

Agreement as Joint Promise

Hanoch Sheinman

Rice University

(June 17, 2008)

I propose a general promissory account of agreements according to which an agreement is an instance of *joint intentional action*, something the parties do together as a group. What agreement parties do together as a group is *promise*. Specifically, each party promises the other to do her part of the agreement, *neither more nor less*. Two promises count as a joint promise just when they can be said to *interdepend* in an explanatorily important way. The interdependence of agreement-promises is primarily a matter of the *causal-motivational* role each party's *belief* about the relation between her own promise and that of the other plays in the giving of her own promise.

When you and I open our umbrellas as it starts raining, we open our umbrellas together; when we dance the tango, we tango together *as a group*. One of the things you and I can do together, as a group, is enter into an agreement. Presumably, we can also carry out the agreement together as a group. However, to explain agreements is to explain the *entering into* agreements. My general hypothesis is that, whenever you and I enter into an agreement, we do something together, as a group, in the same intuitive sense we do something together, as a group, whenever we tango.

What do we do together as a group when we enter into an agreement? According to Margaret Gilbert, agreement parties *decide but do not promise*. (2006: ch. 10), (1996: chh. 12, 15), (2000: 27f, 59ff). I think it's the other way around. Suppose we enter into an agreement that I'll do X and you'll do Y. In that case, it seems to me, I commit myself *to you* to do X, in much the same way in which I would commit myself to you to do X if I *simply* promised you to do X. Similarly, you commit yourself *to me* to do Y, in much the same way in which you would commit yourself to me to do Y if you *simply* promised me to do Y. But when I simply *decide* to

do X, I don't commit myself *to you* in particular. Similarly, when you simply *decide* to do Y, you don't commit yourself to me in particular. This strongly suggests that, when we enter into an agreement that I'll do X and you'll do Y, we in fact *promise* each other to perform these actions.

Objection: While a *personal* decision doesn't generate the kind of relational commitment characteristic to promises, a *joint* decision does. Joint intentional action involves a commitment to carry out the joint intention. Gilbert (2006, 2000, 1996). But this doesn't do the trick. When we jointly decide that I'll do X and you'll do Y, we still only *decide* to do X and Y (rather than promise to perform these actions). So even if a joint action involves a commitment to perform that action, a joint decision involves a commitment to *decide* to do X and Y. An agreement, on the other hand, involves a commitment to *do* X and Y (respectively). I conclude that, when we enter into an agreement that I'll do X and you'll do Y, we jointly promise that I'll do X and you'll do Y.

A joint promise is not *simply* a set of promises. Suppose I promise to meet you for lunch tomorrow, and, a year later, you promise to drive me to work tomorrow. Surely, these mutually independent promises don't add up to an agreement that I'll meet for lunch and you'll drive me to work. To amount to such an agreement, some *dependence relation* must hold between our promises. My hypothesis is that this dependence relation is *symmetric*. I endorse

Agreement as Joint (Interdependent) Promises

An agreement consists of (an exchange of) interdependent promises. When we enter into an agreement, we jointly promise each other to perform our respective parts, and our promises count as *joint* in virtue of the fact that they interdepend in some explanatorily importance sense.

Gilbert has recently argued that promissory accounts of agreements are all inadequate precisely because they cannot explain the interdependence essential to agreements. Every

agreement, she claims, has three basic features: the obligations it generates are (1) mutually unconditional, (2) simultaneous, and (3) interdependent. Each of these features represents a *criterion of adequacy* on accounts of agreements. And yet

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

In particular,

“no pair of promises can meet the interdependence criterion.” (1996: 221f).

I will counter this by explaining how promises can interdepend in an agreement.

Two General Themes

My account of agreements will presuppose a working account of joint intentional action. In particular, my working account of joint intentional action resembles Michael Bratman’s account of shared agency (1999, Part II) in two general respects:ⁱ

Individualism. Shared intentional action is to be explained in terms of interrelated individual attitudes and actions of personal agents. No appeal is made, explicitly or by implication, to a *sui generis* social superagent, supermind, superaction, or supercommitment.ⁱⁱ

Nonobligationism. Shared intentional action may, but need not, generate *obligations* to carry out the shared intention, namely reasons that make the failure to do so *wrongful*. (This is compatible with the claim that joint *promises* issue in obligations to carry out the promises; it implies that joint promises issue in obligations if but only if *promises* do).ⁱⁱⁱ

Finally, like 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.

