

<sup>11</sup> The discussion in Velleman suggests making the promises *infinitely* conditional—and therefore self-referential. I do not see how this can make things better.

<sup>12</sup> But as I have stressed in the introduction, this is not a constraint on what the agreement can require each party to do; it is a constraint on what each party promises *given* what the agreement requires her to do. The agreement can require each party to perform some action if the other performs another action. Then each party unconditionally promises a conditional action.

<sup>13</sup> Respect for causation’s temporal structure requires a tense-sensitive interpretation of condition (2): A believes before  $t$  that B will promise at  $t$ .

<sup>14</sup> Cf. e.g. Bratman's distinction between conditionally intending that X if p and intending that X on the assumption that p (1999: 154-60).

<sup>15</sup> Even if the counterfactual "If A didn't believe *at t2* that B promises at t2, A wouldn't promise *at t2*" is true, it cannot track causation. In this respect, it is no different from a backtracking counterfactual. Cf. Lewis (1986: ch. 17). What must be shown is that A promises at t2 because of her *earlier* belief that B *will* promise. The facts of Accepted Offer (Conversation) do not support this claim.

<sup>16</sup> Moreover, perfectly good agreements are concluded under abnormal communicative conditions. If B sends his acceptance by airmail during a strike, and the letter is delivered, B's reliable prediction to the contrary hardly prevents that letter from concluding the agreement.

<sup>17</sup> As before, respect for causation's temporal structure requires a tense-sensitive interpretation of condition (2): A believes before t that (A promises B at t to (do X) iff B promises A at t to do Y). Notice that the biconditional in the content of the operative belief is not counterfactual but material: A believes that her promise to B guarantees B's promise to her and vice versa. This is consistent with the belief that (A *would* promise B to (do X) even if B didn't promise A to do Y).

<sup>18</sup> Notice that while this complicates matters by conditionalizing the operative beliefs of the parties, it also simplifies matters by making these beliefs identical in content.

<sup>19</sup> I will consider an objection in Section VIII.

<sup>20</sup> Here I am taking a leaf from Setiya (2007: 57f).

<sup>21</sup> See Gilbert (1996, esp. 320).

<sup>22</sup> On a *de re* reading, that is. Rita makes it the case that what Peter intentionally destroys is Tibbles, rather than something else (or nothing).

<sup>23</sup> Hence neither breaks the agreement if neither does anything. Some worry that this would render such an agreement motivationally inert. Cf. e.g. Gilbert (1996: 324ff, 329f); Black (2004: 77, 81). This is not necessarily so. If neither party knows for sure that the other will fail to do his part, the only way each can ensure that she keeps the agreement is to do hers. (And we had better assume that the parties are motivated to keep the agreement; without this assumption it is hard to see how any agreement has motivational force).