Assumptions

Now it seems to me to be *analytic* of agreements that they *commit* each party to the other to do her part of the agreement, and of promises that they commit the promisor to the promisee to do what's promised. It does not follow that these commitments are *normative-ethical obligations*, namely reasons that make action to the contrary *wrongful* in some normative-ethical way. Most people believe that they do, however. I'm happy to take this commonsense belief on board, and assume, for argument's sake, that agreements issue in (so-called performance) obligations to do what's agreed upon, and promises issue in performance obligations to do what's promised. Other than that, I'll treat the notion of promise as unanalyzed.

I. Conditional Promises

We can think of an agreement as exchanged conditional(-in-content) promises.^{iv} Consider the

Agreement as Conditional Promises

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

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

This account threatens incoherence. It seems analytic that, by entering into an agreement, each party commits to do *its part* of the agreement, *neither less* nor more: A commits to do X; B, to do Y. (And we've assumed that an agreement gives each party a performance obligation to do its part, *neither less* nor more: A gets an obligation to do X; B, an obligation to do Y). But it also seems analytic that, by promising, the promisor commits to do what's promised, *neither more* nor less. (And we've assumed that a promise gives the promisor a performance obligation to do what's promised, *neither more* nor less). So condition (1) implies that A makes a merely conditional commitment (and, under our assumptions, incurs a merely conditional obligation), to (do X because/if B promises A to do Y). This is certainly *less* than the commitment (or obligation) to do X.

II. Unconditional Promises

I claim: A promissory account of agreements must clearly and unambiguously require the parties to make promises that are simple or *unconditional in content*: each party must promise to do her part of the agreement, period. Consider the

Agreement as Strictly Interdependent Promises

A and B enter at *t* into an agreement that A will do X and B will 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).

This account gets the content of the parties' promises (and obligations) right, but fails for two reasons. First, it poses a potentially 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). Condition (1)(a) requires that, at *t*, A promise B to do X. But given causation's temporal structure, condition (2)(a) implies that B promises A to do Y *before* A promises B. Yet condition (1)(b) implies that, at *t*, B promise A to do Y. Since A and B promise each other *simultaneously*, neither of them can promise *because* the other does—at least 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. More accurately: A is motivated by *A's belief* that B promises. Arguably, what causes A to promise isn't B's promise, but rather A's own *belief* about B's promise.

III. Belief-Based Accounts

We take care of both problems by modifying the description of the efficient cause of each promise in condition (2). Consider the

Simple Belief-Based Account

A and B enter at *t* into an agreement that A will do X and B will 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.
(b) Because of B's belief that (1a), B promises A to do Y.

These conditions together imply

- (2&1) (a) A believes that B promises A to do Y.
(b) B believes that A promises B to do X.

There is no suggestion that (2&1) must hold at *t*. On the contrary, each party's *operative* belief—the belief because of which that party promises—must be formed sometime *before* the parties promise. Two ideas underlie the new condition (2).

The negative idea. To say that A's and B's promises “interdepend in an agreement,” we don't need to say that each promises because the other does, in the causal sense of “because.” We don't need to posit a *causal loop* between their promises. It is *sufficient* to say that each promise is given because *its maker believes* the other promises—again, in the causal sense of “because.”

The positive idea. To cash out the interdependence imagery, we *must* say that each party promises because she has some *belief* about the other's promise. For the interdependence in question isn't simply causal but causal-*motivational*. To explain the motivational aspect of that interdependence, we must shift our attention from B's actions to A's *beliefs* about B's actions.^v

Addition and Objection

Unfortunately, making one's belief the efficient cause of one's promise does not *suffice* to capture the motivational aspect of the causal-motivational (inter)dependence we're after. One

can perform an action because of one's belief, in the causal sense of "because," without being *aware* of its causal role, 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.^{vi} Freud's unconscious yet causally efficacious belief moves him to perform an action, but does it *motivate* him? Be that as it may, Freud fails to perform a full-blooded *intentional* action. In its central case, *agreement-making* is a full-blooded intentional action. What's missing, I think, is a second-order belief to the effect that the agent act because of her operative belief. We need to add condition

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

(b) B believes (2b).

But this addition gives rise to a new objection.^{vii} While condition (1) requires each party to make a *simple* (unconditional-in-content) promise, together, conditions (2) and (3) seem to render such simple promises *psychologically* impossible. Suppose that

(i) A promises B to do X;

(ii) A gives that promise because A believes B promises (or will promise) A to do Y;

(iii) A believes causal fact (ii).

Put yourself in A's shoes and ask whether the giving of that simple promise *feels* any different from the giving of a conditional promise to (do X because/if B promises A 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, however, the (logical) distinction is one *with* a (psychological) difference. Consider the case of *intention*, where something very much like the distinction I'm relying on is commonly acknowledged.^{viii} Arguably, I know that my intention to jump in front of the steering wheel is caused by my beliefs, say, that the starter is in functioning condition, the battery not empty, and gravity applies. These assumptions are part of the *background* of my intention; intuitively, they don't affect what I intend to do, which is simply to drive off. Now if the car

doesn't start, and my beliefs change, I will no longer have that intention. But this is phenomenologically different from my conditional-in-content intention to drive if there is gas in the tank when I suspect there isn't.

The same principle applies to my intention to promise—and so to my intentional promising. I might truthfully believe that, because of my belief that you will promise to show up, I now promise you to show up. But what I actually intend to promise you, and intentionally promise you, is simply to show up.

Matching Cross Promises

Call the promises A and B make when A promises B to do X, and B promises A to do Y, but their promises do *not* amount to an agreement that A will do X and B will do Y *matching cross promises*. The Simple Belief-Based Account *begins* to explain why some matching cross promises are not agreements. Consider an easy case.

Consecutive Promises. A to B at t1: "Promise to do X." B to A at t2: "Promise to do Y."

Intuitively, A promises B at t1, thereby committing herself (and, under our assumption, giving herself an obligation) to do X later, and B promises A at t2, thereby committing himself (and, under our assumption, giving himself an obligation) 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). A does not promise because she believes B will promise: when A promises, she doesn't *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 t1: "Promise you my car." At the same time, and unbeknownst to A, B emails A: "Promise you \$5k." A and B read these messages at t2.

It is plausible to assume that promises are made (completed) when the promisor *communicates* her promissory intention to the promisee. It is therefore plausible to maintain that A and B promise each other at t2, when they *read* each other's messages. They simultaneously promise each other, moreover, exactly what they would simultaneously promise each other at t2 *if* they entered into an agreement at t2: A promises (simply) to give B her car; B promises (simply) to pay A \$5k. Moreover, A and B simultaneously commit (and, under our assumption, obligate) themselves to do what they would commit (and, under our assumption, obligate) themselves to do *if* they entered into an agreement at t2: A commits (or obligates) herself (unconditionally) to B to transfer the car to B; B commits (or obligates) himself (unconditionally) to A to pay A \$5k. Nevertheless at no point do the promises add up to an *agreement*. Why?

This time, condition (1) is met, but condition (2) remains unmet. Take condition (2)(a). It is true that, by t2, A believes that B presently promises to pay A \$5k (she presently reads B's promissory email). But A does not form this belief *before* t2. Trivially, A would promise at t2 even if she didn't have that belief.

IV. Offer and Acceptance Cases

The Simple Belief-Based Account is better at excluding matching cross promises than at explaining genuine agreements. Consider

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

Intuitively, A and B enter at t2 into an *agreement* for the sale of A's car to B for \$5k. But unlike condition (1), condition (2) goes unmet here. Let me first explain why I think condition (1) is met. A's offer can be said to include her *conditional-in-content* promise to (transfer the car to B if B promises to pay A \$5k). B's acceptance can be said to *include* a simple (unconditional)

promise to pay A \$5k. Of course, it must include more than that. B's acceptance can *also* be said to make the condition of A's conditional promise obtain, turning it into an unconditional-in-content promise to transfer the car to B. So, at t₂, B promises A the car (just like that) *and* A promises B the money (just like so).

One might be tempted to think that condition (2) is also met. Take condition (2)(b). At t₂, B believes that A promises B the car. If B didn't so believe, this would surely be because he didn't say "accepted." But if so, B would not *accept* A's offer, and so would not promise A the money. Now take condition (2)(a). At t₂, A believes that B promises A the money. If A didn't so believe, this would surely be because she didn't hear B saying "accepted." But if so, B wouldn't *accept* A's offer, wouldn't promise A the money, the condition of A's promise wouldn't obtain, and A wouldn't promise B the car.

Unfortunately, this application of condition (2) fails to respect causation's temporal structure. At most, the story I've just told shows that, if one party didn't believe *at* t₂ that the other also promises at t₂, the former wouldn't promise at t₂. Even if this counterfactual is true, it cannot track *causation*. What must be shown is that each party promises at t₂ because of her earlier belief that the other will promise. The facts of *Accepted Offer (Conversation)* do not support this claim. Apparently, at no point before t₂ does A (B) believe that B (A) promise (or will promise) anything at all. For all we know, neither party learns about the other's promise until it is actually given.

The problem becomes clearer when we turn to the more complicated yet common offer-and-acceptance case in which one party learns about the other's promise *after* their agreement is concluded. Consider

Accepted Offer (Email). At t1, A emails B: “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.

It seems to me that, at t4, A and B enter into an agreement to exchange A’s car for B’s money. As before, condition (1) is met but condition (2) isn’t. Here’s why I think condition (1) is met here. At t1, A sends B an *offer-message*, but A doesn’t quite *offer* anything until she communicates the message to B at t2 (when B reads it). At t3, B sends A an *acceptance-message*, but B doesn’t quite *accept* A’s offer until B communicates the message to A at t4 (when A reads it). As in Accepted Offer (Conversation), we can think of A’s offer as a conditional promise to (transfer the car to B if B promises A \$5k), and of B’s acceptance as involving a (simple) promise to pay A \$5k. What we *cannot* plausibly say, of course, is that this simple promise *all* that B’s acceptance involves. As in Accepted Offer (Conversation), it also makes the condition of A’s conditional promise obtain, rendering it *unconditional*. 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.

It is clear, however, that condition (2) is not met in this case. In particular, condition (2)(b) goes unmet, since it is not the case that B believes at t4 that A presently promises him the car. For all we know, B does not learn that A has read his acceptance-message until later. B can accept A’s offer before he believes he does (just as he can hit his target before he believes he does). So B’s promise is not caused by his belief that A promises (or will promise). Roughly, B would promise the money at t4 even if he didn’t believe at t4 (or earlier) that A promises (or will promise) the car.

The Simple Belief-Based Account is too demanding. Often, at least one agreement party does not believe that the other promises (or will promise) until they both promise. That party cannot promise because of his belief that the other does (or will) promise.

V. Another Belief-Based Account

To fix the problem, we make the operative belief (bi)conditional in content. Consider the

Biconditional Belief Account

A and B enter at *t* into an agreement that A will do X and B will 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 will promise B to (do X) iff B will promise A to do Y) A promises B to do X.
(b) Because of B's belief that (A will promise B to (do X) iff B will promise 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 promises, will promise, or will probably promise, to do his part; rather, it is this:

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

The key idea is this. Each party views her promise as part of a *package*, of which *both* her *and* the other party's promise are indispensable parts, and this is part of what causes and motivates her to promise. Under the present proposal, the parties' operative beliefs are *exactly similar* in content, but there is no suggestion that they share this belief in the robust sense of shared agency theory. They share it only in the sense that they both *have* it.

Before I can explain how condition (2) is met, I need to say something about the relation between a person's offer and her agreement-promise. Suppose A's operative belief causes A to make B an offer, and B accepts that offer. As I've suggested, as soon as B accepts A's offer, A gives B a promise to do A's part of the agreement. But now causation is not a transitive relation. So even if A's offer causes A to promise B, it does not follow that A's operative belief causes A to promise B. However, the relation between A's offer and her promise is not simply causal but

also *constitutive*. It is like the relation between A's trying to do X and her doing X. When A tries to do X and succeeds, A's trying doesn't simply cause A to do X; it also becomes part of A's doing X. Similarly, when A offers B to enter into an agreement and B accepts, A's offer isn't simply an efficient cause of A's agreement-promise; it becomes part of A's promise. Such constitutive relations, I think, warrant an inference of *transferred causation*. Whatever is an efficient cause of your trying can be regarded as an efficient cause of your doing. More to the point: *whatever is an efficient cause of your offer can be regarded as an efficient cause of your agreement-promise*.

We are now in a position to see how condition (2) is met in Accepted Offer (Conversation). Take condition (2)(a) 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 promises, and in particular into A's promise to transfer the car. Intuitively, this twofold belief causes and motivates A's offer. Roughly, if A believed she will be the only one promising, 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, she would also not bother making her offer, which is partly designed to *ensure* that B will promise. Given the principle of transferred causation, the same twofold belief can also be said to cause and motivate A's *promise*.^{ix} Now take condition (2)(b). As soon as B hears A's offer, he understands that, if he accepts it (promises to pay \$5k), A will simultaneously promise to transfer the car, and otherwise she won't. Intuitively, this twofold belief causes and motivates B to accept (promise). Roughly, if B believed that he will be the only one promising, he wouldn't bother giving his acceptance, which is partly designed to *exclude* this possibility. And if he believed that A will be the only one promising, he would also not bother giving 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, whether these stories respect causation's temporal structure depends on the subtle factual assumption that the offeror A forms her operative belief that (she will promise iff the offeree B will promise) just *before* she (A) offers. Yet A might form that belief *as* she offers. If so, that belief cannot *cause* her to offer.

There are two complementary ways of acknowledging the truth in the objection but insisting that agreements in offer and acceptance cases exhibit the interdependence 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, then *sends* it, and finally *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. A's sending the offer-message causes B's acceptance—and A's promise. Arguably, A's sending the offer-message, like her offering, is constitutive of her promising. If so, (an analog of) the principle of transferred causation warrants an inference from the claim that A's belief causes A to send her offer-message to the claim that it causes her to promise.

My complementary response to the objection points to the simple fact that, like any action, offering (or promising) is extended in time. 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 the causal-motivational contribution of that belief.^x

VI. Agreement as Joint Promise

I suspect the Biconditional Belief Account is sufficiently weak to explain agreements, but insufficiently strong to exclude certain cases of matching cross promises. Consider

Bogus Reports. A receives a convincing report that B is about to email A: "Promise to pay \$5k," which email A will receive by t2. B receives a convincing report that A is about to email B: "Promise you my car," which email B will receive by t2. These reports prompt their recipients to reciprocate. At t1, A emails B what B expects to read by t2, and B emails A what A expects to read by t2. The parties receive and read each other's 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 arguably, conditions (2)-(3) are met here. Here's why I think condition (2) is met. 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 the car to B. It is therefore plausible to maintain that, when A emails B, A believes that she will promise B the car iff B will promise her the money. And it is plausible to maintain that this belief causes and motivates A to send the message, and so to promise. Since the case is symmetric, if condition (2)(a) is met, so is condition (2)(b). 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. I suggest we leave condition (2) intact and strengthen condition (3). My considered account:

Agreement as Joint Promise

A and B enter at t into an agreement that A will do X and B will 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 will promise B to (do X) iff B will promise A to do Y) A promises B to do X.
 - (b) Because of B's belief that (A will promise B to (do X) iff B will promise A to do Y), B promises A to do Y.
- (3) (a) A believes (2).
 - (b) B believes (2).

The current 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*. It will suffice to consider condition (3)(a). 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 are still *cross* promises.

Notes

ⁱ For defense, see Sheinman.

ⁱⁱ I say “by implication,” because the familiar disclaimer of appeal to social superagents, minds etc. is beside the point; what individualism bans is ontological commitment. Cf. Gilbert’s claim that joint intentional action is a matter of “joint commitment,” which “stands outside” the “singularist conceptual scheme” of “individual human person’s beliefs, desires, goals, commitments, and so on” (2006: 125f). I don’t know how to take such a claim seriously *without* taking it to countenance a *sui generis* social supercommitment of an irreducibly joint superagent.

ⁱⁱⁱ Cf. Gilbert (2000: 25, 54-9), according to whom joint intentional action consists of *joint commitment*, and “obligations with corresponding entitlements inhere in any joint commitment.” “Whenever we understand a shared intention [and so joint intentional action] to be present we understand the noted obligations and entitlements to be present: the connection is conceptual” (2000: 17). As I’ve already hinted, even if this doctrine were cogent, given Gilbert’s claim that agreements are joint decisions rather than promises, it could not deliver the conclusion that agreements obligate the parties to *do* their respective parts. At most, it would deliver the conclusion that agreements obligate the parties to *decide* to do their respective parts.

^{iv} Cf. Velleman’s account of shared intention (2000: ch. 9; 242f).

^v Causation’s temporal structure requires us to reinterpret A’s operative belief that B promises as the belief that B *will* promise.

^{vi} See Velleman (2000: 2-9).

^{vii} I am drawing here on Velleman (2001).

^{viii} 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).

^{ix} I don’t think my account *depends* on transferred causation. We can also redefine “because of” as *causal dependence*, namely a *chain* of causal relations. Unlike causation, causal dependence is transitive. See Lewis (1986: chh. 21-22). Even if A’s operative belief doesn’t cause her to promise, there is clearly a chain of causal relations going from her belief, through her offer, to her promise.

^x Here I’m taking a leaf from Setiya (2007: 57f).

References

- Bratman, Michael. 1999. *Faces of Intention*. Cambridge: Cambridge University Press.
- Gilbert, Margaret. 2006. *A Theory of Political Obligation*. Oxford: Oxford University Press.
- . 2000. *Sociality and Responsibility*. Lanham, MD: Rowman & Littlefield.
- . 1996. *Living Together*. Lanham, MD: Rowman & Littlefield.
- Lewis, David. 1986. *Philosophical Papers II*. Oxford: Oxford University Press.
- Setiya, Kieran. 2007. *Reason without Rationalism*. Princeton: Princeton University Press.
- Sheinman, Hanoah. “Agreement as Joint Intentional Action: Two Accounts.” Under Review.
- Velleman, David. 2000. *The Possibility of Practical Reasoning*. Oxford: Oxford University Press.
- . 2001. Book review, *Philosophical Quarterly* 51: 119-21.